Comment on nhess-2021-97
Andreas Stohl (Referee)

Referee comment on "Modelling the volcanic ash plume from Eyjafjallajökull eruption (May 2010) over Europe: evaluation of the benefit of source term improvements and of the assimilation of aerosol measurements" by Matthieu Plu et al., Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2021-97-RC1, 2021

This paper is a nice comparison of MOCAGE volcanic ash simulations to various types of observations. The volcanic ash simulations were driven with two different source terms, without any data assimilation, and with assimilation of MODIS satellite observations and aerosol lidar data. The paper suggests that the assimilation of MOIDS data has a rather low impact, whereas the assimilation of lidar data improved the results much more. The methods used are in general sound, and the paper is, except for some problems with the use of the English language, well written and organized. However, a few things should be changed before the paper can be accepted.

Major issues:

I think the vertically constant distribution of mass in the Mastin formulation for the source term is quite unrealistic. It is well known that volcanic ash clouds are usually umbrella-shaped. Thus, the unrealistic assumption of a vertically constant ash distribution puts the Mastin formulation at a somewhat unfair disadvantage in the comparison with FPLUME. Assuming a generic shape with a maximum close to the observed plume heights would be much more realistic, and may even outperform FPLUME. I see, of course, the value of using FPLUME, but often these models create a lot of variability that may not match well reality, and a smoother but still somewhat realistic profile may actually lead to better results.

Line 172: How does the variation of fine ash from 0.1 to 5% come about in FPLUME? Shouldn't the size distribution be more a function of the eruption properties, rather than processes in the atmosphere? The latter (e.g., aggregation) may also have some effect but I am surprised these results in such large variations, and I am wondering how realistic these are.

You are assimilating MODIS data but you never show them. However, this is essential for the reader to understand the effect of the MODIS data assimilation. It is surprising that the MODIS data assimilation has so little effect, so this needs some more discussion as well.

The comparison to the aircraft data is somewhat disappointing. A scatter plot of observed
vs simulated values would be much more convincing than just the mentioning of a few values in the text.

The figures are generally well presented; however, the labels in figures 3 and 7 are MUCH too small. On a print-out they are totally unreadable.

Minor issues:

I find it surprising that all simulations are so similar in terms of their FSS values shown in Fig. 6. Does that mean that FSS is not a particularly sharp measure to evaluate the performance of the model?

Line 267, last word: I think this should be right, not left, column

Line 270: What are vertical processes?

Figure 1: Clarify what exactly is shown with number of grid points contaminated with ash. Does this mean the number of vertical levels?

Language:

line 9: dispersion OF the plume

line 12: hundred kilometers downstream of

quite often singular and plural are mixed, e.g. line 26: forecasts REMAIN a challenge; line 52: source termS; line 57: performance IS compared; line 124: of the componentS of ACTRIS; line 127: each of the aerosol layerS; line 192: ash pOcket DOES not show up

line 57: are used IN the case study