Response to Referee 1

General comments

The manuscript describes the response of the surface temperature to major volcanic eruption. The manuscript is in the scope of ACP. The authors applied multiple linear regression analysis (MLR) to retrieve the response from several atmospheric reanalysis products. Several studies of the climate response to volcanic eruptions based on model and observation results have been published before. The novelty of this work consists of the exploitation of multiple reanalysis data sets. However, the applications of similar observation data for all reanalysis products leads to visually similar, patchy and marginally significant response pattern making the manuscript difficult for the readers. The problem is enhanced by a high number of very small panels and to simplified description of the figures. The manuscript is interesting and should be published, but the authors should make it more attractive for the readers. Otherwise, they great efforts will not be fully appreciated.

Thank you very much for your review. We made the panels for geographical distributions much larger for the open discussion, and believe that the current version of the panels available on the website is acceptable in terms of their size. We will also improve the description of the figures – please see below.

Major issues

1. I do not completely understand the motivation of the SST application as a proxy for MLR. Can be part of the signal lost due to this procedure.

The primary motivation of this approach is to remove the component of El Niño events right after the three volcanic eruptions as discussed in the first paragraph of the Introduction. It is common practice to use SST data to describe the El Niño Southern Oscillation (ENSO) time series, but there are actually several indices proposed/used as described in the fourth paragraph of the Introduction and in Table 2. The ENSO is an atmosphere-ocean coupling phenomenon affecting primarily the tropical Pacific. But, as again discussed in the fourth paragraph of the Introduction, there are similar basin-wide atmosphere-ocean coupling phenomena in the tropical Indian Ocean and in the tropical Atlantic Ocean. They are less impactful compared to the ENSO, and thus often neglected for upper air (e.g., stratospheric) studies. However, it is considered that they have a direct impact on the surface temperature over those oceans. Thus, we considered these as well.
Finally, some recent studies have revealed a wintertime teleconnection rooted in the Arctic Ocean that

influences East Asia and North America (as again described in the fourth paragraph of the Introduction.)

These are the reasons why we used SST data. In short: to remove those influences on surface air temperature that can be considered independent from the volcanic influence.

As discussed in detail in the final paragraph of Section 3 (Method), our approach is imperfect. A fraction of ENSO and other coupled atmosphere-ocean variability described using SST data may be inextricably linked to or even emerge from the volcanic response through forced changes in atmospheric circulation. Thus, the "volcanic response" reported in this paper should be regarded as the component of the volcanic forcing that is not mediated by coupled modes of atmosphere-ocean variability.

2. Would it be possible to explain how the statistical significance was estimated? In text it is said that 1 SD criteria was used, but is it not too weak?

We cannot use a rigorous statistical test for this study. Instead, we compare the residual during volcanic eruptions to its SD as a measure of how unusual the values are. This 1-SD criterion is used in previous studies on the surface temperature response to the Pinatubo eruption (e.g., Kirchner et al., 1999; Yang and Schlesinger, 2001).

3. In the introduction I recommend describing expectations for the surface temperature response to volcanic eruptions. Then it will be easier to describe the obtained results.

In the first paragraph of the Introduction, after the first two sentences (at page 2, line 3), we will add the following sentences:

The increased concentration of aerosols in the stratosphere causes a net negative radiative forcing at the surface (Robock, 2000), resulting in cold surface temperature anomalies when averaged globally or over the tropics. The geographical distribution of the surface temperature anomalies is, however, found to be much more complicated. Robock (2000) reviewed observations and theory of winter-time warming over the Northern Hemisphere (NH) continents (or the wave pattern of warm/cold anomalies) that result from changes in the tropospheric and stratospheric circulations after large eruptions. The surface temperature response at the regional scale is thus not only influenced by the direct radiative

forcing but also by the dynamical response of the atmospheric circulation. Studies on the geographical distribution of the surface volcanic response all show complex patterns of cooling and warming (e.g., Kirchner et al., 1999; Yang and Schlesinger, 2001).

Minor issues:

1. Page 8, first paragraph: How to explain similar anomalies during the periods w/o volcanos. Is it the results of similar observations used for assimilation?

For the case of the current study, we can say that very similar observations are assimilated for the full input reanalyses and for the surface input reanalyses, respectively. In this study, we do not attempt to explain the anomalies during the periods without volcanic eruptions.

2. Page 8, line 28: Why for two reanalyses only? Do the others produce similar results?

Yes, other reanalyses show similar results. We will add this information at page 9, line 19.

3. Page 10, first para: It is difficult to read. It looks like text description of the pattern visible from the figure

We described the features of the results, organizing the discussion similarly to previous figures; i.e., starting from the tropics, then NH extra-tropics, then SH extra-tropics, and then in and around Antarctica. These features are of course visible in the figures, but we think such a text description is necessary and useful for some readers.

4. Page 11, line 1: More sophisticated methods are not always more accurate.

Our method is more sophisticated in comparison to others by the removal of the influence of known forcings such as ENSO, Indian Ocean variability, Atlantic Ocean variability, etc. We do not claim it is more accurate.

5. Page 11, second para: Not conclusive. What are the conclusions from the discussion?

We will add the following sentence at the end of this paragraph:

Therefore, the cause of the transient cooling event in 1976 needs further investigation.

(When the ERA5 reanalysis is extended back to 1950 (most probably next year), we will re-investigate this issue.)

6. Page 11: Too descriptive to my taste. I recommend improving the text.

We assume that you are referring to the two paragraphs regarding Figures 6 and 7. They are indeed descriptive in nature. We have already made efforts to keep this part short, so that information is not duplicated and readers can proceed quickly to the next discussion.

7. Page 11, lines 24-26: Please, discuss why the difference between considered periods is so important.

We added this paragraph because this analysis provides a very good opportunity to test the robustness of the results. We will add the following sentence in the beginning of this paragraph:

The 1958–2001 analysis provides us not only the Mount Agung response but also an opportunity to test the robustness of the results following the eruptions of El Chichón and Mount Pinatubo.

8. Page 12, line 11: Maybe because it is not robust.

At least our results for the six reanalyses show very similar characteristics as described in the previous paragraph.

9. Page 12, line 22: “All known externally-forced”. Could be internal than?

We will remove the word “externally” since it could be confusing. What is meant by the “forced component” is described later in the same sentence.