Biogeosciences Discuss., 7, C2257–C2263, 2010
www.biogeosciences-discuss.net/7/C2257/2010/

Biogeosciences Discussions

Interactive comment on "Long-term bare fallow experiments offer new opportunities for the quantification and the study of stable carbon in soil" *by* P. Barré et al.

P. Barré et al.

barre@geologie.ens.fr

Received and published: 11 August 2010

We thank Benedikt Scharnagl for his very stimulating review. We won't come back on his technical corrections (part 3) that will be all taken into account in a further version of the manuscript. Please find below our responses (in plain text) to the specific comments (between " ").

"2 Specific comments (2.1) The pedotransfer function used to estimate missing bulk densities (p. 4895, l. 26, equation number is missing) uses only one predictor variable (SOC). Moreover, it seems that this equation is rather site specific and only regionally applicable. More universal and reliable estimates can be obtained when accounting for

other predictors, such as texture, stone content, soil depth and soil type. I recommend to use the pedotransfer function developed by Martin et al. (Soil Sci. Soc. Am. J., 73:485-493, 2009)."

We thank the reviewer for his suggestion. We will use this pedotransfer function and it will bring some small changes in Table 1.

"(2.2) Why do the authors use informative priors with "very large" dispersion for the model parameters? How large is "very large"? Wouldn't it be more appropriate to apply noninformative (uniform) priors in this case? Noninformative priors would actually better reflect the prior knowledge about the model parameters and circumvent the need for assuming log-normal and uncorrelated informative priors. From my point of view, both assumptions are questionable anyway. In doing so, only the limits of the feasible parameter space need to be specified. Furthermore, the term that quantifies the mismatch between the model parameter estimates and their priors would disappear from Eq. (1). The approach would then no longer be considered Bayesian, of course."

We believe that the use of noninformative priors, as suggested by the reviewer, is not appropriated for the following reasons: 1) By construction we know that the parameters of the different models should be positive as these equations are chosen to represent a specific SOC decomposition process. The use of a log-normal transformation is thus crucial to ensure this condition (positive values). 2) We choose the Bayesian approach including a mismatch between model parameter estimates and their priors, weighted by the prior parameters uncertainty, as a very general approach widely used in inversion problems (refs Tarantola 1987 ; Santaren et al. 2007). Indeed, although our knowledge on the prior parameter is rather small, such formulation allows to account for it (i.e. to account for parameter errors and error covariances). For most sites, we choose "very large" dispersion (i.e. dispersions that are between 2 to 4 times the parameter values) and we verified that the results are rather independent to the choice of these errors. However for some sites and models if we use no prior information (i.e. extremely large errors) the algorithm becomes unstable with a solution oscillating

during the minimisation process (iterative process) between different sets of parameter values. Some of these parameter sets are unrealistic given the physical law that we intended to represent with the model equations. The use of prior information, although rather large regularize the inverse problem and leads to a physically acceptable and stable solution. For these reasons, we thus keep the Bayesian approach with very weak prior information. We will change the text to justify this choice.

"(2.3) In general, the uncertainty in the SOC measurements could be estimated simultaneously with the model parameters. This would avoid making assumptions about the magnitude of the measurement error, which has been reported by the authors to be problematic. In the case of the Askov, Kursk and Versailles data the first guess was shown to be too overoptimistic (that is, too small). Also, including the measurement uncertainty in the vector of estimated parameters $\sim$xb would result in more reasonable uncertainty bounds on the estimated parameters and, consequently, model predictions. In this approach, however, the interpretation of this error measure would change slightly, in a sense that it now includes the uncertainty due to measurement error and the uncertainty due to model error."

The use of an optimization approach that estimates simultaneously the measurement errors is not classical and not widely used. We believe that our attempt to specify the "SOC measurement errors" based on independent and expert physical knowledge brings additional information to the optimization process. It allows to include the information given by the spread of different replicates ("Versailles" site) as a proxy of the measurement error. It also allows to eventually assign different errors to different observations. However, the "data error" matrix that enters the cost function needs to include the model error. Given that we can not assess such error component independently, we have scaled for each site the overall "data error" magnitude to obtain statistical consistency of the optimization process. We scaled the errors so that twice the value of the cost function at the minimum is close to the number of degrees of freedom (Chi-square statistic). Such scaling can thus be seen as a "data error" opti-

mization after a first guess based on measurement. With this approach the estimated uncertainties on the model parameter are thus coherent with the different hypothesis used in the optimization (Gaussian independent errors) and similarly for the model prediction errors.

"(2.4) Which gradient-based algorithm was used in this study? The reference is incomplete here. Gradient-based algorithms have the drawback of being locally convergent. That is, the algorithm might converge prematurely because it gets trapped in a local minimum of the cost function. This also means that it is difficult to judge whether indeed the global minimum was located or just a local, non-optimal solution was found. To gain confidence in the best solution found by these gradient-based algorithms, it is convenient to repeatedly start the optimization from various locations in the parameter space. Did the authors follow this procedure? Also, the expression "we converged to a minimum..." appears inappropriate. It is the algorithm rather that converges."

We have used the so-called BFGS iterative algorithm that follows the gradient of the cost function using an internal approximation of the Hessian. We were aware of the potential drawback of being trapped in a local minimum and we have thus performed several optimizations, starting with different parameter prior values. For all cases we obtain nearly the same set of optimized parameters. We are consequently confident that our results do not correspond to local optima. We will add some sentences in the text to inform on these tests.

"(2.5) For models that are nonlinear in their parameters (which is true for all but the first model used) it is not allowed to average over the parameters. The model output corresponding to the averaged parameters and the average of the model outputs obtained for each replicate will not be the same in this case. I suggest to include all individual measurements taken at a given site in the measurement vector $\sim$x and to fit one model per site only. This results in effective parameters for a given site, that is, parameters that take into account the variability between the various replicates."

We agree that with non linear models, it is difficult to average over the parameters. However, our objective was not to have a model parameterized to best fit the whole data set at each site, but an estimation of the stable C content including the spatial heterogeneity of that quantity, thus allowing for plot-specific optimized parameters. For this objective it is important to determine parameters for each plot as the texture, composition, and history of the soil might differ between plots. We would lose this information by optimizing one model per site instead of one model per plot. We thus use the different parameter sets as an ensemble of plausible estimates of the stable C content, given prior information and observations. For simplicity we did not report all parameter sets and choose to indicate only the mean values and the spreads. In fig. 1, the lines represent the average of the different simulations at plot-scale and not a simulation with average parameters. We will justify our choice in the text.

"(2.6) p. 4900, ll. 26-27: How did the authors test if a parameter was significantly different from zero? Please provide some explanation here."

The parameter was considered as significantly different from zero when zero was not included in the 95% confidence interval since the parameter distributions are assumed to be Gaussian.

"(2.7) How does the sum of the estimated pools compare to measured SOC? This comparison would help to judge whether the estimated pool sizes are realistic, and hence, the conclusions drawn from this model exercise are robust. If the discrepancy is substantially larger than the uncertainty in the SOC measurements, however, this would indicate that something is wrong. The model would be able to nicely fit the data, but for the wrong reasons. The estimated model parameters would be rather meaningless then."

There were no such discrepancies between modelled SOC and data. This is somehow illustrated on Figure 1 where we can see that the model output (representing the sum of the different C pool) fits the SOC measurements. We also checked that the residuals

are equally distributed and that no clear trend can be observed.

"(2.8) Wouldn't it be more reasonable to constrain the model by calculating the size of the most stable pool from the sum of the remaining (estimated) pools minus measured SOC? This approach was adopted, for example, by Paul et al. (Soil Sci. Soc. Am. J., 70:1023-1035, 2006). This would also supersede the assumption of a log-normal prior for the inert pool, as described in the last paragraph of the Methods section. Given the relative small amount of data that is typically available for calibration of the model and the limited amount of information about this pool contained in (some of) the data sets, this prior is likely to have a dominating effect on the posterior estimate. This might also explain why the estimate for this pool reported in this study is substantially larger than previous estimates of the "Inert Organic Matter" pool defined in RothC."

Yes, this could be a second elegant approach. However, we believe that it would not change substantially the results. We agree that the prior on the stable pool concentration can influence the posterior estimate for sites where this parameter is weakly constrained. However, for Grignon and obviously Versailles (truncated or not) where it is best constrained, the posterior value is much more in the order of magnitude of a Century passive pool than of a Roth-C inorganic pool (IOM). For the other sites, we consider that the uncertainties associated to this parameter is more important that its average value.

"(2.9) p. 4905, ll. 18-23: The paper by Lehmann et al. (Nat. Geosci., 1:832-825, 2008) might be worth citing in this context."

Ok, thanks.

"(2.10) p. 4907, ll. 14-22: The first part should be moved to the discussion section. It does not contain conclusions. The last sentence would better fit into the acknowledgements. It is also not a conclusion."

Ok, we may move this part as it is not properly speaking a conclusion to the work

conducted in the paper but we consider it is worth mentioning in a paper on long-term bare fallow.

“(2.11) Tab. 1: The abbreviation “EP” used for potential evaporation should be replaced with “PE” (to make it consistent with the table caption). Moreover, a reference would be useful here since there are several modifications of the Penman equation. What does the uncertainty ranges for initial and final C stock stand for?”

Ok, for PE. The equation comes from Penman H.L. (1948) Natural Evaporation from open water, bare soil and grass. Proceedings of the Royal Society of London, Series A, 193, 120-146. The uncertainties on C stock come from the uncertainties on SOC concentration.

“(2.12) Fig. 1: Why is the model uncertainty the standard deviation times 5? What confidence level does that stand for? Also, I suggest to show the model uncertainty as uncertainty envelopes instead of bars at the bottom of the figures. In my view, uncertainty envelops are a more intuitive representation of model uncertainty.”

We presented standard deviation times 5 for graphical reasons. It would have been better to present sd times 4 to have the 95% confidence interval. It will be done in a revised version. We did not use envelopes as they would have hidden a part of the data points.

Interactive comment on Biogeosciences Discuss., 7, 4887, 2010.