Biogeosciences Discuss., 8, C2322–C2326, 2011
www.biogeosciences-discuss.net/8/C2322/2011/

Biogeosciences
Discussions

Interactive comment on "Modeling the vertical soil organic matter profile using ^{210}Pb$_{ex}$ measurements and Bayesian inversion" *by* M. C. Braakhekke et al.

M. van Oijen (Referee)

mvano@ceh.ac.uk

Received and published: 3 August 2011

GENERAL COMMENTS

This is a very well written paper about parameterisation of a simple but elegant soil profile genesis model, which will hopefully lead to lively discussion. I found this interesting and original work and the presentation, including very informative graphs, made it a pleasure to read. Much of the Discussion is also worthwhile, and the attempt to link the results of the parameter estimation to ideas for model simplification (rather than adding extra processes) warrants wide attention. However, some of the methodology used needs to be explained in greater detail, in particular regarding the inference of multimodality of the posterior distribution for the model's parameters. It is not clear we are really dealing with multimodality here. The methodology seemed at some places

dubious, as commented on in more detail in the specific comments below.

SPECIFIC COMMENTS

1. p.7262. The simulations started at forest planting. The assumption was made that at that time there was zero initial soil carbon. How can that assumption be justified given that we are not talking about desert soils? How does the assumption affect the simulated SOM profile of the mature forests and the relative roles of the three main soil processes in profile formation? Did it not lead to underestimation of the role of slow processes like bioturbation?

2. p.7262. The use of smoothed, average annual cycles of soil temperature, water content and root litter production: will that affect all three soil formation processes equally? Or will it downplay the role of processes that depend on events like big rain showers, such as liquid transport? Can you test that by doing runs with smoothed and non-smoothed weather?

3. p.7270. Please give a reference explaining Hermitian extrapolation.

4. p.7270, Eq.(6). Use common notation and write the likelihood function either as 'P(O|theta)' or as 'L(theta|O)', but not as 'L(O|theta)'. The first common notation emphasizes that the likelihood function equals probability density in data space, and the second emphasizes that the likelihood is a function of the parameters (not the data, which are given), whereas the third notation mixes things up.

5. p.7272, l.26. Why should the transformation factors for two different compounds (root litter, fragmented litter) not sum to 200%?

6. p.7273-7277. What is the evidence for the multimodality? Not enough is given, so the so-called "cases" at present seem purely hypothetical constructs. Note that we need better evidence than contour plots (such as Fig. 12), because interpolation algorithms suffer from sampling error. Multimodality of the parameter distribution is a quite rare phenomenon in simple dynamic models and it needs to be shown carefully that

what we are looking at is not poor convergence of the MCMC or poor intrapolation in contour graphs. The fact that even for Hainich, the site for which multimodality was presumably highest, the dominant mode was several orders of magnitude more probable than the subdominant (p.7279) suggests we are talking about a very slight lack of smoothness in the posterior sample, which is a natural side-effect of using MCMC and may not really be of any importance. If multimodality cannot be demonstrated convincingly, the paper needs much rewriting to remove all the discussion about the "cases".

7. p.7286. The first step of the three-step procedure for deciding acceptance seems superfluous because the upper bounds of the parameters (step (i)) are part of the definition of the prior (step (ii)). Presumably all that is done in the first two steps together is examining whether the proposal has jumped outside the prior bounds of the parameter distribution, which can be checked in one step.

8. Appendix A3 explains the method for establishing convergence of the MCMC, but the method seems dubious. The common (semi-)formal approach is to run multiple chains and declare convergence when all chains end up staying in the same part of parameter space. The authors declared convergence much earlier, namely when groups of chains could be distinguished with similar behaviour while different groups remained in different parts of parameter space. So local rather than global convergence. But convergence is only truly reached when all chains converge together, not just different subsets of them. Being satisfied with local convergence just increases the chances of erroneously believing you have identified local extrema, all the more with the short chains of 20000 that were used and the not very severe requirement of the Gelman-Rubin statistic being less than 1.1.

9. Table 2: why are the marginal prior distributions for the transport parameters not given?

10. Supplemental material, Fig. 1. Vertical root litter input distribution as assumed for

the two sites: should that not be model output rather than input if the aim is to explain or predict soil profiles?

11. Supplemental material, Figs 2 & 5. Modify the headers of these figures, which show more than just the cost function. The cost function as defined in Eq. (8) is part of the calculation of the likelihood and does thus not include a term for the prior distribution. Also, the parameters have distributions and therefore the cost function does as well – so explain that the figures show, for each case, the decomposition for only one parameter vector taken from the posterior distribution, rather than for the whole distribution.

TECHNICAL CORRECTIONS

1. There were small but obvious linguistic errors on (page/line) 7258/10, 7265/10, 7266/4, 7267/12, 7267/15, 7269/6, 7269/25, 7276/11, 7277/25, 7279/9, 7280/7, 7283/2, 7287/5, 7287/15, 7288/24.

2. I believe that the term 'optimization' is incorrectly used in this paper to refer to the Bayesian parameter estimation. Bayesian estimation is just the normalized multiplication of prior and likelihood, following Bayes' Theorem. So the posterior distribution is not the 'optimum' result, it is the only possible result if you are following a Bayesian approach. It is therefore better to avoid any possible confusion with standard methods of numerical optimization, such as are used to tune parameters to find the single parameter vector that gives the highest goodness-of-fit, and use 'estimation' or 'calibration'.

3. The paper confusingly uses the term "sample" in two different ways. In the main text it refers to a collection of parameter vectors sampled from the probability distribution for the parameters. In the three Appendices, a "sample of the parameters" means just one individual parameter vector.

4. The term "likelihood" is also used inconsistently. In Appendix A, the term is used both for the likelihood function itself and for the product of prior and likelihood in the

authors' novel phrase "posterior likelihood". On p. 7272 (from l. 9) the term "maximum likelihood" (which has a specific technical meaning) is incorrectly used to refer to the mode of the prior probability distribution. In this paper I would use the term likelihood exclusively for the likelihood function.

Interactive comment on Biogeosciences Discuss., 8, 7257, 2011.