RC 5 with responses

The study investigates the likelihood of a slowdown in the SPG around the onset of the LIA being linked to the 1600 Huaynaputina volcanic eruption. In order to resolve this issue, the authors attempt to integrate evidence from model-based simulations of past climate conditions with proxy-based paleoclimatic reconstructions and historical records. Despite the inconclusive results, the study highlights both the advantages of adopting an interdisciplinary approach as well as the challenges and limitations of bringing together and interpreting various sources of information.

General comments:

In my opinion, the multi-disciplinary nature of this study represents a considerable strength of this work. In general, the manuscript is well written and the findings are presented in a clear and logical manner. The evidence is interpreted objectively and the authors clearly acknowledge the limits of the analysis. From the presented results, conclusions are drawn to the extent that the simulation, reconstruction and limited observational data from the period allow. However, the unconventional structure is rather confusing since the introduction, methods description and some of the results are all blended together, and this also makes it somewhat difficult to distinguish for example what was done in previous studies and what represents original analysis. The authors should therefore seriously consider whether restructuring the manuscript in a more conventional format would be beneficial.

We would like to thank the reviewer for these generous comments on the manuscript. For the most part, the manuscript adopts a conventional structure: Introduction, Methods, Results, Discussion, and Conclusion. The only exception is an additional section (section 2) explaining the modeling and SPG-shift mechanism in the previous 2017 Moreno-Chamarro et al. studies. This additional section prevented the introductory section from becoming overly long and difficult to follow. It also enabled us to explain the previous modeling for a wider audience, including paleoclimatologists and historical climatologists, in keeping with the interdisciplinary scope of the article.

We propose to clarify at the beginning of section 2 that that section explains previous modeling and results, then at the beginning of section 3 that that section introduces the new methods aimed at investigating whether the previously described mechanism could have been triggered by the 1600 Huaynaputina eruption.

Currently, a large part of the discussion is dedicated to discussing the historical / societal impacts of cold conditions at the end of the 16th and during the early 17th century. Greater focus on integrating and discussing the results of the modeling, proxy and historical datasets in more detail would be helpful.

We can expand the discussion in section 5.1 to review the results in greater depth, including the overall agreement of the reconstructions with each other, the relative contribution of each reconstruction to our picture of climate and environmental conditions ca.1600, and the challenges associated with comparing simulation and reconstruction data (which the reviewer alludes to in the following comment).

Another important point is recognizing and acknowledging discrepancies between model-based simulations with proxy-based reconstructions, which has consequences for understanding uncertainty and the overall reliability of these data sources. This issue is highlighted for example by Figure 6, which shows poor spatial agreement between modeled and reconstructed temperatures. Model simulations are often associated with high uncertainty particularly in relation to post-volcanic cooling and, for example, over-estimation of the magnitude of post-volcanic cooling by some models has been known to occur (e.g. Chylek *et al.*, 2020; Hartl-Meier *et al.*, 2017). Better understanding of some of the shortcomings of these datasets and limitations in their utility within the context of this study could be achieved by exploring a broader set of model simulations or model types to help disentangle the possible influence of model bias and a more detailed examination of the proxy-based temperature reconstructions would also be helpful in this regard.

Some possible sources for discrepancies include: the volcanic forcing itself, which is a simplification and a reconstruction of what happened in the real world; the sensitivity of the model to the forcing; and the impact of internal variability, which might be important enough so that model and data might never be found in agreement for a relatively small ensemble size (which is our case). We will expand section 5.1 to better acknowledge these limitations in the study. However, it would not be within the scope of this study to further test model sensitivity to volcanic forcing.

It is also necessary recognize the potential importance of background climate conditions in modulating the (cooling) response of the North Atlantic to large volcanic eruptions based on the state of the climate system. In relation to this point, the role of internal variability and specifically the potential role of the North Atlantic Oscillation (NAO) in the initiation of SPG weakening and cooler conditions in the north Atlantic sector remains a subject of debate (e.g. Trouet *et al.*, 2009; Lehner *et al.*, 2012). For this reason, some type of examination and discussion of the modes of atmospheric variability in the north Atlantic within this context would be helpful.

We thank the reviewer for drawing this issue to our attention. We propose to expand discussion of the NAO/AO between lines 149 and 150. As discussed in Hernandez et al. (2020) the state of the NAO remains uncertain in this period, but both the Ortega et al. 2015 and Trouet et al. 2009 studies indicate roughly average NAO index values in the 1590s, declining in the decade following the 1600 eruption. Thus, changes in the NAO would not appear to be a strong explanation for the cooling before the eruption; nor do we apparently see the usual post-eruption NAO+ response following Huaynaputina. We can also include a plot of the NAO index during 1590–1610 in the different ensembles (that is, with and without volcanic forcing and with and without an SPG shift) to illustrate that neither phase dominates either before or after the Huaynaputina eruption.

Reference:

Hernández, Armand, Celia Martin-Puertas, Paola Moffa-Sánchez, Eduardo Moreno-Chamarro, Pablo Ortega, Simon Blockley, Kim M. Cobb, et al. "Modes of Climate Variability: Synthesis and Review of Proxy-Based Reconstructions through the Holocene." *Earth-Science Reviews* 209 (2020): 103286. https://doi.org/10.1016/j.earscirev.2020.103286.

One obvious limitation is that most of the presented evidence for the SPG shift is either indirect / circumstantial or entirely model-based. Although the study provides a compelling narrative characterizing anomalously cold conditions in the early 17th century, a certain leap of faith is currently required to link an SPG mode shift to these changes. In any case, more information would be required to clarify the relationship between the eruption, short-term and long-term cooling and how these events and changes relate to the state of the SPG. Ultimately, there are limits to the answers that modeling can provide and additional more direct proxy data would likely be required to better understand the dynamics of oceanic circulation and atmospheric dynamics during this period to more precisely pin down the timing, duration and extent of the purported SPG slowdown. Perhaps then it would be possible to confirm or refute the attribution of the observed longer-term cooling in the early 17thC, and by extension the initiation of an SPG slowdown, to a volcanic trigger.

The evidence for the SPG shift was provided in the 2017 Moreno-Chamarro et al. studies. This included testing against long-term mainly decadal- to multi-decadal-scale paleoclimate reconstructions. Those reconstructions were insufficient to determine whether the Huaynaputina eruption could have been the trigger for an SPG shift. The new data from high-resolution proxies and historical observations examined in this study enables more precise specification of conditions ca.1600, which raised the possibility of examining whether the previously proposed SPG shift could have been triggered by the Huaynaputina eruption. Additional paleoclimate reconstructions may help determine the possible duration, degree, and extent of an SPG shift; and those additional inferences might help determine whether an eruption triggered an SPG shift in the first place. Nevertheless, such an investigation would be beyond the scope of the current study. The simulations do not currently indicate different types of SPG shifts of different timing, duration, and extent, with some types always triggered by an eruption and others not. Moreover, there is a trade-off between the precise, localized, diverse information provided by historical climatology and the more long-term, continuous, homogenous information provided by paleoclimate reconstructions. Thus, the reconstructions previously used in the 2017 Moreno-Chamarro et al. studies to test the presence of an SPG shift in the real world were less suited for determining whether that shift was triggered by an eruption; while the high-resolution proxies and historical observations used in this study to test for an eruption trigger would be less appropriate for examining the duration or extent of an SPG shift.

Specific comments:

L63-72: While it may perhaps be possible for such changes to occur without invoking substantial changes to atmospheric dynamics in the North Atlantic, the background state of the atmosphere, internal variability and the role of the NAO cannot be discounted *a priori*, particularly as these factors may act to modulate the response of the climate system to a large volcanic event.

(Please see the above response to concerns about the NAO state at the time of the eruption.)

L89: The phrase 'possibilities for adaptation' seems a bit vague and it is not clear what this refers to. Please specify / clarify this point.

We could clarify the phrase "nature of societal vulnerabilities and possibilities for adaptation" as "which activities and institutions were vulnerable, and how people could adapt them to changing climatic and environmental conditions".

Figure 2: For easier interpretation of the figure, it may be clearer to also state in the panel sub-headings that the plots are showing temp. / Sv. anomalies.

We thank the reviewer for the suggestion and will correct the figures accordingly.

L233: It is not clear whether this implies that only a 30-yr segment length was used or a range of segment lengths (30-yr+) was examined. If it is the former case, please remove 'minimum' to avoid confusion. Otherwise, please specify the range of segment lengths utilized.

The segment length (minimum number of observations between the changes) in the change point analysis is ≥ 30-years. We will change the language for clarification.

Figure 5: Please specify in the figure caption what the purple dots in top-left plot represent.

The purple dots in top-left plot of Figure 5 are the tree-ring width and maximum latewood density sites that have been used for the spatial reconstruction

L203-210: What was the size of the reconstructed grid cells? Which instrumental dataset was used for calibration? How were the chronologies merged and how was the reconstruction performed (e.g. PCA, nesting), etc.? In general, more detailed information about the development of the spatial reconstruction is needed here (or at least in supplementary materials).

The size of reconstructed grid cells is 5x5° lat/long. The instrumental data used as target field for the reconstruction are May to August monthly temperature anomalies wrt the 1961-1990 period extracted from the HadCruT4 (Cowtan and Way, 2014). The spatial reconstruction was developed using a point-by-point regression (Cook et al., 1999) which accounts for the spatial distribution and relationship of the proxy predictor network to the target field.

References:

Cook, Edward R., David M. Meko, David W. Stahle, and Malcolm K. Cleaveland. "Drought Reconstructions for the Continental United States." Journal of Climate 12 (1999): 1145–62. https://doi.org/10.1175/1520-0442(1999)012<1145:DRFTCU>2.0.CO
Cowtan, Kevin, and Robert G. Way. "Coverage Bias in the HadCRUT4 Temperature Series and Its Impact on Recent Temperature Trends." Quarterly Journal of the Royal Meteorological Society 140, no. 683 (2014): 1935–44. https://doi.org/10.1002/qj.2297.

Figure 6: How does the NVOLC reconstruction compare with N-TREND (and model output) over the investigated period? Currently, only NVOLC is compared to model output, whereas N-TREND is only used for illustration and is not compared to NVOLC or the modeled temperatures. The highly anomalous cooling in SE Europe in the NVOLC reconstruction (Fig. 6a) is rather suspicious and I wonder how robust this feature is. According to Supplementary Figure S3 in Guillet et al. (2017), most of northern Europe and parts of western / southwest Europe calibrate well, whereas calibration / verification statistics are very weak for NW, central and especially eastern and SW Europe. Consider that poor spatial representation of reconstructed temperatures may cause disagreement with modeled temperatures in some areas. Likewise, specific limitations of the model may also lead to disagreement. Such considerations should be acknowledged and discussed.

We will include additional discussion of the limitations of spatial representation of temperature anomalies in the NVOLC. Comparison with the N-TREND reconstruction may create confusion, however, since N-TREND includes many more TRW chronologies (with the problem of memory in series) than the NVOLC dataset, which was created with the goal to detect and quantify volcanic cooling and thus includes temperature-sensitive MXD records.

L286: Why is the NVOLC v2 reconstruction shifted by +0.5 K?

The aim of the figure is to compare the forcing generated by the 1600 eruption in the simulations and reconstruction, rather than absolute temperatures.

L350: I suggest that a more appropriate term to use in this context would be 'support' rather than 'appear to confirm'.

We thank the reviewer for the suggestion and will correct the sentence accordingly.

L368-370: So, considering the timing, might this in fact suggest that the Huaynaputina eruption is rather unlikely to be the cause of the SPG slowdown?

It remains difficult to say. As we explain in the next paragraph of the manuscript, the results would be consistent with a scenario in which the eruption did trigger an SPG shift but colder conditions had already started during the 1590s due to internal variability or a different forcing. Therefore, we could say that the results do *not confirm* a scenario in which the 1600 eruption triggered an SPG shift, but we couldn't say that the results *contradict* such a scenario. In terms of inference, our findings should reduce posterior estimates of the probability for an eruption trigger (per Bayes' theorem) since the likelihood of our reconstruction data given an eruption trigger is lower than the likelihood of getting that data regardless of the eruption trigger. Whether the eruption trigger hypothesis is *a posteriori* improbable -- i.e., $p(h|d) < 0.5$ -- would depend on one's prior probability estimate for the hypothesis.

L371-385: Another possibility could be that a pronounced shorter-term cooling impact of the Huaynaputina eruption was 'superimposed' on the longer-term cooling trend, which may have been initiated prior to 1600 (either in response to the cluster of late-16th century volcanic eruptions or otherwise). Evidence for volcanic-induced short-term cooling is on firmer ground as the results are consistent with this type of response to the eruption (Fig. 6c) and this is also consistent with the duration and magnitude of inferred NH cooling responses to large (tropical) eruptions more generally based on proxy reconstructions (e.g. Esper et al. 2015) and modeling of surface air temperature. In contrast, the mechanism for initiating longer-term cooling / SPG slowdown and attribution of such changes to a particular volcanic event is highly uncertain and rather problematic.

We thank the reviewer for this suggestion. It is, of course, possible that none of the simulations have captured the mechanism for cooling found in the real world, and we will note this at the end of the paragraph on line 375. However, as previously discussed in the 2017 studies by Moreno-Chamarro et al., 8 of the 20 simulations (6/10 with volcanic forcing and 2/10 without) reproduced an SPG shift as well as long-term winter cooling and increased sea ice extent found in previous reconstructions. Thus, we have chosen to investigate this mechanism more closely and to assess different scenarios indicated by those simulations.

L376-377: One could argue that it is uncertain whether this issue could be definitively resolved through modelling alone.

We will revise the sentence to read "First, *further comparison between high-resolution reconstructions and* a larger ensemble of climate simulations could improve…"

L395-397: I would recommend reformulating this sentence considering that, based on extensive paleoclimatic evidence, the occurrence of cool (wet) summers during this period is actually not in question. Therefore, rather than 'confirming' this, it would be more appropriate to state that this study provides further support and a broader context for such conditions at that time.

We thank the reader for this suggestion and will revise the sentence accordingly.

Minor / technical comments:
L95: consider '… activation (or lack thereof) …'

L186/380/384: 'an SPG' rather than 'a SPG'

Fig.4 legend / L233 / (L303): 'ice break-up timing data' instead of 'date data'

L223: consider 'obtained' instead of 'gained'

L238: consider 'recorded' rather than 'left'

L239: 'latter' rather than 'later'?

L243: Suggested wording adjustment: 'These observations were recorded in areas with flat terrain …'

L266: 'did or did not turn back' or alternatively 'ships could pass or were forced to turn back'

L267: 'turn back during a voyage' or perhaps 'terminate a voyage'?

L268: 'cold conditions' or 'cold temperatures' / 'dangerous sailing conditions' or 'danger posed by sailing conditions'

L280: Change 'NTREND' to 'N-TREND'. Also, something is missing here - consider: '… in each year over the 1601-1609 period …'?

L287: should the range be 1593-1650 instead of 1593-1640?
Yes. This apparently a typo.

L290: Consider: 'The analysis in Figure 7 indicating the ice …'

L293: 'can occur' rather than 'can happen'

L303: please change 'wrt' to 'w.r.t.'

L308: 'detected in' rather than 'detected at'?

L362: 'as a Possible Trigger'?

L372: remove 'has'?

L375: 'any role of'?

We thank the review for reading the manuscript closely. We will correct these typos and clarify the phrasing in the lines indicated above.