

## 1 Recommendation

The authors have made a very substantial revision which significantly improved the quality of the manuscript. The title and conclusions are now in a better agreement with the content of the paper and almost all points addressed in the review were considered.

Therefore I now recommend to accepted the manuscript subject to revisions as listed below:

## 2 Specific comments

- As already pointed out in my former review, the authors discuss the situation that there is an ash layer in clear-sky atmosphere. However, in general there are water and ice clouds above/below/within the ash layer, and situations where clouds move above the ash layer. In such situations the ash layer can't be detected by the satellite algorithms (e.g., Prata and Prata, 2012), leading to wrong assimilation input. I expect that in such cases the model run without assimilation will perform much better than the assimilated one. I suggest to (1) add a comment on that in the paper, and (2) discuss how these errors are treated in your equation (8).
- P 2, L 19: Reference to Lu et al., 2016a is not an adequate reference dealing with satellite observations and does not bring added value to the manuscript. But my main criticism is that it is not a good practice to insert unrequested references to own papers during the review process, even without highlighting the change in the "track change" document. See also [http://www.atmospheric-chemistry-and-physics.net/about/publication\\_ethics.html](http://www.atmospheric-chemistry-and-physics.net/about/publication_ethics.html).
- P 3, L 9 – L19: I really appreciate this nice overview on the vertical extensions of ash layers reported in literature.
- P 3, L 20 – 22: The sentence "*Thin ash clouds, by their nature are of less concern to aviation because such clouds would be traversed rapidly avoiding the possibility of particle build-up that might lead to engine failure.*" is problematic as vertically thin layers are not necessarily traversed rapidly in the horizontal direction. Although the statement might be true in many cases it is not difficult to construct a (realistic) scenario where it is not true. Also, "*avoiding*" is a rather strong word here. I suggest to use "*reducing*" or "*minimizing*" in this context.
- P 3, L 20 – 22: "*From a modeling perspective lack of vertical resolution in model wind data makes it not useful to make the cloud depth any less than 500 m.*": I don't think that this argument is conclusive here. The transformation from 2D satellite data to 3D fields is done prior any contact with the model wind fields; as you state "*The outcome of SOO can be considered as preprocessing to the satellite data assimilation system.*".

- P 3, L 20 – 22: To me it still remains a shortcoming that the authors do not consider the full range of observed vertical thickness, 100 m – 3 km. Although I now agree that the vast majority is within the considered range, there is no scientific reason to restrict the analysis on these cases. (In the vast majority of days there is no volcanic ash in the European airspace but nobody would argue that this is an adequate reason to not consider such rare events...)
- P 5, L 8 – 9: "*lowest possible thickness*" should be "lowest expected thickness" or "lowest considered thickness"
- P 6, L 8 – 9: Figure 1c does only report errors of ash load for pixels where volcanic ash is detected. What is about the error of pixels where no ash is detected? There might be ash as well (for example ash concentration below the detection threshold or ash which is hidden by meteorological clouds). I think the latter should also go into the  $ML_{error}$ ?
- P 8, L 22: It is not an objective argumentation to select a small area of the whole region (like Netherlands here), where good agreement is achieved, and to discuss the numbers in a way (observations 3.1 mg/m<sup>2</sup> versus EnSR values of 2.9 - 3.2 mg/m<sup>2</sup>) implicitly suggesting that the assimilation performance is that close to observations in general. This suggestion is even supported by using "*for example*" in L 22. However, the Netherlands are one of very few areas where you achieve that good agreement. There are even regions with similar size like the Netherlands, namely the fjord area of Norway or the area around Norfolk/Suffolk in England where the non-assimilated run is much closer to observation than the assimilated run. These findings should be reported in a fair and objective way as well.
- P 9, L 15 – 25: The aircraft measurements represent PM10 concentration at inlet position (x,y,z,t) while the model gives values which are averages of the concentration for a 0.25° x 0.1° x 1 km grid box. Please justify why is it appropriate to compare these two values.
- Figure 1: While there is a continuous colour scale in Fig. 6b, there are only five colors in the figure. Is the top height retrieval discrete in a way that it turns out only discrete values of 4, 5, 6, 7, and 8 km? I don't think so, as the KNMI height product documentation shows continuous values of top heights.
- Figure 1: This is true for Fig 1a-c, but also several other figures in the manuscript: While zooming into the pdf file it seems that your plot consists of (intransparent) overlapping colored dots which is not the state of art how to plot a 2D satellite retrieval. I.e., the final figure depends on the sequence of the underlying data. I suggest to revise these figures.
- **Figure 6: The measurement values plotted in Fig. 6b are not consistent with the (same) measurement data reported in Weber et al., 2010**, see <http://dx.doi.org/10.1016/j.atmosenv.2011.10.030>. For example, at 10:00 UTC Weber et al., 2010 reports on PM10 concentration values slightly below/above 200 µg/m<sup>3</sup> whereas Fig 6b indicates values at 10:00 UTC of 130 g/m<sup>3</sup> which is an approx. 35% smaller

value. This is important here, as this inconsistency pimps the general agreement of aircraft measurements and EnSR result in the figure.

- Figure 6: Same is true for Fig 6c where Weber et al. 2010, Fig. 8 report on values of approx.  $300 \mu\text{g}/\text{m}^3$  at 13:05 UTC while the authors data report values of  $200 \mu\text{g}/\text{m}^3$ . Please clarify.

### **3 minor comments**

P 4, L 17: "caled" should be "called"

P 7, L 4: broken sentence structure.

P 8, L 8 + L 9: comma seem to be misplaced.