Response to Dr. Raphael Neukom

We thank Dr. Neukom for his constructive comments, which we agree will help improve the manuscript. Below we respond in detail to each comment. The referee comments are italicized:

1) Capability of the model to simulate oxygen isotopes in precipitation over South America. The conclusions of the papers stand and fall with the ability of the model to simulate oxygen isotopes in precipitation (d18Op) in general, not only in response to volcanic events. This is assessed in Figure 6. While the model appears to be quite good in simulating the seasonal cycle, I am not sure whether this analysis is sufficient to be confident about the skill of the model. I am not an expert in this field but I could think of the following options: Literature: It is possible that this has been assessed the literature describing GISS ModelE2-R. If this is the case I suggest including a paragraph reviewing this

The performance of the model, at least the previous version (ModelE-R), and its capability in tracing isotopes through the hydrological cycle has indeed been tested and presented in the literature. Schmidt et al. (2007), LeGrande and Schmidt (2008, 2009), Lewis et al. (2010, 2013, 2014) and Field et al. (2014) have tested the stable isotope results from this model against observations from satellites, IAEA data and proxy records. An earlier version of the GISS model has also been validated specifically over tropical South America (Vuille al., 2003a,b). We have clarified this in the revised manuscript and added a more detailed discussion of this aspect.

GNIP data and volcanic events: The authors state that data availability is not sufficient to perform a reasonable composite analysis for the volcanic events. Is there not sufficient data available to at least show the response to, let's say, the most recent event (Pinatubo)? If this is not the case, which I suspect from reading page 3386, would there be a chance to analyze this based on composites from other years? For example, one could make 18Op composites for the warmest/coldest/driest/wettest years during the period of reasonable data coverage and compare to the model data. This would require composites of reasonable size, so given that I don't know about the exact data situation in the GNIP data, I cannot be sure that this is feasible. But such an analysis would also be helpful to interpret the paleo results (see next point).

This is a valid suggestion, but rather difficult to achieve, given that the GNIP data suffer from substantial temporal gaps, and that data coverage is extremely sparse during the time of the Mt. Pinatubo eruption. Given this, and also the strong unforced variability (e.g., ENSO occurrence coincident with these eruptions), assessing the GNIP post-volcanic imprint yields inconclusive results. In fact, we have tried this approach for multiple events and the spatial structure of the isotope field resembles the well-known ENSO imprint on d18O (Vuille et al., 2003a) and with substantial spread between events.

2) Similarly to the last point, it would be good to simply see how the modeled 18Op responses to climate in the study area (not only during volcanic events). This not only to assess the skill of the model, but also to better understand the results. This could be tested using the control simulation. For example, one could select years with high/low temperature (but normal precipitation) and years with high/low precipitation (and normal temperatures). How do the 18Op anomalies look like? Most probably as expected from theory and described in the text but I think nevertheless it would be helpful to have an illustration confirming this (for example in the SM). This is relevant particularly for the explanation of the seasonal asymmetry shown in Fig. 10, which I think is one of the key findings of the paper. The interpretation provided by the authors (that the strong temperature response masks the precipitation signal in some seasons and regions) could be supported by this analysis.

The suggestion to stratify the model and observations by wet/dry and warm/cold years is interesting. Given the comments here and by R2, we will give much further attention to the model validation segment of our paper for the South American climatology and isotope physics. We agree that, in addition to better citing the relevant literature (as discussed in point #1), a more thorough treatment of how the model handles the isotope field is required.

*3. This is a modeling study. However from the title and abstract this does not become clear and one could still think that proxy data are also used. In the abstract it says ":::
and allows for a direct comparison between GISS simulations and paleoclimate proxy archives". This comparison is not provided in the paper, so I suggest to clarify this (e.g. by saying "future comparisons") and move this statement to the end of the abstract (as kind of an outlook). Even after reading the introduction (with a specific section on reconstructions), one could still expect proxy data to be used somewhere in the text. Given proxy data are mentioned repeatedly in the paper, the reader can hardly wait to see how the anomalies of the proxy data look like following the LM eruptions :::I suspect (and hope) that the authors plan to show this and the proxy-model comparison in a subsequent study, and this should be clarified as early as possible. The importance of this paper for such future analysis can then be stressed (again) in an outlook at the end of the manuscript. I would like to emphasize that I do not think that clarifying this in the title and/or abstract will make this paper less appealing.*

We agree with these recommendations. We have changed the title to: " The influence of volcanic eruptions on the climate of tropical South America during the last millennium in the isotope-enabled GISS ModelE2-R." The abstract has also been changed to make it clear that the purpose of the study is to focus on the volcanic response over tropical South America. Finally we do indeed intend to use some of the results obtained here to inform the interpretation of isotopic signals in high-resolution isotopic proxy records from South America, where we suspect to see volcanic signals, but these analyses are beyond the scope of the study at hand.

4. The paper focuses on tropical South America. Although the entire continent is shown in the figures, the analysis and interpretation is clearly focused on the tropics (and maybe subtropics), which makes sense (e.g. given the distribution of isotopic proxy data). Again, I suggest clarifying this in title and abstract. And again I think doing so will not make the paper less attractive but help the reader to know what to expect. The authors may even consider removing the parts of the paper describing extratropical features, for example Figure 13, to make the paper more focused. Suggest to replace "South America" by "tropical South America" in many instances of the paper.

We agree. As mentioned under point 3, abstract and title have been revised to clarify that this is a modeling study and the focus is on tropical South America. We also agree with removing the discussion of the extratropical aspects. Hence we have removed section 3.2.4 from the revised manuscript.

5. *I suspect that the response to volcanic forcing in tropical South America in the instrumental data is not as clear as described in the text, particularly for temperature. None of the obs-panels in Fig. 4 shows consistent negative anomalies in the region except for Pinatubo in JJA, where the signal appears to be rather weak. A composite analysis could clarify this picture. While Figure 3 impressively shows the consistency in observational and model data in the tropical belt, I think an identical (or similar) figure for tropical South America would be more helpful for this paper (the current fig. 3 could be provided additionally or moved to the SM). The authors may also consider showing an instrumental composite anomaly map for South America to allow a good comparison with Figs 7 and 8. (see my other point regarding figure 4 below).*

We understand the reviewer's concern, but too much emphasis on the regional scale (for the instrumental record) is not helpful in this case. Focusing on tropical South America rather than the entire tropics, will amplify the issue that El Niño tends to mask the volcanic signal over the observational period, given that ENSO has an exceptionally large impact on tropical S. American climate (Garreaud et al., 2009). This problem is not alleviated by compositing events, even if the ENSO signal is removed (e.g. through linear regression). We have made such attempts to remove ENSO, but the signal-to-noise ratio remains very low and the residual signal only perpetuates a false representation of the "typical" volcanic expression and it is not suited to test model performance.

Minor points:
6. Abstract line 4: consider including "instrumental" before "observations" to clarify that proxy data are not used in this study.
7. Abstract lines 5-9: This is a very long sentence. I suggest to split up.
8. P. 3377 line 8: Although this is described in more detail below, I think the statement "most important" should be accompanied with a literature reference (or "see below").

Thanks you; we will include all the suggested modifications in the revised manuscript.

9. P. 3377 line 17ff. Although see Zanchettin et al. (2012) for decadal-scale responses to volcanic eruptions, at least in the North Atlantic sector.

There are several hypotheses that exist for how the decadal-and-longer timescale response to volcanic eruptions manifests itself. Part of the response may be simple mixed-layer physics (McGregor et al., 2015) without the need for appealing to an anomalous circulation or sea ice feedbacks. While this is admittedly an interesting subject of research, we do not consider it very relevant for our paper, and believe that giving it too much attention would distract from the main message of our paper.

10. P. 3379 line 25. Do the authors mean "records" instead of "archives"? The number of archives offering high-resolution proxy data is not increasing that much.

We will change the word to "records."

11. P 3380 Section 1.3 does not describe the climate of the entire continent so suggest to change the title to "tropical".

We agree, and will modify this where appropriate, including the section 1.3 title.

12. Although I somehow like the expression, "rather Mars-like" does not appear to be a very scientific description. I leave it to the editor to decide whether it is appropriate. Given the point above (and point #4), the first paragraph of this subsection could also be considered to be entirely removed.

We will remove the "Mars-like" description following both yours, and Reviewer #2's suggestion. We do wish to preserve a short motivating description of the continent.

13. P. 3381 last paragraph. To be exact, the ENSO response described here is only valid for the SAMS-affected regions. There are parts of tropical SA that have a different (reverse) response (e the Pacific coast area with strong wet anomalies during El Niño events).
14. P.3382 lines 17-21. This is a long sentence, consider splitting up.
15. P. 3383 line 5: One or more References for the amount effect would be helpful.

Thank you; that is absolutely correct. We will modify all text accordingly.

16. P. 3383 line 7: is there an "at" missing after "be"? Or maybe use the word "occur" instead.

Yes, we can insert "at" in the text.

17. P. 3382 line 12 and P. 3383 line 19: I think the use of the terms "Medieval Climate Anomaly" and "Little Ice Age" is generally not appropriate and precise...

We agree; we will instead include an approximate date range for the specific claims in the text.

18. P. 3389 line 2: The linear time trend is later also subtracted from the data to remove the global warming signal or why is it included in the regression?

We do not remove a trend later, we only explain the data using a trend and ENSO as independent variables at each grid point, and remove the (lagged) ENSO effects. As it stands, if the superposed epoch analysis were plotted for a larger number of “prior years” (e.g., year -30 to 0), then the trend would be apparent.

19. P. 3390 line 7: The “cooling over much of the globe” is not really visible in the obs panels (expectations often bias our interpretation. Therefore, I showed the graph to persons not knowing what it shows and they confirmed that it does not visibly show more blue than red). Unless it can be undermined with numbers, this statement should be removed. Potentially, the signal gets clearer if the three events are combined into a composite? This could be added as an additional panel in the bottom of the figures (see also point 5 above).

We thank you for providing yours (and a number of other) eyes to keep our interpretation honest. We will add an additional plot and discuss/quantify our statement more thoroughly.

20. P. 3390 line 22: Please specify what “this” refers to.

We refer to the model simulating the state/amplitude of ENSO at the same “time” as observations. We will clarify this aspect.

21. P. 3392 line 8: What are the composites compared against in the t-test?

At each grid point, we create two lists (“non-eruption” and “post-eruption”) values following the definition in the methodology section of our manuscript. Values for each event are expressed as anomalies relative to the local non-eruption climatology to remove the possibility of low-frequency variations, and the list includes data for all events and all ensemble members to maximize the number of values to perform the t-test.

22. P. 3394 line 1: Although see Greve et al. (2014) regarding the (non)validity of the “dry gets drier” hypothesis.

It is true “dry gets drier” is not applicable everywhere, especially over land. We will clarify this in our manuscript. However, the statement referred to the large-scale tropical/subtropical atmosphere including the ocean, where it tends to be a useful first-order description of the large-scale net precipitation changes under global warming.

23. P. 3396, last paragraph: I think Figure 12 could be moved to the SM. I was missing confidence intervals in Figure 9. These could be inserted by shading the 95% range of the distribution from the random composites in Fig. 12. This would make Fig. 9 much stronger and the additional information in Fig. 12 would then be minor so that it could, in my perspective, be removed from the main manuscript.

Figure 9 and 12 still convey different pieces of information. Figure 9 shows the ensemble spread (and mean) response in temperature/precipitation for each event (the variable plotted against AOD). Figure 12 emphasizes the composite mean (and the likelihood of the composite response being realized by chance in the control simulation). We would like to retain both figures.

24. P3397: I think section 3.2.4 could be removed…

Yes, we agree that this section can be removed, and by extension, Figure 13.

25. Figure 1. Suggest to mark the eruptions that are finally used to create the composites with a different color in the top panel.

We will do this.

26. Figure 3: I think this Figure should contain confidence intervals, so the reader can see what magnitudes of anomalies are significant. A standard approach in superimposed epoch analysis plots is to show the 95% range of years no affected by an eruption.
27. Figure 3: The positive anomalies in instrumental precipitation between ca. 1.8 and 3.5 years after the eruption appears to be about as large as the immediate drying response. Do you think this is an artifact? Is it seen in both eruptions? Any reference to this in the literature? Again, indicating the significance threshold could help here.

Reviewer 2 also raised the same point. We will modify the figure to improve the statistical presentation. There are quite different responses to both eruptions in precipitation, so the composite may amplify/mask anomalies in ways not representative of either eruption.

28. Figures 7,8,10: Include "anomalies" to the color bar caption. I think it is worth mentioning in the caption that only significant results are shown (at least that's how I understand it from reading page 3392)
29. Figure 9. This figure should also include a significance threshold and this could be taken from Fig. 12 as mentioned above (point 23).
30. Figure 12: The blue colors are hardly visible and somehow masked by black in the print version of the manuscript.

Agreed on all points, we will modify the manuscript accordingly, thank you.

References cited

Field, R.D., D. Kim, A.N. LeGrande, J. Worden, M. Kelley, and G.A. Schmidt, 2014: Evaluating climate model performance in the tropics with retrievals of water isotopic composition from Aura TES. *Geophys. Res. Lett.*, 41, no. 16, 6030-6036, doi:10.1002/2014GL060572.

Garreaud, R.D., Vuille, M., Compagnucci, R., Marengo, J., 2009: Present-day South American climate. *Palaeogeogr., Palaeoclimatol., Palaeoecol.,* 281, 180-195, doi:10.1016/j.palaeo.2007.10.032.

LeGrande, A.N., and G.A. Schmidt, 2008: Ensemble, water-isotope enabled, coupled general circulation modeling insights into the 8.2-kyr event. *Paleoceanography*, 23, PA3207, doi:10.1029/2008PA001610.

LeGrande, A.N., and G.A. Schmidt, 2009: Sources of Holocene variability of oxygen isotopes in paleoclimate archives. *Clim. Past*, 5, 441-455, doi:10.5194/cp-5-441-2009.

Lewis, S.C., A.N. LeGrande, M. Kelley, and G.A. Schmidt, 2010: Water vapour source impacts on oxygen isotope variability in tropical precipitation during Heinrich events. *Clim. Past*, 6, 325-343, doi:10.5194/cp-6-325-2010.

Lewis, S.C., A.N. LeGrande, M. Kelley, and G.A. Schmidt, 2013: Modeling insights into deuterium excess as an indicator of water vapor source conditions. *J. Geophys. Res. Atmos.*, 118, no. 2, 243-262, doi:10.1029/2012JD017804.

Lewis, S.C., A.N. LeGrande, G.A. Schmidt, and M. Kelley, 2014: Comparison of forced ENSO-like hydrological expressions in simulations of the pre-industrial and mid-Holocene. *J. Geophys. Res. Atmos.*, 119, no. 12, 7064-7082, doi:10.1002/2013JD020961.

McGregor, H. V., M. N. Evans, H. Goosse, G. Leduc, B. Martrat, J. A. Addison, P. G. Mortyn, D. W. Oppo, M-S. Seidenkrantz, M-A. Sicre, S. J. Phipps, K. Selvaraj, K. Thirumalai, H. L. Filipsson, V. Ersek, 2015: Robust global ocean cooling trend for the pre-industrial Common Era, *Nat. Geosci.,* 8, 671-678, doi:10.1038/ngeo2510.

Schmidt, G., A. LeGrande and G. Hoffmann, 2007: Water isotope expressions of intrinsic and forced variability in a coupled ocean-atmosphere model. J. Geophys. Res., 112,D10103.

Vuille, M., Bradley, R.S., Werner, M., Healy, R., Keimig, F., 2003a: Modeling d^{18}O in precipitation over the tropical Americas: 1. Interannual variability and climatic controls. *J. Geophys. Res.,* 108, D6, 4174, doi:10.1029/2001JD002038

Vuille, M., Bradley, R.S., Healy, R., Werner, M., Hardy D. R., Thompson, L. G., Keimig, F., 2003b: Modeling $d^{18}O$ in precipitation over the tropical Americas: 2. Simulation of the stable isotope signal in Andean ice cores. *J. Geophys. Res.,* 108**,** D6, 4175, doi:10.1029/2001JD002039