Author response to reviews of "Analysis of simultaneous aerosol and ocean glint retrieval using multi-angle observations" by Kirk Knobelspiesse et al.

Feb 8, 2020

We would like to express our gratitude to the reviewers for their thorough and positive assessment. Below are our individual responses to each reviewer, reviewer's text in italics.

We would also like to note that we have added supplementary material at data.nasa.gov at the following location:
https://data.nasa.gov/Earth-Science/MISR_MODIS_AtmCorrection/sg4r-ftwb
This contains figures similar to 3, 4, 5 and 7, but for each of the 7,000+ simulations at various geometries and parameter states. We have added a note regarding this in the 'code and data availability section.

Referee: Hongqing Liu

This manuscript has an interesting topic of assessing the information content of the MISR multi-angle measurements aiming for a potential simultaneous retrieval of aerosol and wind speed. By using AFRT to generate a lookup table and applying GENRA for ICA, several sensitivity tests are designed to reveal the optimal choices of the retrieval strategy. I think this is a good paper, and I don't have any disagreements regarding the analysis and conclusions; but I do have some questions about the technique, especially how to implement the GENRA approach and perform the sensitivity test.

My general understanding (not sure it is right) of this work is that, instead of using real MISR data, the simulated MISR measurements for given sets of parameters (r,f,τ ,w) at the geometries of the test cases (interpolated from the LUT) would be used as "inputs" to derive the posterior PDF of the parameters (r',f',τ',w'), where the prime indicates the "retrieval space" which is much finer than that of the "input" (as in the LUT), so "the result will have n+n dimensions" (line 253-254).

My first question is for equation 2: what is the "likelihood function"? It is said "Pd(y) is the stochastic measurement distribution (d for data), and Pl(y) is the same for the likelihood function" (line 243). Is the "Pl" has the same function form as "Pd" (equation 12), but using different y'? From Vukicevic's paper (2010), her equation 1 has the stochastic measurement distribution (Pd) and model distribution (Pt) terms, and the "likelihood" is the "combining of the model and measurement pdfs" (section 2.3), which seems to correspond to the "Pd(y)" only, since "the data (Pd) is a PDF created from a single node in the LUT and expectations of measurement and model uncertainty" (line 248-249), as defined in the equation 12, the uncertainty variance (σ) includes both the measurement and model contributions. In this paper, the "likelihood function" is mentioned at five places, but it is hard to me to get a clear understanding of this term.

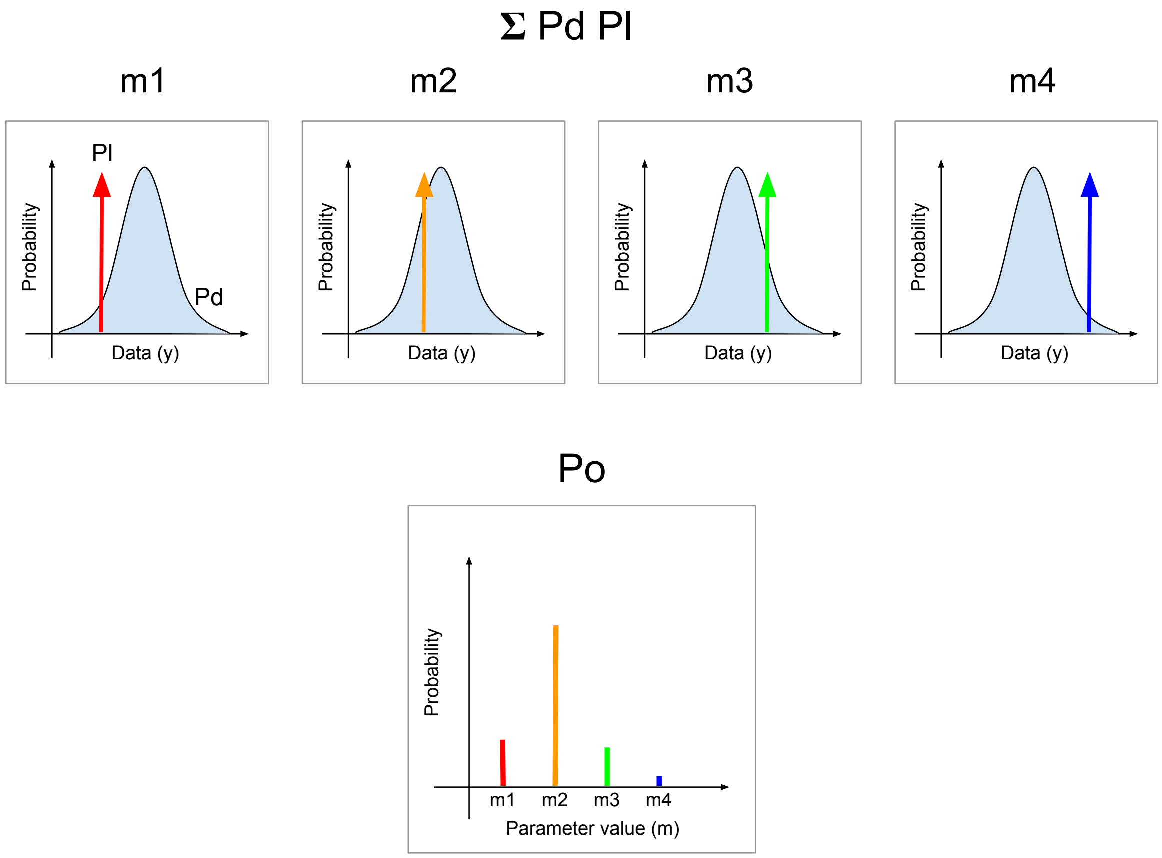
Thank you for these thoughts, and we agree with you that this is a subtle topic, often difficult to transition from the comparatively simple theory to practical application (i.e. code). By way of explanation, consider the likelihood function (p_l) in the absence of model uncertainty. In that case, the function (a PDF) is a delta function with no width. If we assume that the prior (p_r) is uninformative, then the posterior (p_o) for a given m is the probability that the data correspond to the simulation from the likelihood function.

I attempted to draw a cartoon to illustrate this process. Again, making the assumption that the prior is uninformative (has no effect), and treating the likelihood as a delta function (no model uncertainty), then the top row represents four different summations from equation (2), at different parameter values (m). In each, the 'data' are the same, as illustrated by the light blue Gaussian. The radiative transfer model output for a given parameter value (m1, m2, m3 or m4) is the likelihood delta function indicated by different colors for each parameter value. p_d and p_l are multiplied, then the summation is made over all values of y (data). The result is the p_o for each parameter value as indicated by the plot at the bottom. As you can see, we start to get an idea of what the shape of the a posteriori PDF should look like, and parameter value m2 is the most likely value so far given the observation.

The shape of the 'data' PDF p_d comes from the observed value (or in our case, a simulation node in our LUT) and the expected uncertainty (equation 12). In practice, we use a sigma (defining the width of p_d) that is a (squared) summation of all sources of uncertainty, both measurement and model. It makes the code operation more efficient to keep p_l as a delta function, and is mathematically identical. The sparsity of p_o also shows why we must heavily interpolate the calculations that generate p_l, so that p_o is not overly coarse.

Of course, our illustration is for a one dimensional cartoon, in practice this occurs over the four dimensional parameter space. Also, one view angle observation is assessed at a time, and p_o becomes the prior for subsequent assessments.

Hopefully this illustration clarifies the application of GENRA.



My second question is about the last two sensitivity tests (sections 3.4 and 3.5). It appears to me (again, might be wrong) that the simulated MISR measurements (from LUT) are calculated with the plane-parallel and scalar wind speed assumptions. It is a little hard to understand how to test the sensitivity to the conditions whose signals are not in the inputs. For the test of the plane parallel assumption (section 3.4), if the simulated Df and Da camera observations are from the plane parallel calculations, then it would be expected that retrieval with them (9-camera) would have a better result than without (7-camera), since the (extra) information contained in these two cameras are consistent with the retrieval assumptions (both assume plane parallel). For the sensitivity test of the glint, discarding the uncertainty associated with vector wind would only lead to the improvement since the inversion would be more consistent with the input (both assume the scalar wind without penalty). If I didn't misinterpret the result, figure 10 seems indicate that the test without the scalar uncertainty term would have more information content than the test with the uncertainty since the SIC difference is positive.

Taking the above illustration as an example, we performed a GENRA analysis with larger model uncertainties, and compared that to the standard GENRA results. In both cases we estimated a model uncertainty, either from the literature (as was the case with the plane parallel assessment) or by testing the differences between a more complex model and a simpler one (as was the case for the scalar vs vector wind speed). In both analysis we used the same model calculations, but with larger model uncertainty. For the plane parallel case, we also did not incorporate the Da and Df camera observations.

So, you are correct that, without the vector based uncertainty in figure 10, the information content is higher. This indicates that the use of the scalar instead of the vector model does have consequences. However, in most cases (top panel of Figure 10) the difference is quite small, and the cases where it matters are limited to low solar zenith angles (bottom panel). Considering that an additional parameter adds computation expense and retrieval ambiguity, we decided that the difference was not significant enough to warrant a change. This is further illustrated in Figure 11.

Hopefully this answers your questions.

Other than these two questions, I think the paper is well written and the analysis is thorough and clear. So I would recommend this paper be accepted after some clarification about the above questions are made.

Thank you.

One minor issue, for the equations 4-7, should the summation be make over the parameters with prime (i.e., r', f' etc.)?

Actually, no. Consider that we are attempting to reduce an 8 dimensional volume to five dimensions – one for each of the simulation nodes, and the other representing the marginal PDF for a given parameter at that node. Thus summations over three dimensions reduces the volume dimensionality from 8 to 5.