Biogeosciences Discuss., 12, C5507–C5512, 2015
www.biogeosciences-discuss.net/12/C5507/2015/

Biogeosciences
Discussions
Open Access

Interactive comment on "An objective prior error quantification for regional atmospheric inverse applications" *by* P. Kountouris et al.

A. Michalak (Referee)

michalak@stanford.edu

Received and published: 18 September 2015

The manuscript "An objective prior error quantification for regional atmospheric inverse applications" aims to diagnose the spatial and temporal correlation scales associated with the errors in terrestrial biosphere model estimates of net ecosystem exchange. This is an important question because such models are often used as prior information in atmospheric inversion studies, and correctly specifying the statistical joint distribution of their errors is necessary for robust estimation of carbon fluxes and their associated uncertainties.

The strategy used in this study is to characterize the scales of correlation in the difference between model estimates and both tower-based and airborne flux measurements, as a measure of prior errors. The study also explores the use of the difference between predictions by multiple models as a surrogate field for characterizing the statistics of model errors. The study builds on Chevallier et al. (2006, 2012), where the analysis was based only on tower-based flux measurements.

The paper is clearly written, and reasonably easy to follow.

MAJOR CONCERN:

My largest concern with the study is that the spatial resolution / support of the flux observations is substantially finer than the spatial resolution of the model simulations, making flux values at these disparate scales fundamentally incompatible. The authors acknowledge as much in p. 9397 lines 11-13 "While typical inversion systems have a resolution ranging from tens of kilometers up to several degrees (hundreds of km), the spatial representativity of the flux observations is typically around a kilometer." In the Chevallier et al. studies that the authors cite, the analysis of errors was conducted by comparing km-scale flux observations with "a site-scale configuration of the ORCHIDEE model," thereby leading to compatible spatial scales. The resulting error statistics were then upscaled to be representative of the scales estimated by typical inversions.

In the current manuscript, I saw no discussion at all of the scale mismatch issue, despite reading the manuscript multiple times. The scales in this work are as follows:

- Tower-based flux observations: ∼1km x 1km

- Airborne flux observations: 10km spatial windows, but no indication of the "width" of the window (p. 9402 line 6), i.e. 10km x ?km.

- VPRM model: 1km x 1km, and 10km x 10km

- ORCHIDEE model: 0.5o x 0.5o (∼50km x 50km)

- 5PM model: 0.25o x 0.25o (∼25km x 25km)

My concern is that these are all fundamentally incompatible scales, and differencing

BGD

12, C5507–C5512, 2015

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

flux observations at fine scales from flux estimates at substantially coarser scales not representative of errors at either scale. This fundamental concern throws all the study's conclusions into question.

The only analysis for which this concern is not a problem is when differencing estimates from one model from estimates from another model (Section 4), assuming that the models are aggregated to a common resolution first. This approach has in fact already been used previously to estimate prior error statistics (e.g. Chatterjee et al. 2013). For the other analyses, there is no easy way around the scale mismatch issue.

For example, the authors note with surprise that the correlation lengths derived from the model-to-model analysis are substantially longer than those from the model-to-observations analyses (p.9410 line 26 – p. 9411 line 7). But, this is not at all surprising, because here coarse-scale estimates are being compared, and correlation lengths almost always increase with scale. The authors try to "remediate" this by adding unrealistically large random noise to the model predictions (p. 9415 lines 1-20), but this is not in reality a problem to be fixed. Instead, they are simply looking at fluxes that are at a fundamentally different scale than those in the observations.

The authors conclude (p. 9415 lines 18-20) that "This would suggest that for inversion studies targeting scales much larger than the eddy covariance footprint scale, the statistical properties of the prior error should be derived from the model-model comparisons." This is certainly valid from a scale perspective, but this conclusion is contingent on the model-to-model differences being a good surrogate for actual model errors. Although this is tempting and has been done before (Chatterjee et al. 2013), qualitatively I would expect that the statistics of model-to-model differences may be highly sensitive to the choice of the two specific models used. Although this was not clearly the case for the three models examined here, I would refer the authors to Figure 1 in Huntzinger et al. (2013), where the degree of spatial variability (i.e. variance and correlation-length) is very clearly model-dependent. One would therefore conclude that the statistics of model-to-model differences would then also vary as a function of which models were

BGD

12, C5507–C5512, 2015

used in the analysis if a broader variety of models had been examined. See also for example Figure 2 in Huntzinger et al. (2011).

OTHER COMMENTS:

For all of the analysis, it would be important to more explicitly discuss the time scales for which the analyses are conducted, and emphasize that the error statistics computed therein are therefore only valid for that same (i.e. daily) temporal resolution. Both the spatial and temporal correlation lengths will be affected by the temporal resolution of the analyzed data.

Throughout the manuscript, the terms "correlation length" / "correlation time" (approximately 3*tau and 3*d in the authors' notation in eqns. 3 and 4) and the terms "e-folding time" (tau) and "e-folding correlation length" (d) and their variants are used, but due to the number of variations, it is not always clear when the authors are referring to 3*tau vs. tau, and to 3*d vs. d. This should be made completely clear throughout to avoid confusion. Please also pay close attention to this when comparing your numbers to those from earlier studies.

For the airborne analysis, the authors find correlation lengths of approximately 39 days (3 * e-folding time of 13 days, page 9408 line 22). Given that there are only 36 days of data, correlation lengths of much beyond $\sim$18 days (half the maximum separation distance) cannot be reliably identified. This should, at a minimum, be discussed.

In terms of the overall correlation lags, the authors need to make a fundamental choice as to whether they are trying to represent errors at synoptic scales, or errors at seasonal scales. While the numbers that come out of their analysis represent errors at the seasonal-scale, it is important to note that this means that they are assuming that errors at the synoptic scale are very highly correlated. This may not be a valid assumption. Although I understand how these numbers come out of the analysis as it has been designed, some thought should be given to whether these are indeed the scales that are relevant to whatever atmospheric inversions the authors have in mind as an ultimate application, as alluded to in p. 9418 lines 3-8.

Although the paper is very clearly written, there are a lot of details, and a lot of numbers reported. With this in mind, as I was reading through Section 7, it was difficult to easily glean the main take-away messages from the analysis. The author should consider making their main conclusions clearer.

ADDITIONAL LINE BY LINE COMMENTS:

p. 9396 lines 11-12 This statement is not entirely correct. Objective approaches were proposed earlier by Michalak et al. (2004, 2005), and have been applied in a number of studies since. The authors distinguish the Michalak et al. (2004) study as applying a “geostatistical” approach, but fundamentally both inversion approaches rely on characterizing the statistical characteristics of prior errors. I note that the Michalak et al. (2005) study was also for a classical Bayesian approach.

p. 9402 lines 5-7: I disagree with this statement. Even if the aircraft observations were “grouped” into 10km segments, this still does not match the VPRM grid, as the airborne segments are not representative of a 10km “width,” just “length” along the flight path.

p. 9403 eqn. 3 and associated text: A nugget parameter would typically be defined as one minus alpha in the notation used by the authors, as it represents the portion of the variability that is not spatially (or temporally) correlated.

p. 9406 line 13-18: I wonder whether the better correlations at the site scale are simply due to the fact that the models and towers agree as to the overall seasonality of the fluxes. A more representative analysis might be to calculate the correlations after removing an average seasonality.

Please note that as a matter of principle, I try to avoid disproportionately citing papers from my own research group in reviews, and I also prefer to remain anonymous as a reviewer. In this particular case, however, I have cited multiple papers from my own group as I felt that they were genuinely helpful in explaining some of the points that I

BGD

12, C5507–C5512, 2015

Interactive Comment

was trying to bring across. The authors should not interpret this as an indication that I am insisting on citation of any or all of these in a revised manuscript. In order not to hide behind anonymity given the specific perspectives that I have presented above, I am also opting to sign this review.

Best Regards,

Anna M. Michalak

REFERENCES CITED (I am not including those already cited in the manuscript)

Chatterjee, A., R. J. Engelen, S. R. Kawa, C. Sweeney, and A. M. Michalak (2013), Background error covariance estimation for atmospheric CO2 data assimilation, J. Geophys. Res. Atmos., 118, 10,140–10,154, doi:10.1002/jgrd.50654.

Huntzinger, D.N., S.M. Gourdji, K.L. Mueller, A.M. Michalak (2011) "A systematic approach for comparing modeled biospheric carbon fluxes across regional scales",Biogeosciences, 8 (6), 1579-1593, doi:10.5194/bg-8-1579-2011.

Huntzinger, D.N., W.M Post, Y. Wei, A.M. Michalak, T.O. West, A.R. Jacobson, I.T. Baker, J.M. Chen, K.J. Davis, D.J. Hayes, F.M. Hoffman, A.K. Jain, S. Liu, A.D. McGuire, R.P. Neilson, B. Poulter, H.Q. Tian, P. Thornton, E. Tomelleril, N. Viovy, J. Xiao, N. Zeng, M. Zhao, and R. Cook (2012) "North American Carbon Program (NACP) Regional Interim Synthesis: Terrestrial Biospheric Model Intercomparison", Ecological Modelling, 232, 144-157, doi:10.1016/j.ecolmodel.2012.02.004.

Michalak, A.M., A. Hirsch, L. Bruhwiler, K.R. Gurney, W. Peters, and P.P. Tans (2005, "Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions", Journal of Geophysical Research, 110 (D24), D24107, doi:10.1029/ 2005JD005970.

Interactive comment on Biogeosciences Discuss., 12, 9393, 2015.