Biogeosciences Discuss., 6, C546–C549, 2009
www.biogeosciences-discuss.net/6/C546/2009/

Biogeosciences Discussions

Interactive comment on "Improving land surface models with FLUXNET data" *by* M. Williams et al.

Anonymous Referee #2

Received and published: 28 May 2009

This manuscript provides an extensive summary of the role of model-data fusion techniques in improving terrestrial biogeochemical modeling using eddy covariance observation data made available through the FLUXNET network. It also presents a series of examples from the literature and the author's own studies to show how these technqiues can improve model performance. The Figures in some cases are not suitable for final publication and need to re-drawn due to illegible axes titles and legends. The paper repeats some points and the text needs to be tightened up in places, but otherwise it is suitable for publication pending responses to the specific comments below.

Specific comments

Abstract P2787. L7. 'fusing' is a colloquialism and should be put here in single quotes. Ideally it should be defined as well.

P 2787. L17. Note that fusing multiple independent data provide a potential means to

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

limit equifinality. This may not always be so if data sets are not orthogonal.

Introduction P2788. L24. "...has led to a large increase..."

P2789. L7. "data exploitation" is ambiguous. What does it mean to exploit data in this context?

P2789. L27. "...and modeler, and thus MDF." Is this sentence complete or missing part of the last phrase.

P2790. L1-9. Here parameter estimation and state estimation are lumped together. However, I think this is confusing to the reader. They should strictly be mentioned separately as there are different issues to consider if the problem is a state estimation problem or a parameter estimation problem.

P2790. L11. Delete "at hand."

P2793. L6. This is true, but it is also true that the processes may be able to be identified but cannot be parameterized.

P2793. L27-28. Careful consideration of data limitations is important, but it is also more than this. The previous par refers to fast and slow processes. It is important that the timescales of model processes and their observations must overlap otherwise no useful information is conveyed to the model from those data by the assimilation scheme.

P2794. L1-2. It is important to note here that systematic errors will be identified by the multiple-constraints approach via model-data mismatch for one or more of the objective functions. This is a strength of the method. On the other hand, large random errors in observations, while increasing uncertainties of the estimated parameters, decrease the usefulness of the model-data fusion approach because the analysis errors are not much reduced from the original background (prior) errors.

P2794. L19. "LaPlace..." This is unclear and needs re-wording.

BGD

6, C546–C549, 2009

P2794. L23. The flux errors quoted here should be put into context. Perhaps cite their associated coefficients of variation? Same for P2795. L3.

P2795. L9. "...bias annual estimates of net sequestration upwards." Be specific from what value to what other value. Or what is the magnitude (absolute and relative) of the bias?

P2795. L20-25. Specifically what is the important distinction between measured and filled data here?

P2798. L1-25. I find Fig 5 difficult to read. Axes and legends are almost illegible making it difficult to see what the author's are referring to. I understand the importance of wavelet decomposition for examining model performance across multiple time scales, but I find the text about OWT and the wavelet half-plane diagram (Fig 5 e & f) difficult to relate to the rest of the text. This could be removed with little loss of message from the manuscript.

P2800. L15-18. "...important information for subsequent steps of model structural development." Why? This statement needs explanation.

P2800. L27. "filter techniques..." this is not defined and not linked with point #3 on P.2801. Thus it will be confusing to readers not familiar with Kalman Filter. Also, what other techniques are used and does model uncertainty need to be defined for these? More clarity is needed here. In addition, the text associated with point #3 is likely to be unclear for many readers. There are multiple concepts in here (batch versus sequential methods, parameter only versus parameter & state estimation, model and observation uncertainties) offered without explanation. The class needs to be stated in the dot point and the detail provided in the text.

P2802. L1-23. This section is a mini tutorial on least squares/Kalman Filter estimation. I am not convinced that it adds a great deal to the section on the technical implementation of model-data fusion over what can be picked up in a text book. Perhaps this

BGD

6, C546–C549, 2009

section should be placed in a text box as a primer on estimation?

P2803. L12-17. "infinitely narrow priors" is the same as saying "holding these parameters constant". There are ways of assessing the maximum number of parameters that can be estimated given the information content of observations and model: e.g. the condition number of the Hessian indicates whether too many parameters are being fitted.

P2804. L2-9. The issue of model errors in the Kalman Filter has been remarked on previously and so this section partly repeats. The text should be made consistent throughout the MS. L7-9 will be unclear to those not familiar with KF and exploration of posterior distributions. What is the "careful inspection" that the user does?

P2804. L23-29. Sentence repeated.

P2804. L28-P2805. L1. What does it mean to "relax" the model? How is this done? In Figure 7 what are the codes down the left hand side (Presumably flux tower site codes)?

P2805. L7-10. What does "independently" (in quotes) mean here? What does it mean that "compromises" are required from data and model? I think you mean that trade-offs among objectives of a multiple constraints problem is needed so that no single objective is exactly satisfied?

P2806. L6-8. Its not just more information that is required , but new data possessing information that is orthogonal to the information already contained within the existing data. This will lead to reduced posterior uncertainties (ie greater variance explained).

Interactive comment on Biogeosciences Discuss., 6, 2785, 2009.