

Comment on gmd-2021-30

Anonymous Referee #3

Referee comment on "Particle-filter-based volcanic ash emission inversion applied to a hypothetical sub-Plinian Eyjafjallajökull eruption using the Ensemble for Stochastic Integration of Atmospheric Simulations (ESIAS-chem) version 1.0" by Philipp Franke et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-30-RC3>, 2021

The authors present an idealised study of a particle-filter-based inversion system to obtain height- and time-resolved volcanic ash emission estimates using ash column loads from satellite. The idealised study provides a good testing ground for their inversion method, without the complication of modelling errors and incomplete observation datasets which are encountered in real situations. I think it is suitable for publication but could benefit from some improvements to the readability and complexity of the manuscript. It is an interesting study and I'm keen to see future developments on this work.

- I find the manuscript unnecessarily wordy in places and not the easiest to read. For example, the title is rather detailed - is 'sub-Plinian Eyjafjallajökull' and 'version 1.0' necessary here. Another example is the caption of Figure 2 - stating it shows the emission profile should be adequate - I would question the need to include what it is to be used for here. The authors use some complex words (e.g., 'pairwise distinct' (line 91) - does this mean independent? - 'investigated exemplary' (line 306) - what does this mean?). In places, sentences are very long. Section 3 is rather long - could it be split into subsections? Readability could be improved, I feel.
- It would help to have some connection between the theory (section 2) and the practice for this case study.
 - Line 102: c is defined as a default mass of ash in the emission package but just previously on line 92 'a unit mass of ash' is stated. I began to wonder why c was needed if it was one. Reading on, I realised that c is probably a first guess (e.g. from a prior) and that a is some adjustment factor. This could be made clearer. Again, on line 153 there is reference to 'default mass' without this ever being defined.
 - The use of the Nelder-Mead algorithm seems unnecessary since the problem here is linear and the resulting minimization problem is quadratic. The situation described on lines 114 and 115 ('cases where the function to be minimized has discontinuities or the function values are noisy') does not apply here. Again, this seems to overcomplicate the manuscript and is an example of the gap between the theory stated and the case study. Furthermore, the required restriction to only allow integer values for a for the method to be efficient (I presume the use of the word 'effective' is a typo?) is unnecessary for such a simple minimization problem which could be

solved explicitly and will no doubt introduce some uncertainty / errors. I take the authors point that the Nelder-Mean algorithm is more widely applicable but wider applications are not studied.

- Line 133: What is 'the model state'? Presumably it is either the source emissions (a times s used earlier) or the model predictions of ash column loads (both $H M[as]$ and \hat{x} (equation 2) used earlier). Can you standardise the notation? Is x in equation 3 related to \hat{x} in equation 2? If not, perhaps another variable other than x could be used for one? Also, whilst notation is consistent for x and y , the subscripts in equation 8 suggest to me derivatives or (x,y) coordinate components.
- Lines 150 – 151. Does the mention of the capability of using ensemble meteorological members add anything to this paper? This capability isn't used, and its mention could cause confusion to the reader – what does 'ensemble' refer to hereinafter? Lines 191 - 192: 'EURAD-IM comes with the adjoint code of the chemical and aerosol modules for four-dimensional variational data assimilation.' – again, is this relevant to this study?
- Detail on the error covariance matrix B is thin. On one hand you say, 'no fixed assumptions have been made for matrix B ' but later you state that ' B is chosen as diagonal'. How is the optimal value found?
- Presumably the subscript 0 in line 169 is an iteration subscript?
- It would help for sections 2.3 and 2.4 to be linked better. For example, how does the weightings and likelihood relate to the cost function?
- I'm presuming the 'ensemble mean' (line 242) is obtained from the ensemble members that are accepted by the particle filter, weighted according to the weights in equation 5? It would help the reader, to elaborate here. Also, I find the subsequent mention of 'mean' (line 249) and the 'mean' and the overbar (line 251 and equation 9) confusing. Also, on lines 263 – 264, the ensemble mean is denoted by an overbar. Can any improvements be made to help the reader negotiate these apparent different means?
- I would have liked the authors to state early on (perhaps in the abstract) that the study is an idealised study (I think the word 'idealised' is more commonly understood than 'identical twin') and therefore does not consider errors in the modelling (both in the input meteorology and in the model parametrizations) nor incomplete observations. I would also have liked to have seen more discussion of the implications of this work to real life situations (e.g., when errors will exist in the meteorological data and in the transport model and observations may be incomplete (perhaps due to the presence of meteorological cloud)). For example, the authors state that 'an assimilation window of 24 hours is sufficient in order to provide reliable forecasts', but this may not be true for a real case study.
- I suggest some thought is given to the use of the phrase 'data assimilation'. What is meant by the term 'data assimilation'? Some uses of the phrase I would refer to as source inversion methods, rather than data assimilation.
- Lines 253 – 254: Why is the RMAE calculated over points where both the modelled and observed values are above the limit? If one was to compare model forecasts for a given set of observations, the RMAE may be obtained over a different set / number of (model, obs) pairs. If, however, one was to compare the RMAE calculate over points where the observed values are above a limit, one would have a consistent set. Why is the relative error (equation 11) normalised by the model ensemble mean (similarly equation 12)? The RMAE is normalised by the observations.
- I would encourage the authors to add some model runtime information.

Minor points

- Lines 140-141: 'no assumptions of the error statistics of the model state and the observations were made'. Is this true? The likelihood function commonly includes an error covariance term (as in equation 1 and equation 6), hence I would think there are some assumptions made in calculating the weights, even if the error covariance term is the identity matrix.
- Line 26: 'Chemistry transport models have limits in estimating the emission strength'. The emission strength is usually assumed in chemistry transport models. Models are not generally used to estimate emission strengths, except in the context of data assimilation / source estimation methods or by some simple inference from observations.
- Line 31: 'analysis error' – it's not clear whether this refers to the emission estimates or the predicted cloud.
- Line 37: 'and thus making' should be 'and thus make' or 'thus making' or something similar.
- Line 44: Satellite retrieval methods also usually retrieve an estimate of the cloud height.
- Line 52: 'remains to be solved'. I would probably dispute this. Established source inversion methods for volcanic ash use atmospheric wind shear to be able to determine the three-dimensional ash cloud information from two-dimensional observations.
- Line 73: 'but also may the'. Something is wrong here with the English – perhaps remove 'may'?
- Lines 96-98: The work of Stohl et al and Kristiansen et al estimates the source emission profile (from which the volcanic ash column mass loading can be modelled) and the Bayesian method used does provide uncertainty information.
- Line 150: 'suitably' should be 'when suitably'?
- Line 187: 'on our case' should be 'in our case'?
- Line 198: 'Spin' should be 'Spinning'
- Lines 223-224: 'Contrary, vertical and horizontal mixing of volcanic ash emitted may limit the benefit that is gained by increasing the assimilation window length.' I can't see why this would be the case – can the authors explain? I can see that additional observations from increasing the assimilation window length may not provide any further information (particularly in an idealised study) but why the reference to 'vertical and horizontal mixing'? Is it because the ash cloud may be below satellite detection limits when widely dispersed? Does the benefit differ depending on whether the study is an idealised study (i.e., no modelling errors and full view of the ash cloud) or a real case study (with modelling errors and missing observations)?
- Caption Figure 4: 'in approx. 5 km' should be 'at approx. 5 km'. This typo appears in a few other places in the manuscript (e.g., lines 326, 355, 396) – it should be 'at' a height, not 'in' a height.
- Line 229, Figure 4b. It's not clear whether this is height above the volcano (which probably doesn't make sense over a spatial region) or height about ground or height above sea level? How does one associate it with Figure 4a? Does one need to know the height of the volcano? Similarly, line 231. Lines 231 – 232 and Fig 4c require some units / labels for the variables stated / shown (e.g., label on the y axis, units for temperature and pressure). Similarly, for Figure 5 and associated text. Also, height information is not specific in Figures 9-12 and associated discussion.
- Lines 279-280: 'Again, the pattern correlation coefficient does not account for deviations in the strength of volcanic ash column mass loading at locations in which the ensemble mean and the nature run differ in volcanic ash load'. It considers differences above and below the limit applied.
- Line 316: Given the same emission profile is used in each nature runs, why are these

total values different?

- Line 337: 'the mixing of volcanic ash in the atmosphere is too effective'. This study is an idealised case study so the mixing of volcanic ash in the atmosphere is represented perfectly. My opinion is that the second case study does not enable the vertical distribution of the ash emissions to be determined by wind shear and hence the filtering method yields an emission profile which is widely distributed in the vertical compared to the nature run. The first case study has significant wind shear which allows the vertical distribution of emissions to be strongly constrained, but this is not possible for the second case study.
- Line 381: It's not just the uncertainties in the meteorological fields which are neglected in this study, uncertainties in the model parametrizations (e.g., turbulent dispersion, washout, etc.) are also neglected.