GENERAL COMMENT

In the manuscript 'Abrupt termination of the Little Ice Age in the Alps in the mid-19th century: lessons from a multi-proxy tree-ring reconstruction of glacier mass balance' Lopez-Saez and co-authors present seasonal (and annual) mass balance reconstructions for a Swiss glacier since 1802 CE. Authors use several proxies obtained by different methods (total ring width, quantitative wood analysis, and isotopes) and Principal Component Analysis to perform a multiparameter linear regression. The obtained loadings were used to explain and to reconstruct mass balances' variance in the last century (since 1919). Results are statistically significant and pass the tests normally used in dendroclimatological reconstructions. They show variations of the mass balance compatible with known glaciological history in the Alps. Thus, authors conclude that the use of different wood-proxies permits the seasonal mass balance reconstruction of the Silvretta glacier.

The manuscript, in my opinion, is well written and the aims are clearly presented. Authors present exceptional datasets for an overlooked species in the Alps (i.e., *Pinus cembra*). In fact, in my knowledge, they present first isotope chronologies from Swiss stone pine in the area and one of the firsts chronologies of anatomical traits. Scientific design is solid and well presented. Moreover, only few dendroglaciological papers about European Alps were published, thus the manuscript is also characterized by a high level of novelty. However, in my opinion, the use of some methodologies is partly questionable, and both discussion and conclusion lack a bit of control in some parts resulting presented in a bloated fashion and quite speculative way.

SPECIFIC COMMENTS

[#001] Line 25: $\delta^{13}C$ isotope is not utilized for B_s reconstruction.

[#002] Lines 35–36: Consider mentioning that glaciers are one of the Essential Climate Variables (ECV) as reported by the Global Climate Observing System (GCOS) (https://gcos.wmo.int/en/home)

[#003] Lines 57–66: Authors at lines 57–61 state that '*Mass balance modelling based on meteorological series (Huss et al., 2008; Nemec et al., 2009) offers an alternative method to infer glacier mass balance over long time-scales at high temporal resolution but results are not backed-up with in-situ observations before the onset of glaciological measurements and therefore might be biased or may incompletely resolve the relevant processes.*'. At lines 62–66 they state: '*Tree-ring proxies clearly have the potential to overcome these limitations and to extend glacier mass balance series farther back in time. Based on the concept of Oerlemans and Reichert (2000) according to which mass balance series can be reconstructed from long meteorological records, several dendrochronological studies have been developed to demonstrate the reliability of high-elevation tree-ring proxies as reliable recorders of past summer temperature and – to a lesser extent – also winter precipitation (e.g., Büntgen et al., 2005; Coulthard et al., 2021; Carrer et al., 2023; Lopez-Saez et al., 2023).*'. These sentences seem in conflict to me. In the former authors declare that since there are not glaciological measures, reconstruction based on glaciological modelling and meteorological series might be biased, on the other hand, they declare, in the latter sentence, that tree-ring proxies can overcome to this limitation since they are representative of the summer temperature and winter precipitation that can be used to reconstruct mass balance. Thus, tree-ring proxies were used to reconstruct temperature and precipitation (i.e., meteorological series) that were used to reconstruct the mass balance that, however, supply results that might be biased since glaciological measurements are still missing before the start of the monitoring. Please consider clarifying the concepts.

[#004] Lines 77–78: In Cerrato et al. 2020 the authors did not used HISTALP dataset, but an improved version of the database presented in Brunetti et al. 2006 as described in material and methods.

Brunetti, M., Maugeri, M., Monti, F., and Nanni, T.: Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series, Int. J. Climatol., 26, 345–381, https://doi.org/10.1002/joc.1251, 2006.

[#005] Lines 91–93: In the manuscript only one glacier was tested and no comparison with others mass balance was performed. Thus, referring to 'Alpine glaciers' seems a bit pretentious if it is meant as glaciers of the entire European Alps.

[#006] Line 98: Please consider adding the reference period for which the mean equilibrium line altitude is calculated. If it refers to the entire period (1919–2023) please consider adding also the maximum and the minimum since in the last century alpine glaciers withdraw quite continuously and abundantly.

[#007] Line 108: the tree stand is located c. 30 km southeast of the glacier (not southwest).

[#008] Line 149–150: Please check the total number of analysed rings. Between 1968 and 2017 there are 49 years, whereas between 1802 and 1967 there are 165 years, that divided by 5 results in 33 years (34 if 1802 is considered too). Thus, if I correctly understood, the total number of analysed rings is 83 and not 73 as reported.

[#009] Line 164: R_{samp}, samp with lowercase 's' to be coherent with the formula.

[#010] Lines 229–230: Please consider moving the sentence about the isotope at the end of the paragraph (at Line 243) to not interrupt the discussion about the QWA.

[#011] Line 248: the number of the chapter is missing in section heading (i.e., 3).

[#012] Lines 260–261: probable typo, from July 15 to September 11 there are 58 days, not 152 as declared. Moreover, Figure 3A shows the mass balance reconstruction, I think that the figure was moved to supplementary material or totally removed.

[#013] Line 267: see comment [#034].

[#014] Lines 269–270: The sentence is misleading. From Figure S1B very low values of correlation is interpretable, spanning between –0.1 and 0.1. In some situation, as function of the windows width, slightly positive or negative results can be appreciated (usually between (–)0.1 and (–)0.3).

[#015] Line 277: The winter signal embedded in $\delta^{13}C$ chronology is not reported in Table 2 and only partially inferable from supplementary material.

[#016] Line 283–284: In table 3 only one combination for reconstructing B_s and B_w is reported. Consider rephrasing.

[#017] Table 3: I wonder if all isotopes are really necessary in the Principal Component Regression model for B_w. More precisely, one isotope chronology (Carbon) resulted to be most sensible to the summer precipitation (i.e., ablation period, Table 2), and the other (Oxygen) is sensible to the precipitation from November to August (covering not only accumulation period but also quite the first two-third of the ablation season). How different would be the results if only D_{rad} is used? Or, in another way, are the authors sure to include proxies that are sensible to environmental parameters of ablation season in the reconstruction of B_w? It seems quite counterintuitive to me, and clarification are necessary in my opinion.

[#018] Lines 289–293: Even if it is true that the first two Principal Components (PCs) positively correlate with B_w, seems that authors overlooked at the relationships between the original data and the PCs variables. If I correctly understood figure 3, it reported the correlation circle between the original variables (i.e., isotopes and D_{rad}) and the first two dimensions obtained by PCA. Looking at that plot seems that only D_{rad} is positively correlated with the PC's first dimension and basically uncorrelated with the second. I can thus hypothesize that, due to the positive correlation index between the PC variables and the B_w (shown in Table 3), D_{rad} is representative of the environmental condition that drive the B_w. Contrarily, Carbon isotope show a high negative correlation with PC first dimension (indicating a quite linear negative correlation with this dimension) with a (maybe significant) positive correlation with the second PC dimension. This seems to be coherent with previous analysis, i.e., Carbon isotope find its best correlation window with summer precipitation, and being summer precipitation mostly liquid, they enhance snow melting with heat transfer along the snowpack. Thus, probably, the negative correlation that was showed along the first PC dimension is representative not of a B_w, but of a B_s that, judging from figure 3, seems quite well negatively correlated with B_w. Considering the Oxygen isotopes, the series result completely uncorrelated with the PC first dimension being aligned with the axis of the second dimension. Considering that first and second dimension have quite the same explanatory power of the original dataset (i.e., 46.1 and 34.0, respectively) it is plausible that Oxygen represents something different from both D_{rad} (precipitation from November to May and spring-summer temperature), and Carbon (summer precipitation and whole year temperature, this last point is questionable as reported in comments [#015] and [#017]). In fact, from previous analysis it results that Oxygen series is correlated with spring-summer temperature (as well as D_{rad}) but with winter-to-summer precipitation. Maybe the representativeness of this variable of such a long precipitation period appoints it as a second major source of variability in the original dataset. Authors should consider these results, or at least supply more explanation on the motivation that drove them to use proxies sensible to summer precipitation and/or to two-third of the ablation season to reconstruct the B_w bearing in mind that correlation does not mean causation.

[#019] Lines 301–302: Authors hypothesis could be true, however should be noted that in the referred period occurred the last phase of positive B_a in the (Southern) Alps (Huss et al. 2015). Moreover, if the gap in B_w starts in 1984, I wonder how it is possible that the correlation values start to decrease 10 years before (as author stated, and as figure 4 shows, considering that the results are right-aligned, thus the considered 30-year window in 1974 is 1945–1974) and reach their lowest value nearby the years when the modelled B_w starts. Moreover, the lowering in correlation values ended around 2000 (i.e., the 30-year right-aligned window 1971–2000) when around 57% of the data are modelled. Considering this, the hypothesis supplied by authors seems to be not really supported by reported data. Probably, the changes in environmental conditions that bring less negative or even positive B_a, is not well represented by the selected variables for B_w (it is just a hypothesis that should be verified). On the other hand, a decrease of the correlation values in those years is observable also considering the seasonal mass balance reconstruction based on *Imfeld23* (both B_s and B_w) and the wood-proxy based B_s. Also in these cases, the lowering in correlation values starts well before the 1984, thus, in my opinion the lack of measured glaciological data could not be the (only) explanation to the observed behaviour in correlation trend.

Huss, M., Dhulst, L., and Bauder, A.: New long-term mass-balance series for the Swiss Alps, J. Glaciol., 61, 551–562, https://doi.org/10.3189/2015jog15j015, 2015.

[#020] Lines 305–307: To me it is not clear the advantage in using all the variables deriving from a PCA instead of the original data. The PCA was thought to lowering the number of considered variables, creating new variables that 'summarize' the variance of the original data. Variables reduction is obtained retaining only those new variables that explain the largest part of the original data variance (usually 80% but it depends by the aims). If all PC variables are kept, it is equivalent to apply the multiparametric regression using the original data.

[#021] Line 308: consider modifying 'Ghiacciaio del Careser' to Careser glacier, as used before in the manuscript.

[#022] Lines 310–313: From figure 4 it is clear that the increasing trend appreciable from 1950 to 2000 (as exception of the 1980s where a decrease in correlation values is appreciable as commented in [#019]), is reverted to a negative trend since 2000 with correlation coefficient that drops from 0.75 to 0.5 in 17 years (mean decrease of -0.015 $year^{-1}$, analysis should be performed to verify if the trends are significant and if the change is significant too, but I can speculate that, at least the negative one, is significant). Considering this, the decreasing of correlation starts in the 1970s, 10 years before the start of decreasing reported in Cerrato et al. 2020 (in Figure 4 the correlation are right-aligned, thus the 2000 value refers to 1971–2000 time window and this is why the decrease in correlation values starts in 1970s).

[#023] Lines 326–327: Being the correlation obtained using a multiple regression, this statement is not supported by data in this context, even if it is true as reported in cited papers. Moreover, consider to cite also Cerrato et al. 2019, that report data about the divergence between Swiss stone pine MXD and temperature in the high-frequency domain and being the source of data for Cerrato et al. 2020.

Cerrato, R., Salvatore, M. C., Gunnarson, B. E., Linderholm, H. W., Carturan, L., Brunetti, M., De Blasi, F., and Baroni, C.: A Pinus cembra L. tree-ring record for late spring to late summer temperature in the Rhaetian Alps, Italy, Dendrochronologia, 53, 22–31, https://doi.org/10.1016/j.dendro.2018.10.010, 2019.

[#024] Lines 351–353: A reconstruction of an Alpine glacier mass balance at annual scale was already reported by Cerrato et al. 2020 and by Nicolussi and Patzelt, 1996 (in my knowledge, but could be other studies. These studies are already cited in the manuscript even if the latter is missing in the reference list) and both show less negative (or even positive) mass balances around the last peak of the LIA. Please consider rephrasing. Moreover, in the present study, volcanic forcing or radiative data were not considered, thus the sentence, in the present form, seems a bit speculative to me.

[#025] Lines 354–364: Speculative paragraphs. If two reconstructions are available and no verification is possible, it is basically impossible to determinate which is the most correct. It is certainly true that back in time, meteorological series loss representativeness and explanatory power in remote areas, but, considering reported data, also tree-ring proxies reconstruction suffer of a decrease of explanatory power in cold phases (see for instance 1980s where mass balance data, even if modelized, are present) and also after 2000s (even if in this last case the environmental conditions that drive a loss of correlation are hotter than the previously experienced; see comments [#019] and [#022] for more details). Thus, concluding that the B_a based on *Imfeld23* lacks representativeness whereas tree-ring bases reconstruction surely represent the behaviour of the glacier in such bloated form seems a bit speculative. Moreover, authors never consider that their approach using the meteorological data could suffer of a big and simply issue: authors calibrated the reconstruction using an optimal time window based on temperature and precipitation occurred since 1919. Statistics are solid and tests were passed in the considered period. By counterpart, authors are assuming, based on their results, that the length of the accumulation and ablation seasons are the same in a period where the temperature has been proved been lower and also precipitation might be, testified by a different duration of the snowpack (Carrer et al., 2023, cited in the manuscript). Maybe authors should consider that is not the meteorological dataset, but the selected optimal window of a fixed length based on recent environmental conditions that can bias the results, as already reported in Cerrato et al. 2020. However, should be noted that the here proposed reconstruction for B_w shows lower values during the Dalton minimum compatible with previous work that reported more dryer winter during that period (Anet et al., 2014). However, the *Imfeld23* based B_w is not reported, thus it is impossible to evaluate if also meteorological-based B_w reconstruction shows comparable results.

Anet, J. G., Muthers, S., Rozanov, E. V., Raible, C. C., Stenke, A., Shapiro, A. I., Brönnimann, S., Arfeuille, F., Brugnara, Y., Beer, J., Steinhilber, F., Schmutz, W., and Peter, T.: Impact of solar versus volcanic activity variations on tropospheric temperatures and precipitation during the Dalton Minimum, Clim. Past, 10, 921–938, https://doi.org/10.5194/cp-10-921-2014, 2014.

[#026] Lines 372–374: The study can be considered a first step in Switzerland, but not throughout the Alps since both Cerrato et al. 2020 and Nicolussi and Patzelt, 1996 already presented mass balance reconstructions. Consider rephrasing.

[#027] Lines 374–375: Due to the lacks of validation on the correctness of the B_a reconstructions (at the actual state it is impossible to define which is the most correct reconstruction between the wood-proxy based and the *Imfeld23*-based B_a since no comparison with previous reconstruction is performed, neither a comparison between the potential glacier volume with geomorphological and/or cartographical evidence) the sentence seems quite speculative.

[#028] Lines 378–380: This sentence seems speculative. Since it is impossible to validate both wood-based and *Imfeld23* reconstructions (see comments [#025] and [#027]) it is also impossible to be sure of the correctness of the estimated quantity of water equivalent gain (or loss) in period were the reconstructions differ in a more pronounced way (from Figure 4 and 5, the period of less agreement between wood-proxies based and *Imfeld23*-based reconstructions occurred for the entire XIX Century). Moreover, Authors stated earlier that in the earlier portion the used meteorological dataset is not completely reliable, so it is impossible to verify this sentence, maybe the disagreement between the expected and obtained *Imfeld23* B_a is due to the uncertainties of the original dataset, or maybe not.

[#029] Figure 1: Consider inverting the vertical order of the inset A and both B and C.

[#030] Figure 2: Consider to explain the meaning of the purple dotted line (or purple dots, but it seems a line to me) in caption.

[#031] Figure 3: caption, unclear to me, please consider rephrasing.

[#032] Figure 5: please consider to maintain the same y-axis scale among the plots for readability.

[#033] Figure S1: caption is misleading on the time-window information. Moreover, reported information (e.g., standardization method and windows length) does not match those declared in the main manuscript and thus the results are not easily comparable with those reported in the main text. In fine, the addition of contour lines at the significance level of $p<0.05$ will be appreciated.

[#034] Table 2: correlation between $\delta^{13}C$ and temperature: the reported optimal time window is equal in length to the maximum window tested (330 days), probably enlarging the tested windows, other (and longer) 'optimum windows' could be found. Beside this mine speculative consideration, authors in M&M stated that they '*calibrated regression models on temperature and precipitation averaged over 30 to 330-day windows starting on October 1 of the year preceding ring formation (n-1) and ending on September 30 of the year in which the ring was formed (n)*'. Results does not match with declared methods.
Caption: ‘optimal time windows used in annual mass balance reconstructions’ is misleading. If I correctly understood, these are the optimal time windows resulting from the correlation analysis between the tree-ring parameters and the meteorological series. Mass balances are not involved in these results.

[#035] Table 3: consider using B_s, B_w, CWT_{rad}, D_{rad}, $\delta^{13}C$, and $\delta^{18}O$ both in table and in caption to be coherent with the rest of the manuscript.