Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-216-AC1, 2017

Geoscientific Model Development Discussions

Interactive comment on "Bayesian integration of flux tower data into process-based simulator for quantifying uncertainty in simulated output" *by* Rahul Raj et al.

Rahul Raj et al.

r.raj@utwente.nl

Received and published: 3 March 2017

We thank for the constructive and helpful comments for our manuscript. We will carefully consider all comments and these will be incorporated in our revised manuscript accordingly. We have inserted our response to each comment. We use "R1C" for referee #1's comment and "A1C" for author's response to referee #1.

Major concerns of Referee #1

R1C 1: On conclusion that temporal correlation matters A control case without the correlation is missing. How do the results and implications change between accounting versus not accounting for correlations?

A1C 1: We ran the whole procedure without the temporal correlation in the residuals for Experiment 1. We found that both cases (accounting versus not accounting for correlations) led to similar temporal profile of the posterior simulated gross primary production (GPP) and similar values of statistical criteria (Root mean square error and Nash-Sutcliffe efficiency). The fact that the temporal correlation in the residuals is not only responsible for the temporal development of GPP indicated that the representation of dynamic processes within the BIOME-BGC simulator could be improved. We will add and discuss the new results of the control case without the correlation in the revised manuscript.

R1C 2: On conclusion about time varying parameters I do not agree with the applied approach. In the presented study several independently simulated time series are mixed together. Each series includes the impact of changed parameters on the previous state. The parameter set valid for July was applied already to April, May, and June and affected the starting states of July. In my opinion one cannot conclude on time-varying parameters with this approach. The simulator needs to be run for the previous months also with the previous parameter set. The model state of the end of the month must be the starting state for the run of the next month with changed parameters. In an ideal case the entire time series would be run as one forward model and the combined (larger) parameter set would be estimated. A more feasible approach is to calibrate each month separately. For the next month calibration continues from a state of the previous month. This starting state needs to be drawn from the distribution of state vectors from the previous month posterior of states for each run with a new parameter sample. For the currently used method, at minimum, the forward runs that produce the predictive posterior and the fit statistics need to change the parameters across months in each single forward run to discuss seasonally changing parameters.

A1C 2: We are aware of the problems the reviewer identifies. We found that the response of simulated GPP to weather conditions is rather similar among months: The simulated GPP was mainly driven by the meteorology, and much less by seasonal phenology. We then hypothesized that some important state variables (such as LAI and carboxylation capacity) may not have a pronounced seasonal cycle in the model. Because these state variables are updated internally in the model, it is not possible to perform a calibration per month: This would require changing the model code, and more importantly, it is in conflict with the main idea of the process simulator. In the second experiment we therefore calibrated the model to the data of each month separately, as if we had no information on GPP for the other months. If some of the parameters have different optimum values when calibrated to different months of data, then this is an indication that the relation between these parameters and important state variables that (should) change during the season, may require improvement.

By doing this experiment, we were able to identify the process that may require an improved description. We will mention above points in the revised manuscript. We decided to avoid the term 'time varying parameters' in the revision.

R1C 3: On using GPP to calibrate the mechanistic model Net-ecosystem (NEE)-Flux-partitioned GPP is already the output of another statistical model – here the nonrectangular light response curve. This model already makes some strong process assumption e.g. on relationship of respiration with temperature. In effect the mechanistic model is calibrated against the output of another model. This makes it difïňĄcult to interpret the estimated parameters, their distribution and their meaning and process understanding. This needs to be discussed. Biome-BGC also computes respiration and NEE. You can compare these predictions to observations to gain additional insight into the model and the calibration. The flux partitioning also provides seasonally changing respiration at reference temperature and temperature sensitivity. Comparing these quantities to BIOME-BGC predictions lends further insight, which however, may also reveal sub-optimal calibration. A more direct way would be to include the respiration parts of the Biome-BGC model in the simulation and fit the simulated, i.e. predicted NEE to the NEE observations. Probably, this will introduce correlations in the joint posterior parameter estimates. But the weaker correlations in the presented

GPP fit, are only resolved by the assumptions of the NEE-partitioning model that was used to derive GPP. While the presented GPP calibration has its own ground, those aspect needs to be addressed. The study would greatly benefit from a comparison to a calibration that uses NEE instead of GPP.

A1C 3: We would like to mention following points on using partitioned GPP instead of NEE data in this study:

A. Indeed in our approach, the output of the process-based simulator was validated against the output of another model, notably the flux partitioning model. Although this approach has been used in other studies to validate the output of process-based simulator as well (Collalti et al., 2016, Liu et al., 2014, Yuan et al., 2014, Zhou et al., 2016), it could lead to error propagation. We clarify that the flux partitioning model (NRH model, Raj et al., (2016)) was tuned to the Eddy Covariance data in blocks of 10 days. Because the NHR and the relationship of respiration with temperature and moisture were tuned for these short blocks separately, we expect that the GPP still reflects realistic responses to environmental drivers, and does not depend much on model assumptions.

B. Calibration of BIOME-BGC using NEE data is more challenging as NEE is the difference between fluxes caused by two processes (assimilation and respiration) We argue that calibration of such a complex model to NEE instead of GPP may not be a good idea, but calibration to NEE or respiration in addition to GPP is possible. However, we limited this study to the primary productivity, because this was our primary interest. A future study should be done to include both GPP data and ecosystem respiration data (can also be achieved by partitioning of NEE data) in a Bayesian calibration of BIOME-BGC. This may ensure the accuracy of all related carbon budget terms (GPP, NEE, and respiration terms).

C. As far as the comparison of simulated NEE and respiration with the measured NEE and portioned respiration is concerned, this is out of the scope of the present study.

We mainly focused on the simulated GPP and already presented a lot of results.

We will discuss all above points in our revised manuscript.

R1C 4: On hitting the prior bound of residual uncertainty Fig 1f clearly shows that the calibration tries to increase the residual variance and that high residual variances yields lower cost. In the current inversion, the residual variance is only bounded by the prior. This hints to deïňĄciencies in the inversion. I sometimes experienced the same effect because a single equation of the cost (eq. 5) may in some cases not prefer the best ïňĄtting variance but the larger variance together with suboptimal parameters. Prescribing an upper bound is to my opinion not a good solution for this problem. Even fixing the residual variance would be a better option. My recommendation is to use several parameter blocks in a Metropolis within Gibbs sampling (Chib S Greenberg E (1995) Understanding the Metropolis-Hastings algorithm): One block to fit the model parameter conditional on the parameters of the residual statistical distribution and another block to fit the residual distribution parameters conditional on the current sample of model parameters.

A1C 4: We would like to clarify that SD in Fig 1 of the manuscript is not the residual variance. SD represents the effective soil depth and this is considered in this study as one of the parameters of BIOME-BGC simulator. In the figure cation, we have written that “Information about the BIOME-BGC parameters is given in Table 1”. We will specifically mention in the figure caption that “SD is effective soil depth” in the revised manuscript to avoid any confusion.

Further Concerns of Referee #1

R1C 5: The cut of the posterior by the edge of the prior distribution of LFRT and FRC:LC (Fig 1) shows inconsistency in the combination of the model, the prior knowledge, and the observations. This hints to deficiencies of the calibration. It also makes it difficult to interpret the parameter estimates and process understanding. This needs more discussion. The introduction of bias parameters in model drivers or model pre-

dictions could help to resolve the inconsistencies and, moreover, the bias parameters then can be interpreted.

A1C 5: Fig 1 shows the cut of the posterior by the edge of the prior distribution, which is called edge effect, of LFRT and FRC:LC. We agree that this clearly indicates a significant effect on our posterior parameter space of LFRT and FRC:LC by our particular choices of parameter space to be included in the prior distributions of LFRT and FRC:LC. It could be argued that the prior uniform distributions of LFRT and FRC:LC could be made wider in order to eliminate the edge effect. However, we did not do this for the following reasons:

A. We carried out extensive literature review in a previous study (Raj et al., 2014) to compile the information on FRC:LC, LFRT, and other BIOME-BGC parameters. This information led to the characterization of uncertainty in the parameters and helped to define the prior distributions. For the present study, we used this information on prior distributions. We had no further scope to make the prior distributions wider at the study site.

B. In the present study, we have used the upper limit of FRC:LC at 2.15. In our previous study (Raj et al., 2014) we found that higher values of FRC:LC led to disappearance of the forest (LAI=0) due to negative cumulative NEE, and hence no production at the study site. Therefore, we had no other choice of the upper limit of FRC:LC other than 2.16.

Further, we don’t fully agree that the edge effect indicates deficiencies of the calibration. This can also be thought in another way that even if there is the edge effect, a drop in RMSE and enhancement in NSE coefficient (Table 3 in the manuscript) before and after calibration indicated the efficiency of the calibration. We will mention above points in the revised manuscript.

R1C 6: How were the initial states of the model prescribed?

A1C 6: Initial states of the model were prescribed with very low value (≈ 0). Spin up simulation of BIOME-BGC was performed first to achieve steady state condition of soil carbon and nitrogen pools under given climate and site condition. Normal simulation was then started with these steady state condition using daily meteorological data of 2009. We will add these points in the revised manuscript.

R1C 7: Do you have correlations in the posterior parameter distribution, and how to you interpret them?

A1C 7: In the revised manuscript, we will add a brief explanation and a plot showing the correlations in the posterior parameter distribution obtained in Experiment 1. We found a strong positive correlation between the posterior distributions of C:Nleaf (carbon and nitrogen ratio) and FLNR (Fraction of leaf N in Rubisco) with r=0.95 (r is correlation coefficient). This strong positive correlation is in-line with the formulation of FLNR that shows direct proportionality with C:Nleaf (see Appendix A in RajEtal2014, for details). The parameters C:Nleaf and FLNR showed similar negative, but weak (> -0.5), correlation with Wint (Canopy water interception coefficient) (r $\approx$ -0.4). This can be explained by the fact that the simulated GPP is expected to vary inversely with Wint via soil water potential and stomatal regulation and directly with FLNR and C:Nleaf (see Section 5.1 in the manuscript, for details of BIOME-BGC internal routines). The parameter SD (effective soil depth) had similar positive, but weak (< 0.5), correlation with FLNR and C:Nleaf (r $\approx$ 0.4). This can be explained by the fact that the simulated GPP is expected to vary directly with SD (via soil water potential and stomatal regulation), and FLNR and C:Nleaf .Two parameters of any other pair combinations did not show any notable correlation.

R1C 8: Please discuss your finding in the context of other studies that already performed a Bayesian calibration of BGC-models against Flux data. E.g. there is big body of studies using the DALEC model also looking at multiple constraints, model error, and different sources uncertainties.

A1C 8: We couldn't find papers on a Bayesian calibration of BIOME-BGC against flux data that compare with our results directly. We found other papers on calibration of BIOME-BGC, and one study (Hidy et al., 2012) on Bayesian calibration of BIOME-BGC. Because Hidy et al. (2012) focussed on an herbaceous ecosystem, we could not compare directly the outcome quantitatively.

Technical comments of Referee #1

R1C 9: Fig 1: Shows a very slow mixing. One chain needs more than 1000 steps to become uncorrelated with its previous state. Before computing the Gelman-Rubin criterion you should thin the chains by a factor so that autocorrelation or spectral density of the chain gets small.

A1C 9: We agree with the referee that thinning would reduce the spectral density of the chains. We decided not to do it again for the revised manuscript because we expect that thinning will not change the posterior estimations achieved in the study.

R1C 10: Fig 1: shortly explain phi and SD in the figure caption, e.g. "parameters describing variance and correlations of the distribution of model-data residuals (eq. 5)"

A1C 10: Please refer to our comment A1C 4 on the clarification of SD. We will explain phi in the figure caption in the revised manuscript.

R1C 11: Fig 1: Maybe mention, that only the end of the chains after the burnin are shown.

A1C 11: Thank you very much for this suggestion. We will mention this in the figure caption in the revised manuscript.

R1C 12: Fig 3,. . .: Legends are missing. Please, use a different line type so that model and observations can be distinguished without color. Readers would benefit if you indicate months at the time axis instead of or in addition to Julian day.

A1C 12: We will modify the figure in the revised manuscript according to the referee's suggestion.

R1C 13: Fig 6 and associated discussion: For a model with state variables or pools this result is trivial. I suggest omitting or explicitly elaborating on the magnitude of the impacts of state versus drivers on the model output and with witch conditions the one or the other becomes important.

A1C 13: We will omit figure 6 in our revised manuscript.

R1C 14: P9L25ff: More discussion needed on hitting the upper prior boundaries and its consequences.

A1C 14: Please refer to our comment A1C 5.

R1C 15: P10L18: typo percentile

A1C 15: We will correcte this in the revised manuscript.

R1C 16: I cannot agree to the discussion because of the method that actually did not alter parameters across seasons during a single simulation run.

A1C 16: Please refer to our comment A1C 2.

References:

Collalti, A., Marconi, S., Ibrom, A., Trotta, C., Anav, A., D'Andrea, E., Matteucci, G., Montagnani, L., Gielen, B., Mammarella, I., Grünwald, T., Knohl, A., Berninger, F., Zhao, Y., Valentini, R., and Santini, M.: Validation of 3D-CMCC forest ecosystem model (v.5.1) against eddy covariance data for 10 European forest sites, Geoscientific Model Development, 9, 479–504, 2016.

Hidy, D., Barcza, Z., Haszpra, L., Churkina, G., Pintér, K., & Nagy, Z.: Development of the Biome-BGC model for simulation of managed herbaceous ecosystems, Ecological Modelling, 226, 99-119, 2012.

Raj, R., Hamm, N. A. S., van der Tol, C., and Stein, A.: Variance-based sensitivity

analysis of BIOME-BGC for gross and net primary production, Ecological Modelling, 292, 26–36, 2014.

Raj, R., Hamm, N. A. S., van der Tol, C., and Stein, A.: Uncertainty analysis of gross primary production partitioned from net ecosystem exchange measurements, Biogeosciences, 13, 1409–1422, doi:10.5194/bg-13-1409-2016, 2016.

Liu, D., Cai, W., Xia, J., Dong, W., Zhou, G., Chen, Y., Zhang, H., and Yuan, W.: Global validation of a process-based model on vegetation gross primary production using eddy covariance observations, PLOS ONE, 9, e110 407, 2014.

Yuan, W., Cai, W., Xia, J., Chen, J., Liu, S., Dong, W., Merbold, L., Law, B., Arain, A., Beringer, J., Bernhofer, C., Black, A., Blanken, P. D., Cescatti, A., Chen, Y., Francois, L., Gianelle, D., Janssens, I. A., Jung, M., Kato, T., Kiely, G., Liu, D., Marcolla, B., Montagnani, L., Raschi, A., Roupsard, O., Varlagin, A., and Wohlfahrt, G.: Global comparison of light use efficiency models for simulating terrestrial vegetation gross primary production based on the LaThuile database, Agricultural and Forest Meteorology, 192âĂŤ193, 108–120, 2014.

Zhou, Y., Wu, X., Ju, W., Chen, J. M., Wang, S., Wang, H., Yuan, W., Andrew Black, T., Jassal, R., Ibrom, A., Han, S., Yan, J., Margolis, H., Roupsard, O., Li, Y., Zhao, F., Kiely, G., Starr, G., Pavelka, M., Montagnani, L., Wohlfahrt, G., D'Odorico, P., Cook, D., Arain, M. A., Bonal, D., Beringer, J., Blanken, P. D., Loubet, B., Leclerc, M. 5 Y., Matteucci, G., Nagy, Z., Olejnik, J., Paw U, K. T., and Varlagin, A.: Global parameterization and validation of a two-leaf light use efficiency model for predicting gross primary production across FLUXNET sites, Journal of Geophysical Research: Biogeosciences, 121, 1045–1072, 2016.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-216, 2016.