

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 #1

Received and published: 16 May 2018

The model presents the calibration of a forest model, BASFOR, using a Bayesian MCMC algorithm, to data from a variety of sources at a range of European forest sites. The authors seek to identify which datasets improve model fit as well as finding the best way to use data spread across multiple sites. This kind of study is generally very useful in pinpointing the critical data needs for model calibration as well as finding model knowledge gaps.

However, I find that the paper is very poorly written. There is very little detail on the model structure and the way that the data was collected and processed. The results section lacks detail and actual values for model fit providing only vague statements

C1

about the results. The discussion and conclusion are largely a repetition of the results without going into any depth. In any case, any discussion would be hard to follow without knowing more about the model and the parameters which were fitted, a fact which I am afraid is reflected in my detailed comments, which mostly focus on the first part of the paper.

I find the introduction and discussion have a very narrow focus, referring only to process-based forest models, and ignoring the very very large range of studies that calibrate terrestrial biosphere models or land surface models and ask similar or identical questions to the current study.

A few topics that I would like to see discussed more in depth are: Which parameters are constrained by the different datasets used? Are the parameters correlated? Is there equifinality? Are the inconsistencies between the data and models due to the model structure or the data used? Or the fitting algorithm? Several times in the discussion it is mentioned that there are inconsistencies between the different datasets? Why is this? Can this be improved? What is the path forward for future model calibration? What are the data needs?

Detailed comments

P2 l15 I find the discussion of the previous use of Bayesian methods for data assimilation very narrow. Whilst these references might cover the previous studies that constrain forest models there is a much wider pool of studies using terrestrial biosphere model and land surface models which use these techniques.

P4 l15 Was the soil data measured for this study? If so, more details of the methods are needed. If not, please provide a reference.

P4 l16 What is the annual integrated metric you have used? Is it a mean or a sum or something else?

P4 l18 “Data from CEIP and other project (e.g. FLUXNET) databases as well as pub-

C2

lications “ More details are needed here: which other databases? What data was taken from where? This phrasing is repeated in subsequent sections too, my comment applies to all

P5 I1 This model description is inappropriately short. At least some basic equations or a model schematic are needed for the reader to understand how the model and data go together.

P5 I16 What do you mean by replicated? Did you use a mean seasonal cycle from the available data? The common practice is to use climate reanalysis data for time periods when local met data is not available. You at least need to explain why this was not done.

P5 I18 Due to the poor description of the model it is difficult for me to understand why this data is needed.

Sections 2.3.3-2.3.5 These all refer to driver data and should be put into that section, they do not need their own subsection. I do not understand the need for the detailed description of N deposition, much more detailed than for any other data. Does your model have a particular focus on N deposition?

P7 I1 “we chose to calibrate nearly all of the parameters” Which parameters exactly did you calibrate?

P7 I10 The forest types chosen, pine spruce and deciduous, need more explanation. Generally, models would choose to use standard plant functional types. I understand that with the focus on local modelling you might want to use species as categories but then why use ‘deciduous’?

P8 I7 How did you aggregate this data?

Section 3 Phrasings like ‘closer’ and ‘further away’ need to be more quantitative, include for example use the RMSE

C3

Section 8 Conclusions are generally text and not a bullet point list

Figures 1 and 2 Please add y-axis labels with variable names and units. Also, it would be great to see both fluxes and stocks at both sites. Figures 5-8 Figures should be understandable from the caption alone without reference to other section. Also, another missing Y label

Tables 1 and 2 It is common practice to report site name, geographical coordinates and some other site information such as forest age and climate (e.g. mean annual temperature, mean annual precipitation), maybe soil type. References need to take only one column since I assume everything is included in the bibliography and we really do not need all this information in a table

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

C4