Response to Referee #2:

We appreciate the very helpful feedback from the referee. The referee's comments are listed in *italics*, followed by our response in blue. New/modified text in the manuscript is in **bold**.

1. The definition of total number of overlapping pixel polygons used in averaging for grid cell j (formula (4)) is somewhat confusing. To my understanding, S(i, j) is the overlapping area, thus $D(j) = \sum_i S(i,j)$ *is just sum of overlapping area and the unit of* $D(j) = \sum_i S(i,j)$ *is* km^2*. If so, is the unit of D(j) for tessellation in figure 8 and 9 also km2? Is D(j) actually defined as* $\sum_i \frac{S(i,j)}{\sum_j S(i,j)}$ *or* $\sum_i \frac{S(i,j)}{grid\ cell\ area}$ *in figure 8 and 9?*

The scaling of *S(i,j)* does not matter as it appears both in *A* and *B* and will be normalized out when calculating *C* (Eq. 1-3). To clarify *S(i,j)*, we always define it as the fractional overlapping area (a dimensionless number). As such, *D(j)* is always dimensionless (sum of fractional overlapping area), and can be understood as the number of overlapping pixels for grid cell *j*. This is what was presented in Fig. 8 and 9. We clarified the sentences at page 7, lines 13-14 as:

"**When the destination grid is regular with constant grid cell area, it is convenient to normalize *S(i,j)* by the grid cell area, leading to overlapping fractions. We will follow this convention hereafter, and hence *S(i,j)* is always a dimensionless number**."

2. In formula (10), why $\int\int_{grid\ j} S(x,y|i)dxdy$ *is normalized by grid cell area? If S(i, j) is just defined as* $\int\int_{grid\ j} S(x,y|i)dxdy$, $W(i,j) = \frac{S(i,j)}{\sum_j S(i,j)}$ *is the normalized spatial response function for observation i and its spatial integration* $\sum_j W(i,j)$ *is unity. Considering discretization of spatial response function in computation, both definitions of S(i, j) are fine. Physically, should* $\int\int_{grid\ j} S(x,y|i)dxdy$ *be normalized by grid cell area or not?*

As indicated by the referee, whether normalizing *S(i,j)* by the grid cell area or not does not change the oversampling results. We choose to normalize the spatial integral of the spatial response function *S(x,y)* by the grid cell area, because (1) the resultant *S(i,j)* is compatible with the definition in tessellation (see the response to the previous comment) and (2) the spatial integration of *S(x,y)* over grid *j* can be directly approximated by *S(x,y)* evaluated at the center of grid *j*. In this way, *S(i,j)* is always a dimensionless number between 0 and 1. To clarify, one sentence is added to page 11, line 7:

"**Similarly to the tessellation approach, *S(i,j)* is always a dimensionless number between 0 and 1.**"

3. Levi Golston has a good point of what is the extent of the enhanced resolution result physical real (Short comment 1 for this discussion paper, https://www.atmosmeas-tech-discuss.net/amt-2018-253/amt-2018-253-SC1-supplement.pdf). Levi Golston shows an example that "true" value is unity while oversampling result is 1/63. The result is, however, based on using 2-D boxcar spatial response function on observation generation and oversampling. If 2-D super Gaussian *function is used, will it show better oversampling result? For 2-D super Gaussian function, will small k1 and k2 give better result than larger ones for Levi Golston's example? I suggest adding discussion of it.*

Strictly speaking, the spatial resolution is only determined by the satellite (e.g., we say TROPOMI has a higher resolution than OMI). We may choose a very fine target grid (i.e., oversample the spatial distribution) but that does not help reproducing the true concentration distribution once the satellite spatial response is adequately resolved. To avoid confusion, we replaced the term "**grid resolution**" in the manuscript by "**grid size**".

If the true spatial response function of the sensor is used in the physical oversampling, and the spatial response is accurately integrated over each grid (Eq. 10 of the manuscript), the result is always physically real as it represents the "exact" observation of the sensor. However, the exact observation does not equal to the true concentration distribution due to the limitation of satellite pixel sizes. One subsection (section 4.3) and one figure (Fig. 8 of the revised manuscript) is added to clarify the definition of spatial resolution vs. spatial sampling:

"**The difference between resolution and sampling density for 1-D spectral data has been thoroughly discussed in the literature (e.g., Chance et al., 2005). However, for 2-D, spatially resolved data, it is common to refer to both the sizes of the Level 2 pixels and the size of the Level 3 grid as the spatial "resolution" of the data. To avoid confusion, it is emphasized here that the true spatial resolution is limited by the sizes of Level 2 pixels. The size of Level 3 grid only determines the density of spatial sampling, which does little to enhance the true resolving power of the data after reaching a certain point. For example, the oversampling results using synthetic OMI data at 1 km vs. 0.05 km grids are very similar (Fig. 6). Nonetheless, it is still beneficial to oversample, i.e., make Level 3 grid size significantly smaller than Level 2 pixel sizes, as demonstrated by Fig. 8. As the ground truth, an array of 2-D Gaussian functions are generated with FWHM ranging from 1 km to 16 km (the second column of Fig. 8) and peak height of unity, and this true field of concentration is measured by an imaginary sensor whose spatial response function is a 2-D super Gaussian (Eq. 8) with FWHM = 10 km and k_1 = k_2 = 8 (the first column and the white boxes inserted in the third column). The third column shows the oversampling results using 10000 randomly located observations. The fine structures in the ground truth are clearly smoothed, limited by the spatial resolution that is inherent to the Level 2 pixel sizes (10 km). However, by oversampling at a fine grid (0.2 km for the first row vs. 5 km for the second row), the spatial gradients are better recovered, and spatial features finer than individual Level 2 pixels can be identified. Additionally, the details in the spatial response function is better resolved with a finer target grid, which is particularly beneficial when collocating with higher resolution measurements (e.g., a cloud imager). As such, although the spatial resolving power is ultimately determined by the spatial extent of satellite pixels, the physical oversampling approach helps enhancing the visualization of spatial gradient and the identification of emission sources.**"

Figure 8 of the revised manuscript:

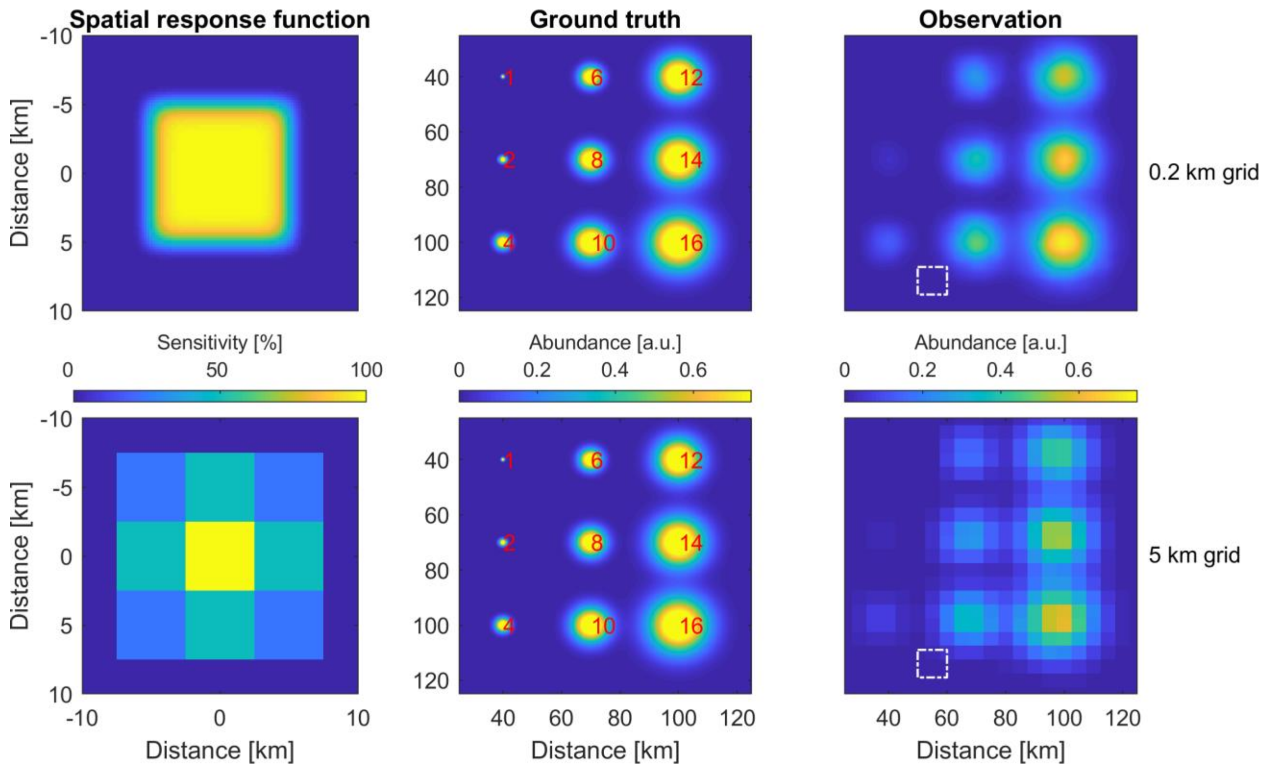

Figure 8. First column: spatial response function of an imaginary sensor discretized at 0.2 km (top) and 5 km (bottom) grid. Second column: ground truth spatial distribution generated as an array of 2-D Gaussian functions of same height (the top and bottom panels are the same). The FWHM of each Gaussian is labeled. Third column: physical oversampling results using 10000 randomly generated observations and discretized at 0.2 km (top) and 5 km (bottom) grid. The pixel size, which determines the spatial resolution, is labeled as the inserted white boxes.