Atmos. Meas. Tech. Discuss., 8, C4–C10, 2015
www.atmos-meas-tech-discuss.net/8/C4/2015/

Atmospheric Measurement Techniques
Discussions

Interactive comment on "Big grains go far: reconciling tephrochronology with atmospheric measurements of volcanic ash" *by* J. A. Stevenson et al.

Anonymous Referee #1

Received and published: 12 January 2015

Volcanic ash layers (cryptotephra) found on the ground far from volcanic eruptions include some very large particles, although ash cloud particle size distributions inferred from spaceborne (thermal IR) remote sensing typically retrieve size distributions which are too small to account for the presence of significant numbers of the largest ash particles. This paper uses examples from Icelandic volcanic eruptions in an attempt to understand the reasons for this discrepancy.

This paper straddles the atmospheric remote sensing, modelling, and *in situ* communities and overall does a (mostly) good job of explaining the jargon inherent in each

(e.g. conventions about reporting particle sizes), which is welcome. I am approaching this review as an atmospheric scientist so hope that other reviewers are from the modelling/ground-based sides, as those are not my strength. The topic is relevant to AMT and of scientific interest and importance.

Because of the blend of fields, this paper required a lot of careful reading to make sure that I understood the main conclusions (not a criticism, just a consequence of the nature of the paper). In a few places I think it would be beneficial to have bullet-point summaries of the main points, to better signpost the main messages. If I have it right, one of the main conclusions (obtained from radiative transfer modelling) is that the thermal IR split-window technique has insufficient information content to be sensitive to ash clouds dominated by larger particles, causing a low bias in the retrieved particle effective radius. Additionally, it seems like the split-window technique becomes less effective at identifying such pixels as particle size increases, pointing to the need for better detection tests. I have some comments on that (below) and think the analysis could be expanded a bit. I would also appreciate some more discussion about the *in situ* data and its associated uncertainties – from my reading it appears that the authors believe that the remote sensing data is what is at fault here? However it is apparently not that simple because airborne measurements also report distributions too small for these largest grains, although there may be a sampling bias (aircraft don't fly through thick ash clouds because it is dangerous).

My technical comments (below) mostly focus on the remote sensing aspect of the paper. With that caveat I favour acceptance after revision for minor corrections, although if other reviewers have issues with the ground-based data and modelling sections then I suggest the editor defer to them in those matters. In these comments PXXLYY refers to page XX, line YY. Should another round of reviews be required then I would be happy to provide another review.

P67L11, 15 and throughout: One of my issues with the paper was the definition of 'cryptotephra'. On line 11 the term appears to refer to all volcanic grains which are deposited on the ground. However, on lines 15 and 16 (and later) it sounds like 'cryptotephra' refers only to the largest particles. The authors should be more clear in the definition and more explicit and careful in their usage of the term. What size counts as 'cryptotephra'? Additionally, I tried to search for definitions online and am unsure why the term 'cryptotephra' is used and not simply 'tephra'?

P69L26-27: if 'basaltic' and 'rhyolitc' are the two main classes of volcanic eruption, it would be good to add an introductory sentence mentioning this (e.g. 'There are two classes of eruptions: basaltic and rhyolitic.'), if it is relevant. So I suggest that this paragraph is reordered.

P69L4: This is the first use of the term 'BTD' and it should therefore be defined. Or the authors could reword or remove this clause.

P69L12: I infer here that 'proximal' refers to locations < 500 km (based on 'distal' being > 500 km) but again it would be good to mention this explicitly here.

P71L20-P72L3: This paragraph highlights an issue that I think could be resolved by including some sort of more direct 'statement of the problem'. Basically, this paragraph mentions some ground-based measurements of ash size distributions which are a lot larger than spaceborne estimates of the airborne ash cloud. But how do we know that it's not a case that the large particles fell out while the smaller ones remain aloft? Or smaller grains that fell were either aggregated into larger particles, or just not detected/sampled (for whatever reason) in the ash measured on the ground?

Surely it could be that there is not in reality a discrepancy, just the two techniques are measuring different things? How do we know there is a problem that needs to be 'reconciled' in this way? It might be better to expand and move this subsection to before Section 1.2.

Section 2.1: This is a place where more detail is needed so remote sensing and modelling people have a good idea about the uncertainties involved in estimating the particle size distribution from ash several years old which has fallen into a peat bog. Intuitively it sounds to me like that would be a nontrivial thing to do. For example, P73L16-18 suggests that sizes were counted manually for the peat samples. For a start, does the sample preparation technique sample the 'true' size distribution well, or can small/large particles be missed; what about fracturing, coagulation, or contamination from non-ash particles? How many grains do you have to look at and measure to get a representative sample of the size distribution? Is there any selection bias when imaging grains? I would suggest inserting a new subsection after the current 1.2 (satellite data) to give an overview of techniques and uncertainties for estimating ash particle sizes from these deposits. The premise on the paper seems to rest on these ground samples being 'correct' so it would be good to strengthen this assertion.

Remote sensing analysis and discussion: The main limitation of this section is the use of Mie theory. The authors acknowledge this. However it does seem central to the results of the study, i.e. large grains don't produce a BTD effect (under Mie theory) so when they are dominant ash detection schemes fail. The authors cite Kylling *et al.*, 2014, which states that irregular particles can produce negative BTD for larger particles than spheres do. It would therefore seem sensible to use those optical properties rather than (or in addition to) Mie theory, to present more directly how large an effect this is. After all the real world does not work with Mie theory so it isn't clear that the simulation results represent the real signals that satellites will see very well in these cases. One option would be to use irregular particle optical properties when simulating the radiance (rather than Mie theory). Then do retrievals on this simulated data using both irregular particles and spheres, and see how the detected area and retrievals change. This is the logical next step to try to solve the problem rather than just illustrating it. Otherwise the simulations seem set up to fail, i.e. if a BTD is observed then Mie theory says it must be from a particle size distribution which is not very large.

The second point that comes up in the discussion is that, because the BTD tends to zero as particle size increases, additional channels are needed for ash detection in some cases (the example given on P83L21 is 8.7 μm). However the authors do not seem to directly assess this. My suggestion is to add this wavelength to the simulation and see how much that changes the results. Or perhaps this is what is meant by 'additional tests' in the caption of Figure 10 – this should be made clear. The last line on P81 talks about tests that 'may' detect additional pixels: were these applied or not? Francis *et al.* (2012) state that 8.7 μm was not used in that work because of SO_2 absorption, so it leaves it ambiguous as to exactly what was done in this section.

I suppose another related point is whether other parts of the spectral domain (e.g. the visible bands on SEVIRI and future sensors the VAACs will use) are of use for ash identification and/or size distribution retrieval. Expansion of the analysis to include solar bands is probably out of the scope of the study but more alternatives to the split window technique could be mentioned (e.g. hyperspectral stuff is mentioned briefly).

P83L25: This highlights another example of an issue I have with the analysis. In situations where there is limited sensitivity to a parameter (for this case, as the authors note, the largest grain sizes), this type of retrieval scheme will be sensitive to first

guesses and *a priori* information. So this will influence the statistics overall, and in this sense the analysis presented is quite algorithm-specific (since the retrieval solution is sensitive not only to the underlying physics but also the minimisation method). Changing the first guess at the solution or *a priori* value/constraint strength may change the results. For example if the *a priori* radius were set to 12 μm instead of 3.5 μm, presumably these retrievals with limited information would sit at 12 microns instead. From Francis *et al.* (2012), which is the algorithm used, it looks like the *a priori* constraint on ash size (10 μm uncertainty) is fairly weak, so my guess is that changing the *a priori* value would result in a larger change to the retrieval statistics in these cases than the *a priori* uncertainty. Although it is interesting to note that the constraint as applied will increasingly work against the retrieval of large particle sizes as the true particle size increases. So the construction of the algorithm itself will lead to smaller retrieved grain sizes in some situations, regardless of the physics.

It is also not clear to me whether the underestimate of size for the largest particles is due to the physics or the retrieval. If the first guess is 3.5 μm (which is what I infer from Francis *et al.* 2012 Table 2 and paragraph 20) presumably the retrieval will iterate towards a higher radius and then stop around 8-9 μm when the BTD signal disappears and there is no further gradient. So the low bias may be because the first guess is on the low side. If the first guess were above the true value perhaps the opposite result would be obtained, i.e. the retrieval would converge at a radius above the true value. Otherwise I see no reason why there should be a low bias as opposed to a high bias in the remote sensing data (it depends on the direction from which the solution is approached). This seems like it would be something fairly straightforward to test. The Optimal Estimation framework the authors are using provides additional statistics (averaging kernels and state vector uncertainty estimates) which the authors could analyse to see whether in fact the 'true' value is bounded by the retrieval uncertainty estimates or not.

Conclusion: This is a place where I think it would be clearer to present the take-home messages as bullet points.

Interactive comment on Atmos. Meas. Tech. Discuss., 8, 65, 2015.