Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC2
https://doi.org/10.5194/amt-2021-103-RC2, 2021

Comment on amt-2021-103

Anonymous Referee #2

Referee comment on "Ash particle refractive index model for simulating the brightness temperature spectrum of volcanic ash clouds from satellite infrared sounder measurements" by Hiroshi Ishimoto et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-103-RC2, 2021

This manuscript proposes a method to identify the best refractive index for reproducing volcanic ash aerosols observations in the thermal infrared. The authors use observations from IASI and compare how the simulation fits the observation with different refractive indices. The one leading to the best fit is selected as the optimal refractive index. The IASI observations used for these comparisons are very carefully selected to avoid (as much as possible) the presence of clouds.

Although this is not a revolutionary approach (to my knowledge, scientists working on volcanic ash retrievals from IASI often select “the” refractive index to use in their retrieval by comparing simulations with observations and taking the best fit), this work has the merit to show it done for a number of volcanic eruptions, using a wide range of refractive indices. The authors also demonstrate that a more “simple/easy” parameterisation through varying the NBO/T ratio (instead of using specific more detailed refractive indices from different measurements) allows obtaining good results.

I have however major concerns about the radiative transfer model and retrieval method used in this work because not enough information is provided. Those are however the way to guarantee the quality of the results and the use made of those results to reach any conclusion. I also have some concerns about those results, which are in contradiction with other results found in the literature. This does not make the current results false, but when a result goes against current knowledge I would like to see the method being precisely described, discussed, and shown robust. More in the detailed comments.

Major Comments

- The manuscript does not follow the usual titles, which is fine by me except for the fact that I would like to see a section shortly describing the different instruments/data: IASI, GANAL, and I would also transfer to that section the RI information and maybe also the radiative transfer code if it is re-used from a previous publication and with only a short description (see also another major comment on the RT itself)
- Please be quantitative when assessing the quality of a result, do not use "good agreement", "good fit", "agree well" or similar but provide numbers (RMS, or other more relevant depending on the case)
- The strength of "V-shaped" spectral feature of volcanic ash aerosols (or the slope on each side of that V-shape) depends on the refractive index, the effective radius, the optical thickness and the plume altitude. In the cited ref Clarisse and Prata 2016, the Fig 10 shows clearly the dependence of the split-window BTD with optical depth and effective radius for a specific refractive index, and in Fig. 3 of the same reference the authors also show the impact of the plume altitude, for a specific refractive index (and others have reached the same conclusions e.g. Maes et al, 2015, doi:10.3390/rs8020103). In this manuscript under review, it seems that the problem is sometimes taken too "lightly", not considering this complex relationship between the 4 parameters and the fact that their retrieval might well not be independent, and multiple solutions may be plausible. For example this sentence, line 62: "The particle size of ash clouds, which can be determined from the negative BTD between two infrared split-window channels" -> yes, if you know the optical depth (or thickness) and refractive index.
- The radiative transfer code: the authors state that they "developed an original radiative transfer code", using a demonstrated approach. This is fine, but it is very unclear to me if the authors re-used the same code, or wrote a new code. In the first case, I would not think appropriate to write the sentence I cited here. In the second case, I would like to see some "validation" of the radiative code. What is stated in lines 88-89 is unclear (I don't understand against what the calculated spectra were compared) and not sufficient to demonstrate that the radiative transfer works as expected. I would like to see comparisons with at least one other well-established RT code.
- Particle size: in the cited reference Clarisse and Prata 2016 it is written that "IR sounders are highly sensitive to the effective radius of the distribution within the range 0.5-5μm" (page 198), and this reference is cited to justify the particle size retrieval done in the current manuscript. Then in the results (Table 3) 6 over 21 cases end up with an average particle size smaller than 0.5μm. How much can we trust these results?
- At the end of page 7, it is said that the method "aims to select pixels showing sparse VACs comprised of small particles". I am unsure to understand this. As mentioned before the large slope in the V-shape does not ensure small particles, it could also be high optical thickness and/or high altitude ash aerosols. Second, why would it target only "sparse VAC"? In the selected spectra, I see some with a very strong ash signature, which I would not refer to as sparse ash (for example fig 9c with about 20K drop in BT along the V-shape)
- The retrieval method (pages 8-9): actually, no retrieval method is explained. It is only mentioned that parameters are "estimated". This part should be much more detailed, as all results depend on how much the retrieval can be trusted. If I understand correctly, there are 2 steps: one estimating the ash parameters (how and from which a priori values?), and then a second step to estimate O3 and SO2 parameters estimate (same questions). In those steps, surface temperature and temperature profiles are maintained constant (not retrieved) to the GANAL values – which are not described, see another comment on this later. In most of the thermal infrared retrievals, at least the surface temperature is a retrieved parameter (because a wrongly assumed Ts has a devastating impact on the retrieval results), why is it not the case here? Then, when calculating the RMS, it is very unclear why "error in the GANAL atmospheric profiles" (what exactly do you mean? Uncertainties or biases? In which profiles?) would have an impact only between 650 and 750cm-1. If there are indeed higher uncertainties/bias in

that spectral range, then why use it at all in the retrieval? Finally, I would like to see a discussion on the possibility to retrieve together the altitude, optical thickness and particle size of volcanic aerosols, linked to the retrieval method use, information content, a priori and constraints of the retrieval. At this point, I am not convinced that there is enough information in the observations to retrieve it all together, except maybe with extremely strong constraints (and then the choice of a priori value is very important). But none of this is discussed.

- Presentation of the results:
  - The supplementary material contains a lot of information, but lacks a short description. For example, I had to guess that S1 contained the RMS. For S2, maybe some graphs would be useful (as in Fig. 7) in addition to the numbers. Those could be also in supplementary information, but I think that just the table with all the numbers is very difficult to analyse quickly.
  - I do not think that it is enough to mention in the paper that the RI leading to the lowest RMS was selected; you need to provide some statistics of that RMS in the paper. I would do that in Table 2: add columns with mean RMS (not total, that does not allow comparing between different eruptions) and standard deviation on RMS, or number of RI leading to a mean RMS within a certain range (e.g. 0.15K = IASI noise from the first estimation) from the "optimal", to show if there were large differences between the results with different RI. In addition, I find much too limited to just take the minimum RMS as criterion (with exception of the additional criterion on size) to select "the best" RI. Indeed, in many cases the mean RMS for different RI is very close, as I will underline again in the specific comments by eruption. The difference in RMS between different RI in that case (when the difference is very small) could just be linked to uncertainties in the other parameters (surface and atmospheric T being the most important), and therefore the conclusion on the selected RI could be wrong.
  - When the selected pixels are less than 5, the statistical significance is pretty low. This is mentioned in the text but should be briefly mentioned in the Table 2 caption
  - In Table 3 the mean parameters for each eruption are listed, but it makes little sense to average SO2 content, ash optical depth/thickness and plume altitude at different locations and different times (even if for the same eruption, those vary a lot with time and location). Only (maybe) the effective radius can be considered to not vary much – although it should vary as the biggest particles fall sooner.
  - Spectra as they are represented do not really allow seeing how big the difference is between observed and modelled. For example, in Fig 6 spectra look really similar but the simple fact that, at such a small scale, we can see the different colours means that actually the difference is probably of the order of 1K, which is significant. I would suggest to plot all spectra as in Fig. 7 (or zoom even a little bit more), with focus on the 2 slopes of the V-shape and not showing in full the CO2 O3 and H2O bands.
- In most results Figures (from Fig. 6 onwards) the legend is pretty difficult to read

Specific Comments

Line 107: is an interval of 10cm-1 enough to reproduce the fine features of the used RI?

Line 112: “and artificial weak absorption features were added to them” -> this reads very weird; clearly I understand what the author means, after reading the complete manuscript, but when reaching this part it is very unclear what is meant. I think that the authors should here either detail what they mean, or clearly refer to a further paragraph.

Line 235: is the BTD for the selection of the IASI scenes calculated using only 2 channels? If not, the details should be given; if yes, then I strongly suggest using the average of a number of channels, to reduce the impact of noise. Indeed, at those wavenumbers, the IASI spectral noise was reported to be ~0.15K at the early years of the instrument (Clerbaux et al., doi:10.5194/acp-9-6041-2009) and a bit later ~0.25K (Hilton et al, doi:10.1175/BAMS-D-11-00027.1). Therefore, a simple BTD without any averaging may bear an error of up to 0.5K, which is 25% of your threshold.

Line 250 and following -> only day-time IASI (because MODIS is used to analyse the imagery) but then Line 270 also night-time data is included in the analysis. This is unclear. At the end, is it only day-time IASI pixels or also night-time?

Lines 261-263: “As the temperature of the MC layer is generally lower than that of the sea surface at the same geographic location, a VAC above an MC layer tends to have a lower infrared brightness temperature than a VAC with no MC if the cloud parameters of the VAC are the same” -> This is confusing and should be rephrased. The VAC will have the same brightness temperature in both cases (unless the temperature profile is changed). The observed spectrum is not the BT of the VAC (unless it is optically thick and the surface underneath is not seen at all), but the surface emission (either of the sea surface or of the MC) followed by the atmospheric impact of all gases and aerosols between the surface (again, sea or cloud) and the instrument. The observed BT (not the VAC BT) at the satellite is indeed usually lower if the ash plume is above a MC than above sea, on the condition that the ash cloud is not optically thick (and this happens).

Line 266 and further: TB_obs(nu_a) is only at one wavenumber? Then as for the spit window do not forget the spectral noise. In addition for the calculation of TB_clr don’t forget the model uncertainties (those linked to RT should be minimal, hopefully, but there could be significant uncertainties linked to the surface T and the T profile)

Lines 270-272 Please provide a reference and more details on the GANAL data.

Line 279: “artificially”? = empirically?

Line 382: the sentence starting with “on the other hand” reads weird, it gives the impression that the MP_A should have been selected, but I guess that you mean it was better than the original but still not the best RI?

Line 383: “high accuracy BTS simulations …” -> the accuracy is currently never discussed in this manuscript: you would need to provide some RMS values, and discuss those with respect to IASI instrumental noise (and even better with respect to uncertainties in the non-retrieved parameters but maybe that is out of scope here)

Line 477: the data from Ventress et al given in the () is actually the a priori of their retrieval, not their result. The result of their OE was a particle size of about 1µm and a plume top altitude of about 3.5km, which are I guess the values that the authors read in Fig. 5. Please rephrase correctly the whole sentence lines 476 to 479.

Line 514: “no systematic bias was apparent” -> bias with respect to what?

Lines 522 to 526: it should be stated (as in Fig 7 caption) that the other ash parameters were kept constant.

Line 527-528 “Therefore, selection of the proper RI model is essential, especially for the estimation of Reff. This result also suggests that RI model selection may have a strong influence on the estimation of the ash column density, …” -> I am not sure to understand; the word “especially” in the first sentence is to emphasize that the impact is targeted on that variable, then a second sentence says that another variable is also highly impacted. I would remove “especially”. Second, in your results, the retrieved optical thickness does not seem to depend on the selected RI (Fig 7c), in contradiction with this second sentence.

Figure 7: in the caption there is something weird, the d and e are described twice, and I think the first is wrong. In addition, in those Figs 7d and 7e, I would like to see the spectrum with the VAC parameters leading to the minimum RMS (the one “selected”).

Eyja results: the mean RMS on all pixels for all RI range from 1.039 to 1.388K with 12/21 RI leading to a mean RMS below 1.1K, so within 0.06K from the “optimal RI”. Considering the IASI spectral noise of 0.15 to 0.25K (see before) and the possible bias linked to surface temperature and atmospheric parameters, can we really conclude that only one (or two) RI is “correct” in this case? [This is connected to my Major Comment 8.2]

Line 579: “In addition to Eyja…” -> This reads weird as a start for the new section, what do you mean?

Grimsvotn results: the top altitude seems really low with respect to what is reported in the literature; indeed Moxnes et al report an ash top height below 4km, but actually with ash up to about 4km (and matching all measurements, not only IASI). Therefore the plume

altitude reported here (top at 1.6km) does not really match the cited literature, on the contrary to what is stated in line 599. A wrong retrieved ash altitude could be linked to a biased temperature profile; this is worth a check when results do not match anything previously published. For that eruption, the mean RMS on all pixels range from 0.806 to 1.198K for all RIs, with 11/21 RI leading to a mean RMS below 0.9K; this leads to the same remark from me as for the Eyja.

Lines 624-626: I am not convinced of that conclusion, because of all the comments done on the results.

Calbuco results: there is no comparison with literature here; the Calbuco A mean RMS are all much higher than for other eruptions (except for the rejected RI which had a mean RMS of 1.572K, the other RI give a mean RMS between 1.823 and 3.131K) while for Calbuco B only 1 RI model leads to a mean RMS above 1K. In addition, when looking at the spectra for A and B, it is very clear that the situation is completely different. To me, the A spectrum (Fig. 9c) could hardly be for an optical thickness of only 0.14 even if at high altitude (a drop of more than 20K in radiance…) – in comparison the Kelud shows an optical thickness of 0.6 at 11km altitude for a similar radiance drop. And the different mismatches in the A spectra give me the feeling that something is not correct in the atmospheric composition, either some remaining ice clouds, or another problem in the atmosphere setup or surface temperature, or even in the radiative transfer. In any case, I think this is worth a discussion (the complete difference between A and B situations for plumes coming out of the same volcano at the same eruption) and some additional analysis to be able to trust the conclusions based on both groups of pixels.

Kirishimayama and Nishinoshima results: again, no comparison with literature to assess the obtained plume properties; those 2 eruptions lead to very good results in terms of RMS, with almost all RI leading to a RMS below or around 1K; therefore I am again not so convinced that the “optimal” is a “True optimum” and that we can really draw conclusions on the optimal RI based on these results; basically almost all RIs would be acceptable.

Kelud results: no comparison with literature; the RMS are overall higher with a mean between 1.131 and 3.182K and here maybe the “optimal” RMS makes sense because differences are larger, but on the other hand statistics on only 4 pixels are doubtful.

Puyehue results: the mean RMS is very high (1.989 to 7.08K) and the displayed spectra show that clearly something is wrong/missing in the model (atmosphere or RT), as for Calbuco A; again I am skeptical about the 0.33 optical thickness for 30K to 50K radiance drop. Such a small optical thickness in the plume centre does not match literature (e.g. the cited Klüser 2013, or Maes et al, doi: 10.3390/rs8020103, or Bignami et al, 2014 doi: 10.1109/JSTARS.2014.2320638). In addition, the retrieved particle size here is very small and does not match the literature at all (e.g. Bignami 2014 mention a particle size of 4 to 5µm). This needs discussion.

Results for the 4 remaining eruptions: I am not sure that there are enough pixels in the analysis to be relevant.

I find the final paragraph of the conclusions a bit strange, describing additional work done but not presented. This should either be removed or placed in the paper sections discussing the results for each eruption.

Overall the conclusions should be rewritten in light of other comments and changes to the manuscript.

Minor Comments

Everywhere: please use BT as abbreviation for Brightness Temperature, not TB (mostly in equations)

I find it slightly confusing to call the ash plume a "cloud", although the difference is clearly made with Meteo Clouds. I would suggest the use of "plume", and maybe also of "aerosols" instead of "material" when referring to atmospheric ash. At least, I would avoid "cloud parameters of the VAC" and just use "VAC parameters". However, this is just a suggestion and if the authors feel that they really prefer the terms they first selected, I do not object.

Line 22: "volcanic silicates" … here I would remove "volcanic", because all silicates have that feature (also mineral dust, for example)

Line 28: "other infrared channels": this is unclear to me, what do you mean?

Lines 28 and 29: the 2 lists of references are certainly not exhaustive, maybe use "e.g." ?

Line 97: providing CPU time without the system on which it was run has very low significance

Kelut -> Kelud

In the conclusions, I would define again the acronyms (and the RI data sets, shortly) so that people can read abstract and conclusions only and still understand. But we may also leave this call to the Editor / Typesetting crew.