Interactive comment on “Inferring the size distribution of volcanic ash from IASI measurements and optimal estimation” by Luke M. Western et al.

Anonymous Referee #1

Received and published: 29 May 2016

This paper deals with the quantification of airborne volcanic ash from infrared satellite observations. The authors attempt to simultaneously retrieve five parameters, but the emphasis of the paper is clearly on the retrieval of two size parameters: the effective radius and the spread of the psd. Retrieval of the latter is the innovative aspect of the paper, as most other satellite retrieval schemes retrieve only the effective radius. The first part of the paper analyzes the information content of IASI spectra using forward simulations, especially again with respect to the size parameters. The second part then illustrates the retrieval on two case studies. While the topic is worthwhile and the paper is well written, the text and analysis is very short and lacks depth in several places. In my opinion the main result sections (3.3 and 5) need to be reworked/expanded and reviewed again before the paper can be considered for publication.

Section 3.3 on the information content relies on the averaging kernel being a good measure of information content. However, as the authors point out, the averaging kernel depends on the a priori covariance matrix, which was chosen here to be (very) large; e.g. $x_a = 500 \, \text{hPa}$ and $S_a = 500^2 \, \text{hPa}^2$ for the aerosol cloud top pressure. The DOF, in the Rogers formalism expresses the gain of information, and if one starts out with almost no information (large $S_a$), the gain will always be large. Looking at the limiting case $S_a \to \infty$ and thus $S_a^{-1} \to 0$, we have that $A \to I$. Therefore, getting large values (here 5) of the DOF is not so difficult if one starts out with a large $S_a$; but this does not necessarily mean that the information is contained within the spectrum. All of this should be discussed in more detail. I would also urge the authors to find a better way of presenting/analyzing the information content. Perhaps either by looking at the relative increase of the DOF, what I believe was suggested by the other referee in the technical report, or to quantify the information content in another way, e.g. looking at the retrieval error covariance matrix, which is easier to interpret than the averaging kernel, and still provides information for large $S_a$.

Figures 1 to 5 are almost not discussed, although there is potentially a lot to talk about. Without proper discussion, there is really no point in having them. It is not easy to present results on a 2 x 5D space, and it would perhaps be easier not to discuss the underlying cloud case or the gamma distribution (which does not seem to bring anything to the table anyway). Discussing the triplet “effective radius, spread, mass” and how information on these three parameters can or cannot be independently retrieved would be much more interesting. Looking at the on and off-diagonal elements of the retrieval error covariance matrix and its corresponding correlation matrix would be useful.

The presented case studies are not very convincing, and again not sufficiently discussed or explored. Please choose a different colormap for figures 11 and 12, the different shades of green are really hard to distinguish. The colormap used in figures
1-4 would do. Some of the retrieved values seem very large. If the results are what they seem, I really doubt they make sense. Very large values seem to indicate that the retrievals were not sufficiently constrained. But again, it is practically impossible to read the data from the colormap and apart from the spread, all the other retrieval results are also not discussed at all in the text. It would also help if some of the retrieval results (e.g. altitude) are compared with external information (e.g. CALIOP or other satellite retrievals). Another way to deepen the analysis and impact of the paper would be to repeat the retrieval for a fixed lognormal spread and subsequently compare the retrieved masses and heights (with and without retrieving the spread).