Rebuttal

Dr Fred Prata, 6 July 2017.

The revisions required do not seem to be of great significance to my paper. There seems to be a desire to have more statements on increased uncertainty in retrievals. My paper is not claiming any particular improvement on accuracy of retrievals or retrieval methods, in general. The results presented are for Grímsvötn and were validated. But the paper does not depend in any way on exactly how accurate they are.

Responses

"I would like to see more detail on the uncertainties due to cloudiness and lack of thermal contrast. In the Response to Reviewers, the authors suggest that I am under the 'common misconception' that ash clouds with high concentrations are not detected. They are detected as some kind of cloud due to their increased optical depth and lower brightness temperature, but cannot be identified as volcanic ash and a retrieval cannot be made. What is not clear from the current manuscript, nor from other similar publications e.g. Prata and Prata (2012), is how such pixels are identified and incorporated into mass loading estimates. Is it a manual process? Are they excluded from calculations? In which case, mass loadings must be considered minima."

Response: There are plenty of papers and technical reports that go into detail about ash detection – which seems to be the reviewer's concern here. The interested reader can find a series of technical reports on ash detection, validation and errors here: http://vast.nilu.no/project/deliverables/. It is not the purpose of this paper to examine ash detection methods in any detail. This would require a very major revision of the m/s and a complete change of emphasis. Suffice to say the Grímsvötn retrievals were validated and found to fall below the commonly accepted errors of 40-50%. Kylling states 45-50% discrepancy for shape but really we don't know how this compares to observations because it is a theoretical study. This could all be bias and could cancel with other biases – it is simply not known. What he means is that if you use some irregular shapes then you can get 45-50% differences in retrieved mass loadings – but what shapes should one use? What composition? (Composition and shape are correlated). Has any of it been validated?

In my last response I was quite clear that optically thick pixels are not used in the retrieval and in fact no retrieval is possible. The mass loading is determined on a pixel-by-pixel basis. The retrieval for each pixel could be high, low or just right. Why should they be considered minima? For optically thick pixels no retrieval is made – no statement can be made about whether it is high, low or just right, because no retrieval is made. Likewise, for optically thin pixels, when a retrieval is made it could be too high – in fact that is more likely because often the optical thickness may lie within the noise but the retrieval is made based on other information – for example the spatial context. It cannot be stated that the mass loadings should be considered minima. Perhaps, the total mass retrieved is biased low. There are other complexities. To demonstrate, consider this: for a MODIS pixel (1 x 1 km^2) suppose I retrieve a mass loading of 1 g m^{-2}. Now I have assumed (implicitly) that the pixel is completely covered by ash – but that's an assumption. It could be that the ash is actually confined to an area of 100 m x 100 m. The total mass then would be 10 kg. But for the MODIS pixel observation, the 10 kg is spread over 1 km^2. So the "real" mass loading is 10 kg km^{-2} or 1 x 10^{-2} g m^{-2}. In other words I have overestimated the mass loading by a fact of 100. My retrieval is hardly a minimum. You could ask if the total mass were actually 10 kg then how could I retrieve 1 g m^{-2}? This is complicated because it depends on what else is in the pixel and exactly where the ash is. The instrumental response is not uniform across the pixel and indeed since we use two different channels, the fields-of-view (fov) need to be co-aligned (they never are). If the edge of one fov observes cloud, while the same edge of the second fov does not, then the brightness temperature difference between the fovs could be negative without any ash at all in the pixel. Of course, the problem here is one of heterogeneity and spatial resolution – I would need to address this as well as all the other complexities, such as sub-pixel cloud, "mixels", misalignment of instrumental field-of-views, calibration non-linearities (these affect "cold" pixels), slant-angle effects, overlapping pixels, the modulation transfer function … if I were to properly address uncertainties in satellite retrievals. But see Fig. 14 of Clarisse and Prata (2016) (referenced in this paper) where fovs from three different IR sensors are collocated. It is a complex problem.

"Just because someone wins the lottery doesn't mean that the odds weren't millions to one!"

Response: I really don't know the meaning of the comment. The data are what they are – no luck was involved. The validation shows that the Grímsvötn retrievals were better than to be expected. It is a reasonable admission that errors, in general, could be larger. This is a paper about Grímsvötn ash dispersion.

"The brightness temperature difference method assumes that particles are dense spheres, which only exhibit the BTD effect when the size distribution is dominated by particles <10 μm diameter. Thus, any pixel displaying a BTD signal will be interpreted as being dominated by particles <10 μm diameter. If nonspherical and bubbly particles cause a BTD signal at larger grainsizes (as demonstrated by Kylling et al., 2014), and ash grains are not dense spheres (as demonstrated by hundreds of tephrochronology studies), then the grainsize will be underestimated."

Response: Actually no such assumption is made. The reverse observation due to ash happens – it is observed in the data. The model makes various assumptions but that is a different matter. The paper is not claiming any significance about grain size retrieval. Kylling says the range is increased from 5 μm to 10 μm (I think it is larger than 5 μm for dense spheres). Kylling does not address "a cloud of particles" where the radiation from individual scatterers interacts. If you look at his Figure 5 you will see that he has brightness temperature differences as low as -20 K. These are never observed. It does not match reality. Further, his particle shapes are also "idealized". Real particles have asperities, aggregates and are also compositionally complex. The radiative transfer for these particles may be quite different to Kylling's particles and may even suggest that treating them as dense spheres works just as well if trying to estimate mass loadings. The only way to tell is to validate the theory. Prata and Prata (2012) did that. The retrievals for Grímsvötn were validated – the theory could still be wrong (probably is wrong) but my paper is not trying to suggest this nor do the results depend on this.

The "hundreds of tephrochronology studies" are irrelevant to very fine ash dispersing in the atmosphere. (They are in fact orthogonal studies as they are studies of **exactly** what is not in the atmosphere).

"The only logical way to argue that this doesn't apply is to present evidence that Kylling is wrong, or that volcanic ash grains ARE dense spheres when airborne."

Response: Well it is absolutely clear that you cannot show that Kylling et al. is right. Indeed in the Karl Popper sense I have already shown that Kylling et al. is not right – I validated the dense sphere assumption against independent data and found it to be reasonable. Their study is theoretical and they provide absolutely no data to validate their findings. They are therefore just one of several possibilities that can only be shown to be wrong by experiment. (It can never be proved to be right). Suppose we use odd shaped particles with vesicles. The retrievals are going to produce 45-50% larger mass loadings and disagree with the ground-based data, the lidar data, the aircraft measurements, and all other satellite retrievals. So some other assumption must be wrong.

Kylling et al. also state that: "*It is noted that ash particle shape is not usually known for an on-going volcanic eruption. Thus, for operational monitoring of ongoing volcanic eruptions it is preferable to assume spherical ash particles and rather increase the uncertainty in the mass estimate.*" "… not usually"???? They mean "never". I have never seen a single study, published or otherwise, that reports particle shapes in dispersing ash clouds. Furthermore, what is the distribution of shapes? This must be important too.

"Neither the response to reviewers nor the updated manuscript addresses the additional source of uncertainty described in Stevenson et al. (2015), namely that the mathematics behind retrieval algorithms biases them towards solutions involving smaller grainsizes. For a given observation, the algorithms prefer solutions with low concentrations of optically active (small diameter) particles."

Response: The algorithms have been validated. I'm not sure why the reviewer says mathematics? He may mean physics? I have added a paragraph stating what I think the Stevenson et al. (2015) paper is reporting.

"The method of Prata and Prata (2012) is not a 3-parameter retrieval like that of Francis et al. (2012), but instead uses a lookup table for a specific cloud-top temperature. In this scenario, there is only one possible combination of effective radius and optical depth that matches a given observation. The choice of cloud top temperature can therefore bias retrievals towards higher or lower grainsizes. I would like to see discussion of how the cloud top temperatures were chosen and the contribution of varying this to the uncertainty in the retrievals."

Response: The Prata and Prata (2012) paper underwent peer review and was published after revision. If the reviewer wishes to critique that paper then he should submit a comment to JGR. I'm not prepared to enter into a discussion of that paper and it is unusual to expect this.

I'm not sure I understand the meaning of "…*a 3-parameter retrieval like that of Francis et al. (2012)*"? Does he imply this is somehow superior or correct? The Francis method uses optimal estimation. It is necessary that the parameters being retrieved are uncorrelated. In the case of Francis et al. they are not. They retrieve plume altitude, effective particle radius and mass loading. Two of these are correlated. Plume altitude is extremely difficult to estimate and I would argue is a large source of introduced error into the retrieval scheme. A key point: if you have two measurements then you can only retrieve two parameters, unless you add constraints (assumptions).

I am not at all sure why I am being asked to comment on other papers and other methods. This is largely irrelevant to the current paper.

"There are some reasons why we cannot be sure that aggregation is the sole driver of a partial collapse." I'm not clear what argument you are making here. Are you saying that there are many course particles and so the plume would have collapsed anyway? Can you rephrase?"

Response: Yes this has been re-phrased to:

"Because there is not that much very fine ash in the column to begin with to generate a sector wide plume collapse we cannot be sure that aggregation is the sole driver. The particles in the 100 μm size fraction contain less than 10% of the mass erupted at any one time, so that even if all of this ash forms aggregates, the mass fraction is still small compared to the total mass."

Minor comments.

References
Clarisse and F (2016): who is second author? Changed to Clarisse, L. and Prata, F. (2016).
Line by line comments
3:10 - Did you mean that 16 microns is the largest size? Yes. "smallest" changed to "largest".
3:24 - Spelling: glacier Corrected.
6:9 - Standard atmosphere is important, but the ice/water at the vent at Grímsvötn were probably a bigger factor and are not accounted for in Mastin. Agreed. Changed to "..., without the inclusion of ice/water at the vent, which is likely a large factor in the case of the Grímsvötn eruption.
7:30 - Spelling: grain sizes Changed "size" to "sizes".
17:5 - Do you mean fine ash or very fine ash? The two terms seem to be used interchangeably in consecutive sentences. Clarify by repeating definition. OK. I have defined my use of the term very fine ash and replaced all instances of "fine ash" with "very fine ash".
18:6 - Kylling (2014) found errors of 40%, which is greater than the 10-30% cited here. Kylling found this to be 40% but as this differs from someone else's theoretical study, one is driven to the conclusion that the models have some inherent discrepancy. I have therefore decided to rephrase this entire sentence to read: "Estimating precision in retrievals is difficult because of the uncertainties in the input parameters, such as the complex index of refraction, the size distribution and the shapes of the particles, although shape is generally found to result in the smallest discrepancy of the input parameters with theoretical simulations showing differences in the range of 10-40% (Yang, 2007, Kylling et al, 2014)."
20:29 - Spelling: Grímsvötneruption Corrected.
22:7 - Adding meltware removes energy from the plume Changed to "...in the form of hot water vapour (steam) and also contributing ..."

Co-Editor comments

Errors. (See my comment above). I have added this paragraph in Section 4.5:
"Retrieval methods are being continually improved and there is an international effort (http://cimss.ssec.wisc.edu/meetings/vol_ash15/) to inter-compare retrieval schemes

and help reduce uncertainty. At the current time no firm conclusions have been made about retrieval accuracy as no robust validation has been made. Uncertainties can only be assessed against independent observations and so far independent measurements of mass loading are extremely sparse, let alone independent measurements of atmospheric ash particle size distributions, shapes and composition."

I have also made this change: "In the case of the ash retrievals for Grímsvötn, the error estimates are within the expected range, giving an error of ±0.1 Tg or roughly 20--50% of the estimated mass of very fine ash. It is emphasized that this is not the total mass emitted by the volcano, which is typically a few percent of the total mass. It is however, the mass fraction that is dispersed by the winds and the very fine ash that can cause damage to aircraft jet engines."

I have also fitted in a reference to Stevenson et al. (2015) but as indicated above this is contentious and any detailed discussion of that work requires a more considered approach. Also I don't understand why I should need to discuss someone else's work in my paper – it is now referenced. After (Section 4.5), I have added: "The error (precision) in estimating very fine ash mass …"

"Stevenson et al. (2015) discuss potential errors in satellite retrievals by using cryptotephra data to speculate that larger particles exist in dispersing ash clouds (although no atmospheric observations are presented) and claim through modelling studies that current retrieval schemes (all of them) underestimate mass loadings because of the dense sphere assumption and lack of sensitivity to particles with diameters > 10 μm."

Figure 4. Caption. Added: "Isolines (contours) of brightness temperatures are shown in white to highlight the location and expansion of the top of the column."

Figure 7. Font size increased.