Report on the article 'Makela, Susiluoto,Markkanen,Aurela,Mammarella,Hagemann, Aalto: Constraining ecosystem model with Adaptive Metropolis algorithm using boreal forest site eddy covariance measurements'

The work presents an approach for studying parameters of an ecosystem model. Especially, the focus is the site level parameter optimization of the JSBACH model, the land surface component of the Max Planck Institute MPI-ESM model. JSBACH simulates the water and carbon storages and fluxes of the ESM model. The motivation of the present work is to correct observed biases of the model, especially in evapotranspiration (ET) over continental areas and gross primary growth (GPP) under water limitation. The authors study how well the selected cost functions, featuring ET,GPP and LAI (maximum leaf area index) constrain the respective model parameters. The key idea is to use a Monte Carlo sampling method, adaptive MCMC, instead of a direct optimization of the parameters. This approach allows for a comprehensive study of the parameter identifiability. The authors are able to improve the model fit in several aspects, but can not remove the bias of an extremely dry season.

The work is certainly professional and worthwhile to publish. However, the novelty of the approach should be more clearly given, e.g., the benefits of the chosen MCMC approach should be clearly described. Also, a typical reader of the journal is most likely not too familiar with the details of the JSBACH model, and would need the authors giving more background and insight of the model and parametrizations of it. Some more specific comments below.

Major Comments

1) To estimate the distribution of parameters B of a model F based on data Y given by experiments X, connected by the standard expression 'Y= F(X,B) + eps', the distribution of the measurement error 'eps' should be known. But here the authors give almost no information of any of these to a reader not already familiar with JSBACH and the measurements. Certainly it is not possible to give all details, but the basic parts of the underlying modeling and numerical solution should be described, maybe in an Appendix, not to leave F(X,B) as a fully black box for the reader. See comments 4), 6) and 7) below.
2) To optimize the model parameters the Adaptive Metropolis (AM) method is chosen. It is, however, a sampling method rather than optimization. The motivation and benefits of the choice should be given: instead of a point estimate, samples of the full distribution of possible parameter values are obtained, together with (nonlinear) correlation information, sensitivity, identifiability of parameters, etc.
3) The parameter estimation is based on the two cost functions on p. 4 and 5. But no info is given here on the assumed statistics of the expressions, only a hint on Gaussian distribution later on p. 7. Usually, the sum of squares of the residuals is divided by the respective estimated variance of measurement error. Here, the residuals are normalized by the observations. This can be quite

acceptable if no 'true' error statistics is available, and the sampling is done in the spirit of studying the identifiability and correlations of the parameters. However, this should be done explicit in the text.

4) For the general audience (not familiar with JSBACH) at least the basics of the numerical approach used in JSBACH should be given, together with the CPU demands of the runs. Now only an implicit statement (' …interval is looped over to generate a 30 year spin up …', line 30 , p.3) is given that would indicate that JSBACH is a dynamic model that has to be initialized or run into a (quasi)steady-state to compare with observations ? Or is this due to the uncoupled version used here? The concept and use of spin-up should be clarified.
5) How much does the uncoupling impact the results in general? The authors mention (P.10, line 13-15) that the lack of coupling of the LSM model to atmosphere generates an erroneous energy balance. This aspect should be discussed or commented more explicitly.
6) The discussion in Section 2.5, parameter posterior distribution vs PCA, is not clear. The authors 'perform a PCA analysis transforms of the covariance matrices …' – but do not tell what covariance ? My guess would be that they actually mean the matrix of the AM samples of parameter vectors, and compute the PCA of it to get the eigenvectors of the least identified parameter directions. This can lead to correct conclusions, assuming that the nonlinear correlations between the parameters are not too strong. That, on the other hand, is typically indicated by plotting the 2D scatter plots of parameter marginal distributions. So I would recommend the authors to show them as well, and clarify the discussion on how PCA was used.
7) No information is given on how the studied parameters appear in the model. It is well-known that the parametrizations strongly impact the identifiability. A good example is the logistic function, where centering and scaling typically removes correlations. So this point should be made explicit by showing the formulas, at least in case of the LoGro phenology model where high correlations appear 'since the parameters are intimately connected' (L.30, p.7).
8) The measurements consist of the CO2 fluxes as given by the eddy covariance method. But the cost functions are given in terms of 'observed' and modeled GPP,ET and LAI. The connection between CO2 fluxes and those cost function expressions should be given.

Minor comments

1) In addition to the PCA/MCMC analysis of least identified parameters, the authors study which parameters are the most relevant for the change of the cost function. They introduce an OAT (one-at-a-time) method of their own (?). The relation of it to well-known methods such as the MOAT (Morris-OAT, see the reference below) could be make more clear. Also, it is not clear what the 'tuned parameter' (p.5, Step 1) is: the mean of the sampled values, or the maximum likelihood (minimum cost function) value? I would gather that the 'reference value' is the initial/default value of optimization. These points should be made clear.

2) Only the cost functions are given in the text, not the likelihood used in the sampling. If it is Gaussian as indicated on p.7, it should be mentioned that the 'f' function of step 2., p. 4, actually is the exponential function of the (negative) cost function.
3) P.4 line 15: The sentence 'A sample in the parameter has a value …' could be removed. Instead, the term 'chain' could be explained for a reader unfamiliar with MCMC.
4) P.4, line 16: edit the sentence 'The algorithm is used …' something like 'The algorithm can be used' or 'is used here', since the basic form of the AM algorithm is or a single chain. Maybe add a reference to parallel chain adaptive MCMC
5) P. 6, line 20/Step 3: 'Initial covariance' means the initial proposal covariance for MCMC sampling ?
6) P. 6, lines 17 and 27: it would be good to know here how many parameters were used for the 10000 sample long chains.
7) P. 7, Section 3.3: motivate why only maximum LAI is retuned for Sodankylä.
8) P.8, line 20: edit 'Given into account' to 'Taking into account'
9) P.9, lines 18-25: clarify the discussion. As the tuning aims at the 'best parameter', how could they be different?
10) The contents of Table 3 should be clarified, preferably in the bulk text where PCA is discussed. While the meaning of 'weight' is OK, the way the two most dominant parameters are calculated should be told to the reader.
11) Overall, the English language could be double-checked ('the' added in several places, etc)

A possible example reference for OAT methods:

Morris M (1991) Factorial sampling plans for preliminary computational experiments. Technometrics 33(2):161–174