Clim. Past Discuss.,
doi:10.5194/cp-2017-26-RC2, 2017

Climate of the Past Discussions
Open Access
EGU

Interactive comment on "Blue Intensity based experiments for reconstructing North Pacific temperatures along the Gulf of Alaska" *by* Rob Wilson et al.

M. Rydval (Referee)

rydval@gmail.com

Received and published: 9 May 2017

General comments:

This manuscript investigates the climatic utility of tree ring data from the Gulf of Alaska. The study experiments with a range of chronology variants utilizing various tree-ring variables (including ring width - RW, latewood blue intensity - LWB and delta blue intensity - DB) and standardization options in an attempt to examine the sensitivity of chronology structure to these choices and to identify the most suitable options for the purpose of developing reconstructions that are minimally biased and accurately represent past climatic variability. The manuscript primarily focuses specifically on the ability of these tree ring variables to accurately represent lower frequency trends. The analy-

Printer-friendly version

Discussion paper

sis is performed on tree ring series from a species and a region previously unexplored with respect to the Blue Intensity parameter. In particular, the use of the DB parameter has thus far only been minimally explored in experimental studies. In this sense this work represents an important contribution towards understanding the behaviour of Blue Intensity data and the suitability of such data for reconstructing summer temperatures. On the whole this article is well written, logically structured, without any major fundamental problems and supported by a set of generally clear and relevant figures and tables. Nevertheless, there are a number of specific issues that could be addressed in order to further improve this article and a range of suggestions are detailed below. Some of the limitations of this work include the relatively low series replication for individual sites, limited exploration and discussion of certain aspects of standardisation and chronology development, and some methodological considerations particularly in relation to the BI parameter which could be elaborated on in more detail.

Specific comments:

Considering the experimental nature of the LWB and particularly DB parameters, it would have been useful to develop even a limited MXD dataset on at least part of the samples (e.g. from one site) in order to enable a direct comparison of the lower frequency trends in the BI data. Although it is argued that the structure of mountain hemlock wood makes it more difficult to prepare and measure these samples for density, it is not impossible and has been done in past studies. This would have been helpful in evaluating and constraining the utility of differently detrended BI chronologies and therefore considerably benefited this study in further strengthening the case for the use of DB as a better, less biased parameter relative to LWB and a suitable alternative to MXD for this species, especially since this is the first study to measure BI on mountain hemlock samples. Was this option at all considered?

I am somewhat surprised that a higher number of samples was not used for the individual sites. According to Table 1 replication should ideally be 12-28 series for LWB and 14-36 for DB depending on the site. In several cases, the actual maximum number

of series used is below (and in some cases well below) this optimal level. Is this not a problem? The weaker signal strength of BI data and the need for higher replication in order to develop 'robust' chronologies is acknowledged (e.g. L368-369). Also, the relatively low replication may even affect the RW data as stated in L229-230. As stated in the text, a subset of samples was selected for this study from earlier work so why not aim for 25-30 samples per site? That would have at least reduced uncertainty about the representativeness of some of the BI site chronologies, particularly in earlier periods when replication is likely even lower. It would be nice to see a replication plot over time (and EPS plot) for separate sites as well as for the 'all series' pooled version (perhaps as an SI figure) to give a better indication of which periods might potentially be affected by low replication.

L96-98 and L217-219 – What is the rationale for this statement? To my knowledge this issue has not been investigated in any previous study. Presumably a higher correlation between EWB and LWB would imply that EWB expresses in part the same information as LWB, but does that necessarily mean that this information is related to climate? How do you define 'weakly correlated'? Or in other words, what correlation would be acceptable and what would not? Ideally, this statement could be supported with an example and actual data. If nothing else, I would suggest elaborating further on this statement to more clearly express the justification for this claim.

L102-105 As a general comment, some of the limitations of BI (specifically LWB) have already been explored in other studies. Clearly DB is a major improvement, although I wonder just how well DB resolves these issues and specifically whether DB could still have some problems at lower or other frequencies. It is interesting that in some cases the calculation of DB weakens the common signal, suggesting that information which should ideally be preserved is to some extent being removed in the process, yet in other instances the strength of the common signal remains relatively unaffected or even shows improvement. I suppose these questions can only be answered by various future studies that will further explore DB and I would not expect this to be fully

Printer-friendly version

Discussion paper

covered here. But perhaps a statement could be included somewhere to caution and emphasize that considerable uncertainty remains with respect to the performance of DB and more work is required in this area.

L109-111 - Is there any indication to what degree early instrumental biases could be a limitation (if at all) in achieving the stated aim?

(L133-138) Is there actually any need to detrend DB series? What is the justification for this? Hypothetically, if both LWB and EWB contain the same ontogenetic trends then by the nature of the DB calculation this trend would be automatically removed. I do not know whether or not that is true. This may be a more complex issue - perhaps only the LWB contains this trend or the LWB and EWB trends related to age differ in some way. But is it not possible that by detrending the DB data some lower frequency climatic information may be unnecessarily removed? Was the development of DB chronologies without performing detrending considered or explored in the analysis? The DB chronology in Figure 5 actually looks like a reasonable chronology variant and so I wonder how non-detrended DB chronologies would perform in terms of calibration and validation statistics relative to the detrended versions.

L146-147 – This is a fairly short validation period. Why not choose an equally long calibration and validation period which has been a common approach in other similar studies? How sensitive are the results to this choice? L244-245 acknowledges that this may be an issue. For example, would a different period affect the significance of any results in Table 4?

L147-148 and top panel in Figure 8 - Why not perform a nested PC reconstruction? By excluding even just one or two of the shortest sites this reconstruction could be extended to the mid or early 18th century, providing more information for comparison with the other reconstructions.

164-173 - Was the variance stabilized to account for changing replication / r-bar through time? Even just a look at the changing variance with replication in Figure S4 in the

supplementary material suggests that this should be performed. Either I have missed this or it is not stated in the methods section.

L171 – Figure S4 suggests that linear detrending may not be the most appropriate choice to detrend for example the DB data. Based on the initial increase in the juvenile period of growth in the DB results, perhaps a somewhat more flexible detrending alternative could be explored to account for this initial increase? Additionally, Figure S4 raises some intriguing questions about non climatic trends in the BI data. I suggest that a fourth panel showing results for EWB should also be shown here. It appears that the initial juvenile trend is present in the EWB only. Does this not suggest that the EWB should not be used to 'correct' LWB in this initial ∼30-50 year period?! Because it appears that this initial trend is not present in the LWB data, but is introduced into DB by the EWB data making it necessary to then remove this trend again from the DB series. Due to the experimental nature of this TR variable, several methodological considerations such as the one discussed here remain unaddressed and have not been explored in this study. Considering the nature and aims of this manuscript, I would not expect a detailed evaluation of DB, however, methodological issues such as this and the need to explore them further should at least be acknowledged and more clearly highlighted in the text.

Is it reasonable to develop a RW + DB chronology / reconstruction considering the difference in the seasonal response of RW and BI data and the acknowledgement that RW may not be primarily controlled by summer temperatures in this region (L349-351 and L379-380)?

L187-189 – Maybe already refer to Figure 2 at the end of the first sentence.

Perhaps specifically state somewhere in section 3.3 that JJAS was selected for further analysis for BI (presumably because this is the optimal season).

L196-198 – What might be the reason for such a broad seasonal response in the RW data in this region? Has this been discussed in any of the previous studies (e.g. Wilson

et al. 2007 or Wiles et al. 2014)?

L216 – Just an observation, it is a little surprising that RW agrees more strongly with DB and LWB than DB does with LWB.

L219-220 - Maybe worth clarifying here that EWB would not be expected to contain an actual climate signal in the first place.

L241-242 – Is it possible that the failed validation could be related to the quality of instrumental data for this early period?

L278-280 - This may simply indicate that the use of LWB and EWB to calculate DB is an imperfect procedure. It would be necessary to look at data from other locations (and also other species) to identify whether the DB trends in Figure S4 actually reflect inherent properties related to age (and should therefore be detrended) or are related to other factors. This is an important issue and it is unfortunate that this paper does not or cannot investigate this type of issue in more detail. But perhaps at least a bit more discussion could be included in relation to this?

L294-295 - The results in Figure S4 already indicate that a linear detrending curve can lead to a serious underestimation in the early parts of the series so the poor performance of LINsf is not surprising.

L314-317 - But as discussed in the text (and in relation to the results in lines 286-291) it is apparent that LWB is inherently biased. So why even consider it as a feasible option?

L330 – Maybe include '(Figure 6 and 7)' in the bracket since this reconstruction appears in both figures and is discussed in relation to Figure 7 in the previous paragraph.

L329-331 – It may be worth stating here something along the lines that the 'best performing' PC and extended reconstructions are shown here and compared with Wiles and the glacial advance record - i.e. state the reason why these reconstructions are shown in Figure 8.

L358 – Or for species which do not have this colour difference to begin with.

Figure 3 – It is interesting that LWB calibrates and validates more strongly than DB in terms of the strength of the relationship. Why might this be the case? Could this difference be related to replication?

Figure 6 - It is somewhat difficult to identify the trend of the LINres curve for LWB and especially DB - is it possible to improve the visibility of these curves? Also, in the figure caption (L596-597) maybe consider changing 'low plots' to 'lower set of plots'.

Figure 7 – Why not also show r2 rather than r for the 1850-1900 period?

Table 1 - Is there any real meaning in including N-EPS information for EWB data? Presumably these data do not (or should not) contain any common climatic signal and there would be no point in developing a chronology from these data that would be of much use.

Table 2 – Minor detail, but why not arrange the site order from west to east?

Table 3 – Why is EWB positively (though weakly) correlated with DB?

Technical corrections:

L32 – affecting instead of effecting

L71 – 'as they are a measure of ...' instead of 'as they measure' may be a more accurate statement.

L109 – reconstructions instead of reconstruction

L126 – Please specify the exact calibration target type as there are different versions. IT8.7/1 is a transparency target whereas IT8.7/2 is a reflective target. I assume that the latter was used.

L211-212 – Consider rewording 'potentially optimal' to something along the lines of 'more optimal compared to a PC approach'.

L218 – Consider specifically stating that the correlation between EWB and LWB is not significant. Also, ‘of’ missing in ‘the utilization DB to’.

L288 – change ‘particular’ to ‘particularly’

L637 – Is there a better word than ‘dominant’ which could be used here?

Interactive comment on Clim. Past Discuss., doi:10.5194/cp-2017-26, 2017.

Printer-friendly version

Discussion paper