Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2018-303-RC3, 2018

Biogeosciences Discussions
Open Access
EGU

Interactive comment on "Optimal Inverse Estimation of Ecosystem Parameters from Observations of Carbon and Energy Fluxes" *by* Debsunder Dutta et al.

P. Rayner (Referee)

prayner@unimelb.edu.au

Received and published: 11 September 2018

This paper demonstrates a method for assimilating site-level flux observations into a terrestrial biosphere model. Its novelty lies in breaking the assimilation into short windows to capture high-frequency variations in the parameters it estimates. given the variety of journals within the Copernicus family, I wonder whether this article is better suited to GMD than BG (see comments below) but this is mainly a question for the editor. the paper is also clearly written, verging on the tutorial at times.

I have one significant concern with the paper and one general request for more analysis. My concern is the analysis of the results. This is quite thin. The only commentary

Printer-friendly version

Discussion paper

I can see on the results in the discussion section is: "There is strong evidence from measurements that under normal conditions LAI and photosynthetic parameters have seasonal variability [Wang et al., 2008; Wilson and Bal- docchi, 2000; Wilson et al., 2000] which correlate with observations of energy fluxes. Our model inversion results are in alignment and agree well with these observations." this seems quite a poor scientific return from a difficult and well-executed piece of work. I would recommend particularly using the posterior simulation to look at some other observables. Do you do a better job matching the high SIF values over the corn site? If so, why, e.g.which parameter, Vcmax or LAI is mainly responsible? What temporal resolution of the parameters is necessary to capture the important variations? I suspect these questions only scratch the surface. I stress that this is potentially a good paper. What it does it does well but I believe it needs more scientific content before publication. If the authors wish to maintain it near its present form I believe it is better suited as a demonstration of a new methodology and hence to GMD.

My request is to delve a little deeper into why the system works better at some places than others. I note there seems less analysis of the Niwot Ridge results which were, in general, also less successful (lower correlation for example). Remember that a less successful assimilation is *not* a failure but rather a useful probe into model performance. It says definitively "we have a problem here and it isn't the choice of parameters". This is even clearer in this case where the parameters are allowed to vary in time.

Minor comments P14 In fact the Jacobian doesn't quite show the problem is non-linear, it could be that all the variation is a result of different forcing.

P16L3 The choice of observational error is quite important in DA, hoefuly this is checked later.

P16L10 I doubt the size of observational vector has much impact on computational efficiency, can you comment why it would?

P16L20 The choice of time resolution is also important and yours seems very short. This is likely to lead to parameters which can vary fairly rapidly in time but which are also quite uncertain as they are constrained by fewer observations. Hopefully you can comment on whether parameters change significantly, i.e outside their uncertainty limits.

P16 Eq. 12, this should have a term from the prior included I think. Unless there's no prior.

P19L7 "reasonable and realistic" is a little vague, perhaps some references would help

P20L10 be careful about describing correlations as describing how parameters move since these are uncertainty not signal correlations. the sentence above makes it clear you understand this difference but many of your readers will be less clear.

P20L14 but here you do confuse signal and error, this correlation does NOT indicate they are changing in sync

P20 in general you seem to be quoting rˆ2 but claim this can be negative. You probably mean r.

P21 I'm not sure that the figures showing your algorithm works are necessary, especially in a journal like biogeosciences where you should focus more on the science and less on the algorithm.

P22L3 as noted earlier the diurnality is not a measure of nonlinearity

P22L10 don't quote improved correlation as a measure of fit, you could have a great correlation and terrible performance if, for example, diurnal variations had great phase and terrible amplitudes. rms is a better though not perfect statistic.

P25 See earlier comments on signal and error correlation.

P25 can you explain further why a strong negative correlation means you need to optimise both, the step from "you can't see them separately" to "you must do both of

Printer-friendly version

Discussion paper

them" isn't so clear to me

P25 I hope you go on to compare the performance at the two sites, one of them seems much harder than the other.

P27 I'm betting you originally tried to fit LAI at NWR and couldn't. That's not a failure, it's interesting information so is probably worth discussing. You're only fitting in 3 day windows so neither site really knows about the evolution of LAI from one window to the next so why does one work well and the other not, provided I'm guessing correctly.

P29L13 This site analysis doesn't seem as well developed as the others, e.g. quality of fit etc.

P29L30 do you mean changes in the temperature dependencies or more simply that there *is* a temperature dependence?

P30L20 In what sense is the approach "stepwise"? This term was previously used by Bacour et al. (2015), doi:10.1002/2015JG002966) to describe optimising for one observable then using its posterior parameters as priors for the next observable. They would describe your method as "all at once", what do *you* mean by stepwise?

This paper demonstrates a method for assimilating site-level flux observations into a terrestrial biosphere model. Its novelty lies in breaking the assimilation into short windows to capture high-frequency variations in the parameters it estimates. given the variety of journals within the Copernicus family, I wonder whether this article is better suited to GMD than BG (see comments below) but this is mainly a question for the editor. the paper is also clearly written, verging on the tutorial at times.

I have one significant concern with the paper and one general request for more analysis. My concern is the analysis of the results. This is quite thin. The only commentary I can see on the results in the discussion section is: "There is strong evidence from measurements that under normal conditions LAI and photosynthetic parameters have seasonal variability [Wang et al., 2008; Wilson and Bal- docchi, 2000; Wilson et al.,

Printer-friendly version

Discussion paper

2000] which correlate with observations of energy fluxes. Our model inversion results are in alignment and agree well with these observations." this seems quite a poor scientific return from a difficult and well-executed piece of work. I would recommend particularly using the posterior simulation to look at some other observables. Do you do a better job matching the high SIF values over the corn site? If so, why, e.g.which parameter, Vcmax or LAI is mainly responsible? What temporal resolution of the parameters is necessary to capture the important variations? I suspect these questions only scratch the surface. I stress that this is potentially a good paper. What it does it does well but I believe it needs more scientific content before publication. If the authors wish to maintain it near its present form I believe it is better suited as a demonstration of a new methodology and hence to GMD.

My request is to delve a little deeper into why the system works better at some places than others. I note there seems less analysis of the Niwot Ridge results which were, in general, also less successful (lower correlation for example). Remember that a less successful assimilation is *not* a failure but rather a useful probe into model performance. It says definitively "we have a problem here and it isn't the choice of parameters". This is even clearer in this case where the parameters are allowed to vary in time.

Minor comments P14 In fact the Jacobian doesn't quite show the problem is non-linear, it could be that all the variation is a result of different forcing.

P16L3 The choice of observational error is quite important in DA, hoefuly this is checked later.

P16L10 I doubt the size of observational vector has much impact on computational efficiency, can you comment why it would?

P16L20 The choice of time resolution is also important and yours seems very short. This is likely to lead to parameters which can vary fairly rapidly in time but which are also quite uncertain as they are constrained by fewer observations. Hopefully you

Printer-friendly version

Discussion paper

can comment on whether parameters change significantly, i.e outside their uncertainty limits.

P16 Eq. 12, this should have a term from the prior included I think. Unless there's no prior.

P19L7 "reasonable and realistic" is a little vague, perhaps some references would help

P20L10 be careful about describing correlations as describing how parameters move since these are uncertainty not signal correlations. the sentence above makes it clear you understand this difference but many of your readers will be less clear.

P20L14 but here you do confuse signal and error, this correlation does NOT indicate they are changing in sync

P20 in general you seem to be quoting rˆ2 but claim this can be negative. You probably mean r.

P21 I'm not sure that the figures showing your algorithm works are necessary, especially in a journal like biogeosciences where you should focus more on the science and less on the algorithm.

P22L3 as noted earlier the diurnality is not a measure of nonlinearity

P22L10 don't quote improved correlation as a measure of fit, you could have a great correlation and terrible performance if, for example, diurnal variations had great phase and terrible amplitudes. rms is a better though not perfect statistic.

P25 See earlier comments on signal and error correlation.

P25 can you explain further why a strong negative correlation means you need to optimise both, the step from "you can't see them separately" to "you must do both of them" isn't so clear to me

P25 I hope you go on to compare the performance at the two sites, one of them seems

Printer-friendly version

Discussion paper

much harder than the other.

P27 I'm betting you originally tried to fit LAI at NWR and couldn't. That's not a failure, it's interesting information so is probably worth discussing. You're only fitting in 3 day windows so neither site really knows about the evolution of LAI from one window to the next so why does one work well and the other not, provided I'm guessing correctly.

P29L13 This site analysis doesn't seem as well developed as the others, e.g. quality of fit etc.

P29L30 do you mean changes in the temperature dependencies or more simply that there *is* a temperature dependence?

P30L20 In what sense is the approach "stepwise"? This term was previously used by Bacour et al. (2016, doi:xxx) to describe optimising for one observable then using its posterior parameters as priors for the next observable. They would describe your method as "all at once", what do *you* mean by stepwise?

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2018-303, 2018.

Printer-friendly version

Discussion paper