Biogeosciences Discuss., 9, C5928–C5940, 2012
www.biogeosciences-discuss.net/9/C5928/2012/

Biogeosciences Discussions

Interactive comment on "The relative importance of decomposition and transport mechanisms in accounting for C profiles" *by* B. Guenet et al.

M. Braakhekke (Referee)

mbraak@bgc-jena.mpg.de

Received and published: 23 November 2012

Maarten Braakhekke

This paper describes a modelling study of soil carbon cycling at a long term bare fallow experiment performed on a Chernozem soil in Russia. Soil carbon profile measurements from several time points up to 58 years into the experiment as well as an unmanaged site were used to calibrate soil carbon models. Two formulations of organic matter decomposition and three formulations of soil organic matter transport, were combined in a factorial model calibration experiment. Calibrations were performed in a Bayesian framework.

This study probably represents the most comprehensive comparison of model formulations for soil organic matter transport done so far. Also the comparison of a representation of limitation of SOM decomposition by labile substrates with a simple first order kinetics model is relatively new. Especially the combination of vertical transport combined with this new decomposition model is very innovative. This topic is also relevant since the combination of the two may have important implications for long term soil carbon cycling, although this still quite uncertain. Long term bare fallow experiments represent a very useful "test bed" for studying these processes because stopping of fresh litter input may lead to increased limitation of SOM decomposition.

Unfortunately, the execution of the experiment is done rather poorly to my mind. The authors have performed the calibration in a Bayesian framework–which I commend–but several unnecessary simplifying assumptions were made which reduce the merit of the results. Furthermore, there are several issues with the results which suggests that mistakes were made. The exact details of the study setup cannot be fully checked because the methods section is lacking much important information. Finally, I also believe that too strong conclusions are drawn from regarding the importance of the nature of SOM decomposition.

It seems unavoidable that the calibrations have to be redone. Furthermore, although I classified the paper as in need of major revision, it may be necessary to submit a new manuscript, in view of my many criticisms. I'll leave this decision to the editor.

For clarity I divided my comments into several levels of importance labelled "suggestions", "strong suggestions", and "required". Suggestions are points that I believe would improve the paper but do not strictly need to be followed. Required means that the authors should follow the advice, or provide adequate reasons for not doing so. Strong suggestions are in between these two.

Finally, a paper describing a similar calibration study I performed is currently under review for BG and available in BGD (Braakhekke et al., BGD, 9, 11239–11292, 2012).

It has many similarities with the current study and might be helpful to elucidate some of my comments.

General comments

- It is not clear how the posterior parameter estimates were derived. The usual approach for models that cannot be analytically inverted is to use a Markov Chain Monte Carlo method to sample the posterior distribution. However, no posterior distributions are shown and, aside from the dashed lines in figures 4-5 (which are incorrectly derived to my mind; see below), no information is provided on posterior parameter uncertainty. This suggests to me that the authors simply used a classical gradient method (e.g Levenberg-Marquardt). If this is correct, I would strongly suggest redoing the calibration with an MCMC approach. Contrary to what the authors assume (page 14154, lines 8-10) the posterior distributions may be quite non-Gaussian and may even be multi-modal, since the models are complex and likely over-parameterized w.r.t. the available data. This can only be evaluated by approximating the full posterior distributions. (strong suggestion).
- The authors assume that the errors on the parameters and the observations are uncorrelated (page 14155, line 1-2). While the ignoring correlations in the observations may be acceptable, I believe that this assumption is unjustified for the parameters. Since the amount of data used in the calibration is quite limited compared to the number of optimized parameters, all models are presumably over-parameterized, which means strong correlations between the parameters can be expected. This must be considered, also for determining the predictive uncertainty (cf. my point related to Figs 4 and 5, below). If the posterior distribution is sampled using MCMC, the posterior sample will automatically reflect these correlations. In case a gradient approach is used to derive the mode, the

linear correlations can be derived from the Jacobian at the posterior mode (see e.g. Omlin & Reichert, Eco. Mod., 1999). (required)

- The conclusions drawn from the model results regarding the importance of labile substrate limitation of SOM decomposition for long term soil carbon dynamics are too strong, considering the uncertainty in the models,a the limited amount of data used in the calibration, and our limited understanding of soil carbon cycling. The superior fit of model formulation FS2 is to be expected since it has one additional parameter compared to FS1 (cf comment first on Results section below). But even if model FS2 would score better om a statistic that considers the number of parameters, this would still not be hard evidence for the importance FOM-SOM feedbacks. I would recommend constructing an additional first order kinetics model with three pools instead of two. If this model would also perform worse than model FS2, this would lend a bit more support for the priming formulation. However, I believe that the amount of available data in this study in general precludes the possibility of such hard conclusions. The wording of several sentences in the discussion needs to be adjusted accordingly (page 14158 lines 5-8 and 24-27). (required)
- It is incorrect to refer to the limitation of SOM decomposition by labile substrates as the "priming effect". The priming effect, as I'm sure the authors are aware, is simply the observed increased (heterotrophic) respiration under certain conditions. This can be caused by many factors, including stimulation of microbial activity due to increased availability of labile substrates (see Kuzyakov et al., SBB, 2000). Calling the dependence of SOM decomposition on FOM priming is mixing up cause and effect. A more appropriate description could be "substrate interactions". This should be corrected in various locations in the MS. Calling model formulation FS2 a "priming model" may be acceptable, because one could argue the model is designed to capture the priming effect. (required)

- The term “priming”, “priming model”, and “priming effect” are used with double quotes throughout the paper. Double quotes are used to introduce a new term or to signal that a term is term is used in an unusual way. This only needs to be done only the first time this term is used. Please remove the quotes for all the other times the term is used. Possibly the quotes could remain in both the abstract and the introduction. (strong suggestion)
- Throughout the paper “over estimation” and “under estimation” are spelled incorrectly. Please write “overestimation” and “underestimation”. (strong suggestion)

Abstract

- Line 1: Please write either “Soil is the major terrestrial reservoir” or “soil is one of the major terrestrial reservoirs”. (required)
- Line 10: “advection or diffusion”, replace “or” with “and” (suggestion)

Introduction

- Page 1414, line 7: please write: “...first soil layers which is/was considered...” (suggestion)
- Page 1414, line 11: “an increasing attention”; please remove “an”. Attention is an uncountable noun. (required)
- Page 14147, lines 18-23: Bruun et al. (SBB, 2007) compared a model with dispersion (diffusion) and advection to a model with advection only. Therefore I don’t

think it’s really correct to write that no clear comparison between transport representations has been done so far. Just indicate that in the current study the comparison is done more formally and comprehensively and adds the diffusion-only representation, and the two representations of decomposition. (strong suggestion)
- Page 14148, line 18: please remove “has” and “for years”. (suggestion)

Methods

- The paper is missing much important information on the modelling and optimization techniques. Although some things may be guessed or deduced, it is better to provide enough explicit information so that the study could in principle be reproduced by others. Certain information could included in an appendix or online supplementary material (I’ll leave that to the authors’ discretion) but the methods section definitely needs to be extended. (required)

  Specific aspects related to the models that should be included are:

  - the top and bottom boundary conditions of used; depth of the bottom boundary
  - the model solution; i.e. what numerical scheme is used? What are the thicknesses of the layers of the spatial discretization; the same as the measurement depths?
  - the precise setup of the simulations; e.g. the initial conditions and the simulation length. Is the model run until steady state to acquire the estimate for the steppe soil?

  Also the algorithm for doing the calibrations and its precise setup should be described.

- Please consider writing equations more professionally, e.g. with MS Word Equation Editor, LaTeX, or similar software. Single line equations are difficult to read, particularly Eq. (14)-(15). (strong suggestion)

- The units of the variables in the methods section are used inconsistently: C stocks are expressed in t ha^{-1}, the diffusion coefficient is expressed in cm^{2}yr^{-1}, and the advection rate is expressed in mm yr^{-1}. I would strongly suggest using consistent units, preferably SI. For graphs the quantities can be converted if necessary. (strong suggestion)

- Page 14150 lines 3-8: how was the soil in the LTBF site kept bare; just by ploughing? Please explain. (suggestion)

- The explanation of the method used to account for compaction of the LTBF soil with respect to the steppe site is not clear. What exactly is meant by the "floor of the steppe"? Was this correction applied on the model results or on the measurements? (required)

- Here it is written that the soil was sampled in 10cm depth increments down to 150cm, i.e. 15 data points per profile. However, section 2.3 and Fig 4 & 5 indicate 12 points per profile. Why the difference? (required)

- Equation (1): in the text it is written that the C stocks are expressed in t C ha^{-1}. However the units of the variables in this equation lead to units of g cm^{-2} for the C stock. Please correct. (required)

- Fig 1: there are arrows leading from the SOM to the FOM pool, but this does not correspond to any process or feedback explained in the model description. (required)

- Page 14151, line 5: this sentence, up to the comma is awkward. Consider revising. (suggestion)

- Equation (3): this equation is incorrect because the change of the FOM stock is also affected by input and transport, which are not included. Consider defining a decomposition flux (e.g. F_{dec}), and later defining the rate of change of the stock as the sum of all fluxes, including input, transport and decomposition. Alternatively, one could write the rate of change of a stock by one specific process PR as follows:
  $\frac{\partial FOM}{\partial t}\big|_{\mathrm{PR}} = \dots$
  This applies also to equations (4)-(7). (strong suggestion)

- Equations (4)-(5): There seems to be a mistake related to the parameters r and e. It is written that these parameters partition the decomposition flux of FOM into a fraction flowing to SOM and a fraction lost as CO_2. First, this means that these two quantities are fully dependent according to $r = 1 - e$. Thus, really only one parameter needs to be defined here. Second, in Table 1. it is indicated that r is not included in the calibration and fixed at 0.4, while e *is* included (but for some reason the posterior estimate is 0.5 for all models). This seems incorrect: either both of them must vary, or both of them must be fixed. In any case the sum of the two quantities must always equal 1, otherwise mass balance errors will occur. If this is a mistake, the calibration must be redone. (required)

- Page 14152, line 6: Where does notation "FS" for the decomposition models come from? A more intuitive acronym might be better. (suggestion)

- Eq. (7): This equation assumes that the microbial biomass always in equilibrium with FOM. Please mention this explicitly in the text. (required)

- Why were the 20 and 26 year BF data not included in the calibrations? To leave some data for validation? I would suggest using all profiles, since the amount of data is quite small as it is, and it seems that the value of the 20YBF and 26YBF profiles for validation is limited. (suggestion)

- Why was the decomposition rate of FOM (k_{FOM}) not included in the calibration? Was there enough a prior information to fix this parameter? If this cannot be adequately defended reruns should be done in which this parameter is included. (required)
- Page 14154, line 7: please "contrasted" with "contrasting" (suggestion)
- Page 14154, line 12: please write "least square*s*" approach" (required)
- Page 14154, line 15: "the Fick's coefficient"; remove "the" (required)
- Page 14154, line 19: replace "than" with "as" (required)
- Page 14154, line 18-20: were the prior distributions also Gaussian? (required)
- Page 14154, line 20-21: if the parameters should be "as free as possible to fit to the data", uniform priors (possibly for all positive reals) would more appropriate. I would remove or rephrase this line. (strong suggestion)
- Page 14154, line 21: replace "possibly" with "possible" (required)
- Page 14154, line 19-20: please provide also a reference for the justification prior estimate of the c parameter of the priming model. (required)
- Page 14154, line 22: "observational error" refers only to the variance of the measurements. I would suggest using the phrase "variance of the model-data residuals". (required)
- Page 14154, line 25-27: "to fulfil some statistical hypothesis...". Please explain this more clearly, possibly with a reference. Also, the "cost function" has not been explained previously. Finally, it is not discussed in the results whether this hypothesis is actually fulfilled. (required)

- Page 14155, line 6: please write "mean square*d* deviation". (required)
- Page 14155, line 5: the reference should "Gauch et al., 2003", not "Hugh" (required)
- Page 14155, line 11: replace "then" with "the". (required)
- I believe the SB, NU and LC are not widely used quantities, so they could use a bit more explanation as to their precise interpretation. (strong suggestion)

Results

- The number of calibrated parameters ranges between 3 and 5 for the 6 models. However, the statistical indicators used here do not reflect this, nor is it mentioned in the discussion. It is not surprising that a model which has more parameters has a better fit. When comparing the models please use a goodness of fit statistic that considers the no. of parameters such as the Akaike information criterion, or (better) Bayesian information criterion. (required)
- Fig 3: There seems to be a problem with the calibration results: the fit for model III is worse than that for models I or II, which should not be possible. If the number of parameters is not considered, the fit of model III should always be at least as good as for models II and III because it includes both the processes included in the other two models (diffusion and advection). This may be caused by lack of convergence of the fitting algorithm. (required)
- No distinction is made between profiles that were used in the calibration and those that weren't, neither in the graphs nor in the discussion. I as mentioned above, I think it would be better to include all data in the calibration. If the authors

prefer to use only a subset of the profiles, the distinction should be clear in the discussion of the model results. (required)

- Table 1: I assume the columns with the “values for formulation” indicate the posterior modes. Please report additional information out the posterior distributions, i.e. the uncertainty and correlation structure. If possible show graphs of the marginal distribution of each parameter, or otherwise box plots. Also compare this with prior uncertainty. (required)
- Fig 3: It is impractical to have to refer to Fig. 2 to for the interpretations of the roman numerals. I would suggest writing the meanings of the letters a-e as titles above the bar graphs and explaining the roman numerals in the caption (possibly with abbreviations for the model formulations). Also the font size should be increased. (suggestion)
- Figs 4-5: Similar to previous point: rather than labelling each graph with “C stock” it would be better to label each graph with the corresponding model formulation and the site. Also here the font size needs to be increased. (suggestion)
- Figs 4-5: there seems to be a misconception related the predictive uncertainty as indicated by the dashed lines. It is not fully described how these are derived but I guess that the authors made three simulations per plot: one with all parameters at the mode and two with all parameters at the mode$\pm$1 standard deviation. This not the correct way to assess the predictive uncertainty because the correlations between the parameters are ignored. Likely strong correlations exist (cf my 2nd general point, above) which means that the simulations where all parameters are increased or reduced simultaneously are not meaningful. In some cases the dashed lines cross, which the authors interpret as the model not being realistic. I strongly suspect that this in fact caused by ignoring the parameter correlations. Instead, the predictive envelopes should be determined from results of an ensemble of simulations with parameters sampled from the posterior distribution, in

which the correlations are considered. A sample derived with MCMC could be directly used for this. Alternately, one could sample the parameters from a multivariate normal distribution around the posterior mode, but again the parameter correlations must be taken into account. (required)

Discussion and conclusions

- The SB, NU and LC indicators are useful quantities for evaluating model performance. However, although these indicators are reported for the different models, their interpretations are not discussed (except for the last line of the conclusions). Simply reporting these quantities without discussing what they mean in terms of the model-data comparison does not really add any insight. Either add additional discussion on this, or simply report and discuss only an overall goodness of fit statistic. (strong suggestion)
- No attempt is made to interpret the transport equations in terms of the underlying processes bioturbation and transport with the liquid phase. I think the discussion would benefit from this. This also relates to my next point. (strong suggestion)
- The cease of fresh organic matter input in the LTBF experiment will negatively affect the soil fauna, which means bioturbation will be reduced. This in turn is expected to reduce mostly the diffusion coefficient. This means that D likely is different before and after the start of the bare fallow experiment, which has not been considered in this study. This may have been a necessary simplification but it should be discussed, particularly since this concerns a Chernozem soil which are thought to be strongly influenced by bioturbation. (required)
- Page 14158 line 6: “shows off”; incorrect usage. (required)
- Page 14158 line 19: replace “when” with “where” (required)

- Page 14158 line 24: remove the last “s” in “decomposers” (required)
- Fig. 6: what is meant by “carbon input”? Since the LTBF sites show also input here, I assume it includes also the input due to transport. Is this then net input (difference between flux at the top and at the bottom of the layer)? Also, please indicate which model formulation was used to derive the results in this figure. (required)
- Page 14159 line 15: “in face of”; incorrect language. (strong suggestion)
- Page 14160, “the oldest is the SOM...” This conflicts with figure 3e. This is the oldest bare fallow site but here the differences between the T_A and T_{AD} are largest. But even if this was correct, I don’t think it can really be said that it has been shown that “different transport mechanisms are identified for different pools of C” (lines 17-19). (required)
- Page 14161 line 18: replace “this” with “the” (strong suggestion)
- Page 14162 line 7: add “a” before “crossing” (required)
- Page 14162 line 9: remove “the” before “some”. (required) Consider revising the whole sentence.
- Conclusions section: this section does not really have any conclusions. The points made are more appropriate for the discussion or even results section. Please summarize the study, and draw real conclusions here. (required)