

## ***Interactive comment on “Forward modelling of tree-ring width and comparison with a global network of tree-ring chronologies” by P. Breitenmoser et al.***

**Anonymous Referee #1**

Received and published: 18 August 2013

### **1 Summary and General Comments**

Breitenmoser et al. perform a validation study of the VS-Lite (VSL) model of tree-ring width by simulating a global network of tree-ring chronologies using the CRU reanalysis product, and comparing the results to observations. The analysis is performed both for individual chronologies, and also for “aggregated” chronologies, which are constructed to represent an average growth index for trees within a common region and sharing a common climate signal. The authors report that the model is generally able to simulate the observations skillfully, with increased skill using the aggregated series. This result indicates that the aggregation procedure increases the signal-to-noise ratio

C1763

(where “signal” implicitly refers to the time series variations captured by VSL). An alleged motivation for the study is to explore the use of VSL as an observation operator for use in paleoclimatic reconstructions proceeding by data assimilation (DA) methods. However the authors’ analyses and discussion stop just short of mapping out a way forward toward this goal.

The analysis presented in the paper is well-founded. However, with just a small extension to the work already presented, the authors could make a much more novel, interesting, and impactful contribution by providing explicit direction about building an observation operator from VSL for DA-based paleoclimate reconstructions. The method by which they aggregate temperature- and moisture-related tree-ring data separately up to the scale of the CRU grid is already strongly suggestive of a way to construct such an observation operator. At the very least, the authors should explicitly discuss how the comparison they do between simulated and observed aggregated TRW could be expanded or extended for such a purpose. But the authors might even consider refocusing the entire paper toward this aim. Revision along these lines would require a quantitative formulation of the discrepancy between the simulated and observed aggregated TRW series (eg. what is the size of the model error at each gridcell? Is the model error spatially and/or temporally correlated?). In fact this type of analysis should be included in the forward-model validation context already presented.

Whether or not the authors choose to shift the focus to paleoclimatic data assimilation, the paper requires reorganization to clarify the main points of the study. The goals of the work as outlined in the abstract and introduction should be clearly revisited in the Results, Discussion, and Conclusion with the related findings for each goal. The methods, results, and discussion are also interleaved at present, which adds to the confusion and blurring of the main findings. I suggest separating the “Results and Discussion” sections (where “Results” should be an objective report of the findings, while “Discussion” should contain interpretation and speculation, and discuss potential applications and connections to existing literature). The aggregation procedure described

C1764

in 3.4.1 is quite important, and should be described in the "Methods" section. Most of the specific comments below are suggestions for re-structuring and focusing the paper.

Additionally, in many paragraphs throughout the paper, the analysis is described in the order it was performed, rather than in the most logical presentation order. More specifically, often a key result, number or method comes only at the very end of a paragraph, rather than at the top where it would be more clear for the reader. Several of these instances are also noted below.

## 2 Specific Comments

### 2.0.1 Abstract

- line 9, Goal a): suggest making this goal more specific. The authors want to examine the relations between simulated and observed series for what purpose?
- line 10, Goal b): as VS-Lite is a *forward* model, it has no intrinsic potential for reconstructing past climate on its own, as paleoclimatic estimation is an *inverse* problem. Again, a more specific phrasing may be all that is needed here, eg. "...to evaluate the potential of the VSL model as an observation operator for data-assimilation based reconstructions of climate from tree-ring width."
- line 11-13: Is the result about the parameter estimates a key finding of the study that belongs in the abstract? Constraining the model parameters is not listed as a main goal of the study. If it is, add it to the list of goals. If not, eliminate this sentence (and consider condensing section 3.2).

C1765

### 2.0.2 Introduction

- pp. 4067, line 16: see also work by N. Steiger for relevant paleo-DA research, <http://www.atmos.washington.edu/~nathanjs/>
- pp. 4068, line 10-11: As discussed above, the VSL model is NOT a candidate in itself for paleoclimate reconstruction, as the reconstruction problem is an inverse problem and VSL is a forward model. As above, revise for specificity and accuracy (eg. make it clear that the authors mean as an observation operator in DA contexts, or more generally as a link between the proxy data and climate variables in any reconstruction methodology that can support use of a forward model).
- pp. 4068: line 12-15: Here 3 main goals of the paper are listed, which conflicts with only two main goals listed in the abstract. Decide what the main goals of the paper are and outline them clearly and unambiguously.

### 2.0.3 Data and Methods

- pp. 4069, line 11-16/Figure 1: To support the discussion of the relative density of the TRW and instrumental networks here, it might be useful to include a subpanel or subpanels in figure 1 showing the locations or density of instrumental coverage at one or more points in time.
- pp. 4069, line 19: Version 2.3 (not 2.2) of the VSL model is accessible from the link given.
- pp. 4071, line 4-6: Describe the Bayesian parameter approach in detail here, unless it is the estimation procedure of Tolwinski-Ward et al (2013), in which case cite this reference at first mention.

C1766

- pp. 4071, line 14: unclear what correlation constraint is imposed and referenced here.
- pp. 4072, line 14: Start the first paragraph by describing the number of chronologies in the network that were actually used in the study, and the state in which they were actually used (eg. 2287 standardised chronologies). Explain the criteria that were used to filter from the ITRDB after (perhaps more briefly than is presently done).
- Section 2 should also include a description of how aggregated chronologies were constructed from the individual site chronologies;
- Section 2 should also include a brief but comprehensive description of the structure of the study or experimental design. List the sets of analyses run, and explain that these were run on simulations of two sets of data: the individual-site chronologies, and the aggregated chronologies.

#### 2.0.4 Results and Discussion

- Strongly suggest splitting this section into two sections (Results and Discussion) for clarity.
- subsection 3.1: It might be interesting to show the VSL skill as a function of some of these different indicators of climatic signal (EPS, Rbar)
- pp. 4074, line 20-27: Was independent noise assumed, or AR(1) noise structure at each site? Analysis and reporting of estimated noise variance or signal-to-noise ratio also seems conspicuously absent.
- pp. 4075, line 6: note that consistency across intervals is also indicative of stability in TRW response to climate over time.

C1767

- p. 4075, lines 10-18: Note that there is some dependence between results here. Tolwinski-Ward et al (2013) drew on the conclusions of Körner (2012), Korner and Paulsen (2004), and Rossi et al (2007) to develop priors for their study. If the uniform priors used in this study are identical to the defaults for uniform priors in the algorithm of Tolwinski-Ward et al (2013), then the choice of prior support has also been influenced by these studies.
- pp. 4075, lines 19- 26: if the differences are really insignificant in the statistical sense, then the authors should follow the procedure that makes the most sense before seeing the results, and the discrepancy need not be discussed. However, why not correct site-by-site rather than making an average correction, which will improve things at some sites but degrade the realism at others? Note too that the parameter estimation should be carried out after any corrections to temperature inputs for elevation.
- pp 4076, line 1-10: These species-specific results are very interesting, especially since the effect of species on parameter choice was neglected by Tolwinski-Ward et al (2013). Can the differences in parameter estimates by species be accounted for by the differing distributions of species at different elevations? Do statistical tests of differences between empirical CDFs show significant differences between the marginal posteriors across species?
- Since the authors have used a Bayesian approach to carry out this analysis, statements about the difference in threshold parameters between species (eg. line 6) can and should be made more quantitative by reporting posterior median and credible intervals for the difference.
- The statements about joint posterior relationships between parameters (discussed pp. 4076, lines 16-24) should also be supported graphically; the authors might consider including 2D scatterplots or contours of posterior probability for all pairs of parameters being estimated.

C1768

- pp. 4077, line 6-7: the parameter tuning at each site should take care of this problem (if the correction for elevation was correctly applied before the parameter estimation). . .
- pp. 4078, line 18-20: This comment about simulations under projections of future climate change is an example of a discussion point that properly belongs in a "Discussion" section.
- pp. 4079, line 10-11: Fair assumption, but note that an interpolation of the CRU field to the particular location of the TRW site might be an even more defensible assumption, especially for TRW sites that fall near the boundaries of a CRU gridcell.
- pp. 4079, line 14-16: "We further demonstrated that joint influences of temperature and precipitation on tree-ring growth is implemented within the VS-Lite model, and hence, no separation of the climatic influences on growth is needed"—unclear what this means; sounds like a reference to a conclusion that has not yet been drawn on the basis of analysis performed within the paper.
- First paragraph of section 3.4.1: This is another example where what the authors are about to do/explain should be stated in broad-brush terms at the top of the paragraph (eg. "We perform aggregation of temperature and moisture sensitive chronologies up to grid-scale in order to reduce the signal-to-noise ratio") before describing details. The details of the aggregation (pp. 4079, line 24 - pp. 4081, line 19) should go in the "Methods" section.
- pp. 4079, last paragraph: Another paragraph that would benefit from re-ordering. Before launching into details of search radius, explain overall goals: searching for what within each radius? How will the results of radial search be combined? The reader should have a sense of how the authors plan to "aggregate" before

C1769

the term is used on line 10 of pp. 4080, as well as a broad-brush idea of how the "principle climatic drivers on growth" will be determined.

- pp. 4080, line 16-17: Suggest accounting for the weighting by the relative daylength ( $g_E$ ) within VSL when "counting" the moisture vs. temperature-limited months. At most sites in the Northern Hemisphere, for example, January will be "temperature-limited", but the growth response from January will contribute very little to the overall signal variability because it gets downweighted by the relatively short daylength. If this is not accounted for, the scheme you suggest will likely bias locations to be classified as temperature-controlled.
- pp. 4081, equations 6-9: Yet another example where re-ordering is necessary. On first read, it was unclear what the weights would ultimately be used for, as the authors do not simply state up front that the aggregation would be performed by a weighted average of chronologies within the given radius. Suggest moving the last sentence of this subsection to the top of the paragraph (before the equations for the weights).
- pp. 4082: Lines 4- 6 ("the 600 km search radius improves. . .") is a result that belongs in a "Results" section, while the information in the following sentence ("We. . . introduced the condition that at least one chronology. . .") belongs in the separate "Methods" section.
- The main result of this paragraph is again stated in the last sentence, but should be moved to the top ("ATRW<sub>VSL</sub> and ATRW<sub>ITRDB</sub> show spatial coherency and capture the main climate signals. . ."). Additionally the presentation of the result in its current form is unconvincing, in part because the figure is not constructed to show the results clearly. I suggest rearranging figure 5 (perhaps 2x2 subpanels, rather than 4x1, so that the size can be increased and the reader can see what is being discussed without squinting; a change in color scale is also strongly suggested so that negative and positive correlation coefficients can be easily

C1770

distinguished by eye; perhaps statistically insignificant correlations can be shown as open circles with no color; also all color axis labeling and titles should be in larger font.) In a revised figure I would hope to be able to easily see the stated result that the increased search radius improves the comparison between simulation and observation, as well as the other spatial features discussed pp. 4082 lines 9-20. Consider adding additional figures as well to support the claims in this paragraph.

## 2.0.5 Conclusions

- pp. 4083, line 22: "performs well" should be replaced with a more specific finding, perhaps "produces parameter estimates that are stable with respect to the choice of calibration interval."
- Again, the authors should consider whether parameter estimates of  $T_1$  from this study should be a focal point. If not, and the result is actually more of a discussion point, then it should be excluded from the abstract and conclusions.
- pp. 4084, line 5-6: not clear that any figures or results presented support the claim that VSL shows "notable skill at locations with a less extreme climate due to the model's explicit consideration of joint temperature and moisture controls on modelled tree growth." If the authors want to keep this statement in the conclusions, it should be supported with evidence and discussed in the body of the paper.

## 2.0.6 Figures

- Figure 3: It is generally preferable to show a kernel smoothing of the numerical output to visualize the posterior probability densities than to joint the tops

C1771

of a histogram (as through the function `ksdensity` in MATLAB, for example). As stated above, since the joint posterior relationships are also discussed though (pp. 4076, lines 16-24), the authors might consider including 2D scatterplots or contours of posterior probability for all pairs of parameters being estimated.

- Figures 1 and 5: Suggest change of color scale that makes negative/positive correlations more easily identified by eye.
- Figures 3, 5 and 6 need larger font sizes.

## 3 Typographical and Stylistic Comments

- pp. 4073, line 2: Suggest "These tests and Pearson correlation analysis. . . LEAD US TO favour a hierarchical approach. . ." so as not to anthropomorphize the tests and analysis.
- pp. 4074, line 8-9: "Typical tree-ring chronologies have a mean segment length of. . ." How are the set of "typical" trees determined over which to take the mean? Or do the authors just mean the mean provides a measure of the typical segment length over all trees?
- pp. 4077, line 21 and elsewhere: suggest changing "verification" to "validation." The former word has its roots in the Latin word for "truth", and of course one can't expect a model to exactly represent the truth, but only to provide a *valid* approximation to the truth.
- pp. 4078, line 12: re-word "the high-elevation site is negatively correlated with. . .", as it's not the *site* that is correlated with growth, but the simulated time series at the site in question.

C1772

- pp. 4083, line 12: "promote to" should be "promotes the"?
- pp. 4084, line 1-3: "Moreover, our results demonstrate the VSL model skill to simulate in response to climate variation. . ." Revise for proper English.
- pp. 4084, line 14-15: ". . .by using in data assimilation approaches. . ." Revise for proper English.

## References

- Körner C (2012) Alpine treelines: functional ecology of the global high elevation tree limits. Switzerland: Springer Basel
- Körner C, Paulsen J (2004) A world-wide study of high altitude treeline temperatures. *J Biogeogr* 31:713–732, DOI 10.1111/j.1365-2699.2003.01043.x
- Rossi S, Deslauriers A, Anfodillo T, Carraro V (2007) Evidence of threshold temperatures for xylogenesis in conifers at high altitudes. *Oecologia* 152:1432–1939, DOI 10.1007/s00442-006-0625-7
- Tolwinski-Ward S, Anchukaitis K, Evans M (2013) Bayesian parameter estimation and interpretation for the vs-lite model of tree-ring width. *Clim Past Discuss* 9:615–645

---

Interactive comment on *Clim. Past Discuss.*, 9, 4065, 2013.