Geoscientific

Model Development

Discussions

[GMDD]

Interactive

comment

Interactive comment on "Land surface model photosynthesis and parameter calibration for boreal sites with adaptive population importance sampler" *by* Jarmo Mäkelä et al.

Anonymous Referee #1

Received and published: 2 April 2019

Overview:

1. The authors apply adaptive population importance sampler and a simple stochastic optimization algorithm to optimize the parameter sets of six different stomatal conductance models using measurements from 10 FLUXNET sites. For the validation, the experiment period is split into the optimization period and validation period at the six study sites. The remaining four study sites were used only for the validation. The reproducibility of GPP and ET was investigated with the optimized parameters. For the drought event at one flux site, the effectiveness of additional parameter optimization for water use efficiency was also investigated.

[Printer-friendly version]

[Discussion paper]

2. The results indicate that the optimization scheme presented in this paper successfully improve the estimation of GPP and ET, even for the drought event. The model was also modified to use a delayed effect of temperature for photosynthesis activity.

General comments:

Parameter optimization is essential for the model development. The methods proposed in this paper successfully optimize the parameter sets which control carbon flux or water flux. The procedure of this paper seems generally adequate, and I think the paper is relevant for GMD. However, the manuscript is needed to be improved from the two aspects:

1. The readers of this paper may not understand and reproduce the experiment because some procedures in the paper are not clear. In addition, the descriptions are sometimes too much redundant or too much simple, and the argument becomes unclear. The authors need to improve the manuscript carefully according to the specific comments.

2. The authors indicate that some of the settings are not appropriate for the USO model to simulate the drought event. Nevertheless, the authors concluded that the estimation with USO is one of the "best" results. The authors should run the experiment again with the appropriate setting if they would like to use the USO result for the discussion.

Specific Comments:

(1 is related to general comment 1, and 2 is related to general comment 2)

1-1. P2 Lines 3-8: The explanation for soil drought is confused. I think it is better to explain "general" soil drought first, and then emphasis the importance of soil draught at the boreal forests. This section is important to explain why the authors chose boreal sites for the experiment. I also do not understand the sentence "reversing the development".

1-2. P2 Lines19-20: I do not understand the sentence; "However, it can be

C2

. . .conductance formulations."

1-3. P2 Lines 25-27: "We will assess the inter-site variability ...one site." and "We will provide an assessment of ... descriptions." are about the validation. I think it is better to explain the validation process explicitly starting with e.g. "The validity of the optimized parameters is assessed ...". The explanation should include the two points: 1) At the six flux sites, the experiment period is split into the optimization period and validation period, and the reproducibility of GPP and ET with the optimized parameters were investigated. The remaining four sites were used only for validation. 2) The drought event at one flux site is also investigated with some of the optimized (fixed)model parameters and with additional parameter optimization for water use efficiency.

1-4. P2 Lines 27-28: I think the sentence about the optimization method, "We utilize the adaptive population importance sampler ... (peak of high probability)." is too much short. One paragraph may be needed to explain this part. Many optimization schemes have been used for land ecosystem models. Therefore, it is better to review some of them, and the authors should explain the difference between APIS and well-known methods (e.g. MCMC and some other optimization methods). There are also several studies which estimated model parameters using flux site measurement. Therefore, advantages of this study also needed to be explained comparing to those previous studies.

1-5. P3 Lines 2-5: I do not understand the procedure clearly. - What is the time resolution of the model? - What is the difference between "half-hourly" and "daily" time series? How these different time series are used in the experiments?

1-6. P3 Lines 7-10: I think the explanation is confused. - Please explain the experimental setting more intelligible and clearly (refer to 1-3 of my comments), and the detailed explanations written in these sentences should be added. - "(with measurements separated into successive time period)": How many years are used for optimization and validation respectively?

1-7. P5 Table2: Definition and range of g1: "Table3" instead of "-" may be better to understand. I could not find the explanation about initial distributions for these parameters. How do the initial ranges in Table 2 relate to the black lines in Fig. 1?

1-8. P6 Lines 7-8: "However, coniferous evergreen trees do not ... following spring." I do not understand the connection before and after this sentence.

1-9. P7 (2.4 Parameter estimation) – P11 (2.9 Cost functions): I do not understand the procedure for the parameter optimization clearly. I think it is better to add the section for "overview of the experimental settings" between section 2.3 and section 2.4. Then, Section 2.7 may better to be included in this overview section. The procedures for initial parameter settings, spin-up, parameter optimization, and validation should be easily understood. A process diagram may help the readers to understand.

Especially, I do not understand the relations between the "parameter estimation" by adaptive population importance sampler (section 2.5) and "parameter optimization" by the simple custom stochastic optimizer (section 2.6). - In my understanding, the APIS is used to estimate model parameters roughly. Then using the estimated distributions by APIS as the initial state, parameter optimization is done. Is that correct? The overview section should include more detailed explanation for this point. Introduction (P1 Lines 2-3) also needed to be corrected so that the readers can understand the procedure. - How many years observations are used for these two different optimization methods? Are the observations for the two optimization methods same? - I do not understand how to merge the optimizes parameters as shown in Table4, because in my understanding, the parameters are optimized separately at each study site.

It is also necessary to clarify the role of "2.8 Simulation analysis" and "2.9 Cost function" at the overview section. In my understanding, the first half of 2.8 indicate the parameter sensitivity, and the latter half of 2.8 indicates the validation of the result using the observation. Cost function (2.9) is used both for parameter optimization (APIS and a simple stochastic optimizer) and for validation of the results (e.g. Table 4, Table

6). These descriptions also should be included in the overview.

1-10. P8 Line 23: What is "M=40 proposals"?

1-11. P8 Line 30 – P9 Line 3: I do not understand the procedure (see, 1-9 of my comment).

1-12. P9 Lines 30-33, P10 Line2: I do not understand these sentences (see, 1-9 of my comment).

1-13. P11 Line 29: What is "sampling limits set"?

1-14. P12 Lines 1-18: This may be better explained in introduction (see my comment 1-4). The parallel mode is not used in this study, therefore this advantage (parallel simulation) is not suitable for this study setting.

1-15. P12 Lines 29-31: "The actual soil moisture ... unreliable and even unrealistic.": Then, what is the recommended setting for the future study? Is it OK for this experiment?

1-16. P14 Line 25 – P15 Line 2: Some descriptions are redundant. Improvement of description is needed so that the readers understand the Fig. 2 clearly.

1-17. P16 Line 9: "We optimized the model for individual (calibration) sites as well.": I do not understand "as well". I thought the model was optimized at the individual sites.

1-18. P16-P19: I do not understand the arguing point in this section (3.3). I think it is better to explain Table 6, Fig.3 and Fig. 4 first, and then more detailed discussion should be done. - P16 Lines 26-32: Too much detailed and complicated. First, the categorization of the optimized (fixed) parameters and the parameters for further optimization (for WUR) should be explained using Table 4 and Table6. What is the most important different between these parameter groups? The detailed settings for the fixed parameter may better be explained in Appendix. - P16 Line 33 - P17 Line 4: This paragraph is important to describe the parameter optimization for the drought event.

How many years WUR optimization was done? Is the optimization procedure different only for cost function calculation? Are the observations for year 2006 repeatedly used?

1-19. P17 Line 16-18: The parameters are just optimized in this experimental setting, and the "true value" is not known. Therefore, I think "optimal value" should not be used here. Authors can just say that "the optimized parameter set for WUE greatly improved the simulation results (Fig. 3)".

1-20. P17 Line 14 – P 18 Line 18: The detailed explanations for each parameter are too much complicated and I do no understand. The paragraph of P18 Line 19-26 should be placed before these paragraphs. Then, the relationship between the results in Fig. 3 and the estimated parameters should be discussed as below: - Which parameters are the important to control WUE in this experiment? - How do these parameters affect WUE? - Are the estimated parameters reasonable compared to the previous studies? If not, why?

1-21. P18 Lines 22-23: I do not distinguish "the actual drought" in Fig. 3. I think it is better to add the period of the drought in the Fig. 3.

1-22. ïijř18 Lines 27-33: I do not understand this paragraph because I do not understand Fig. 4. Does the lower panel show USO results? What is "Medlyn"? I also do not understand how is the β-function for the observation calculated.

1-23. P18 Lines 34- P19 Line 5: The authors did not show the experimental setting and result explicitly, therefore I do not understand the purpose of this experiment. What is the difference between this experiment and the parameter optimization in section 3.3?

1-24. P19 Lines 17-20: I do not understand these sentences.

2-1. P17 Lines 7-13: The authors explain that the setting for USO is not appropriate. Then the results should not be used for further discussion after this paragraph. If authors would like to use the result, they should perform the experiment again with the appropriate settings.

Technical corrections:

1. P1 L11: "correctly time and replicate" -> "correctly reproduce"

2. P2 Line 27: Abbreviation "APIS" should be placed here (this is the first appearance).

3. P9 Line 5: "from 2006" -> "in 2006" (only one-year optimization).

4. P10 Lines 17-18: "high", "average", or "low" effectiveness value: this explanation should be the same as Table 4.

5. P11 Lines 19-20: Description of the Supplemental materials is needed at under each figure.

6. P15 Table5: some values are different form the Fig. 2 (i.e., r2 of Bethy).

7. P18 Line13: "disregardin" -> "disregarding"

8. P17 Line 14 "The most noteworthy" what? (modified word is needed)

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-313, 2019.