Response by authors to the review by Katherine Allen

General comments

This paper led by Higgins examines the ability of eight New Zealand tree species to reflect volcanic dimming. As the authors point out, little is known of the impacts of past volcanic eruptions on Southern Hemisphere climate, and there is a discrepancy between models and palaeo-data. The authors found that there are differences across and within species in terms of their apparent reaction to dimming. The authors also present a summer temperature reconstruction from the tree-rings that reflects influence of dimming. This same response is detected in the 7-model ensemble response. Previous studies on Southern Hemisphere trees have not identified a volcanic impact. This study is therefore of considerable interest and importance. Overall, the manuscript is quite well written, although some work is required to improve clarity and succinctness, and to better emphasise main findings.

Thank you for your thoughtful and detailed review. In our opinion, your review raises two major points to be addressed and several more minor points. The two major points, which are reiterated over several questions, are:

1. The use of regional dimming index for event selection and whether this then biases the results of our study, explaining why previous studies have not found a Southern Hemisphere volcanic temperature impact.

2. Why biological persistence in ring-width records does not appear to impact the volcanic signal in the temperature reconstruction, particularly compared to Northern Hemisphere studies.

All answers are provided below the questions in blue text, or, where appropriate, reference the response to another review that has asked the same question. Please note we have renumbered your points because the original numbering was corrupted in the uploaded version.

I wonder if a rearrangement of the material would help emphasise the key findings of this study a little better, better follow on from the introduction and also provide a basis for stronger and more substantial conclusions. One suggestion might be to start with showing a response in a new [more fully described] temperature reconstruction (or reconstructions) composed of a multi-species network of temperature sensitive chronologies to regional volcanic dimming, and a comparison with the CMIP ensemble before delving into the details of species used and species-level responses? The reason I suggest this, is that the introduction (l. 40 – 49, 65) seems to indicate that the model-paleo discrepancy will be an important aspect of the paper, but this doesn't really come through (might the choice of a regional dimming index be relevant? – see below). I think that finding a volcanic response is the first 'big' result of the study. I am guessing the authors view it as more useful to build the case from the sites/species first and then to look at the temperature reconstruction but nevertheless ask that they carefully consider what aspects of their work deserve greatest emphasis.

We believe there are two 'big' results from this study, and both deserve to be highlighted in the paper.

a) opposing volcanic signals can be identified in chronologies from the same species, and

b) there is a clear Southern Hemisphere volcanic dimming signal.

While the first result may be predominantly of interest to dendrochronologists and the second to the wider paleoclimate community, we don't believe either is of greater importance. As you have acknowledged, the paper is structured to build a case for temperature reconstructions from the site/species information. It is also structured this way to highlight both significant results. We hope that in resolving some of the other issues you have raised and revising the discussion in response to your and the other reviewers' questions, the conclusions may be emphasised without a paper restructure.

The authors' use of a regional dimming index relevant to New Zealand latitudes/longitudes rather than a selection of events based on eruption magnitude may be a key reason for their results. I think this needs to be discussed in more detail – it seems to be the elephant in the room when the authors are focussed on reasons for differences in responses amongst species. The fact that a relationship with volcanic eruptions has been identified in temperature-sensitive Southern Hemisphere trees is highly newsworthy. Why/how did the authors find this when other studies haven't? It would be good to place considerably more emphasis on this in the Discussion/Conclusions.

To test whether the event selection is the reason why our study finds a significant volcanic response when others haven't, we looked to the studies of Allen et al. (2018) and Cook et al. (1992, 2002) both of which produced a temperature reconstruction, assessed them for volcanic impact, and concluded the relationship between reconstructed temperature and the volcanic response was not clear/significant. The November-April temperature reconstruction of Cook et al. (2000), based on a single Huon pine chronology from Mt Read, Tasmania, was downloaded from the NOAA paleoclimate database. We then ran SEA using both the 21-event list and 13-event list and plotted the results. The figure below shows a significant response in year t+1 to the largest volcanic subset (SAOD > 0.08) but no response to the entire event list (SAOD > 0.04). Therefore, while a volcanic response is evident in the Mt Read reconstruction, it is more sensitive to the magnitude of the volcanic dimming than the New Zealand reconstruction. Without undertaking a detailed analysis, there are several reasons why this may be the case – Tasmania may be more influenced by the maritime climate than New Zealand, Huon pine may not be a strong responder to volcanic events, using multiple chronologies in our reconstruction may strengthen the volcanic signal compared to the single-chronology Mt Read reconstruction, and/or the difference in target seasons (summer versus growing season) may be an important factor.

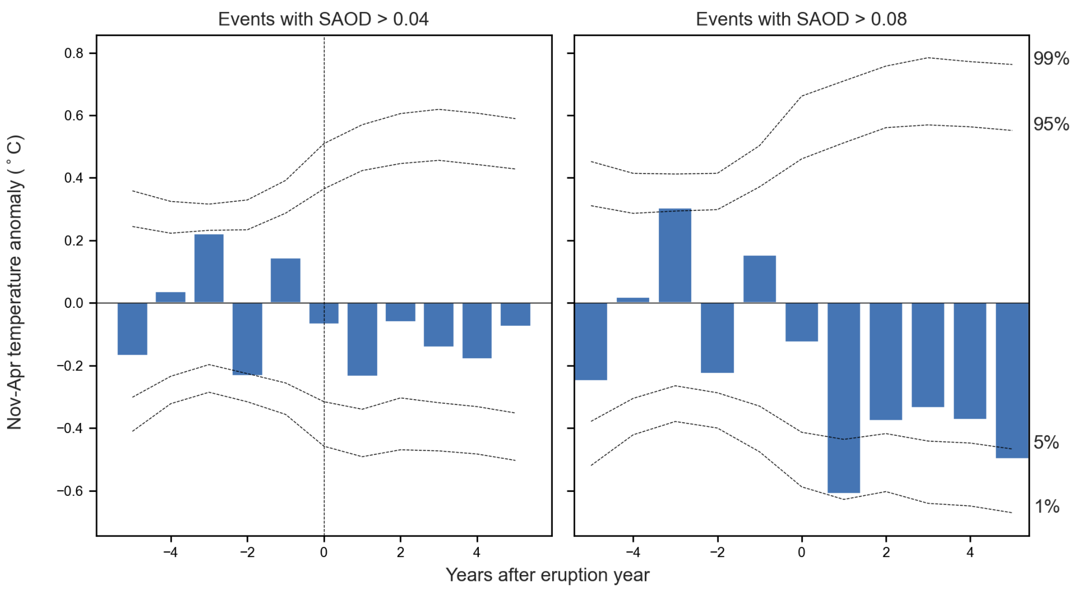

Figure 1 - Mean anomalies five years before and after 21 eruption years with SAOD > 0.04 (left column) and 13 eruption years with SAOD > 0.08 (right column), for the Tasmanian temperature reconstruction of Cook et al. (2000).

Why no response was identified by Cook et al. (1992) is impossible to analyse as the paper does not provide an event list or specify the methodology used to assess the link between volcanoes and reconstructed temperature. However, it can be seen in Figure 1 that only events resulting in quite significant Southern Hemisphere volcanic dimming can be identified in the Mt Read reconstruction. Event selection is therefore likely to have contributed. In the last 20 years, many research efforts have focused on reducing the uncertainty in the dating of volcanic events (e.g., Toohey & Sigl (2017)).

Therefore, the influence of volcanic dating errors in the analysis of Cook et al. may also be a factor in their lack of a significant response.

Allen et al. (2018) produced two reconstructions of Tasmanian summer temperatures, one based solely on *Lagarostrobos franklinii* and the other based on several endemic conifer species. It would be ideal for testing these reconstructions against our event list, as they targeted the same season (DJF) and used the same reconstruction methodology as our current study. These reconstructions are not publicly available but can be requested from the authors and are an obvious aspect of a follow-up study.

See also the response to Question 30 regarding previous NZ temperature reconstructions.

The authors rightly mention the moderate temperature response of the New Zealand species. This also applies to the ring widths of other Southern Hemisphere species. Do the authors consider that their use of the regional dimming index might 'compensate' somewhat for this moderate temperature response?

Using a regional dimming index to select events rather than an eruption magnitude almost certainly has contributed to identifying volcanic signals in this study. This was the intention of the event selection method, and we believe it was the right choice for this study. However, we don't think of the selection method as 'compensating' for a moderate temperature response. All tree ring-volcano studies select a threshold for which events to include, and this is known to be a significant source of uncertainty. Whichever the chosen threshold, studies aim to select events that would have resulted in a climate response in their study region/hemisphere. We used two different event thresholds and bootstrapped confidence intervals to account for this uncertainty in our results. The 90th percentile bootstrapped confidence intervals show that event selection could substantially impact the observed temperature response.

Please also see the response to RC2, where we demonstrate the effect of using a hemispheric rather than a regional dimming threshold.

In relation to the more detailed analysis of species and sites, it would also be useful to show more information on the actual sites. It is not really sufficient to state that the meta-data for sites can be found in Palmer et al. 2015. It is difficult to adequately consider some of the points made by the authors in the discussion without having the sites put into context much earlier. For example, the information in Table 2 (along with references to the supplementary Figure 2) could be presented in Section 2.1. A summary (possibly pictorial and perhaps in the Supplementary?) of the various species' sites by altitude/location would also be very helpful to better guide the reader through the results/discussion. Would such a figure help when providing some detail on which sites within a species did not have a strong volcanic response (Section 3.1)? Are there common factors – like altitude for example – that play a role in nonsignificant response within species? Could it be linked in any way to the reason for previous studies not finding a relationship with volcanic eruptions? (e.g. the authors discuss elevation and latitude).

We are happy to include the site metadata to a table in Supplementary – this was also requested by RC1, and we have included the proposed example in the response to RC1.

A summary of the temperature sensitivity of the different species can also be added to Section 2.1 with reference to the supplementary figures. However, there are limits to the details available or that can be provided (e.g., exhaustive site details were not always recorded).

Also, the authors comment in the conclusions that only a subset of the temperature sensitive chronologies show a response to volcanic eruptions. Figures S2-6 show that a number of the chronologies that are not temperature sensitive.

a) If the argument is that volcanic eruptions affect temperature and it is this that then impacts radial growth, why not use this information to exclude chronologies from the reconstruction and/or the analysis for a volcanic signal in the first place?

While there are broadly consistent temperature sensitivities for each species group, as the reviewer highlights, different chronologies show different strengths in these relationships (Figures S2-S6 – please note only chronologies extending until 1990, and thus considered for the temperature reconstructions, were plotted on these figures).

Considering the full suite of 96 chronologies in this study, only four (1HUI.r, 1MOE.r, 1MWL.r, 8WKT.r) shows no significant ($p < 0.05$) correlations to NZ average temperature in any month over the two growing seasons. As this is the likely criterion to be used to exclude chronologies from volcanic analysis, the change to the results would be negligible.

b) Some better links between this information and discussion of which sites do and do not show a volcanic response (or show a range of responses – i.e. cedar) may be warranted. Ditto in terms of positive/negative responses – does a dominant current season [positive] temperature response equate to a negative response to volcanic eruptions? Does a dominant prior season [negative] temperature response equate to a positive response to volcanic eruptions? (for eg)?

A general consideration of these points can be added to the discussion, although with the small volume of data available, it is not possible to tease out all these relationships in full. Figure 2 summarises the temperature-volcanic response relationships for the three species with the largest numbers of chronologies: kauri, pink pine, and cedar. The species responses are clearly identified, positive volcanic response for kauri, negative response for pink pine, and mixed response for cedar. Note that the significant positive volcanic response in pink pine is TOS.r, a site with substantial disturbance, so confidence in this result is lower.

Considered over both growing seasons, there appears to be a relationship between the strength of the temperature response and the strength of the volcanic response at most sites (as expected). The correlation between current growing season temperature and volcanic response is complex, with significant relationships in all four quadrants. The relationship between previous growing season temperature appears more straightforward, with a negative temperature response associated with a positive volcanic response and vice versa.

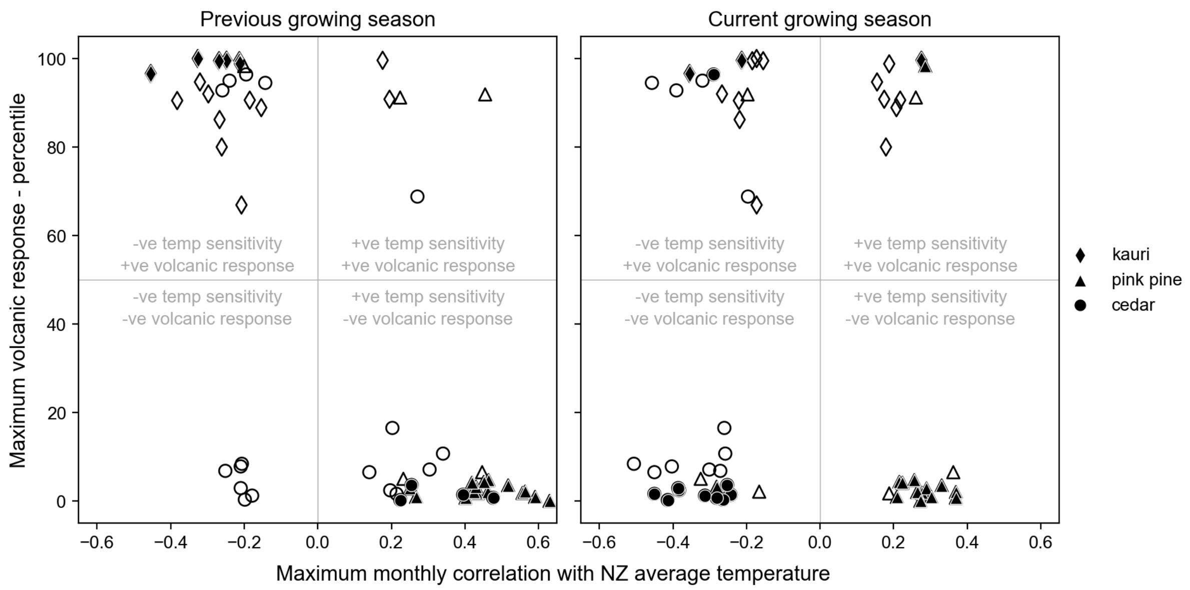

Figure 2 – Summary of the relationship between sensitivity to temperature and volcanic response for three species: kauri, pink pine, and cedar. Left: maximum temperature correlation in any month of the prior growing season against maximum volcanic response in the five years following an eruption for the 13 largest events. Right: same plot but for the current growing season. Filled markers indicate that a site has a significant temperature correlation ($p < 0.05$) and a significant volcanic response ($< 5^{th}$ or $> 95^{th}$ percentile of bootstrapped responses). Open markers are not significant for either the temperature or volcanic response, or both.

c) Any comment on seasonal window of temperature response and its relevance (or not) to volcanic response?

The seasonal sensitivity of tree growth is potentially an important determinant of the volcanic response, but we don't have sufficient evidence from the results of this study to support a discussion. It would be very interesting to investigate whether the whole-of-season sensitivity of pink pine causes more or less sensitivity to climate disturbance compared to the other species with narrower temperature sensitivity, but this is a question for future research. A greater number of (updated) sites with two or more species (i.e., pink pine and cedar) would allow a comparison of the role of seasonal sensitivity largely without other confounding factors.

Loosely linked to this point, the authors make the case that lower elevation sites have a temperature response related to moisture stress (l. 432-434, 503). While this is not an unreasonable suggestion, the authors need to be more careful about how they state this – they have not shown data to support the statement, so comments should be more cautious when it is mentioned.

Yes, the temperature-moisture stress response at low elevation sites is a hypothesis and should be qualified as such. Please also refer to the response to RC2, where we provide additional evidence for the kauri-moisture stress hypothesis based on dendrometer band studies.

At face value, there seems to be an inherent contradiction in the authors' discussion around 'stress tolerators' and delayed responses and their later comments about the lack of a lagged response to volcanic eruptions in the temperature reconstruction. It is important to clarify this given that the memory in tree-rings has been found to be an issue in the response of Northern Hemisphere trees to volcanic eruptions. The authors should carefully consider what they are implying in their discussion of lagged response (as shown in Figure 3 and S1) as opposed to their comments about the 'lack of memory' in the temperature reconstruction. Why are there apparently lagged responses of varying magnitude across the chronologies that are not apparently reflected in the reconstruction? Some careful consideration needs to go into this.

Firstly, the suite of tree rings used in this study have lower biological persistence compared to published Northern Hemisphere ring-width records (Esper et al., 2015). 80% of the predictors used to develop the NZall reconstruction only have significant lag-1, or lag-1 and lag-2 autocorrelation after standardisation. So, while we see lagged effects after eruptions in the composite for the stress tolerators due to persistence, there is less lag compared to arctic trees. Please see also the response to RC1.

Reconstruction methodology also plays an important role in the persistence of the final temperature reconstructions compared to the predictors used. We refer to the study of Büntgen et al. (2021), which compares 15 temperature reconstructions developed by different research groups with different methodologies but using the same set (or a subset) of chronologies. AR1 persistence in the final reconstructions varied from < 0.4 to > 0.9 due to methodological decisions, and there was a substantial variation in response to volcanic cooling between the reconstructions (Figure S2, their study). The important elements of our reconstruction methodology which likely contribute to the final autocorrelation structure are a) pre-whitening of both the tree rings and temperature data prior to reconstruction and b) testing predictors for significance in both year t and t+1.

Why were the specific 7 CMIP models selected (Table S2)? Why not other models?

These were the models that could be freely downloaded from the ESGF@DOE/LLNL, and met the condition of having both a past1000 and historical run.

There is quite a bit of repetition (almost the same sentences in some cases) between the Results and Discussion. This should be minimised, especially given that the authors are presenting a range of interesting results across and within species groups, and a temperature reconstruction and its response that is compared with a model ensemble. It is important to draw the threads together as coherently as possible.

Agreed.

A minor point: the title implies Southern hemisphere, but the study is focused on NZ. Perhaps it would be pertinent to include "A case study from New Zealand" or similar in the title.

Agreed.

Specific comments

Introduction

1. 47-9 Might seasonality of signal matter?

   Yes, see the response at line 165.

2. 52-5 "Tree growth…." Yes, but this almost sounds like the vast majority of the SH trees should be ruled out simply based on location and lack of serious elevation. It also seems to differ from what some of the results suggest (low elevation sites in mid-latitudes apparently sensitive).

   This is not the intention of the paragraph. We wish to highlight that it may be harder to identify volcanic signals from less temperature–limited trees, as indicated by similar studies in the Northern Hemisphere.

3. 66 – 7 – This sentence doesn't follow previous. Why would understanding whether a site is likely to have a volcanic response necessarily be relevant for studies of hydroclimate?

   We are not sure to which sentence the reviewer is referring. Sentence 66-67 reads, *"This knowledge will benefit future studies of hemispheric temperatures and help identify which species and/or regions should be prioritised for future proxy development."*

4. 73-74. Reword a little – awkward to read.

   A proposed change to:

   Land clearing has resulted in the loss of most lowland forests and nearly all forests from the eastern drylands of New Zealand. The most common remaining forest types are wet conifer-broadleaved forests and montane to alpine southern beeches (Nothofagaceae) dominant forests (McGlone et al., 2017).

Methods

5. 104 maybe reword this first sentence slightly to improve clarity.

   A proposed change to:

   Event selection is a significant source of uncertainty in tree-ring studies of volcanic cooling.

6. 113 10 and 18 eruptions. On next page on l. 126 and also l. 118, 13 and 21 events? Seems to be an error here? In any case, this is confusing.

   This can be clarified. 10 and 18 refer to the number of events between 1400 and 1900 CE, including the three events subsequently happening in the 1900s, bringing the event lists to 13 and 21.

7. l.116 "Between 1900 and 1990, we selected the three largest.." Where these the largest based on the same criteria for selecting the historical eruptions? (ie. based on the regional dimming index?)

   The dataset we used for the latitudinally modelled SAOD (Toohey and Sigl, 2017) doesn't extend past 1900. Therefore, event selection from 1900 was not based on the same dimming index. Please refer to the response to RC2 to see the impact on the results from using a different dataset to set the thresholds.

8. 128 "Species-level…". This again makes me wonder if it would be wise to first screen out those sites for each species that do not have a strong temperature response?

   See response at line 125.

9. l.153 DJF Maybe point to Table 1 as justification – but include actual months of sensitivity in Table 1 – see below for comment.
   Agreed.

10. 165 – 169. Presumably for DJF so comparable with the tree-ring reconstruction

    Yes, it can be clarified.

Results

11. 173 – 175 rewrite this a bit to be as clear as possible.

    A proposed change to:

    Superposed epoch analysis was used to analyse the volcanic response of the eight tree species to the 13 volcanic eruptions with SAOD > 0.08 between 1400 and 1990 CE (Figure 3). Two composite responses are shown for each species; the response averaged across all sites ('All chronology composite'), and the response calculated only from the site chronologies which individually showed either a significant (positive or negative) response to volcanic eruptions ('Sensitive chronology composite').

12. 178 Which species had a neutral response? List here.

    Agreed.

13. 180 – 195 Include references to lagged responses shown in Figure…compare with later comments on the 'lack of memory' in the NZ trees compared to the NH trees. Maybe also better link this with the nature of the temperature responses (Figure S2-6).

    See response at line 180.

Section 3.2

14. I am tempted to suggest that this could potentially go in Supplementary material (or even be omitted altogether) to simplify the paper and amplify the main points of the paper. I think it is useful to look at this, but not a key point. Also, the discussion here is a little confusing. In some cases it is the difference between the species at the various sites that is noted but not really described fully, but in other places, both species record a negative response. It would be useful to discuss both the nature of the responses of the individual species at these sites and then if they differ from the other species.

    We could agree to move Section 3.2 to Supplementary material; however, we think the opportunity to compare different species at the same site directly is an interesting and valuable addition to the paper and should not be omitted. It could be clarified that at some locations, the climate response is not sufficient to result in a response in either species, but at other sites we can observe differences between species.

Section 3.3

15. 218 It would be good to preface this section with some statement about why focus on cedar (why not do similar analyses for other species? – i.e. be explicit). One suggestion…begin with a comment about how the cedar average showed a generally muted response, but this masks very different individual site responses…and hence why this section of the results is useful. While it is certainly understandable that the authors wish to consider this material in the main manuscript, it may be worth considering whether at least some of this information could go in the Supplementary? (so as not to distract too much from the bigger messages in the paper).

    With Section 3.2 moved to Supplementary and the addition of an introductory sentence/s explaining why NZ cedar was selected for analysis that references Supplementary, we think the results of Section 3.3 are important to be included in the main paper.

16. 279 models – not CMIP models? Just be clear which models (reconstruction model or CMIP) is being referred to.

Ok.

17. 284 – 287 So there isn't a substantial issue with memory in the temperature reconstruction, but there are lags in the species-level responses. Be careful how this is discussed throughout the Results and Discussion.

    This can be discussed further.

Discussion.

18. As mentioned earlier, there is considerable repetition here (representation of Results) that clouds the text somewhat.

    We can revise the discussion to remove the repetition of results.

19. Section 4.1 – Compare with the above. Why doesn't this play out in the temperature reconstruction? Would be good to comment on.

    See response at line 180.

20. 340 – 345 This seems a little confused. Why separate silver beech from the other stress tolerators in this section? This section could probably be tightened up a bit.

    In this paragraph, we compared the responses of the two beech species, as distinct from the remaining species, which are all conifers. We found it interesting that the beech species showed such different response characteristics. However, this discussion can be incorporated into the previous paragraphs if that is clearer for the reader.

21. 350 – 354. So how does this relate to strong responses in years 0 and 3?

    We propose that the initial response in year 0 is foliage production, and in year 3, a secondary response is triggered by the cooler than average summer temperatures in year 0. Whether this response is due to a quasi-biennial flowering cycle or mast seeding or is related to normal cladode lifespan is a hypothesis. We favour the cladode senesce hypothesis due to our own, albeit limited, field observations.

22. 369 "…around 1000mm" This still seems relatively high, but how does it relate to the needs and distributional range (with respect to precipitation) of the species?

    This section could be clarified by discussing soil moisture balance rather than precipitation. In average years, there is a summer soil moisture deficit at Takapari because evapotranspiration exceeds precipitation, whereas in Westland, there is almost always a precipitation surplus (see Fig. 3). We therefore hypothesise that we see a positive summer moisture response for the Takapari chronologies but not the Westland chronologies because they do not experience summer moisture stress.

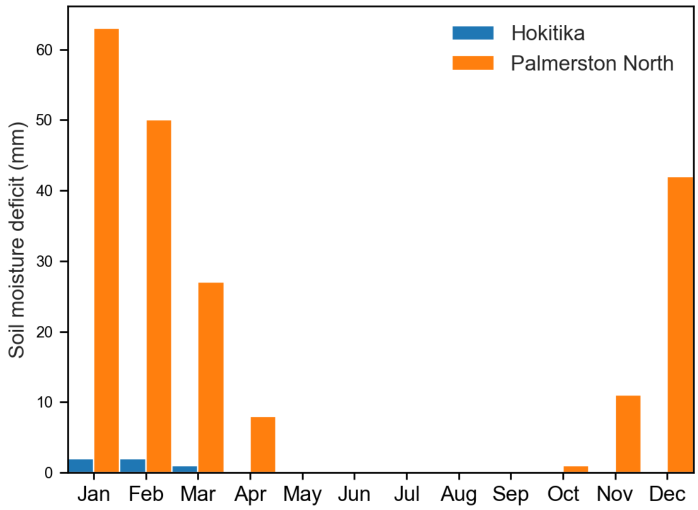

Figure 3 – Average monthly soil moisture deficit (mm) for a soil moisture capacity of 150 mm at two selected gauges: Hokitika, Westland, and Palmerston North, Manawatu-Wanganui. (Data sourced from: Chappell, 2015; Macara, 2016).

23. 370 – 373. Again, mention of possible link to role of moisture. This seems to suggest that perhaps moisture-related variability should have been considered in this study? However, the low number of samples is of some concern, and perhaps this section should be shortened accordingly.

    As New Zealand precipitation is very spatially variable, this analysis would need to be undertaken on a site-by-site basis. However, the locations of precipitation gauges are often not representative of conditions at the chronology sites due to distance (sometimes > 100 km), differences in altitude, orographic effects of the ranges, etc.

24. 392 I'm not sure this is counter-intuitive response given the negative response to temperature (Supplementary). L. 397-99. This mention of seasonality of growth again makes me wonder if this should have been more fully described for all species much earlier (the longer climate response window of pink pine for eg is interesting)?

    We think the combination of cool and dry spring/summer conditions is counter-intuitive rather than just the temperature response, but we can revise this wording.

    We also believe the different seasonality of growth has been addressed sufficiently in Tables 1 and 2 and Figures S3-S9.

25. Section 4.2 Needs considerable tightening up. Reference to moisture-related responses is speculative (but not unreasonable), but it needs to be couched that way. Also, while the results and discussion for cedar in Section 3.3 are suggestive, I don't think they should be presented as being THE causes of differences. They may well be, but further work, and closer examination across all the species would provide more evidence for this. At l. 433, other factors are mentioned. This reference should perhaps come earlier in this section to better set up the discussion around the evidence presented in Section 3.3. Especially when the authors go on to discuss location in the landscape (l. 440 – 446). This isn't discussed in relation to the PCA results in Section 3.3. If these factors are so important, they should be mentioned in that section – do the PCA results reflect this?

    The section can be revised. Please also see the response to RC2 for additional evidence for the findings of this section.

26. 422 " …sensitivity to temperature, including volcanic cooling…" So why include sites/chronologies not sensitive to temperature (or not at their limits)? This gets back to the locations of many of the SH tree-ring chronologies, and the relative lack of 'choice' compared to the NH.

    See response at line 125.

27. 450 "..large on site-related…". Only Elevation and latitude really mentioned.

    The wording can be revised.

28. 450 – 452. While it is great that the authors produced this new reconstruction and tested it for volcanic response, I have two points to make about it:

    a. The main features of this (these) new reconstruction are not really described in the study. How does it differ from earlier reconstructions? (Obviously the climate target may differ, but does it show similar features? If not/if so, where….?). Is it different enough to be the potential reason for volcanic dimming being detectable here but not in previous reconstructions?

Here we present a comparison of our reconstructions to three previously published temperature reconstructions from New Zealand (Fig. 4). The Cook et al. (2002) and Palmer & Xiong (2004) reconstructions were downloaded from the NOAA paleoclimate database, and the authors provided the Duncan et al. (2010) reconstruction. All reconstructions show approximately consistent high and low-temperature periods and an increase in average temperature since ~1950.

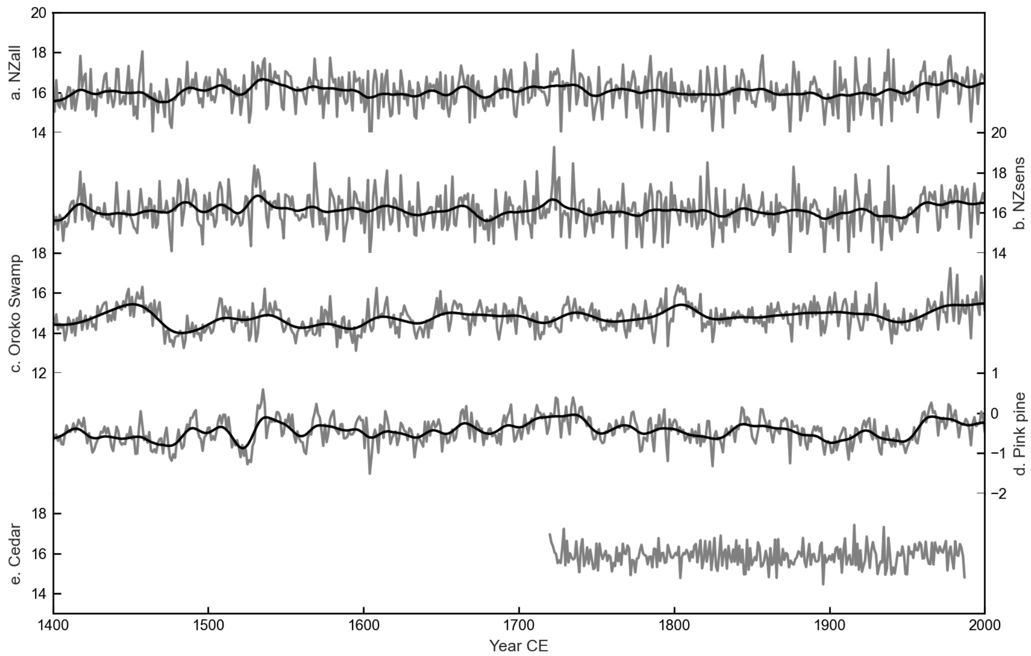

Figure 4 - New Zealand summer temperature reconstructions. a, b) DJF New Zealand average temperatures, this study; c) January-March temperature at Hokitika, Westland (Cook et al., 2002); d) Annual average New Zealand temperature (Duncan et al., 2010); e) February-March average New Zealand temperature (Palmer & Xiong, 2004). Reconstructed temperature is shown in grey, and the 20-year running mean is in black.

Despite different climate targets and seasons, and large differences in the number and geographical distribution of the predictors, all reconstructions are significantly correlated (p <<

0.001). The highest correlation is between our reconstruction NZall and the pink pine reconstruction of NZ annual average temperatures by Duncan et al. ($r = 0.62$). Should the paper be recommended for publication, we propose briefly describing the reconstructions' main features and adding the figure above to Supplementary.

b. If previous reconstructions were compared with the regional dimming index in the same manner would the same result be produced? Has it been the compilation of volcanic eruptions based on their magnitude which has been the problem in the past? Or is it the combination of chronologies used? The target season? The season of the eruptions vs seasonality of tree growth without due attention to regional and global circulation patterns?

None of the studies discussed above considered a volcanic response, nor did any of the other published NZ temperature reconstructions for which the data are not available. Thus, we suspect the main reason volcanic responses have not been previously identified in New Zealand tree-ring temperature reconstructions is that no analysis has been done until now.

Palmer & Ogden (1992) and Norton (1992) are the only studies to our knowledge that looked for volcanic signals in tree rings from New Zealand. Both studies consider only the Tambora eruption of 1815, and their methodologies simply identify whether there were narrow rings in the years following the eruption event. Both conclude that the evidence is not sufficient to identify a volcanic signal. Our analysis and the results of Palmer et al. compare quite well: cedar chronologies declined for several years following an eruption, similar to the lagged and persistent pattern we observed over 13 eruptions, and the two years following the eruption show increased average ring width in their kauri chronologies. Similarly, our results for the two beech species correspond well with the results of Norton. The issue faced by these researchers is that they were seeking common responses across species, whereas our analysis, with the benefit of much more data, shows responses vary widely between species.

Without undertaking a detailed analysis, the global investigation of Krakauer & Randerson (2003) likely suffers from the compositing of the many different chronologies, which results in some of the volcanic signal cancelling as noise. This is the reason we caution against compositing too widely in our conclusions.

Conclusions

29. I still think the big news is that a volcanic signal was identified. The ‘next big news’ is related to the species-level responses.

These points can be rearranged.

30. 488 “..proxy selection…”. This almost sounds like a choice amongst corals, trees, speleothems etc. Do you mean site selection?

Perhaps ‘chronology’ is better here.

31. 492 “….plant life history traits…”. What is meant by this? Not really discussed in the manuscript. Maybe just be explicit to avoid confusion.

It can be replaced explicitly with stress tolerator/fast responder.

32. 500 “We found that….” Not convinced this is a major finding when it depends heavily on Section 3 and then later observations not related to the analysis in Section 3.3.

Ok.

33. 503 " …summer moisture…" I think this is too speculative to include in the conclusions.

    It can be removed from the conclusions.

34. 513-515. But this study seemed to indicate that it didn't matter so much whether a subset of the most sensitive sites was used or not (especially for the temperature reconstruction). Reword.

    Ok.

35. 518 – 521. Yes, this applies to other types of reconstructions as well. Note that several large databases include composites that are then used by modellers who may not appreciate these types of nuances.

    Ok.

Figures and Tables

36. Figure 2. The orange and red are quite close to one another. Might it be useful to darken the red so there is a clearer visual difference between the two?
    Agreed

37. Figure 5 – colours for G4 and G5 difficult to tell apart for some. Change one of the colours.
    Agreed

38. Figure 8 – is it possible to use different symbols for the different species?
    Yes.

39. Table 1 Could this table be merged with Table 2 that simply summarises the nature of the response. Maybe also note which response is stronger, prior or current season?
    Yes, the Table 1 column 'reported climate sensitivity', which summaries the referenced publications, could be replaced with the calculated responses from Table 2.

40. Table 2 – I think this could be merged with Table 1, but in discussing Table careful references to Figures S2-6 should be made.
    Agreed

Supplementary

41. Figures S3 and S4. It is unclear why master chronologies are included just for these two species. Perhaps more could be made of the differences between 'master series' and individual series for all species in the main text? It would actually be good to see master series for all species included here given that species wide averages have been used in the main manuscript (Figure 3).
    The master chronologies provided in S3 and S4 are the published master chronologies downloaded from the ITRDB and associated with the references in Table 1. The published master chronologies may differ from the average chronology across sites used in our analysis because they are based on a subset of the chronologies. Certainly, we can also plot the temperature correlation of the average chronology for each species (i.e., shown in Figure 3). The addition of the table of site metadata should resolve any confusion.

Technical comments

Abstract

42. 28 "The has…" This has.

Agreed

Introduction

43. 88 amongst rather than between

Agreed

44. 89 Should “proxy” be “site”?

Perhaps ‘chronology’ is better to capture both sites and species.

Methods

45. 98 should “species depth” be “sample depth”?

Possibly. The chronology was excluded because it is the only chronology from that species, not because the chronology itself had insufficient sample depth; thus we chose ‘species depth’. We are not wedded to the current wording.

46. 133 “…two species..” Maybe insert ‘different’ between these words?

Agreed.

Results

47. 175 “averaged across….” All sites of a species, not just all sites. Ditto in relation to sensitive chronologies.

Agreed.

Discussion

48. 438 “…altitudinal range..” Altitudinal limit?

Yes, it should be the altitudinal limit.

49. 478 MDX – should be MXD?

Yes, this is a typo.

References used in this response

Allen, K. J., Cook, E. R., Evans, R., Francey, R., Buckley, B. M., Palmer, J. G., Peterson, M. J., & Baker, P. J. (2018). Lack of cool, not warm, extremes distinguishes late 20th Century climate in 979-year Tasmanian summer temperature reconstruction. *Environmental Research Letters*, *13*(3). https://doi.org/10.1088/1748-9326/aaafd7

Büntgen, U., Allen, K., Anchukaitis, K. J., Arseneault, D., Boucher, É., Bräuning, A., Chatterjee, S., Cherubini, P., Churakova (Sidorova), O. V., Corona, C., Gennaretti, F., Grießinger, J., Guillet, S., Guiot, J., Gunnarson, B., Helama, S., Hochreuther, P., Hughes, M. K., Huybers, P., … Esper, J. (2021). The influence of decision-making in tree ring-based climate reconstructions. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-23627-6

Chappell, P. R. (2015). The climate and weather of Manawatu-Wanganui. *NIWA Science and Technology Series*, *66*, 40. http://docs.niwa.co.nz/library/public/NIWAsts66.pdf

Cook, E. R., Bird, T., Peterson, M., Barbetti, M., Buckley, B., D’Arrigo, R., & Francey, R. (1992).

Climatic change over the last millennium in Tasmania reconstructed from tree-rings. *Holocene*, *2*(3), 205–217. https://doi.org/10.1177/095968369200200302

Cook, E. R., Palmer, J. G., Cook, B. I., Hogg, A., & D'Arrigo, R. D. (2002). A multi-millennial palaeoclimatic resource from Lagarostrobos colensoi tree-rings at Oroko Swamp, New Zealand. *Global and Planetary Change*, *33*(3–4), 209–220. https://doi.org/10.1016/S0921-8181(02)00078-4

Duncan, R. P., Fenwick, P., Palmer, J. G., McGlone, M. S., & Turney, C. S. M. (2010). Non-uniform interhemispheric temperature trends over the past 550 years. *Climate Dynamics*, *35*(7), 1429–1438. https://doi.org/10.1007/s00382-010-0794-2

Krakauer, N. Y., & Randerson, J. T. (2003). Do volcanic eruptions enhance or diminish net primary production? Evidence from tree rings. *Global Biogeochemical Cycles*, *17*(4). https://doi.org/10.1029/2003gb002076

Macara, G. R. (2016). The Climate and Weather of West Coast. *NIWA Science and Technology Series*, *72*, 40. http://docs.niwa.co.nz/library/public/NIWAsts72.pdf

Norton, D. A. (1992). New Zealand Temperatures, 1800-30. In C. R. Harington (Ed.), *The Year without a summer? : world climate in 1816* (pp. 516–520). Canadian Museum of Nature.

Palmer, J. G., & Ogden, J. (1992). Tree-Ring Chronologies from Endemic Australian and New Zealand Conifers 1800-30. In C. R. Harington (Ed.), *The Year without a summer? : world climate in 1816* (pp. 510–515). Canadian Museum of Nature.

Palmer, J. G., & Xiong, L. (2004). New Zealand climate over the last 500 years reconstructed from Libocedrus bidwillii Hook. f. tree-ring chronologies. *Holocene*, *14*(2), 282–289. https://doi.org/10.1191/0959683604hl679rr

Toohey, M., & Sigl, M. (2017). Volcanic stratospheric sulfur injections and aerosol optical depth from 500BCE to 1900CE Matthew. *Earth Syst. Sci. Data*, *9*, 809–831. https://doi.org/10.1108/eb058541