Biogeosciences Discuss., 6, C3745–C3749, 2009
www.biogeosciences-discuss.net/6/C3745/2009/

Biogeosciences Discussions

Interactive comment on "Reducing impacts of systematic errors in the observation data on inversing ecosystem model parameters using different normalization methods" *by* L. Zhang et al.

Anonymous Referee #2

Received and published: 30 December 2009

The study of Zhang et al. addresses the topic of systematic errors in observation data used for model optimization. This topic is important as in previous studies the systematic error was usually simply assumed to be zero and not mentioned, although it is obvious that model optimization results must be biased if systematic errors are in the used data. Zhang et al. use synthetic LAI data, including various types of systematic errors, to optimize vcmax in a process-based ecosystem model. To reduce the influence of the systematic error they apply three different normalization methods and find that the z-score normalization performs best in retrieving the true vcmax value. They conclude that the z-score normalization should be applied for parameter estimation, especially when potential systematic errors are unknown.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

General comments:

Although the results show a great improvement in the parameter estimation, there are limitations of the method that need to be addressed, not only in the discussion but also in the general conclusion:

1) Is the artificial experiment of the study really representative for real world parameter estimation problems? On the one hand the errors in the data could be more complex, on the other hand usually more than 1 or 2 parameters are optimized. A comparison of the model output that is used as “true” values with observed LAI could be insightful to understand whether the synthetic error mimics the real world error well.

2) The information content of the observations is reduced by the normalization, e.g. the information about the absolute values is lost and only the information about the relative variability of the variable remains. This must have an effect on the optimization, probably on the number of parameters that can be constrained, e.g. equifinality could occur for a smaller number of parameters included in the optimization. This is not addressed in the manuscript, but is a major limitation of the proposed approach. For instance if the model was only y=ax+b, where x is a driver, y the model output that we want to fit to observations, then neither the parameter a nor b could be estimated applying a z-score normalization. Thus it depends on the model structure, whether the approach is applicable or not. An additional analysis could be to compare the sensitivity of the LAI model output to variations of different parameters with and without z-score normalization. For the linear model y=ax+b, varying a and b does not change the normalized output. A comparison of the correlation structure between the parameters when using the normalized and not normalized cost function could indicate whether normalization increases equifinality issues.

These limitations need to be addressed, either by additional analysis or by a more differentiated conclusion.

It should be further discussed why the methods work and why not. The reason why the z-score normalization works for the linear errors is that for a linear model y=ax+b that is assumed for the synthetic errors the normalized output does not change. Other error structures may not have this property.

Specific comments:

1. P.10448, l. 24-25: the normalisation is not only applied to the observations, but also to the model ouput, please rephrase.

2. P. 10448 l. 26: remove “especially”, the normalization should be applied only then, if the errors are known they should be removed.

3. P. 10449,l. 18-20: Maybe the sentence is incomplete? The variations at one site could be used as uncertainty of vcmax. Please rephrase.

4. P. 10451,l.22-26: you use synthetic data, the exact distribution of this vegetation type does not matter. Please remove this sentence. If the coupling between LAI and vcmax is different for other vegetation types it would be helpful to include them to support the general conclusion.

5. P. 10453: It would be good to have a description of the phenology module here to understand what causes the spatial variability of LAI and how it is coupled to the photosynthesis and vcmax.

6. P. 10454, l. 15: do you use the LAI of one specific year? It should be clear from the beginning that you use August LAI and that the seasonality of LAI is not included.

7. P. 10456, l. 28: what are the true observations? This chapter should be extended and parameters also estimated for different types of errors.

8. P. 10459, l.5: in the figure it looks like a1 is hardly constrained.

9. P. 10459, l. 17: direction means positive or negative? please rephrase the sentence it is difficult to understand. Explain why the z-score transformation works fine for linear errors.

BGD

6, C3745–C3749, 2009

10. P. 10459, eq. 7: i can't follow the transformation

11. P. 10460,l. 14: normalization can't change the spatial distribution.

12. P.10461, l.2-4, does this mean most sensor have only linear errors? Consider that usually the "observation" is not a pure measurement it usually involves models, for instance to derive the LAI, or in case of eddy covariance data complicated corrections need to be applied. Nonlinear errors can arise in this step.

13. P.10461, l.5: you showed only that the method can be applied to observations with a linear error model, you cannot conclude that it can be applied to "any other observations"

14. P. 10461, l. 24-26: you did not show this for an increased numer of parameters, not even for the example of 2 parameters. If you would show that the uncertainty of the 2 parameters do not increase using the normalization this would support this statement, but still equifinality problems could arise for a higher number of parameters.

15. P.10468 please remove the c unequal 0, it is zero in most cases.

16. P. 10469 please add, that it is LAI in August, here a comparison of the "true" values with observations would be interesting to see whether the assumed errors are realistic.

17. P. 10471 please add the not normalized cost function.

18. Fig 5a: why does the uncertainty decrease for min-max and max normalisation? Please, add without normalization

Technical comments:

1. P.10448, l.2: "modelling carbon cycle", please insert "the"

2. P. 10448,l. 21-24: I don't understand the sentence.

3. P. 10450, l. 26: please change: "How systematic..." into "How do systematic..."

4. P.10450, l. 27/28: please change: "Do the potential impact of systematic errors on

parameter estimation can be . . .” into “Can the potential impact of systematic errors on parameter estimation be . . .”

5. P. 10451, l.1: please change “Whether the three normalization methods is effective. . .” into “Are the three normalization methods effective. . .”

6. P.10453, l. 3: add “or simulation”

7. P. 10470 please explain a,b,c and add uncertainties, if no random error is added, bootstrapping can be used to derive the parameter uncertainty. It will be interesting to see, that for the systematic errors the true values are not within the uncertainty of the estimate.

8. P. 10471: what is the dotted line?

Interactive comment on Biogeosciences Discuss., 6, 10447, 2009.