

Interactive comment on “Calibrating a process-based forest model with a rich observational dataset at 22 European forest sites” by David Cameron et al.

Anonymous Referee #3

Received and published: 11 June 2018

General comments:

The authors calibrate the BASFOR forest model using various data and analyze what data best constrains the posterior predictive model uncertainty. The type of work presented is important even though new methodology is not presented. The novelty of the paper comes from using more data and in a more varied setting than was done before.

I will comment on the calibration part in these comments, since I'm not an expert of forest modeling.

Whereas the work does have merits, it is seriously lacking in detail, and many design choices seem rather arbitrary. There are no formulas describing what was done - even

C1

though Bayesian model calibration is a mathematical exercise.

There are several other issues as well, the most important ones being lack of cross validation, model and model parameter descriptions are virtually missing, and although Bayes' rule is described, a probabilistic explicit observation model is not given. Details of the MCMC experiment are not properly described either.

Posterior parameter values are not described, not even in the supplements. That belongs to the main text. I would want to see full descriptions of the parameters, priors, and posteriors in the main text, and a discussion referring to model equations about what the results of the calibration mean.

The manuscript is too long - there is a lot of repetition, and many of the conclusions are quite obvious. Such repetition should be reduced to a minimum. In addition to this, there are still very obvious editing errors, which need to be corrected. Not all of them are listed below, since there are too many.

For these reasons, a major revision is still needed.

Some more detailed comments below:

Specific comments:

1. One general issue with model calibration is, that once calibration is done, it can be used to inform where the model can be improved - where it does not perform well and where it does not. I don't understand why this aspect is not discussed. Please add this aspect throughout, especially in the discussion.
2. Please make the abstract more concise and articulated: what's new in the work and why it is important. At the moment it is a little long and unclear. And please check the language/style.
3. One enormous source of uncertainty is biases in input data. Could you please comment on this aspect.

C2

4. p 4 l 30: The 1m rooting depth seems pretty arbitrary, and I'd guess it does affect the results. Why was this value chosen, and if it's not known, why isn't it included in the calibration? How much are the results affected?
5. p 5 section 2.2. You need to give the model equations, pointing out the parameters you are calibrating in those equations. This is crucial - at the moment it is impossible to say what happened in the calibration, when the parameters are not described at all. The manuscript should be self contained in that by reading it one does understand what happened.
6. p 5 l 16: "replicated backwards...?" what does this mean? Please describe in a detailed and compact way what you did, preferably with equations.
7. p 5 l 22: planting density assumption: yet again an important source of potential bias / uncertainty that is not discussed at all. Please include in a discussion of the input data uncertainty.
8. p6 l24: What parameters were used for the Beta distribution? Why? How did you use literature data to obtain the priors? Tables in supplements should be moved / summarized in the main text, with references to how the priors were chosen.
9. p6 l28: Why the covariance structures are not shown? The pairwise posterior marginals contain the most interesting information: What directions in the parameter space are constrained and what are not; are the correlations linear; are the distributions unimodal and close to Gaussian etc. An important part of the analysis should be model parameter (and hence process) identification and finding information about what data constrains what processes. The covariance structure is an important key to this analysis. Please include a figure of the 2d-marginals with probability contours (using e.g. Gaussian KDE), and discuss.
10. p7 l16: In a manuscript describing Bayesian calibration of a model, the prior values most definitely belong to the main text, not to the supplement. Please include it here.

C3

I'd like to read the prior values for different forest types from a table. Justification of the prior distributions used should be included.

11. p7 l26 "Prior uncertainty was set at 20% of mode value..." But isn't your "prior uncertainty" set by the parameters of the Beta distribution? Also, in any case the 20, 30 and 40% values are arbitrary, and this should be stated explicitly as a potential source of error. Please clarify.
12. p7 l30 ..."at a given parametrisation" should rather be "at a given point in the parameter space" or something along those lines
13. p7 l31 "uncertainty about random data error", do you mean "data uncertainty". The data uncertainty should be discussed more widely, see later.
14. p. 8 l. 3. What is the observation equation (like $y_t = M(x_{0,t};\theta) + \epsilon$, where $\epsilon \sim$ some distribution)? How does the averaging come into play in that equation? And again: "measure of random error about ith data point". Do you mean measurement error? Or something else? Please discuss the merits of the chosen likelihood function, what does it mean? (also see next item)
15. (same) It looks to me like you are treating the residuals as independent. Is this a reasonable assumption? An alternative would be to fit a time series model (like an AR/ARMA/ARIMA etc model or such) to produce wider posteriors. You should analyze the residuals and verify that whatever probabilistic observation model description you are using, your residuals in the end conform with your error model (please show histograms of the residuals and autocorrelation functions in the supplements). If you are not able to do this, you must acknowledge that all the ranges in the figures and the scaling ("steepness") in the posterior probability distributions is arbitrary. At this point, what you have left, is then the covariances of both parameters and predicted errors / quantities, and discussing those would still be valuable. In the best case scenario, you should add both.

C4

16. p8 l7: So you are predicting monthly values? Clarify and see previous point.
17. section 2.4.5. Details are missing. What was the proposal distribution for MCMC? Where is the picture of the chain? (please include). How long was the burn-in? What was the acceptance ratio? Effective sample size? And was there a reason to go with just Metropolis? Usually e.g. Adaptive Metropolis works a lot better (unless if you happen to a priori know what the optimal proposal is)
18. p8 l. 24. The lonely sentence could be left out. Throughout the text there are lots of these types of sentences, and many of them could be just removed.
19. p 10 l1. I'd really like to see the corresponding contours on the pairwise 2d marginals instead / in addition.
20. Generally about results: Of course a fitted model fits data better. To know anything of how successful / good the calibration is, a k-fold cross validation (with a suitable k) should be done instead and all the results should be reported for those. That would give information about how good the calibration is for `_predictive_` purposes. At the moment the results read a little like "the model was fitted to data and after fitting, the model fits the data better". With cross validation the results would be significantly more valuable and interesting.
21. p. 10 l. 18: "Model output uncertainty is reduced from prior to posterior". This is not interesting without the cross validation - the result could be very site-specific and for predictive purposes the result could be in many places different. This being said, there is some value in the analysis, that calibrating the model using data x reduces error in variable y . But without including the model equation this is hard to see. Please, discuss these types of results with references to the model equations. Also, I would like to see what the expected value of the change of this error is.
22. p. 11 l. 2 Are you sure this is the right metric? What if you get some really bad outliers but mostly good behavior? Is this then an acceptable model? If such behavior

C5

- never occurs, the approach can be ok.
23. section 4.1 would be meaningful with the cross validation, but not so much otherwise
24. p. 12 l. 1-2: I have difficulty understanding the sentence.
25. Section 5.1. Would love to see what these results mean with some short and concise discussion. A time series of the predicted values before and after calibration would be valuable.
26. General: the observation set is advertised to be "rich" a few times too many
27. p 15 l. 20. In practice there are always inconsistencies with real data. I'd say such results tell about weighting of the different variables in the likelihood function (choice of error model).
28. p. 15 l. 21, 25: We will return... & We will discuss... could compactify the text a little here and leave these out.
29. p. 15 l. 28 "mathematical probabilities", rather say "a probabilistic model" and describe the model properly as I mentioned earlier.
30. p. 16. Lots of text for the content. Simplify and remove repetition.
31. sect. 7.1.1 generally: without the equations it is difficult to say much about whether the results are just obvious or if there is something interesting here. The discussion is not very helpful either. Please refer to model equations to explain model behaviour.
32. p. 18 l. 32. You recommend using a hierarchical model for describing the parameters, but you are not doing a hierarchical modeling yourself. Why? Also, if you were to hierarchically model the parameters, you'd need to have a way of predicting the parameters for predictive purposes. How would you do that?
33. p. 19-20. I would like to see a more compact Conclusions section, and also I'm not

C6

sure that the bullet point style is a good idea.

34. p. 19 l. 22. What's the role of the prior?

35. p. 19 l. 27. It is unclear to me how model structural errors were represented in your work. How did you quantify that "uncertainty is not reduced inappropriately"?

Fig. 2: The wavy shapes look strange - where did that come from?

Figures: Merge 5&6, 7&8, 9&10, 11&12, 14&15; check legend style/formatting and maybe optimize the general presentation, with improved captions. Explain in words what is seen in pictures.

Technical corrections:

grammar: e.g. p1/l12 inclusion...were => was

p4 l 29 www address to references? (check journal style, but it does not look good.)

p6 l6 on: sentences don't have verbs.

p6 l19: "probability calculus" => Bayes' rule

p8 l 26 "of" missing most likely

p13 l30 2nd sentence, sounds a little strange.

Interactive comment on Biogeosciences Discuss., <https://doi.org/10.5194/bg-2018-156>, 2018.