

# ***Interactive comment on “A Bayesian Ensemble Data Assimilation to Constrain Model Parameters and Land Use Carbon Emissions” by Sebastian Lienert and Fortunat Joos***

**Sebastian Lienert and Fortunat Joos**

lienert@climate.unibe.ch

Received and published: 6 April 2018

We want to thank the reviewer for the time and effort for the careful and very insightful review. In the following, we respond to the reviewer point by point, with our **responses in bold** and *quotations from the updated manuscript in cursive*. Please also consider the updated manuscript with track changes in the supplementary. Also note that we expanded the discussion section to include a paragraph on a potential bias in the fossil fuel emissions used for the deconvolution by including non-fuel uses.

**Point-by-point response**

Printer-friendly version

Discussion paper



The study presents an approach to constrain a DGVM with multiple observational streams of carbon stocks, gross and net fluxes. The authors rely on a latin hypercube stratified sampling to perturb model parameters and create several 1,000-member ensemble simulations of the terrestrial carbon cycle for the historical period. Results focus on the estimation of land-use and land-cover change emissions. This study is quite innovative in the context of the global terrestrial carbon cycle as model parameters are constrained globally.

### Thank you

Unfortunately, I have troubles following the method description. The description of the way, the parameter distributions are updated is remains fairly unclear. I recommend that the authors devote a special section in the Methods section to clarify a couple of points:

#### **A separate section on prior selection is added to the manuscript.**

a) how was the prior distribution of the parameters derived (literature ranges typically only allow to assume uniform distributions); b) how exactly is the ensemble updated after the metrics are calculated. Is it the probability distribution of each parameter, which is updates? This would lead to a new LHS set to be produced, and subsequent new model runs? Figure 1 would suggest that this is the case, but in this case, the new set would be dependent on the metric and metric weighting, which contradicts the statements made on P2 (also, it's computationally probably prohibitive). Or is each LHS sample weighted according to the model performance, and this weight then used to calculate the PDF of a modelled output? If that is the case, I don't understand the iterative nature of the LHS sample updates? Also, in this case, it would be good if the authors would elaborate on the way they've estimates the posteriori distributions.

**We introduced a new subsection discussing the method used to obtain the prior parameter distribution and reordered the section for clarity. Please refer to section 2.3.2 (p.6 l.9-p.8 l.7 in the manuscript).**

BGD

Interactive  
comment

Printer-friendly version

Discussion paper



Figure 1 in this reply shows the evolution of the median parameter values and ranges of the ensembles with 200 and 300 members (T1-T6) and large ensembles with 1000 members (E1-E3), discussed in the new section in the manuscript. Only the parameters used in the final ensemble E3 are shown. In the small ensembles, different model configuration and parameters were tested. For instance, in ensemble T2 nitrogen limitation was not considered and thus the nitrogen cycle related parameters were not sampled. The factorial simulation and small ensembles informed the choice for the prior of ensemble E1, which is then iteratively improved to arrive at the prior of E3.

Given that the authors highlight the ability to change the cost-function and weighting as a key strength of their method, it would be also interesting if they would add a discussion point as to how robust they believe the posteriori parameter distributions are against their choice of metric weighting.

**We qualitatively assessed the robustness by reevaluating the ensemble for a subset of the observational targets. For this purpose we created multiple hierarchical weighting schemes, each missing one of the observational targets or a category of targets when compared to the default version, and looked at the induced changes in the parametrization of the best guess version. We added a section discussing this approach to the manuscript:**

*“We investigate the dependency of the constrained ensemble on the choice of the observational constraints by reevaluating the ensemble for a subset of observations. We created 19 weighting schemes, each missing one of the individual observational constraints (Figure 2 and table 2) and otherwise identical to the default scheme. Then the median skill weighted parameter values of these ensembles are compared to the best-guess values of  $M_{net,net}$  (section 3.3). The relative change in parameterization is less than 1% for 15 out of the 19 considered alternative weighting schemes. Leaving away the global vegetation and soil carbon constraints lead to moderate changes, notably to a change in the parameter for mortality ( $mort_{max}$ ) of 4% and 2% respectively. Not*

[Printer-friendly version](#)[Discussion paper](#)

including the soil carbon distribution in high latitudes lead to an increase of the parameter for the dependency of soil respiration on temperature ( $E_{0,hr}$ ) of 2%. The largest changes in parameterization were observed when not considering the atmospheric deconvolution, most notably the sapwood-heartwood turnover time  $\tau_{sapwood}$  decreased by 5%. When omitting entire categories in the benchmarking scheme, the changes in parametrization are larger than for omitting individual constraints, with parameter changes of up to 1% for the fluxes, 5% for the inventory and 6% for the transient category. This shows that the final parameterization is not overly sensitive to the inclusion or omission of a single observational product.”

I have a number of further suggestions to improve the clarity of the manuscript: P1 L17: in the context of a data assimilation paper, the use of assimilated here is confusing. replace by stored?

**Done**

P2 L3: Add “Amongst others,” at the beginning of the sentence

**Done**

P2 L4: unclear what uncertain prescribed LULCC processes are meant to be, perhaps give examples, or clarify that it's the representation of these processes that is uncertain

**Done, the sentence now reads:**

*“In addition to uncertainties in the prescribed LULCC forcings and the representation of LULCC and other processes in DGVMs, the values of the applied parameters are subject to substantial uncertainties.”*

P2 L9 DA “should” be an integral part of model development, but unfortunately it is not always.

**Replaced “is” by “should be”.**

P2 L10: Is the Houwelling reference appropriate here? This does mostly relate to

Printer-friendly version

Discussion paper



inverse atmospheric modelling

## Removed Houweling reference

P2 L12: Not sure that I understand sequentially correct here. Most DA methods would assimilate different data sources simultaneously. Also, I think cost-function is the more common term for metric in this context

**Revised sentence to read:** *“A drawback of these methods is that the sampling process is dependent on the choice of the cost function, the design of which is not trivial when assimilating multiple observations simultaneously.”*

P2 L14 This sentence is a bit out of context in a paragraph on alternative DA methods, because benchmarking does in general not imply DA. It seems more logical to merge this sentence with the Paragraph starting in L25, and move the entire paragraph to L6 after Le Quere et al. 2016.

## Moved, the paragraph now reads:

*“Amongst others, Dynamic Global Vegetation Models (DGVMs) can be used to assess the contribution of LULCC to the terrestrial carbon budget (Le Quéré et al., 2016). The assessment of the performance of a given model version using observational benchmarks has been actively discussed in the literature (Hoffman et al., 2017; Peng et al., 2014; Kelley et al., 2013; Luo et al., 2012; Blyth et al., 2011; Randerson et al., 2009) and different frameworks have been proposed. The selection of observational targets is vital to a successful assimilation of observational data. In order to constrain the contemporary carbon cycle, 14 data products are used, ranging from global inventories of carbon (Ciais et al., 2013) to spatially resolved satellite estimates of photosynthetically absorbed radiation (Gobron et al., 2006). The goal of the data set selection process was to have observations capturing the magnitudes of fluxes and inventories in the carbon cycle, as well as its transient response to anthropogenic perturbation. In addition to uncertainties in the prescribed LULCC forcings and the representation of LULCC and other processes in DGVMs, the values of the applied parameters are*

BGD

Interactive  
comment

Printer-friendly version

Discussion paper



*subject to substantial uncertainties. We use a Monte-Carlo-like data assimilation approach (Steinacher et al., 2013; Steinacher and Joos, 2016; Battaglia and Joos, 2017) to sample 15 key model parameters and construct a 1000-member model ensemble to investigate this parameter related uncertainty in the DGVM LPX-Bern. Furthermore, we establish a new reference version of the model.”*

L2 L18: As noted above, I have troubles following here: LHS simply provides a set of parameter combinations, in which each parameter is sampled given a specified distribution and notably, ensuring that there is no correlation amongst any of the parameters. LHS does not imply any model metric per se. The way the posterior distribution is derived from the prior distribution and the model metrics is unclear. How many iterations would be needed to arrive at a stable solution, what is the stopping criteria, and why is it possible to change the metric during the DA procedure? This would change the posterior distribution, and therefore impede convergence.

**Please see the answer to the major points.**

P2 L31: I wonder if the flow of the introduction would be more logical if one would first talk about the LULCC processes as in this paragraph, then about the benchmarking in the preceding paragraph, and only then about data assimilation?

**We have restructured the introduction but slightly deviated from the reviewers suggestion to improve text flow. Please also see the attached manuscript with track changes.**

P2 L31: While the (add) “net” land-atmosphere flux can “to some extent” be . . .

**Done**

P2 L32: add “residual” terrestrial carbon sink?

**Done**

P4 L21: I think it is worth highlighting that the strength of LHS over other sampling

Printer-friendly version

Discussion paper



techniques is that the set of parameters in uncorrelated.

**Done:**

*"..to generate an uncorrelated parameter ensemble of a given size."*

P4 L20: The text confuses MC parameter sampling techniques, which are independent of any purpose the sampling is made for, from MC Data assimilation techniques, which are not?

**Changed "Monte Carlo sampling techniques" to "Monte Carlo data assimilation techniques"**

P4 L26: the description of  $\alpha_a$  should correspond to table 1, it is not FAPAR!

**Done:**

*"The fraction of photosynthetically active radiation assimilated at ecosystem level relative to leaf level,  $\alpha_a$  .."*

P4 L7: Literature range only allow to give uniform distribution. How where the nonuniform distribution parameters obtained / estimated?

**See answer to major points and attached manuscript.**

P4 L7: I have trouble following from here on. Maybe this would become clearer, if first all the metrics and data sources were explained, and then the way the distributions are updated is clearer presented.

**Revised this section (See answer to major points)**

P8 L10: which winds were used for the transport? I assume that the winds were not interannually varying?

**Yes, the transport matrix does not include interannual variability. Added:**

*This method does not include the interannual variability of the transport. Additionally, we added an explanation to Figure 8 for clarity:*

**BGD**

[Interactive  
comment](#)

[Printer-friendly version](#)

[Discussion paper](#)



*"As expected, the interannual variability in seasonal amplitude of CO<sub>2</sub> is not captured as the atmospheric transport model TM2 does not represent interannual variability in mass transport."*

L9 L5: Inversion typically refers to the inverse modeling of atmospheric transport, whereas here - as far as I understand this, you simply take the land flux as the residual of the fossil fuel emission and ocean uptake.

**We changed all occurrences of inversion to deconvolution.**

P9 L10: Are these data sources not redundant with the global maps of total and soil C storage described earlier?

**While the information of the global carbon content is also contained in the maps, we feel the inclusion of the additional, well established, global target is warranted by the importance of these targets. This is effectively increasing the weight of these targets.**

P9 L27: I don't understand the reasoning for the duplication of ensembles with gross transitions. Please motivate.

**We did not repeat the procedure to improve the prior distribution (See updated manuscript) for  $M_{gross,gross}$  and as such the prior and posterior distributions do not converge. Consequently we do not feel comfortable to use  $M_{gross,gross}$  as the basis for our estimates for  $E_{LUC}$ . As a compromise we introduced  $M_{gross,net}$ , retaining the confidence in the performance of  $M_{net,net}$  and simply adding the important processes of shifting cultivation and wood harvest.**

P9 L31: As noted above, I have difficulties following this description.

**Section revised completely, see the answer to major points.**

P10 L5-8: Is material for the introduction, not the results section

**Removed Paragraph**

BGD

Interactive  
comment

Printer-friendly version

Discussion paper





P10 L 8-11 can be safely removed.

### Removed Paragraph

Section 3: When giving numerical estimates, please add either range or standard deviation, whenever the number is based on the ensemble. I also think that the more logical arrangement of the Results sections would be to first talk about model performance, and then about the attribution of the net land flux to LULCC and residual.

**Added the skill weighted 90% confidence interval for every reported number, except when reporting the median difference between two ensemble configurations. We agree that the suggested order of the result section is more logical, however we feel that the results on LULCC are of more interest to a broader range of readers, and thus prefer to lead with those results.**

P13 L4: Why would an underestimation of the ELUC not affect your conclusions about ELUC?

**The net land-atmosphere flux is underestimated because  $E_{gross,net}$  features additional processes that lead to an increase in  $E_{LUC}$ , while the residual land sink remains constant. However if only considering  $E_{LUC}$  we expect the magnitude of the residual land-sink and net land-atmosphere flux to be less important than for instance model performance in the vegetation carbon benchmarks (Li et al., 2017). For clarity we revised the sentence:**

*“A caveat of this choice is that the net land-atmosphere flux is underestimated in  $M_{gross,net}$  because the residual land sink only responds to the lower  $E_{LUC}$  of  $M_{net,net}$ . However if only considering  $E_{LUC}$  we expect the magnitude of the residual land-sink and net land-atmosphere flux to be less important than model performance with respect to vegetation carbon (Li et al., 2017) and other benchmarks.”*

P17 L6: is the use of the word significant appropriate here?

**Changed 'significant' to 'relevant'**

BGD

Interactive  
comment

Printer-friendly version

Discussion paper



P17 L8: why not?

**Using the vegetation carbon distribution directly, would have been a valid choice. Exchanging the total carbon distribution for the vegetation carbon distribution in the hierarchical weighting scheme reveals that the median parameter values used for the best guess version change less than 0.5%.**

P 18 L 3: Why is this different from the approach described in Section 2?

**Sentence shortened and clarified to read:**

*"We compare the total land-atmosphere exchange flux to the results of the atmospheric CO<sub>2</sub> deconvolution in Fig. 11"*

Conclusion Section: There is no need to repeat details of the methods or approach undertaken

**Shortened conclusion section by removing sentences which go into too much detail.**

Figure 1: Ensure all lines are visible

**Adjusted legend**

Table 1: Check units and definition for E<sub>0</sub>. This seems more like an activation energy to me (not a temperature sensitivity What are the units of the k<sub>la:sa</sub>? Is this simply a scalar?

**E<sub>0</sub> is defined according to Lloyd and Taylor 1994, which considers the effect of an activation energy which is varying with temperature. Using their representation it has the unit [K] and is strictly speaking not an activation energy. As such we find the definition appropriate. k<sub>la:sa</sub> scales the PFT dependent leaf area to sapwood area and is as such unitless. The leaf area to sapwood area has the units [m<sup>2</sup>/m<sup>2</sup>]**

**BGD**

Interactive  
comment

Printer-friendly version

Discussion paper



Please also note the supplement to this comment:

<https://www.biogeosciences-discuss.net/bg-2018-62/bg-2018-62-AC1-supplement.pdf>

---

Interactive comment on Biogeosciences Discuss., <https://doi.org/10.5194/bg-2018-62>, 2018.

**BGD**

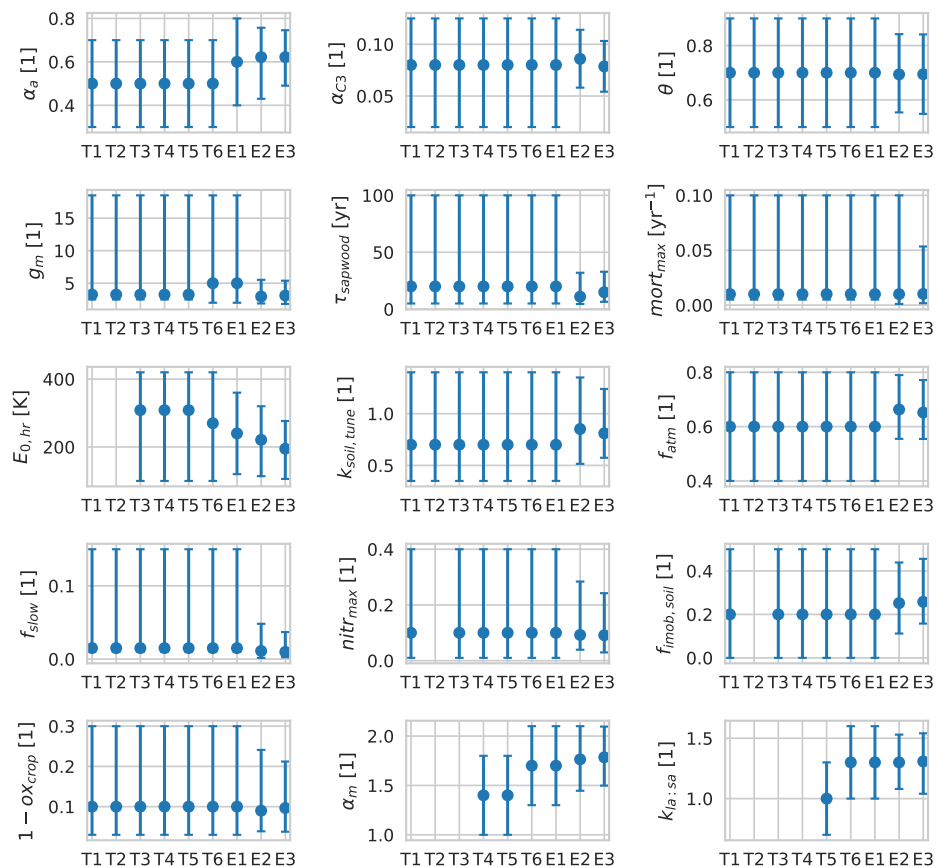
---

Interactive  
comment

Printer-friendly version

Discussion paper





**Fig. 1.** Median and 90% confidence intervals used for the prior distributions of the parameters of 9 ensembles. T1-T6 are ensembles with fewer members and E1 and E2 were precursors of the final ensemble E3.