

Biogeosciences Discuss., referee comment RC2  
<https://doi.org/10.5194/bg-2021-119-RC2>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on bg-2021-119

Milos Rydval (Referee)

---

Referee comment on "Evaluating the dendroclimatological potential of blue intensity on multiple conifer species from Tasmania and New Zealand" by Rob Wilson et al., Biogeosciences Discuss., <https://doi.org/10.5194/bg-2021-119-RC2>, 2021

---

### General comments:

The manuscript, which investigates the dendroclimatological potential of blue intensity parameter datasets from a broad range of Australasian conifers, represents an important addition to research into the potential utility of the underutilized blue intensity tree ring parameter in a dendroclimatological context. Overall, this study provides valuable information about the behavior, climate response, and reconstruction potential of blue intensity in parts of Australasia and helps open up a new frontier in the utilization of this climatically sensitive and affordable tree ring parameter in that region and the Southern Hemisphere more broadly. The manuscript is generally well written, nicely structured and logically organized, although some additional context or clarification would be helpful in certain parts of the text. However, there are several important aspects of the manuscript, which could be improved, addressed or more thoroughly discussed (detailed comments are provided in 'specific comments' below):

- It is unfortunate that lower frequency trends were not assessed, as this would have greatly enhanced the significance of the study and the lack of resin extraction or other form of sample treatment represents a substantial limitation of this study. While this limitation is acknowledged and addressed by adopting a very flexible detrending approach with all analyses focused on high-frequency relationships, as a result, the conclusions that can be drawn are limited. For example, the full potential advantages of the DB dataset remain unexplored. With that in mind, interpretation of the results and some of the statements in the discussion could more explicitly acknowledge this limitation.
- Considering the relatively high BI replication requirements to generate a representative chronology, it is not clear whether and to what extent some of the identified relationships could be affected by the relatively low and variable replication between sites and sub-optimal EPS.
- The quality of the gridded instrumental dataset with respect to its local representativeness over space and time in some locations could possibly affect the results. Also, the aggregation of the gridded data across New Zealand could potentially

be problematic and should be examined in more detail.

- Although the importance of site elevation / latitude in relation to the treeline is highlighted in the introduction, there appears to be no consideration of such factors in this study, which could certainly affect the RW results and may also play a significant role in influencing the climatic response of the BI parameters. Large differences in elevation between sites in Tasmania and New Zealand along with the large latitudinal range (and possible climate response gradient from north to south) of the New Zealand sites may also play an important role and should be considered.
- While a consistent EWB response does genuinely appear to exist throughout this network, the broad and diverse range of relationships between all of the parameters across species and their responses suggests that (at least w.r.t. assessing climate response) it may be more appropriate to investigate and discuss such inter-connections in more detail at an individual site or species level. Combining all data together and extracting the overriding signals using multivariate techniques may overlook other potentially useful climatic signals and relationships present in subsets of the full dataset, especially since multiple species and parameters are involved. This could be evaluated with a more nuanced approach.
- Although the BI-based dataset calibrates very well and is indeed comparable to the QWA-based reconstruction within the context of this particular study, some of the statements in the manuscript may somewhat exaggerate the capabilities of BI in relation to QWA even without considering the lower frequency limitations of BI.

Overall, the most significant limitations of this study should be addressed in some way. Detailed comments and recommendations on a range of aspects are provided below and this feedback will hopefully provide useful input that will guide the authors in further improving the manuscript.

### **Specific comments:**

L17: the use of 'blue reflectance intensity' is somewhat confusing and unnecessary. I would suggest using either the term 'blue reflectance' or 'blue intensity', or otherwise reword to 'blue intensity from reflected light'.

L25: In terms of response, DB results appear similarly mixed and temporally / seasonally variable when considering New Zealand and Tasmania sites on an individual basis. The stronger temperature response for Tasmania over New Zealand really only appears in the multivariate PCA results (Fig.5) – please amend the sentence to reflect this.

L27 (& L424-425): This is somewhat misleading as the '52% to 78%' values refer to the (~50-yr) early / late split period calibration segments rather than the full period calibration results. The statement should instead include '55% to 64%', the temporal span of the calibration period should also be stated and the sentence reworded accordingly.

L32: WA should be defined in the abstract and other relevant sections of the text. I would also strongly suggest using the term quantitative wood anatomy along with the acronym (QWA) throughout the text as this would be more consistent with the literature.

L45-46: 'unexpected' in what sense? Please specify / elaborate.

L50-51: While this is true, there are also limits to this and there are other considerations / limitations that can also affect density data (and by extension BI) – e.g., Divergence phenomenon.

L104-105: The last part of this sentence should be reworded since latewood density / LWB does not represent the density of the cell wall itself, but instead it reflects the proportion of cell wall to lumen area.

L105-107 / L435-436: I suggest reformulating these sentences as the statements may overstate the merits of BI, particularly when making generalizations based only on the specific context of this study. While the high frequency comparison of the BI and QWA chronologies is certainly encouraging, I would, however, urge more caution in comparing / equating the information that can be extracted using BI with wood anatomical data bearing in mind that QWA offers higher resolution (individual cell level), intra-annual information, an extensive range of parameters, etc. Since this study only examines the high-frequency component, I would expect QWA to have an additional advantage and an edge over BI in the development of reconstructions not restricted in this way, particularly considering that the utility of high-frequency reconstructions is limited. Furthermore, note that anatomical parameters can even exceed density / BI parameters in terms of climate sensitivity and the climatic response of other anatomical parameters such as maximum radial cell wall thickness (compared for example to mean cell wall thickness / cell diameter) may provide an even stronger signal (see e.g. Björklund et al., 2020).

L108-109: This statement indirectly suggests that some kind of evaluation of wood anatomical properties was performed, which is not the case. Presumably, the intention here is to use climate response info to pre-screen and highlight sites, which could be the focus of dendroanatomical / paleoclimatic research in the future? Please modify the statement to clarify this point.

L111: what is meant by 'low performing'? Relatively poor / weak climate signal? Consider rewording.

L118: It is not immediately clear that this means four species each from Tasmania and New Zealand (i.e. 8 species in total). Please modify for the sake of clarity.

L126-127: Can you specify which wood anatomical properties?

L133-134: In addition to ENSO, is there potential to capture / reconstruct past variability of any other atmospheric circulation phenomenon (e.g. Southern Annular Mode) using any part of this network? If so and if relevant, this could be mentioned in the discussion for example.

L118-134: Most of this paragraph would perhaps be more suited for the introduction rather than the methods section.

L146: Was the option to dismount and chemically treat the samples considered? What was the reason for not attempting this? Also, does the resin extraction of samples from only one site (AHA) influence the site / species comparisons in this study in any way?

L154: Please include a table in the supplementary information section specifying sanding grade, scanner type and scanning resolution for each dataset as well as which parameter settings were used to derive the EW / LW BI data (at least specify the percentage of light / dark pixels used for the EWB / LWB / DB calculation).

L160: Specify the reason (i.e., different behavior of LWB in Australasian tree samples compared to Northern Hemisphere)

L172-173: The rationale for producing DB to try to mainly correct the heartwood / sapwood color transition issues is somewhat contradictory considering that at the same time flexible detrending and retention of only high-frequency variability is used here to bypass the color transition issue. In other words, this form of detrending may eliminate some or even most of the potential advantages of the DB parameter.

L196: It may be worth mentioning that the CRUTS 4.03 datasets start in 1901.

L194-199: This paragraph should also include a statement mentioning that the instrumental series were filtered in the same way as the tree ring series.

L205-207: I am not entirely convinced that such an approach is adequate, particularly considering the relatively large distances involved. The good agreement could simply be due to a limited number of available stations used to produce the gridded dataset (particularly in earlier parts of the 20<sup>th</sup> century) rather than actual good agreement over these spatial scales. It may be helpful to examine the temporal availability of data and location of stations that contribute to the New Zealand gridded dataset (i.e. the three gridboxes used) over time. As spatial biases linked to the availability of instrumental data

could be an issue, it may also be useful to perform the same comparison using only instrumental / remote sensing data from recent decades with better / more representative local spatial coverage to help substantiate this high degree of agreement. In any case, if appropriate, this type of limitation should also be mentioned in the methods or discussion sections.

L210 / L365-366 / Figure 5: Does it actually make sense to use the CE statistic in the context of this study considering that only the high-frequency component is explored? The issue is that the CE stats may appear more favorable with higher values compared to a calibration / verification assessment that would also contain lower frequency variance. The use of the CE statistic may therefore not be as stringent as it may seem considering that the means of the calibration and verification periods essentially do not differ in this case. In other words, even though CE is usually seen as more rigorous compared to the reduction of error (RE) statistic, in this particular case, RE would likely yield very similar values in this context and so it could just as well be shown instead of CE. This should be addressed and clarified in the text.

L223: Although EPS may arguably mostly reflect and perhaps be overly sensitive to high-frequency relationships, the replication requirements to meet the  $EPS = 0.85$  threshold could potentially be even higher for unfiltered / less flexibly detrended series than the 20-yr spline detrended series. This may be worth mentioning.

L257-258: Please specify / expand on why these results are 'extremely encouraging'.

L260: It may be helpful to include a reminder that this refers to non-inverted LWB data.

L261-262: Be careful with the formulation as the current wording may suggest that there is also a negative relationship between MXD / temperature. Consider rewording to avoid any confusion.

L260-270: The interchangeable use of site names / codes and species names is a bit confusing. Please be more considerate and consistent in their use throughout the text. For example in L261 either state the site code and species in brackets or vice versa, but not both. Also, L262-263 starts by referring to sites and the latter part talks about species.

L265: Could simplify to latewood anatomical parameters.

L268-270: Such behavior would typically indicate the influence of moisture limitation. Although the physiological behavior of Australasian species may indeed be very particular, I would still recommend also testing the response of chronologies using a relevant drought

index (e.g., scPDSI) to check that growth is not reacting to moisture limitation (rather than just examining the response to precipitation).

L285: Looking at the response of the EWB and LWB Pink pine (DPP) parameters, the seasonality is somewhat different, however both parameters are positively correlated with temperature and their seasonal response overlaps considerably (just broader / stronger for EWB), so the relatively high inter-parameter correlation is perhaps not that surprising.

L291: 'Disappointing' in what sense? Weak / variable / inconsistent climate signal? Please specify.

L291-294: I disagree with this statement since the very flexible detrending of the DB data in this study precludes the potential ability of this parameter to correct for the heartwood / sapwood transition (or any other multi-year color-related deviations). Hence, by only considering the high frequency, the full potential benefits of calculating DB could not really be assessed in this study and so this is a matter that will have to be investigated further. This also raises a broader point about the relevance of calculating DB while applying very flexible detrending more generally. Perhaps a more species-specific / individual site approach would be required to assess in more detail whether or to what degree DB can be helpful in the Australasian context. Please adjust the text with this in mind.

L297-298: I wonder whether the quality / spatial representativeness of precipitation data could play some role in this. In some cases, could the weak precipitation response possibly be partly related to poor local representation of instrumental precipitation in some locations?

L307-308: It would be interesting to see if the response of DB would be even better compared to EWB / LWB if lower frequency trends were also considered.

L310: This probably suggests that rather than combining data from a wide range of species over an extensive spatial range, a more selective approach targeting specific species with similar parameter responses may be required in order to truly unlock the full potential of these datasets.

L324-325: The degree to which this affects results could be tested. Would the response be just as strong or would it be closer to that of the other species if the two pencil pine sites were treated independently?

L389: Is the elevation range of the Allen et al. dataset roughly similar to this study or are there major differences? Such factors could affect the calibration strength.

L391-392: This implies that only the final reconstruction of Allen et al. was filtered with a 20-yr spline rather than detrending the individual series (as was done in this study). Could this affect the calibration strength in any way?

L397: The Berkeley temperature dataset could be introduced earlier in the manuscript (i.e., methods section).

L402: Why not also detrend the Berkeley dataset with a 20-yr spline?

L409-410: This is undoubtedly the most important factor. However, other issues that have not been examined may also be important and should be explored. The heartwood/sapwood issue does not necessarily represent the only color bias that may be relevant and so correcting only the heartwood / sapwood shift may not necessarily resolve all of the color-related issues in BI series.

L436-455: In time, perhaps more advanced imaging techniques can also be development to reduce or remove the influence of color on these datasets and so more effectively resolve such issues.

L447-453: Such approaches could certainly represent an interesting solution to this problem. However, at the same time, the effectiveness of such techniques will need to be thoroughly evaluated and it is possible that such an approach may only be applicable to specific instances of color variation even in the context of heartwood/sapwood transitions (i.e. it may resolve some, but perhaps not all such issues effectively).

L455-457: Although again, factors such as elevation / latitude / distance from the treeline, the spatial representativeness of instrumental data, may play a significant role in the sense that mostly just one site per species (or two in a few cases) were evaluated and so it was not possible to evaluate whether the same species in a different location may respond differently. I believe this is an important point that should be mentioned.

Figure 1: Are the coordinates for South / Central NZ correct (they should probably be switched)? The 'o' symbols should also be corrected. However, I suggest adding the lat. / lon. info as a grid directly onto the maps instead.

Figure 2: EPS info in panel B is difficult to see (please use larger font size). The figure shows that series with a higher CV generally also tend to have higher  $\bar{r}$ . However, this relationship is not as clear for the lower RBAR / CV range (i.e., the relationship within that

lower range is considerably weaker and perhaps even absent. This could be pointed out more clearly in the text.

Figure 2 / Figure 3 / Table A1: Table A1 clearly shows that the BI parameters are nowhere near  $EPS = 0.85$ , so how might this generally low replication affect results? Could this affect the representation of climate signal strength in some of these climate response assessments (particularly Figure 3 and A3) and since the replication seems to vary quite substantially (6-16 trees, 10-22 series), do more highly replicated sites have an advantage (i.e. stronger climatic relationships) and is it really a fair comparison? One possible way to check the potential impact of variable replication between sites would be to look at how the results change when all sites are limited to the same nr. of series.

Another related question is how and to what extent the results would differ if the  $EPS = 0.85$  threshold was met by all chronologies? The point is that the lower replication introduces additional uncertainty in relation to the response results and so this should also be clearly stated as one of the limitations of this work.

More generally, could the relatively high replication requirements for some sites affect the attractiveness of BI in some cases (perhaps it would be easier to collect and process five samples and develop QWA series instead of 100+ BI series in some instances)? In any case, this is a matter that should be discussed.

Figure 3: Consider adding Tasmania / New Zealand on the left side next to the site codes to more clearly distinguish the two groups of sites. The species names could also be included to help interpretation. Just as a suggestion out of curiosity, it would be interesting to see the persistence properties of the parameter chronologies (e.g. first order autocorrelation in the SI).

While differences in the response of Tasmanian sites are rather subtle and generally do not appear to be systematic, in contrast, the New Zealand sites may possibly express a gradient in the strength, seasonality, type of response (RW / EWB / LWB) from north to south (i.e., stronger / broader / later / seasonal response towards the south). This may be related to the large latitudinal spread of those sites and should be taken into account in the analysis and interpretation of the results. Furthermore, differences in site elevation and distance from the treeline between the New Zealand and Tasmania sites as a whole may also need to be accounted for, particularly considering that sites in Tasmania are generally located at high elevations compared to the much lower-elevation New Zealand sites. This could make it difficult to directly compare the two areas in terms of climate response as major differences in elevation may contribute to some of the observed differences in climatic response. These factors are not currently accounted for or discussed.

Such factors along with the broad range of species examined could also partly be responsible for or contribute to the mixed DBI response (perhaps predominantly a



reflection of the broad range of LWB response). This may indicate that the behavior of BI parameters should be examined in more detail within a more localized / species-specific framework. Overall, the role of elevation / latitude should be considered in this study.

Figure 4: It is interesting that DB outperforms EWB in Tasmania, but performs very poorly when looking at the New Zealand sites. Does this perhaps suggest that it may be worthwhile to focus on different parameters when working with sites / data from New Zealand vs Tasmania (also keeping in mind that it might be more beneficial to use DB than EWB when lower frequency trends are also included)? It would also be helpful to show the PC loadings for the significant PCs based on the parameter / site variables used in Fig. 4 (A) and (B) as a supplementary table.

Figure 5: It may be somewhat easier to navigate the statistical info presented in the figure if for example all  $ar_2$  stats were displayed at the bottom and CE at the top of the panels. The last approach (variant 3) appears to produce the most consistent pre-instrumental agreement compared to the other two variants. Is this an indication that it would perhaps be more appropriate to move forward with variant 3 rather than variant 2?

Figure 6: Please add a 'correlation' or 'correlation coefficient' label next to the scale in the bottom right of the figure. Why not compare either using Berkeley or CRU rather than both datasets in this figure?

Figure 3 / Fig. A3d / Table A2: The climate response of DB appears to be related to its correlation with EWB / LWB. Higher correlations with EWB while low with LWB tend to yield a stronger positive relationship of DBI with temperature and vice versa. Considering the very low variance of EWB data (Table A2) and generally high negative correlation of LWB and DBI (and since only the high-frequency is retained), would it be reasonable to say that the DB chronologies are predominantly just an inverted version of the LWB chronologies? In relation to this point, I also wonder whether there is actually any real benefit to including the DB data in the PC regression / reconstruction exercise. In fact, presumably the full benefits of using DB are not apparent and cannot be assessed in this study since only the high-frequency is considered.

Interestingly, the relationship between DB and RW seems to be rather strong for most sites (and stronger than RW vs LWB and EWB). Could there possibly be a RW size-related influence that could be affecting the calculation of DB (*sensu* Björklund et al., 2019)?

Figure A1: It would make it easier to visually interpret the figure by including axis labels. Also, consider adding a vertical bar or series of vertical lines to show the position of the heartwood/sapwood transition (e.g., as in Fig. A2).

#### **Minor comments:**

L34: 'parameters' instead of 'variables'?

L35: 'properties' instead of 'features'?

L55: 'climate-limited'?

L56: consider: '... where the application of traditional dendrochronological parameters is less reliable.'

L69: consider adding 'considerable experience' to the list.

L73: consider: 'MXD and BI essentially measure wood properties that are integrally linked'

L83: This may simply be related to the way the font is displayed, however, please change 30oN to 30°N.

L86: ... at temperature-limited locations / sites

L92: delete 'the' in '... on the New Zealand's ...'

L101-102: derived from maximum intensity values of the whole-ring reflectance spectrum

L108: change 'confers' to 'conifers'

L125: please change 'and has not been' to either 'and have not been' or 'and this species has not been'

L131 / L392 / L401: Please add '.' to 'Blake et al 2020' and 'Allen et al's'

L166-167: consider '... proposed a (statistical) procedure that could potentially correct ...'

L185: 'more gradual' may be more appropriate than 'slower'

L226: Perhaps consider changing to 'The common signal is particularly weak for the following species: ...'

L245: Specify 'full 1902-1995 period'

L283: add 'respectively' at the end of the sentence

L297: Consider changing 'these chronologies' to something like 'these RW and BI chronologies' or 'the examined chronologies'

L303 / (L316): Consider changing to '... PCA identified three (RW), two (EWB), two (LWB) and two (DB) principal components ...'

L329: 'Region-wide'?

L350: 'The Tasmanian modelling was performed against ...'?

L366: 'ar2' should be defined somewhere

L380: 'high-pass'

L387: 'regression-based'

L394 / 407: 'BI-based'

L425: 'higher' rather than 'greater'

L432: It may be helpful to reword slightly simply to make it clearer that 'this inverse relationship' refers to the opposite response of SH sites with temperature compared to the relationship of NH sites with temperature (i.e. not an inverse relationship of SH site LWB with temperature).

L469-470: consider changing 'grey bars' to 'grey shading' and '1<sup>st</sup> / 2<sup>nd</sup>' to 'first / second'

L475: change 'Heartwood/Sapwood' to lowercase (heartwood/sapwood)

L514: 'to undertake preliminary analyses'

Table 1: (caption) please change '5' to 'five'. (table heading last column) 'Period - 4 series' to something like 'Period > 3 series' or 'Period min. 4 series'

Figure 4 (caption) / L312: '... for all / individual parameters (A) and species (B).' Maybe also specify somewhere that the species calibration tests in (B) involve the full range of parameters.