We thank the reviewer for his thorough review of our paper and his critical and constructive comments. The reviewer states that "*While the subject is clearly of interest and related studies merits publication in ACPD, there are some substantial shortcomings which need to be removed before final publication".* Following his suggestion on the structure and the content of our paper we have totally revised it. Moreover, we carried out additional model runs and focus on a more detailed evaluation of our time lagged ensembles.

Rather, some type of an analysis of the event is provided. Further, probabilities of threshold values were calculated by time lagged forecast ensembles. The way they are constructed by an ad hoc sequence of intermitted forecasts, some account on meteorological uncertainty is given, which are presumably not the parameters of lowest skill in the context of volcanic eruptions. While the variable emission source strength and height are markedly more uncertain, along with sedimentation and wet deposition velocities and microphysical parameters of the particles, a properly designed ensemble with a generation mechanism, which reflects the poor state of knowledge is lacking. In contrast to the explicit claim in the abstract, only lagged forecasts ensembles cannot provide probabilistic information used for decisions with fundamental consequences like closed air spaces. Hence, probabilities or uncertainties provided do not trustworthily reflect the uncertainty of the forecast. Rigorous applications of stochastic simulation principles are yet a matter of ongoing research. Nonetheless, the claim made in this study should at least acknowledge the limited frame of the probability accomplished.

We can follow the reviewer's comments to a large extent. We know and demonstrate in our revised paper that the source strength causes the highest uncertainties. These uncertainties are however considerably reduced when observations are used to calibrate the model. Then the calculated probabilities of threshold violation are mainly determined by uncertainties in the meteorological forecasts. As we are using an online coupled model system uncertainties in precipitation are transferred into uncertainties of washout etc..

.

In its present form the description of the system set-up is not sufficiently complete and precise to allow their reproduction by fellow scientists.

We improved the description of the setup of the model runs.

The abstract provides a concise, probably too concise summary. It is however too simplistic with judging the simulation in a simulation in a "nearly perfect agreement", which is of little avail to estimate the model skill after the later tuning efforts.

We agree and replaced nearly perfect agreement by the correlation coefficient.

1. In the abstract : What are the “measured data” used for calibration?

We have included an extra section to explain the calibration.

2. Page 2, last paragraph: it is claimed what has been adapted in the forecast system during the volcanic eruption. Please add citations of the work, supporting the claim, or remove paragraph.

We changed this.

3. Page 2, 2nd paragraph: The parameters to be forecasted include volcanologic quantities like size distributions and vertical distribution of effective source heights. These parameters, while in the realm of volcanology, cannot be forecasted with usable accuracy for atmospheric simulations. Please reformulate or remove this paragraph.

We followed the reviewer’s suggestions.

4. Page 3: 3rd paragraph (1. in section 2): The definition of online vs offline; is there a feedback from volcanic ash loads to the meteorological parameters? If not, the claim of being online is obsolete, as “off-line” configured systems can reproduce the same results. Please explain.

Also a treatment of the feedback would be possible we switched it off due to a lack of data that would allow e.g. the modification of cloud droplet number. Still we think that it makes senses to use the phrase online coupled as in most offline coupled model systems an update of the meteorological fields happens in time intervals of at least an hours while in our case it is update every time step.

5. Page 4, subsection 2.2: ad the web site of VAAC London. Consider the time delay and limited accuracy and temporally coarse resolution (6 hours) of available ejection heights. Keflavik radar was the most reliable information in those days, but they were not directly available. Explain your configuration set-up for operational use.

This is better explained in the revised version.

6. page 5, 2nd para.: Give a reference of the DWD forecast set-up. Cite the reference paper for COSMO-ART here.

We have added a more detailed description and added a reference.

7. page 5, last parag. of section 2: I do not understand (e.g. “hindcast after 9 days”). Please reformulate.

We have reformulated this section.

8. Page 5, discussion Fig 5.:"captures : : : quite well". Could you please be more precise on this?

We have added an additional Figure with a quantitative comparison.

9. Page 6 discussion of Fig. 3 in relation to Fig. 1 Gasteiger et al 2011. Please combine these information in a single graphics. It is hard for the reader to follow the explanation.

We have added the Figure of Gasteiger et al..

10. Page 6, 3rd parag.: Subsidence would be better discussed by considering the stability of the air and moving of stable layer patterns associated with the high pressure system. A clear distinction of an Eulerian and a Lagrangian viewpoint for rating the subsidence is crucial. See also the discussion in the conclusions page 10, end of 1st paragraph.

The reviewer is right. We have reformulated this section.

11. Figure captions are generally kept too scantily.

We have improved the figure descriptions.