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.

We thank the reviewer for their thoughtful and constructive comments. Our response follows:

Specific Comments

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

Thanks, we changed "grid" to "grid box" whenever it means a cell/pixel throughout the paper.

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.

Thanks, we included both studies, but it is also important to acknowledge that the downscaling methods heavily rely on the performance of CTMs. In order to characterize the errors associated with the models using measurements, we will need to deal with the spatial heterogeneity issue down the road. We do not think there is a shortcut way to avoid this fundamental problem.

"It is because of this reason that several validation studies resorted to downscaling their relatively coarse satellite observations using high-resolution chemical transport models so that they could compare satellites to spatially finer datasets such as in-situ measurements [Kim et al., 2018; Choi et al., 2020]. Nonetheless, their results largely arise from modeling experiments which might be biased."

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

Thanks for the precise comment. Yes, we boldfaced it, and added:

"Under this condition, the vector of ***h*** becomes scalar (*h* = |***h***|)."

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

Addressed above.

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.

Sorry for the confusion, only c0 is a constant. We removed a and b and explicitly mentioned:

"where *a* and *b* are fitting parameters."

We also removed this part to remove further confusion:

~~In case of two samples, the semivariogram might be explained by a line with no offsets (i.e.,~~ $\cancel{\gamma(h) = a_o h}$ ~~or a constant function~~ $\cancel{(\gamma(h) = b_o)}$.

Line 140, Eq. (7): Does Z_o means $Z(x_0)$ (true value at x_0)? Please define x_0 and Z_o.

We clarified it in the text:

", x_o is the location of estimation, x_j is the location of samples, , and $Z(x_j)$ is point data (i.e., samples)."

,and:

where Z_0 is point observations ($Z_o = Z(x_j)\,, j = 1,2,\ldots,n$),

Line 145, Eq. (9): Do γ_{j1j2}, γ_{j1o}, γ_{oo} mean $\gamma(x_{j1}-x_{j2})$, $\gamma(x_{j1}-x_0)$, and $\gamma(x_0-x_0)$, respectively? Please define them explicitly in the text.

Thanks, yes, we clarified it:

where γ_{j1j2} is the spatial covariance between the point observations and γ_{j10} is the spatial covariance of between the observations and the estimation point. The spatial covariance is modeled by a semivariogram.

Overall Sect.2.1: There are too many subscripts $_0$ 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_0. l_0 is a constant weight. Z_o is (probably) the true value at a point x_0. Some of these subscripts are relevant but others might not. Removing unnecessary 0 subscripts may help readers understand better.

We are grateful for this precise comment. We got rid of all "o" subscripts and replaced them with "0". We redefined some of the parameters to remove any confusion.

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

This comment is a bit unclear to us. From 200 samples, we can create 19900 paired values (i.e., $\binom{200}{2}$), the paired values were binned to 100 evenly binned distances depending on the min/max of distances.

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.

The gaussian function fit to the semivariograms converges to a fixed value (it is a half-bell shape). As for C2, the range tends to be large (~95) making the half-bell converges to a fixed value in much further values beyond the observed distances.

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.

We agree with the reviewer that targeting a relative error might be more suitable if the goal is to better represent the shape rather than the absolute values. But it is entirely subjective and the primary goal of kriging is to accurately reconstruct the absolute values of Z given discrete samples. Therefore we decided to leave the figures alone.

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

We replaced it with the suggestion.

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.

The reason why we did not use the super Gaussian in the theoretical experiment was to simplify the analysis. We now mentioned this: "For simplicity, we consider $S = \frac{1}{m^2} J_{m,m}$."

Moreover, the super Gaussian function described in Sun et al. [2018] is dependent on the viewing geometry of a real sensor. The utilization of this slit function is more relevant for real-world experiments. The idea behind experimenting with theoretical cases was to show the significance of the problem for simplified scenarios. We now have included an example of the spatial response function for OMI-Pandora comparison:

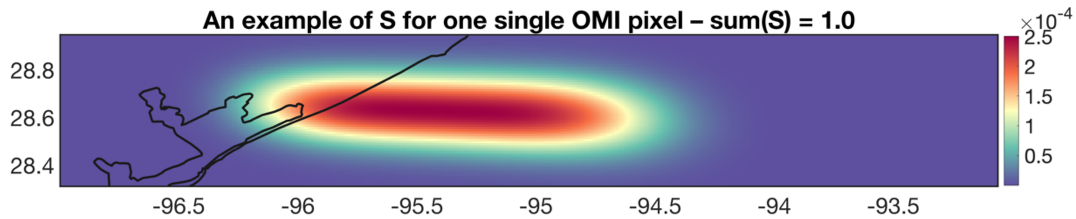


Figure S2. An example of the super Gaussian spatial response function described in Sun et al. [2018] for a given pixel over the region of interest.

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.

Absolutely right, we mentioned it here:

"For simplicity, we consider $S = \frac{1}{m^2} J_{m,m}$; this spatial response function results in averaging the values in the grid boxes"

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.

Yes, we added "where a superscript of 2 denotes squaring."

Regarding the notation, the equations (13,14) are a very standard presentation of a convolution process which can be seen in many different image/signal processing books.

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.

Sure, we now have added the synthetic measurements in the supplementary material (Figure S1).

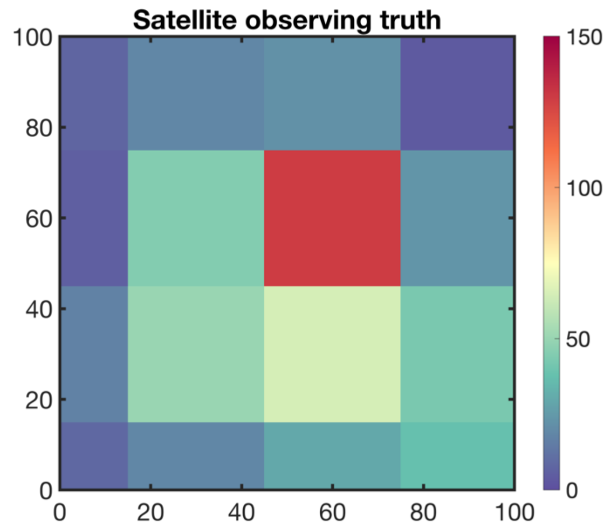
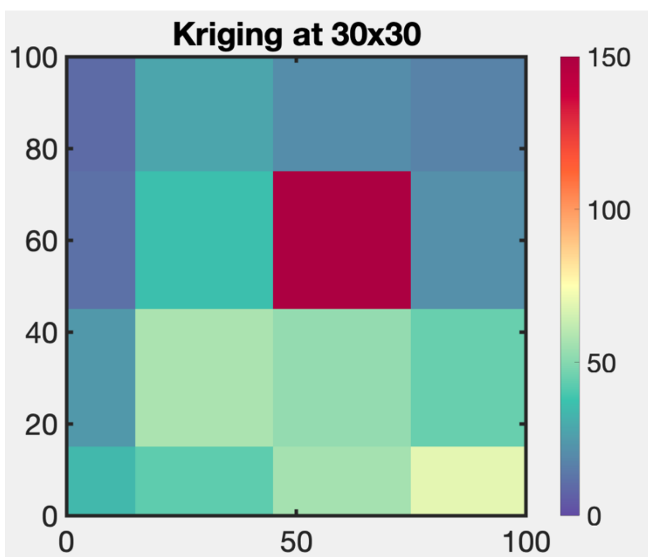


Figure S1. The synthetic satellite observations of the field of truth at the resolution of 30×30 pixels.

Regarding the comparison of non-converged kriging at 30x30 and the synthetic observations, we can observe in Figure 5 that the kriging estimate substantially differs from the truth at coarse resolutions indicating that the estimate should compare poorly with the synthetic satellite observations too. The next figure (for reviewer only) shows the kriging estimate at 30x30 grids:



We are reluctant to include this figure in the paper as our analysis in Figure 5 strongly suggests that we should optimize the resolution of kriging map before comparing/convolving with other datasets.

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.

For consistency, we replaced "o" with "0".

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

Sure, we reworked the figure with larger dots. Due to the fact that we used a random function, the results are slightly different for the random and stratified random cases compared to the original draft.

New figure:

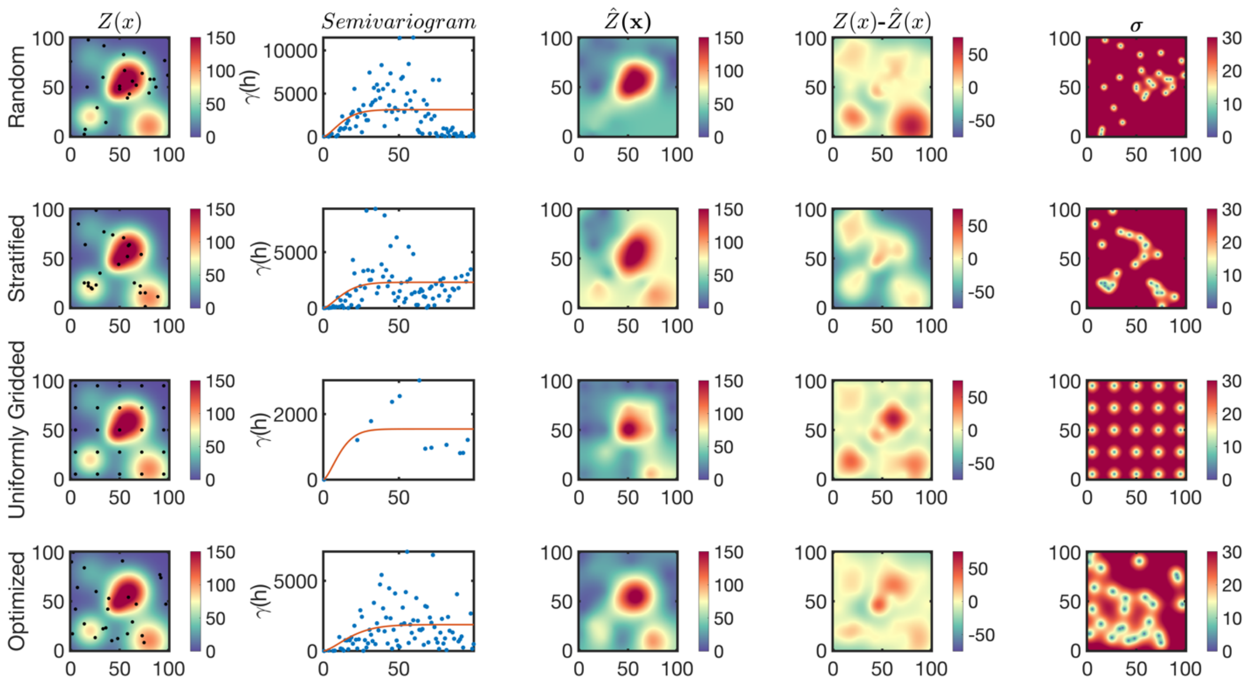


Fig. 12: The resolution is 0.2°x0.2° (approximately 20x20 km^2), instead of 20x20 km^2.

Thanks, corrected.