Biogeosciences Discuss., 7, C813–C816, 2010
www.biogeosciences-discuss.net/7/C813/2010/

Biogeosciences Discussions

Interactive comment on "A geostatistical synthesis study of factors affecting gross primary productivity in various ecosystems of North America" *by* V. Yadav et al.

A. Desai (Referee)

desai@aos.wisc.edu

Received and published: 30 April 2010

THe authors apply a novel geostatistical approach to explain variability of flux-tower observed primary productivity observations against environmental variables, using a combined maximum likelihood regression and Bayesian model selection algorithm. They find the explanation of variance of GPP is strongly mediated by timescale and by site. This is not a surprising finding, but the purpose of this paper is obviously to introduce a technique and discuss the nature of relationships and model selection. Of all the geostatistical approach papers I have read from this group, this is one of the clearest explanations of the technique and I applaud the author for a thorough discus-

sion of the nature of this technique. The model selection and "best" model uncertainty technique is a significant advance for flux tower analysis. As such, I feel that the paper could be published with relatively minor revision. That said, there are at least a few issues that I think are worth discussing in more detail.

- Flux towers observe NEE not GPP. While the authors use a consistent algorithm to infer GPP using the MDS method, there are assumptions made by the MDS method that is likely to modify the estimates of GPP. While MDS does not use functional forms as other environmental variable regression methods (which would certainly complicate interpretation of results here), it does assume that the NEE data can be partitioned by a variety of environmental factors. My worry is that these assumptions bias your results of environmental controls, perhaps over or under estimating an effect. I recognize that the authors are not flux tower operators, but I think at least a discussion of the role of flux partitioning algorithm on the robustness of the relationships is worth discussion. The paper does mention Desai et al (2008), which is more about differences in GPP estimates by method and the importance of using a consistent method when estimating GPP across sites; that paper did not directly address this issue of regression environmental variables by method.

- My other concern, and perhaps this comes from a lack of understanding, is the insistence on using a linear additive model as the only type for model selection. I understand the nature of model selection process somewhat forces this constraint here, since this is in essence a sophisticated multiple linear regression, but we know that factors that influence GPP are not additive, but rather multiplicative (limiting factors). The Montieth light use efficiency style equations are GPP = LUE * f(temperatre) * f(light) * f(etc...), and here instead the authors fit GPP = f(temperature) + f(light) + ... I'm not convinced the two are interchangeable, but I think it might be worth convincing (dense?) people like me that there is some level of equivalence between the two.

- While discussion of differences between sites are discussed, I'm perplexed why the authors didn't go to the next step - namely trying to jointly fit sets of sites with the same

BGD

7, C813–C816, 2010

approach. It would seem to be a natural thing to do and allow one to estimate what are the larger scale controls on GPP. In one sense, this comes from my feeling that if one works hard enough, there is always some variable somewhere that "looks" like GPP. Why not include CO2 concentration or tower-observed fPAR, for example?

- Similar to my comment about flux tower, there is some importance to discussion of "know thy data" that is missing in the discussion of the environmental variables. Granted, this is reasonable given the nature of the analysis here, but I'm curious about the gap-filling of environmental data and the strong correlations of LAI, EVI, NDVI, as they are all derived from the same satellite, using a similar combination of spectral bands. For example, I don't really understand a mechanistic explanation for why one site may be better explained by EVI and another by NDVI. I though EVI was NDVI on steroids? I would suggest discussing a bit more on the provenance and inherent self-correlation among the data, and also what role does error in data play - how can it be incorporated into the maximum likelihood approach. Does the model selection and regression account for self-correlation of data? I think it does, but I suggest the authors make the case in the paper.

Minor points: To truly make this geostatistical approach applicable to the research community, I suggest the authors consider including pseudo-code or sample scripts as a supplement. While well written, it is not trivial to replicate/implement these equations from scratch.

- Wavelet analysis is one approach that has been used. I believe Dennis Baldocchi also has some earlier papers on Fourier-style analysis of flux tower observations. I suggest digging those up and citing.

- P 1458, line 15 - The authors mention desirability to reduce the number of candidate models and describes a nice method for doing so. I'm not sure why precisely this was needed for this study, given the number of candidate models and sites. Can you discuss the computational needs/time of this approach that requires this pruning, or is

BGD

7, C813–C816, 2010

there another reason?

- The writing, captions, and conclusions drawn are well written, so I don't feel a strong need to provide more minor comments.

Interactive comment on Biogeosciences Discuss., 7, 1445, 2010.