

Interactive comment on “Tree-ring responses to extreme climate events as benchmarks for terrestrial dynamic vegetation models” by A. Rammig et al.

Anonymous Referee #1

Received and published: 8 April 2014

This study links TRW and LPJ NPP in Europe using coincidences as an analytical framework. I must say I am confused on the prominence of the "benchmarks" idea. I don't really see any benchmarking here (sensu Luo et al., 2012, which you cite). It feels oversold, especially in the title. I would think it might be a useful add to the conclusion (that TRW can serve as a reference, with the caveats you cite). Also, you state (2548) "Hence, an in-depth investigation as to how these differences can be attributed and to what extent TRW can be used to benchmark dynamic vegetation models for responses to extreme events is necessary". I thought that was a goal of the study given the two questions you pose (2542) and the use of "benchmark" in framing the Introduction? I am also quite curious about coincidences between TRW, P and

C890

T (all three at once). Did you investigate that at all, apart from Fig. 3? Also, and acknowledging the effort to run an LSM, did you ever think to corroborate LPJ with another LSM. LSMs have, generally speaking, low skill, but also exhibit large gradients in skill. It would be enlightening to know whether the coincidences you've found here are model-specific. Apart from this I have a few minor/specific comments below. Many relate to language. I would strongly encourage an overall tightening of the language.

2539/4: Try "reductions of net primary productivity"

2540/4: No comma before which

2540/7: Try "response, e.g., to drought events (Schwalm et al., 2012), may" Comma issues throughout (I've highlighted two here). Minor, yes, but I would encourage a native or native-like speaker to proofread the text.

2541/9: Try "alone as a proxy"

2544/2: Are the TRW available online? I would encourage the authors to add this as a supplement to the study.

Section 2.1.4: This reads very ad hoc. Why/how were the thresholds chosen? MODIS goes back to 2000 but you speak of multi-decadal time scales? The 1901-2001 reference for example. What are the "the connected phases of GSob"? I feel this section needs to be better substantiated.

2549/8: What are "climatically attributable extreme years"? Also I am unclear what was done here? Did you take extreme years, z-score them, and then do a histogram? So a histogram of values in the most extreme deciles that were then z-scored? I think so but I am unclear why you did this? This might be semantics on my part but I think I'm struggling with z-scoring extremes and calling these values extremes? Some values will always be below your sigma thresholds when you z-score?

2549/20: What is a carbon storage product NPP? I think you mean that modeled NPP behaves differently than TRW based on some aspect of how LPJ is put together, how

C891

LPJ simulates NPP?

2549/24: Try "reducing" That or comma placement?

2552/1: Does it not also highlight the difficulties in matching a point to a pixel? I think that's what you are saying but my concern here is from the "other " side. I don't want downscaled data per se. Rather, I want to understand how scale-mismatch degrades coincidence (or any validation study where a point is matched to a modeled pixel).

Figure 6: I find this hard to decipher. The transparency is quite subtle, and the double legend seems off. Also, you mention zonal patterns in the text. But that is more wrt latitude as opposed to climatic space. Also, what is n in each bin?

Figure 7: I find the discussion of this in the text very hard to follow. For example, 2553/1 "Hence,..." implies some 3D surface that is never displayed. The last sentence in the caption is also confusing. I had thought that you had the same time scale for both TRW and NPP? That you pulled the matching pixel and time span from your transient LPJ run from 1901 to present?

2553/11: I think you need to be careful here. TRW is, trivially, based on NPP. I think you mean that the way NPP is modeled is less cumulative than TRW. That is, the model is instantaneous and summed to get a GS value whereas TRW has lags naturally embedded?

2553/25: The final sentence of this paragraph really does not make sense to me.

2554/7: "triggers" Not sure I buy this. Triggers implies a certain degree of causality that I do not find here.

Last para in Conclusions: Your a, b and c. Did we not already know that? I don't think this study is the first to show this. That's not bad, I'm all for contextualization and linking back to established findings. But you might add what specifically your study tells us that we did not know beforehand.

C892

Interactive comment on Biogeosciences Discuss., 11, 2537, 2014.

C893