

Response to the Anonymous Referee #1

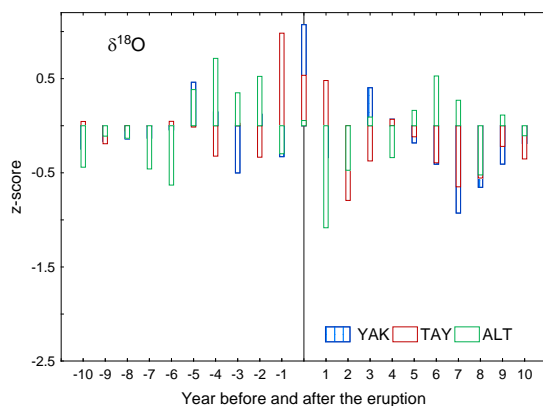
Referee #1: No superposed epoch analysis is presented (despite the mention of it on line 305) to establish a best estimate of the "typical" response to a large eruption, instead the volcanic epochs are only considered individually. The statistical testing on the individual events is therefore of limited statistical power because the size of the volcanic signal has to be equivalent to the 5 (or 10) percentile of the estimated variability of the timeseries for it to have even a 50% chance of being identified as significant. [Also, Fig. S1 where these tests are reported is illegible at the size and resolution provided – I'm surprised this wasn't picked up by the journal technical staff for correction prior to starting the review process.]

It is good to see the responses to individual events – and I like the overall design of Fig. 4 for this purpose – but it is also necessary to see the composite behaviour because of the additional statistical power that compositing (superposing) the events will bring and the different statistical testing that would then be applied. While the purpose may be to illustrate the varied behaviour after each event, it is first necessary to see the aggregate behaviour. Once this is established, the heterogeneity can be considered.

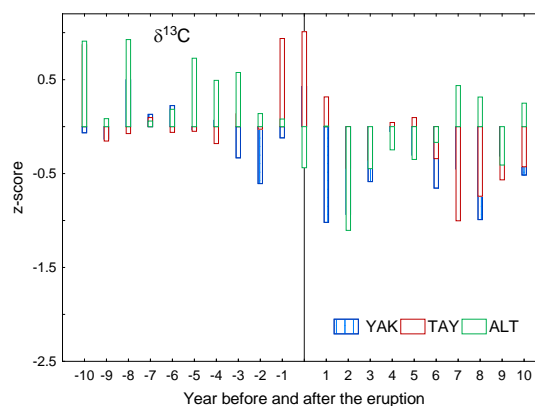
Answer: We are thankful to Referee #1 for this valuable comment. In fact, we used superposed epoch analysis for our proxies and presented the results in the supplementary material (Fig. S1) in the form of probability density functions. These pdfs included all proxies for all periods (n=221) studied after the major volcanic eruptions. In the revised version, we provide new and larger figures with increased quality for better visualization. In Fig. 2, we present only 10 years before and 10 years after the event to increase the readability, even more so as we illustrate 5 different tree-ring proxies ($\delta^{18}\text{O}$, $\delta^{13}\text{C}$, CWT, TRW, and MXD).

However, and to take the referee's suggestion into account, we performed additional analyses, displayed in Fig. 3 for the $\delta^{18}\text{O}$ (a), $\delta^{13}\text{C}$ (b), CWT (c), TRW (d) and MXD (e) chronologies. In the new version, data is displayed separately by combining the volcanic eruptions in CE 535, 1257, 1641, 1815 and 1991. Please find the new Fig. 3 below and in the revised manuscript Fig. 3 (see P. 18-19, L. 360-369).

a)



b)



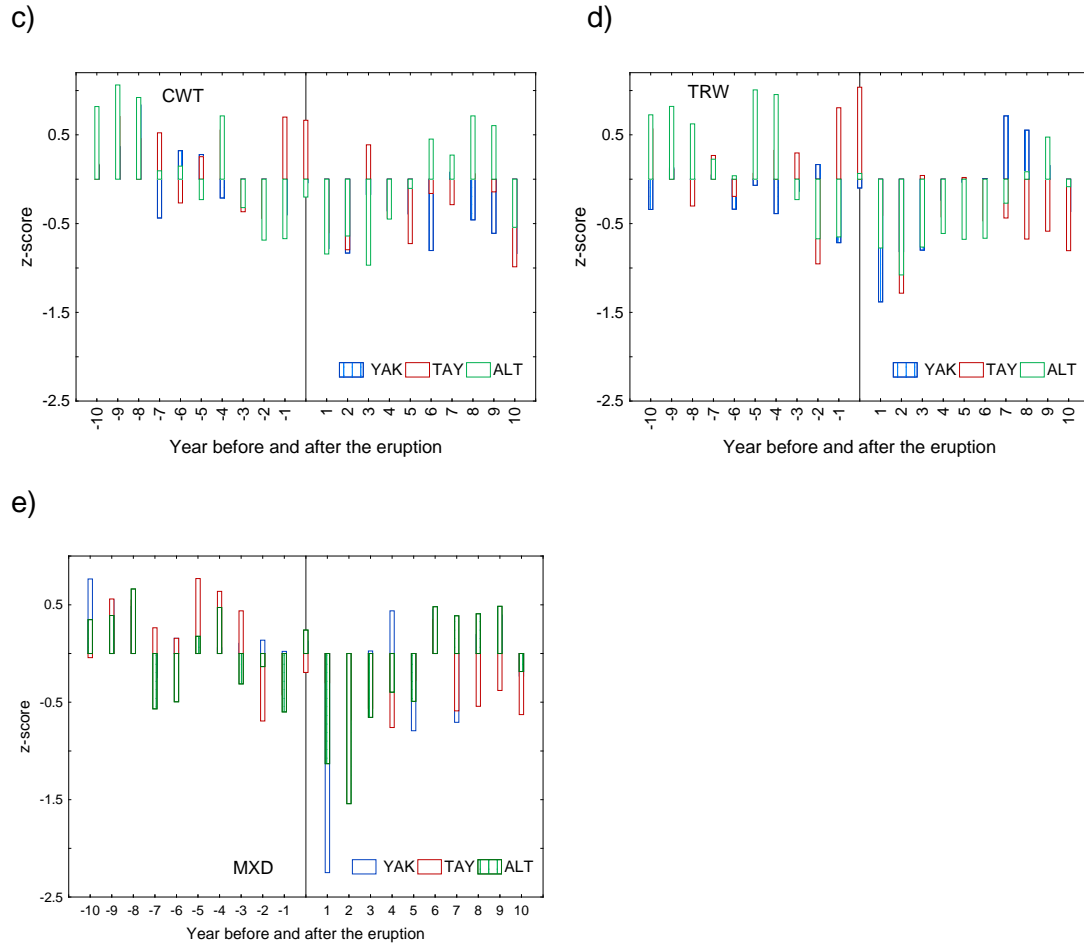


Fig. 3. Superposed epoch analysis of $\delta^{18}\text{O}$ (a), $\delta^{13}\text{C}$ (b), CWT (c), TRW (d) and MXD (e) chronologies for each study site and for the major volcanic eruptions in CE 535, 1257, 1641, 1815 and 1991.

We show that trees growing at YAK responded mainly during the first year after the eruptions, whereas a two years delay occurs at TAY and ALT (P. 3, L. 61-62).

Referee #1: There is no discrimination between (i) different responses to each event and (ii) "random" sampling variability that will make each case appear different anyway. If the purpose is indeed to demonstrate the heterogeneous responses between events (instead of, or in addition to, the differences between sites and between tree-ring/climate parameters) then the analysis needs to consider how to discriminate truly different responses from sampling variability. Different values will occur due to internal climate/weather variability as well as error variance in the data. The statistical tests compare each value to the time series variability to see if they are significantly different from zero – but what is needed is to see if they are significantly different from either each other or from the composite mean (see point 1). That would demonstrate the heterogeneity between events is real and not just down to sampling variability.

Answer: We aimed at showing differences in the climatic signal captured after stratospheric volcanic eruptions by the $\delta^{18}\text{O}$, $\delta^{13}\text{C}$, CWT, TRW and MXD chronologies separately, and show that stark differences exist between sites. We applied unpaired t-test statistics to check for

significance between each proxy and site. We find significant differences ($p=0.014$, $df=40$, $n=21$) between averaged $\delta^{13}\text{C}$ chronologies of the YAK and ALT sites. We now provide this information in the revised version of our manuscript (P. 18, L. 357-359).

Referee #1: Errors and limitations in the presentation and discussion of the results in Fig. 4. The results are hard to follow and their description/discussion is not presented concisely. In part the structure leads to duplication, e.g. the process-based description of stable isotopes is split between 2.6 before the results and 4.1.2/3 after the results but the results don't really confirm or alter our understanding of these processes so it seems unnecessary to split this into the two sections (that would be appropriate if it was to say what our understanding was before this study and then to present the updated understanding after it, but that isn't really the case here). I wonder if the sample sizes are too small for the CWT and isotopes to really be confident in the findings – can anything further be done to show whether the sample sizes are adequate? Are there enough samples to calculate RBAR and EPS as a measure of the common signal and chronology confidence?

Answer: Thank you for this very valuable feedback. We revised the discussion and removed unnecessary repetitions in sections 4.1.2 and 4.1.3 parts. In the revised version, we restrict the discussion to a single 4.1 chapter.

We have calculated RBAR and EPS values of stable isotope chronologies for the period 1950–2000, for which individual trees were analyzed separately, and show that a common signal indeed exists with an $\text{EPS}>0.85$. In previous studies, like Sidorova et al. (2008 JGR Biogeosciences; 2010 Global Change Biology; Climatic Change), pooled material was used for this analysis. For all other tree-ring parameters, EPS also exceeds the threshold of 0.85 (P.12, L. 212-215).

Referee #1: L58: "triggered by" implies a causal link, which hasn't been established by the analysis here.

Answer: we removed „triggered by“ and replaced it by “..led to..”

Referee #1: L85: Briffa → Briffa et al.

Answer: corrected

Referee #1: L120: Why do you expect increased humidity?

Answer: We expect a decrease in the carbon isotope ratio as a result of limited photosynthetic activity and higher stomatal conductance, which results from higher moisture conditions (i.e. increased relative humidity). To clarify our statement, we changed the sentence as follows: “Depending on the study site, a decrease in the carbon isotope ratio can be expected after stratospheric volcanic eruptions due to limited photosynthetic activity and higher stomatal conductance, which in turn would be the result of decreased temperatures, VPD and a reduction in light intensity (P. 6, L. 120-123).

Referee #1: L. 146 Why 6 eruptions and not more to increase the sample size? Why these particular
6? L180-181 says these are not the top 6, why not choose the top 6 in terms of stratospheric sulphur injection?

Answer: In this study we focused only on very large volcanic eruptions with a Volcanic Explosivity Index (VEI) exceeding 5 (**P. 6, L. 136** in the revised version; **see Table 1**). Eruptions exceeding this threshold are considered to be strong enough to affect tree growth over large areas. However, we fully agree that future work could include weaker eruptions as well, which was beyond the goals of the present work.

Referee #1: L190: virtual → virtually

Answer: corrected

Referee #1: L191, 238-239: give the actual sample size for each site, period and eruption in a table (perhaps adding to table 1) rather than just "at least 4"

Answer: We agree that the sentence in the original version of the ms was not very clear. We changed it as follows "Unlike TRW, which could be measured on virtually all samples, some of the material was not available with sufficient quality to allow for tree-ring anatomy and stable isotope analysis. We therefore use a smaller sample size for CWT (n=4) and stable isotopes (n=4) than for TRW (n=12) or MXD (n=12). Nonetheless, replications are still comparable with those used in reference papers in the fields of CWT and isotope analyses (Loader et al., 1997; Panyushkina et al., 2003; Fonti et al., 2013) (**P. 9, L. 183-189**).

Referee #1: L193: "perfectly" isn't needed

Answer: removed

Referee #1: L229: did you consider using CWT averaged only over the latewood, so it is closer to
MXD?

Answer: In our work, we used a common annual resolution for all tree-ring parameters. In line with this decision and in order to keep the approach uniform, we used cell-wall thickness values that were averaged per ring for each of the volcanic eruptions considered. We agree that the idea to use latewood CWT is very interesting, and might be applied later for intra-annual studies of climate-growth relationship. However, in this study we did not consider this parameter.

Referee #1: L304-305: It doesn't look like you subtracted the mean prior to the eruption – this is standard for SEA.

Answer: A standardization procedure (normalizing according to the mean value) was applied for each proxy and each period separately.

Referee #1: L306, 345: 15 years? Only 10 are shown in fig 2

Answer: To increase the readability and clarity Fig. 2 displays only 10 years before and after each event (see P. 9, L. 178, P. 18 L. 342, revised version) – the text has been corrected accordingly.

Referee #1: L325: -4.4 sigma is NOT less pronounced than -1.8 sigma

Answer: We corrected the sentence as follows “Decreasing MXD values for ALT (-4.4 σ) and YAK (-2.8 σ) were observed in CE 537. However, for TAY, we found less pronounced patterns of variation (Fig. 2).” (P. 16, L. 321-322).

Referee #1: L327: -3.9 sigma CWT for YAK is not visible in fig 2, which eruption do you mean?

Answer: We clarified that this refers to CE 541.

Referee #1: L334: no this is not the "response" to the volcanoes. This misconception is a repeated weakness of the manuscript: the complexity show may be due to a fairly stable response to the forcing, but with local climate variability and errors in the chronologies (sample size is small for some parameters) superimposed to give individual realisations that differ from case to case.

Answer: We believe that events are different from one another. The climatic effects of the different stratospheric volcanic eruptions reached the high latitudes represented by the chronologies illustrated here at different times within the year and varying magnitudes in terms of impacts. As a consequence, tree's response to these climatic effects are apparently quite different and also contain some local effects. Many studies based on tree-ring width or maximum latewood density only report about the cooling effect induced by volcanic eruptions. Our work, by contrast, shows that volcanic eruptions may induce changes in precipitation, vapor pressure deficit and sunshine duration. Such differences can be obtained from the same trees but by using a unique multi-proxy approach.

Referee #1: L340: Fig. 2 has a gap in the CWT series for TAY in 536. Explain why and what this means. Frost rings?

Answer: The remaining sample discs from TAY, which were not used for the stable isotope analyses, were broken when shipped back to the laboratory for anatomical analyses (V.N. Sukachev Institute of Forest, Krasnoyarsk, Russia) Thus, it was impossible to produce a clear image of the CE 536 ring from this material. As a result, wall thickness information is missing for TAY. Only one sample was left but we do not present it here for reasons of consistency and sample depth.

“Unfortunately, the remaining YAK sample size was too small for anatomical analyses. Thus, it was impossible to produce a clear anatomical signal for the CE 536 ring from existing material. As a result, cell wall thickness is missing for this year at TAY (Fig. 2).” (P. 12-13, L. 231-234).

Referee #1: L362 onwards: Climate analysis comments:

You only analyse correlations between individual months but later you suggest Jun-Aug temperature response for some parameters. Showing correlations for a 2 or 3 month seasonal mean might be useful, it might give a stronger correlation because intra-seasonal variability would be reduced.

Did you detrend the climate data before calculating correlations?

You need to consider cross-correlations between the climate variables, since this could explain why some tree-ring parameters are correlated with multiple climate variables. e.g. L389, could the negative correlations with MJJ precip arise because precip in these months is negatively correlated with temperature at ALT?

Answer: The seasonal mean in temperature response shows significant statistical relationships with tree-ring width and maximum latewood density for YAK (June-July $r=0.43$; 0.45); TAY (July-August, $r=0.57$; 0.59 ; June-August $r=0.44$; 0.48); and ALT (June-July $r=0.51$; 0.54 ; June-August $r=0.43$; 0.56), respectively for the period from 1950 to 2004 only. Averaged seasonal temperature against stable isotope data and CWT does not show higher correlations.

Meteorological data are available for a relatively short period (1966-2004 for precipitation and 1950-2000 for temperature) only and visual trends do not exist for these periods. Therefore, we did not detrend climate data prior to statistical analyses. The CRU data cannot be used for our purpose due to a lack of representation back in time. No significant correlations exist between precipitation and temperature datasets for ALT.

Referee #1: L372: Fig. 3 axis labelling is too small and blurry.

Answer: Thank you for this hint. The axis labeling of Fig. 3 (**Fig. 4 in revised version, P. 21**) has been enlarged and the quality of the illustration has been improved in the revised version of the manuscript.

Referee #1: L410: CWT label is missing for YAK.

Answer: corrected

Referee #1: L468: "strong" seems, err, too strong, when correlations are only 0.3 to 0.5. Also make clear it is SUMMER temperature.

Answer: We clarified "...is a proxy for summer temperature reconstructions" (**P. 26, L. 476**).

Referee #1: L483-484: First time that the NUMBER of cells is mentioned, similarly for frost rings.

Answer: We now provide a description in the section of results as: "We observe a strong decrease in CWT in CE 536 at YAK where only two layers of cells were formed in CE 536 (as compared to an average 11-20 layers of cells)." (**P.16, L. 318-319**). "...formation of frost rings in ALT (CE 536-538, 1259) has been shown in our study". (**P. 25, L. 489-490**).

Referee #1: L509: what is "gs"?

Answer: This is g_l (stomatal conductance). We removed this section to avoid repetition according to the Reviewers advice.

Referee #1: L510: at TWO sites

Answer: added

Referee #1: L514: positive correlation with VPD is significant at only ONE site

Answer: added

Referee #1: L525: or explained by a delayed/sustained climate response or by the aerosol forcing persisting and perhaps taking some time to reach the highest latitudes?

Answer: We modified sentence as follow "The delayed signal could also reflect the time needed for the dust veil to be transported to the study sites" (P. 26, L. 506-507).

Response to the Anonymous Referee #2

We thank Reviewer #2 for the advices to improve the manuscript and considered them as follows:

Referee #2: The title needs a change since the chosen six volcanic events do not represent the largest volcanic eruptions, as stated in lines 179-180.

Answer: We modified the title of our manuscript as follows: "Siberian tree-ring and stable isotope proxies as indicators of temperature and moisture changes after major stratospheric volcanic eruptions". The six volcanic events selected in this study are the largest eruptions over the past 1,500 years.

Referee #2: The term Common Era seems to be inconsistent with the studied span 'the past 1500 years'.

Answer: We replaced to "the strongest eruptions of the past 1,500 years: CE.." (P. 4, L. 83).

Referee #2: Lines 72, 92, 119, 181 and Section 3.1 line: I am not sure about the wording 'Stratospheric volcanic eruptions'.

Answer: Not all volcanic eruptions can be classified as stratospheric. However, in our case, we considered only stratospheric volcanic eruptions; therefore, we highlighted this point in more detail in our manuscript.

Referee #2: Lines 99-101 and lines 470-472 seem to be contradictory. Please check it.

Answer: We clarified these sentences in lines 103-112 and removed the sentence that was in lines 470-472 in the original submission.

Referee #2: Lines 126-127 are not clear and need rephrasing.

Answer: To avoid confusion and misunderstanding we left this sentence away.

Referee #2: Fig. 1: five volcanic eruptions (vertical lines) are indicated but one eruption is missing.

Answer: We added a reference to the event in 540.

Referee #2: From the map, the two eruption sites (two black circles) are located in tropical areas. It would be more clear to point out in the text and abstract.

Answer: We included information about two eruption sites in the figure legend (P.8, line 165-166) and added as requested this info in the text and abstract (P. 3, line 52-53). More information about volcanic eruptions can be found in Table 1.

Referee #2: Line 184: the authors stated each studies segment is 'around ± 10 years', but Fig.2 caption says 'the specific periods 15 years before and after the eruptions'. It is confusing.

Answer: We corrected the text to 10 years (P. 17, Fig. 2, L. 339-341, legend).

Referee #2: Sections 2.6 and 2.7 can be combined together.

Answer: Based on the reviewer's suggestion we combined both section 2.6 and section 2.7, and modified the subsection as following "2.6. Stable carbon ($\delta^{13}\text{C}$) and oxygen ($\delta^{18}\text{O}$) isotopes in tree-ring cellulose". (P. 13, L. 244).

Referee #2: Superposed Epoch Analysis (SEA) results need significance tests to enhance scientific rigor of the relevant descriptions (for example, section 3.1).

Answer: We applied unpaired t-test statistics to check significance between each proxy and each site and provided new (Fig. 3, P.18-19, L. 361-370).

Referee #2: Section 3.1 and section 3.2 should be swapped.

Answer: We disagree on this point, as, in our opinion, data should be presented first, followed by statistical relationships with climatic parameters. Therefore, we preferred to keep this order as it was originally.

Referee #2: Fig. 2: the gray line is not clear. It is suggested to change the line color to enhance visibility.

Answer: That's a relevant point. We have accordingly replaced grey color with pink, to enhance visibility, as suggested (P. 17, revised Fig.2).

Referee #2: For TRW and MXD, they are affected by both temperature and precipitation (see Fig. 3). It is difficult to separate temperature signals alone. However, the authors chose them as only temperature indicator.

Answer: Absolutely, it is difficult to separate temperature from precipitation signals in temperature-limited environments. We attempted to do so by presenting the most meaningful relations (Fig. 4, P. 20, L. 382-386) with precipitation. However, significant correlation with precipitation are clearly lacking during the summer for TRW and MXD, for example, for YAK.

Referee #2: Lag between volcanic events and response in tree rings is easy to understand for carbon isotope and ring width. But for oxygen and MXD, it is generally accepted that there is no legacy effect from previous year. The authors argued for findings of the heterogeneity with different volcanic eruptions each but a potential associated mechanism is missing, for example, climate response to 1815 Tambora eruptions.

Answer: The $\delta^{18}\text{O}$ chronology displays lag effects after volcanic events that might be due to permafrost availability and mixed source water (atmospheric precipitation and thawed permafrost water) recorded in tree-ring cellulose. This signal can be stored in melted permafrost water and captured back based on climate conditions (cooling, warming anomalies). Another reason could be that volcanic eruptions cause changes in tree physiology (e.g. damage to roots because they are not well supplied by fresh assimilates). This could affect not only growth and carbon isotopes in the following years, but also oxygen isotopes, because they also depend on root conditions and physiological properties (leaf water enrichment depends on transpiration rate and such signal is reflected partially in tree-ring cellulose).

Since MXD represents late-summer climate conditions we can expect that the Tambora eruption, which took place only in April 1815 would not be visible in the 1815 ring.

In a study by Esper et al. (2017) based on 20 Northern Hemispheric MXD chronologies, a strong and coherent post-volcano signal was seen, mainly in the year after the eruption.

Esper, J., Büntgen, U., Hartl-Meier, C., Oppenheimer, C., Schneider, L.: Northern Hemisphere temperature anomalies during 1450s period of ambiguous volcanic forcing. Bull. Volcanology. 79, 41, 2017.

An absence of widespread and intense cooling or reduction of precipitation over vast regions of Siberia over the past half millennia may result from the location and strength of the volcanic eruption, atmospheric transmissivity as well as from the modulation of radiative forcing effects

by regional climate variability. Further studies are needed to understand the mechanisms and causes of these differences.

Referee #2: If available, additional evidence such as historical documents and long instrumental observations are much needed to strengthen the results.

Answer: There are no longer instrumental observations because our study sites are located in remote and hardly populated regions. Gridded CRU data is not representative back in time for our sites. We provided all available evidence for historical documents discussed in detail in papers by Myglan et al., 2008; Büntgen et al., 2016; Guillet et al., 2017.