Responses to Reviewer 2

We truly appreciate the sincere point of view of the anonymous Reviewer 2. We found these comments very valuable, insightful and challenging at the same time. We hope we have fulfilled her/his expectations.

****The original comments appeared as one single paragraph in the file that was accessible to us. We separated the paragraph on the main ideas we believe the Reviewer was interested in.***

This manuscript is a case study for reconstruction of streamflow based on tree-ring growth data in Chile. Just to put things in perspective, and I do not mean any negativity here, there is nothing new in this particular manuscript with regard to dendrohydrology. There is almost 50 years of literature in this area; the same regression models, the same way of reconstruction, and the same variable (streamflow). The authors reconstructed summer flows, instead of annual flows, but reconstruction of seasonal variables rather than annual ones happened also several times before but perhaps not in Chile. So, from my point of view, there is not any aspect of novelty in this research. It is important to state this, at least to give the authors a chance to clarify in case I missed something, and I apologize if I did.

RESPONSE: We truly appreciate your comment because it gives us the opportunity to better stress the importance of our study. This summer reconstruction is important for three main reasons: (1) Chile is one of the countries undergoing strongest precipitation decreases in the last century and (especially the region between 37° and 42°S) it is where there is more agreement among models on further reductions by the end of the 21st century (Koirala et al 2014); understanding how these changes translate into streamflow is needed in order to provide accurate information for developing adaptation and mitigation actions. (2) Although there are many studies about streamflow reconstruction, as the reviewer correctly points out, our work is different and novel relative to those because we explicitly aim to provide an evidence-based criticism of current and proposed water rights regulations and practices. As we posit in the manuscript, the Chilean model is fairly unique and has been studied at the international level for several decades, but to our knowledge this is the first time a dendrohydrological study goes into using this kind of data to discuss the implications for water resources management in detail; we hope that our study encourages other groups to begin a evidence-based discussion on these matters. (3) From a technical standpoint, reconstructing summer streamflow is challenging because relative variations (in percentages) of the average streamflow can be very large at the interannual scale. This way, as the first study focused on this season we believe our work is major step toward understanding base flow dynamics on a multi-century scale.

The section that is most interesting to read is section 4, which the authors called “Discussion and conclusion”. In this section, the authors argued hard for the utility of tree ring-based reconstruction to identify droughts that are more sever or more frequent than those inferred from the instrumental record. This argument can be found in so many of the dendrohydrology papers published in the last decade, and again nothing here is new. However, I have a fundamental issue with the scientific foundation of the argument, and unfortunately this applies also to several other papers published in this field. The authors reconstruct almost 300 years of streamflow data and start comparing it with

instrumental record of 60 years. Ob- viously, they find droughts in the 300 years with characteristics that are different from those in the 60 years, of course! But for water resources engineers, whatever you find in 300-year record MIGHT be a 300-year drought. It is unfair to compare it with the 60-year record. Engineers would fit a statistical distribution to the 60-year record, estimate 100 or 1000-year droughts, then fit a distribution to the longer reconstructed record, and again estimate whatever drought quantiles you want, then compare. Oth- erwise, engineers and water resources planners never use just the deterministic short instrumental record of flows. If you can prove, based on the analysis I suggested, that the reconstructed flows lead to significantly different frequency or severity of droughts, then you made the case about the utility of reconstruction.

RESPONSE: We agree that there are several papers on reconstructions in different parts of the world. But we insist that there is very few research explicitly proposing approaches for using this information in water management such as our manuscript. In South America, there are only 7 papers published on streamflow reconstruction, a small number considering the large and complex river network across the continent. About the other criticism regarding the unfair comparison between observations and reconstruction, we fully understand the concerns of the reviewer. We want to be very clear here, we are not dismissing the importance of available records and if our writing in some ways suggests that, we assure you it was unintentional. Yes, we agree it is unfair, but it is relevant to keep in mind that the records for the last 60 years have shown important streamflow reduction (e.g. Garreaud et al 2017). We are convinced that our study provides long-term context for these recent fluctuations and our aim is to provide evidence for further discussion in both, the hydroclimatic and water management communities. Thus, our results suggest that the post-1980 period is a fairly unique dry one (for summer) in the context of the last 300 years. In the context of the instrumental record, the post-1980 period represents ~50% of the composite instrumental time-series, while it only represents ~10% of the reconstructed period; we believe this highlights the uniqueness of the post-1980 period.

Other specific points: Page 2, Lines 8-10: If you have the future projections based on the CMIP5 results, why don't you investigate if projected future droughts are more sever/frequent than the past ones?

RESPONSE: We appreciate the suggestion but our objective in the paper is to provide context for the low flows that have already occurred. The section you mention is part of our literature review. We haven't utilized CMIP5 projections in this study but it is important to point out that Garreaud (2015) presents a figure where drought recurrence is calculated using CMIP5 output. In that document, it is clear that in the RCP8.5 droughts become more recurrent and that even by the end of the century the definition of drought becomes irrelevant. We will include this explanation in our discussion section.

Just an idea; Page 3, Line 27: How did you get the "natural regime"? Did you account for irrigation abstraction?

RESPONSE: Thanks for the suggestion, in the new manuscript we supplement our explanation of natural regime with a text similar to this: In this reconstruction, we selected stations from rivers where the water has not being diverted for irrigation and hydroelectricity. The three selected stations fulfill these criteria. As matter of

fact, there are two protected areas here: Reserva Nacional Malalcahuello and Parque Nacional Conguillío.

Page 3, Line 29: This sentence is not clear. How did you use the double mass curves to determine the calibration window of time?
RESPONSE: Thanks for catching this up. We believe we need to improve our writing and word choice here in order to explain this idea more clearly. We utilized double mass curves to determine periods that more closely follow precipitation and thus provide us support to detect unreliable records. We eliminated records that didn't fulfill this criterion during the first few years of the time-series.

Page 4: I feel that I miss proper information about the hydroclimatology of the region. How wet or dry is it? How much is the rain and its variation? Just provide some background in- formation;
RESPONSE: We appreciate your suggestion. We have included a climograph (Figure R2.A) using data from Temuco for the period 1980-2010. In this figure you can see that January and February get the lowest amounts of rainfall.

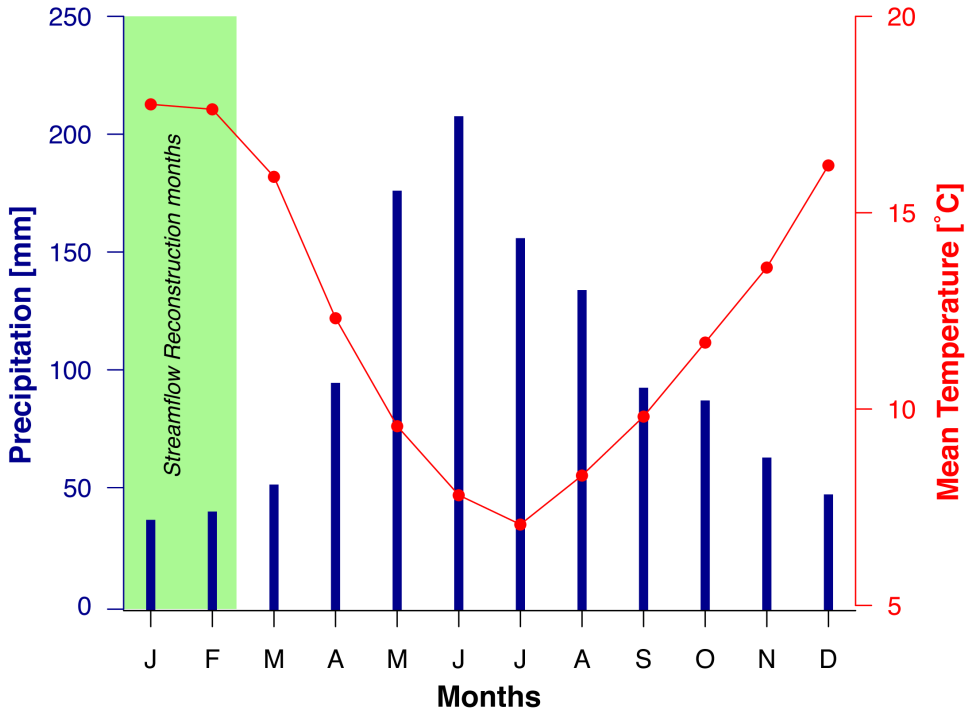

Figure R2.A: Climograph representing the conditions for Temuco for the period 1980-2010.

Page 4, Line 17: You are referring to Table 3 before Table 2, please reorder the Tables;
RESPONSE: We already answered this comment to the Reviewer 1: "We do not agree with this suggestion. Table 2 follows Table 1 in the sense the Table 1 presents the instrumental record and Table 2 provides the information for the analysis of that instrumental record. Then Table 3 appears because it is about the tree-ring chronologies. If we do the change suggested, we feel the manuscript loses readability and that our line of argument weakens".

Page 4, Lines 24-26: peak over threshold is usually used for floods, but here you are doing drought analysis. Do you mean flow below threshold or something like this? You also need a reference for portion;
RESPONSE: Thanks for catching this up. We will improve the description in that section. In effect, we use that method (fully described in Mudelsee 2010) to identify percentiles below the 20% during the whole reconstruction. Our objective was to determine how frequent this percentile has been in the reconstruction. This kind of analysis/representation has been utilized in papers analyzing the BioBío river (Muñoz et al 2016), PDSI reconstructions (Christie et al 2011), and in instrumental records (González-Reyes et al 2017).

Page 5, Line 25: What are these percentage numbers (54.06% and 74.12%)? Do you mean m3/s?
RESPONSE: These are percentages with respect to the average flow. We have modified the text here in order to make it clear the meaning of these numbers.

Page 6, Line 2: "that" should be "than", and "here" is not clear. Do you mean your manuscript? If so, why is the reference, it is confusing;
RESPONSE: Thank you for catching this up. We have modified that section (Lines 1-3) for the following: "Two of three Araucaria araucana tree-ring chronologies extended 800 years or more (PAG and LYV; Fig. 3 and Table 3), which corroborates findings from a previous study (Mundo et al 2012) on the potential of this species for providing long paleoclimatic reconstructions."

Page 6, Lines 11-17: These results are not really good (especially, RE of 0.36), I know they are typical in many dendrohydrology studies, but they should at least make the authors a bit more humble and lighten the assertion tone that is coming later in Section 4;
RESPONSE: We agree with this comment. We recognize this value does not look too good; we have modified the text in order to explicitly specify that this results is good in the context of dendrohydrology. The Reduction of Error (RE) accounts for the relationship between the actual value and its estimate. This RE is however good for dendrohydrology studies. A classic paper in this field is Woodhouse (2001) about a reconstruction of streamflow in the Colorado Front Range where the RE was 0.277. In another important paper, Sauchyn et al (2015) report a RE of 0.73 for a reconstruction of the Atabasca River.

Page 6, Lines 23-24: What does this sentence mean?
RESPONSE: Thanks for this question. We have modified this sentence in order to show more clearly at what temporal scale the dry years calculated from the instrumental record fit into the reconstructed streamflow: "In order to assess the uniqueness of this recent period of extreme summer streamflow, we (a) divided the tree-ring reconstruction into continuous periods of one, five, 10, and 20 years; and (b) we ranked those periods according with their departure from the mean. According to this classification, the dry period 1996-2000, one of the driest in the instrumental record, ranks fourth in the reconstruction, closely followed by 1987-1991".

Page 6, Lines 25-27: Is there any meaning for these windows of 5-year, 20-year,..etc? Of course, every time you change the window, you can get different results, but what are we supposed to learn from this?

RESPONSE: We understand the concern of the reviewer and we clarify this in the new version of the manuscript. Although it is true this time windows are arbitrarily specified, our interest here is to provide context for the occurrence of extreme flows along the whole study period in way that is easier to understand for water managers. Yes, different windows will show different results, but the stress that our intention here is to provide context for the driest summers in the record. This method helps us in perform a more robust comparison between extreme years showing up in the instrumental record and those from the long-term reconstruction. We consider the use of these windows as reliable tool because have allowed us to determine the uniqueness of the post-1980 period.

Page 6, Line 27: You cannot really use reconstructed flows to comment on extremely high streamflow. Look at your Figure 4 (top left) and you will agree with me;
RESPONSE: We understand this concern and that is the reason why we briefly describe high flow in that section and instead we focus on low flows. In fact, in the discussion and conclusion section we already had developed a possible explanation for this behavior based upon our results and literature review (Page 9, Lines 7-29). With the new division of sections as requested by the Reviewer 2, we have modified this section in order to better express our ideas.

Page 6, Lines 29-33: I cannot understand this portion;
RESPONSE: We concur with the Reviewer this section does not read well. We have modified those lines with the following text "Since 1980, years in the lowest 20th percentile of the reconstruction have become more frequent. We calculated the return period of these low flow years in different periods of the reconstruction. We found (a) that during 1709-1750 and 1940-1960, events with streamflow below the 20th percentile had a 20-year return period; (b) a 5-year return period in 1750-1880; c) a predominantly 2 to 3-year period for 1880-1930; and d) a trend toward a 2-year return period since 1960 (Fig. 5)."

Page 7, Lines 1-5: What does this argument imply? Rain and streamflow are different! So, how did you conclude that it is a pluvial system? I think you need to elaborate;
RESPONSE: We appreciate this comment, which allow us to further clarify the correlations we performed in our study. Yes, rain and streamflow are not the same, but it is important to remember that the basin we are analyzing is in a temperate climate. In this region, streamflow data clearly shows that river regime goes from nivo-pluvial regime (high elevation) to pluvio-nival and purely pluvial in lowland areas. Thus, it is expected that streamflow in lower sites are more correlated with rainfall. Thus, the correlation with the rainfall reconstruction validates the record (similarly as for the double mass curves), but it additionally demonstrates the pluvial character of the streamflow at this location and corroborates that our reconstruction is skilled in representing the hydroclimate of the region.

Pages 7-11: Almost half of the paper came under one section called Discussion and Conclusions. This is a style and format issue that does not look good. You need to include more analysis with the Results section, then not very long Discussion section, then a separate Conclusions section, this will be better.
RESPONSE: Thanks for this suggestion. We have split discussion and conclusion sections, shortened the former and added some new analysis in the results

according to previous suggestions by both reviewers (e.g. IPO analysis and climographs).

Page 7, Line 31: Usually trend analysis is misleading. Have you looked at the trend of the entire reconstructed record?
RESPONSE: Thanks for this observation. Earlier in our research we calculated the slope for the entire record and it wasn't statistically significant. We did a new analysis considering the period 1709-2015 (according to other analyses suggested by both Reviewers) and found a non-significant slope of -0.0003. For reference, we will include this value in the first part of the results that describe findings from the tree-ring reconstruction.

Page 8, Line 5: I got confused, was that SAM work done in this study or taken from other studies?
RESPONSE: We apologize if this does not read clearly in the manuscript. We will modify it in order to state more clearly that the SAM indexes have been drawn from public sources (e.g. NCAR-NCEP) and databases associated with peer-reviewed publications (e.g. Villalba et al 2012).

Page 8, Lines 10-12: On what basis was this statement made? Looking at Figure 6 and the correlation numbers does not give me the same impression that the authors have;
RESPONSE: We understand the concern of the reviewer. We think this is an issue of word choice in our text; we wanted to state that the significant correlation with SAM indicates that our reconstruction captures characteristics of the regional hydroclimate. The use of the article "the" gives the impression that we are claiming that the reconstruction captures "all" features of the regional hydroclimate. We changed that statement for the following: "The tree-ring reconstruction showed a statistically significant correlation with the SAM reconstruction presented in Villalba et al. (2012), especially for the long-term trend; we consider this result confirms that our reconstruction captures characteristics of the regional hydroclimate." We have also modified the caption on Figure 6 to indicate that the asterisks represent statistical significance.

Page 11, Lines 21-24: I cannot find proof for this in the manuscript, perhaps the authors need to rewrite this;
RESPONSE: We agree with your assessment. We have deleted this sentence from our manuscript.

Figure 2 (but also a general comment): On what basis was the selection of Jan-Feb only? Why not March and April too? They also seem part of the low flow season, especially that it may not be a good idea to call two months a drought; The authors also need to note that averaging the streamflow of three stations may reduce the variability in individual gauges, and make the reconstruction easier (nevertheless, the reconstruction accuracy is not high anyways). So, you need to justify this;
RESPONSE: We selected January and February as representing summer because they the only two months that fall completely in this season. In the Southern Hemisphere the summer runs from December 21st to March 21st. I addition, these two months correspond to the lowest rainfall of the year, as shown in the Figure R2.A
Concerning the averaging if the individual gauges and the implications for the reconstruction, we agree with the reviewer that averaging the stations may reduce

variability of individual gauges, but the objective of tree-ring reconstructions is to generate regional (for instance basin-scale) hydroclimatic time-series rather than simulate individual records (gauges). An individual gauge station can be affected by local features, but a composite time-series has more chances to average-out local particularities and thus provide a common regional signal that can be compared with the regional signal of a (or several) tree-ring chronology (ies). Despite this, we consider important to recall Figure R1.A (response to Reviewer 1) where we show that our reconstructed streamflow correlates significantly with observed streamflow of each individual station, and also each individual station correlates significantly with the composite time-series produced in our study. In addition, in the response to Reviewer 1 we further explain, citing Figure 1, that the reconstruction compares well with the instrumental records utilized for generating the composite.

Table 1: Please report the standard deviation also or better, the coefficient of variation to see the variability of each series;
RESPONSE: Thank you for this suggestion. We will add the coefficient of variation to that table.

Table 3: What are those LAN and VILL in the Table title?
RESPONSE: Thanks for this comment. We added the description of these chronologies in the caption of the new Table 3.

Table 4: The autocorrelation of the tree ring chronologies are quite high, and is usually transferred to the reconstructed annual flows. I find this unrealistically high for annual flows, can you comment on this and its impact on the reliability of the reconstructed flows? Can you compare it with the autocorrelation of the instrumental flows?
RESPONSE: We agree with the reviewer and we believe it is important to include information on autocorrelation in the manuscript. Now in the manuscript we include the following text: "The tree-ring reconstruction presents high autocorrelation, consistent with the fact that tree-growth has a temporal memory associated with the water reserve and the soil moisture that remains and is captured by the tree. Some of the statistical procedures applied to the tree-ring chronologies are meant to minimize these effects, but it is virtually impossible to eliminate all. In our case, the autocorrelation is 0.56 to 0.49 for each individual chronology, while it is 0.248 for the reconstruction (1709-2005). This autocorrelation is still high considering that the instrumental record is essentially free from autocorrelation (-0.093)".

New References

Christie, D.A., Boninsegna, J.A., Cleaveland, M.K., Lara, A., LeQuesne, C., Morales, M.S., Mudelsee, M., Stahle, D.W. & Villalba, R.(2011) Aridity changes in the Temperate-Mediterranean transition of the Andes since AD 1346 reconstructed from tree-rings. Climate Dynamics 36, 1505–1521.

Garreaud, R. D., Alvarez-Garreton, C., Barichivich, J., Boisier, J. P., Christie, D., Galleguillos, M., LeQuesne, C., McPhee, J., and Zambrano-Bigiarini, M.: The 2010–2015 megadrought in central Chile: impacts on regional hydroclimate and

vegetation, Hydrol. Earth Syst. Sci., 21, 6307-6327, https://doi.org/10.5194/hess-21-6307-2017, 2017.

González-Reyes, Á., J. McPhee, D.A. Christie, C. Le Quesne, P. Szejner, M.H. Masiokas, R. Villalba, A.A. Muñoz, and S. Crespo, 2017: Spatiotemporal Variations in Hydroclimate across the Mediterranean Andes (30°–37°S) since the Early Twentieth Century. J. Hydrometeor., 18, 1929–1942, https://doi.org/10.1175/JHM-D-16-0004.1

Koirala, S., Hirabayashi, Y., Mahendran, R., & Kanae, S. (2014). Global assessment of agreement among streamflow projections using CMIP5 model outputs. Environmental Research Letters, 9(6): 064017. doi:10.1088/1748-9326/9/6/064017.

Mudelsee M (2010) Climate Time Series Analysis: Classical Statistical and Bootstrap Methods. Springer, Dordrecht.