Biogeosciences Discuss., 7, C4501–C4505, 2010
www.biogeosciences-discuss.net/7/C4501/2010/

Biogeosciences Discussions

Interactive comment on "A quantitative approach for comparing modeled biospheric carbon flux estimates across regional scales" *by* D. N. Huntzinger et al.

Anonymous Referee #1

Received and published: 23 December 2010

General

The authors use the spatial autocorrelation of the net ecosystem exchange predicted by three biospheric models to compare and evaluate the models. This is done by a combination of methods from spatial statistics and multimodel inference.

Given the sparsity of carbon flux measurements, models which work on a regional or global scale can not be compared against data. It is an important and difficult question to assess the different models. Several studies addressed this problem with different techniques. The use of the spatial autocorrelation of the predicted fluxes as starting point of the analysis is new.

The methods used are variogram analysis to quantify and model spatial autocorrelation, variable selection based on Bayes Information Criterion, and geostatistical regression with generalized least square. These techniques are well known and standard in spatial modelling and comparing against real data. If this is an adequate and valid methodology also for the problem of model comparison in this article needs additional proof (see below).

With some exception (see below) the paper is clearly written and well structured. The presentation is good and the scientific quality is solid. I am not so positive about the relevance: given models with given simulations runs are compared with known techniques. The starting question is interesting and important, but the results are mainly indications to the main differences between the models.

Special

As an example of the results let me take Section 3.2.1. The only variables selected in all three models are evapotranspiration and MDBF. On the other hand these two variables explain only a very small part of the spatial variability. This means that model specific variables have a large influence and thus point to fundamental differences in the models. This is certainly an important finding, but it may not come to a big surprise for people acquainted with the models. In the course of the analysis the found differences have to be explained by expert knowledge and model understanding. All this is said and done correctly, but it leaves me with the comprehension that the presented method can only point to potential differences between the models.

I have some points of criticism concerning the methodology. (1) Variogram analysis is certainly the appropriate tool to study spatial autocorrelation, however it may not be adequate to compare spatial autocorrelations. At least mathematically two completely different random fields can have the same spatial covariance structure.

(2) More important is the connection between spatial features of the covariates and those of the modelled fluxes. The authors attribute the spatial differences to the dif-

ferent models. However, since the modes are driven with different data sets derived from different sources and hence expressing different spatial pattern, I don't think this is valid without a test of the spatial properties of the driving data. The authors should take the following steps to show that the differences can be attributed to the models and are not solely caused by the resolution and spatial properties of the driving data: In the table listing the driving data, there should be information added about the the native spatial resolution of the data and information which model is using which data set. Table 1 is misleading since it does not list the modis active fire product as a driving data which it certainly is and similar for the NDVI, since here is no mentioning that two different variants derived from different sensors are used as mentioned in the text. Also the authors should make a similar analysis (list variograms) for the sets of driving data. (Some of this could be done in an appendix.) I actually expect that the variograms listed in figure 5 are similar to the variograms of their driving data sets (the ones where the models differ). If this is the case the whole study of limited scientific value, since it does not introduce new methods. Until the authors have not proven the effect of the models on the spatial variability, this effect remains at least questionable. For example for the two CASA variants it would be sufficient to investigate the spatial pattern of the modis active fire product, and the two different NDVI products used in the analysis. I would find it actually very interesting to compare the spatial properties of the driving data directly with the spatial properties of the models.

(3) A minor point is: spatial autocorrelation can be a complicated matter and we hope to capture its features with the variogram. To reduce everything to one single number h_0 may be possible in some cases, but I see no proof that this is the case here. Furthermore the classical parameters of the variogram have a natural interpretation, which is not the case for h_0. In addition the choice of γ_{max} makes its value rather arbitrary.

Regarding the presentation, I have three objections. (1) The method section occupies 10 of the 25 text pages and is too long. It is appropriate to give a short review of the three models, as this is needed to understand the discussion section. However, much of the description of the statistical techniques is dispensable; e.g. the formula for the drift coefficients is well-known to experts and of no help for laymen. On the other hand things like the used variogram model or probability assumptions for the calculation of the likelihood have to be mentioned. The combination of variogram analysis, variable selection, and geostatistical regression was explained in several papers of the University of Michigan group and need not be repeated — just cite it.

(2) The goal of the paper "[..] introduce a set of methods [..] and adapt them for use as tools for comparing model estimates" is too vague and as such not sufficient. Is it an assessment, comparison, and evaluation of models or of methods? Both is said/done in the text and for both I see deficiencies. For the moment I read it as an approach with *one* combination of methods, that gives indications about potential differences of the models, which in the following have to be explained with expert knowledge and model understanding.

(3) The emphasis on a *quantitative* approach is a bit misleading. Of course the statistical methods are quantitative, but the results are qualitative and give only indications to the main differences between the models. I am also not happy with the title: most readers will expect a comparison of the fluxes (totals or spatial distribution) and not a study 'only' of their spatial autocorrelation pattern.

It would be desirable, as also said by the authors, to have a joint analysis of the temporal and spatial autocorrelation. But this would mean asking too much in a review.

Technical

p. 7916, l. 24: Do you really want to refer to (Eq. 5) here?

p. 7918, l. 7: I see the minimum in September.

p. 7919, l. 2: From where do you get these lengths?

p. 7943: Add a unit for separation distance in the axis label.

BGD

7, C4501–C4505, 2010

Interactive Comment

Interactive comment on Biogeosciences Discuss., 7, 7903, 2010.

BGD

7, C4501–C4505, 2010

Interactive Comment