Geosci. Model Dev. Discuss., 7, C2275–C2280, 2014
www.geosci-model-dev-discuss.net/7/C2275/2014/

Geoscientific Model Development Discussions
Open Access

Interactive comment on "Global sensitivity analysis, probabilistic calibration, and predictive assessment for the Data Assimilation Linked Ecosystem Carbon model" *by* C. Safta et al.

Anonymous Referee #1

Received and published: 12 November 2014

Overall:

In this paper the authors work through an example workflow of calibrating and analyzing a simple model (DALEC). The objective of the paper is to develop the example workflow, not to learn about DALEC, and there are indeed many novel aspects to the workflow presented. That said, the idea of the workflow presented is not actually novel (i.e. starting with model sensitivity & uncertainty analysis, doing data assimilation, assessing change in information).

Unfortunately, there were a few substantial problems in the paper. First and most important is that model error was ignored during calibration. Second, prior information on

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

model parameters was also ignored, which leads to a misleading uncertainty analysis and potentially biologically implausible parameter estimates (as an aside, the plausibility of model posterior estimates is never discussed or compared to data). Third, lacking an estimate of model error, predictive distributions are done by propagating observation error into the forecasts, which is inappropriate. Finally, overall the paper was too long, though in many places I felt like I was being hit by a barrage of indices that were all Methods and Results without Discussion.

Details:

Pages 6895-6897: Too much background. There's been multiple recent reviews of data assimilation in ecosystem models that can be pointed to for readers that want all the nitty gritty, so instead you should focus on your message/context.

Page 6897, line 25: Unlike the last 2 pages, this comment is unsupported. Indeed, there have been a number of recent papers performing detailed sensitivity and uncertainty analyses specifically in the context of leading up to data assimilation. This is worth mentioning because it is important to note that while the details of the workflow the authors put forth is unique, the general workflow they are following is definitely not, and that's a part of the literature that DOES need to be discussed if you're going to claim that what you are going is novel.

Page 6898, line 27: This bit is very important and much more rare (indeed, I know of examples where teams have taken one approach or the other, but I'm not aware of a paper that compares the two explicitly), however in the context of this paragraph I had no idea what you were talking about. It was only when I got to the methods that it became clear that you were discussing the alternatives of assuming spin up to steady-state vs. including the IC in the assimilation. This needs to be explained and highlighted more in the intro.

Page 6898, L10-12: Extraneous. That you're developing this workflow to apply it to more complex models like CLM is relevant, but tweaks to the model can be left to the

GMDD

7, C2275–C2280, 2014

methods.

Pg 6899, L12: Great, but the reader has no idea what UQTk is – is it a project, a model, a conference, a piece of software? What does UQTk stand for? Given that the answer is that it's a piece of software, push this into the methods, be more explicit about it's use, and make the overarching workflow for performing this analysis public (even if the source code of the toolbox isn't). Having just that toolbox is insufficient to allow this analysis to be reproduced.

Pg 6901, L9: I fundamentally disagree. A sensitivity analysis (even a global one) connects PERTURBATIONS in the model inputs to perturbations in the model output. What you describe is an uncertainty analysis. The fundamental difference is that to perform an UA rather than a SA you need to know the uncertainties in the parameters. But you don't have those uncertainties, you just have arbitrarily assigned uniform ranges. Personally, I don't think Sobol' indices are an appropriate technique for SA – since they are variance-based their interpretation only really makes sense for UA.

Pg 6902, L15-20: The logic here is completely backwards. You state you "are only given prior information on parameter bounds", but it is you that is only giving yourself that information. If you wanted to give yourself more information you easily could since Mat's made the model and has been working with it for a decade. More to the point, for almost every parameter in DALEC there is more prior information available in the literature about what these biologically meaningful parameters are than is captured by a broad, uninformative Unif distribution. As others have argued (e.g. LeBauer et al 2013), using uninformative priors will lead to incorrect assessments of what parameters are driving model uncertainties and increases the chances of data assimilation selecting biologically implausible parameters. Imagine that you were running DALEC at Harvard outside of the data assimilation and prior to observing the NEE data– are you really saying that sampling randomly from these uniform distributions is the best you could do and all you really know?

GMDD

7, C2275–C2280, 2014

Pg 6903, L6: This info really needs to make it into the legends for these figures – I looked at the figures first and was scratching my head since neither grey nor white was in the legend or caption

Pg 6905, L16: The assumption that NEE data are independent is not plausible and will result in considerably overconfident posterior distributions

Pg 6905, L18: The decision to neglect model error is completely inappropriate, will lead to incorrect posterior distributions, and renders your predictive intervals meaningless. This assumption implies that you believe that your process model is perfect and the only reason for deviations from observations is due to observation error, which is untenable. The inclusion of model error as a fit parameter is neither conceptually difficult or computationally costly – updating the model error doesn't require model runs and the prior could easily be chosen to be conjugate to your Normal likelihood, allowing the update to be done using Gibbs Sampling.

L6906, L14: This is awesome. However, you don't state what your priors are. As I note above in the discussion of informed priors, you should actually be able to construct fairly informative priors for most of your C pools since the biometric data for Harvard Forest is pretty good and all public through the LTER.

P6906, L20: Could you use more meaningful acronyms?? Also, these acronyms are not defined in the tables and figures that include them

P6907 L1: "D is replaced by a random vector of NEE observations" – This bit doesn't make sense to me, why would you be randomizing your NEE observations such that their was no pairing between observations and model predictions? Based on my reading there's no requirement for this randomization in Fisher Information approach.

P6907 L12-13: Given that Sigma is just a constant with respect to the different model parameters (thetas), this standardization by range (which isn't a standard component of Fisher's Information) appears to be why & how you can interpret this in terms of

GMDD

7, C2275–C2280, 2014

uncertainties. Specifically, because of the special case of the Unif, where the variance is directly proportional to the range, this now ends up [to a constant] equivalent to the Taylor Series variance approximation, dm/dtheta1*dm/dtheta2*Cov(theta1,theta2), which has been used elsewhere as a means of variance decomposition (e.g. LeBauer et al 2013). Also, as LeBauer notes, if the variance of the thetas is large this ends up a poor approximation, which is something to consider here. Given this equivalence to variance decomposition, it is no longer "interesting to note" (Line 16) that the results of the FIM are very similar to the Sobol based GSA as you would expect them to be identical (to a constant) if the model was linear. Therefore what's surprising is that the linearized, derivative-based analysis gives largely the same answer as the global sensitivity analysis. This observation is worth noting more explicitly

Pg 6909, L8: high-lighted

P6910, L9-12: This is a really neat trick.

P6911, L5-6: This is also a new and useful contribution (at the least, it's an approach that I haven't seen before in the ecosystem modeling literature)

P6912, L26-28: Neat. I've never seen a piecewise posterior before, but this is a great example of how input accuracy and "if" statements can impact posterior inference.

P6915, L12-14: I don't think this analysis makes sense. First, observation error should not be propagated into a forecast. Model error, which wasn't estimated, DOES need to be propagated into the predictive distributions. Given that, I'd strongly recommend that the authors restrict the current analysis to looking just at the model's credible interval (parameter uncertainty) not predictive intervals (unless they end up quantifying model error as well)

P6916, L1: "computed with KDE" is unclear

P 6916, L14: CRPS is also new to me. I think this is cool, but in the end there's not much interpretation / discussion of the results. Needs to be a more clear set of

GMDD
7, C2275–C2280, 2014

take-home messages in this section

P6917, L14: Likewise, CRPSS is new to me, and while it is interesting, like with CRPS it isn't explained or interpreted enough. If the only take home message is that there was 40% improvement, then you should make this section much shorter in order to get to that point more quickly.

Figure 2: what is NPP2 and why does leaf biomass not affect LAI?

Figure 3-6: please put the months in order! Start with January, end with December rather than starting with November.

Figure 4: Why is DALEC so sensitive to leaf fall in December, which is 2 months after leaf fall occurs?

Interactive comment on Geosci. Model Dev. Discuss., 7, 6893, 2014.