

## ***Interactive comment on “Improving terrestrial CO<sub>2</sub> flux diagnosis using spatial structure in land surface model residuals” by T. W. Hilton et al.***

**T. W. Hilton et al.**

hilton@unm.edu

Received and published: 27 February 2013

Author responses to Anonymous Referee #4

Referee comments in boldface, author responses in normal typeface.

**This study used NEE observation from 65 NA eddy-covariance towers to estimate the parameter set for a simple diagnostic terrestrial ecosystem model, VPRM. On this basis, it calculated the differences between tower NEE (data) and VPRM-simulated NEE (model), and further attempted to quantify the spatial structure existing in the data- model residuals by the means of semivariogram. The main findings of the study, as claimed by the authors, are: 1) the estimated VPRM parameter values do not clearly separate from each other by plant function type**

C8570

**(PFTs); 2) the “range” of the spatial semivariogram of the VPRM-data residuals is about 1000 km. However, I found the methods and the results presented in the current paper are not convincing. My main concerns are as follows:**

**1) The first part of the analysis, the parameter fitting of VPRM, is relative straightforward. The finding that the estimated parameters have similar values among different PFTs is possibly valid and as mentioned by the paper, has been reported by previous studies. However, statements such as “our results suggest that may not be the most useful predictor of a land area’s carbon cycle dynamics, that alternative partitioning schemes may be more skillful (P14, Line 446-448)” seem to exaggerate the implication of this finding. This is because VPRM is a highly simplified model and many important processes (e.g., carbon pools) are not included in its formulation. For instance, a forest/non-forest mask can be very useful in mapping biomass distributions from remote sensing. Therefore, the close values of lamda ( $\lambda$ ) may only indicate that the light-use efficiency of green leaves is similar among PFTs. Also, some VPRM parameters are by definition not PFT-related. For instance, the  $\beta$  parameter of VPRM represents a base respiration rate that is not affected by climate but mostly by site disturbance history.**

Thank you for this comment. We agree that this finding is not directly applicable to more complex model structures. It is also a good point that not all of VPRM’s parameters should be expected to vary among PFTs; we will add a sentence noting this to the text.

Because light-use efficiency models remain a popular method of estimating gross primary productivity (GPP), we believe that the similarity of optimized lambda values across starkly different PFTs is still a point worth noting.

We had not seen the similarity of parameter values among PFTs as one of the main findings of the analysis, but rather as an interesting and mildly surprising result that came out of the parameterization process. We felt it was worth mentioning in the ab-

stract because it belongs to a different “genre” of results than the geostatistical results of the residual analysis. We will revise the abstract and lines 446 to 448 on page 14 to make it more clear that we see it as a limited statement regarding the use of PFTs in optimizing parameters for a certain class of land surface models, and not blanket statement that PFTs should be avoided.

**2) Though the concept of semivariogram is clearly explained in the paper, its application in the analysis is quite confusing. First and most important, it is unclear that whether the 65 sites are enough to render a robust estimate of the semivariogram. This is particular true considering that quite many of the sites are co-located, which further reduced the number of free samples (Fig. 1). Second, though the authors tried to test the above question by Monte-Carlo simulations and AIC tests (P11, 2-3 paragraphs), the results seem biased towards the “pure nugget” model as only 7% generated with underlying exponential model are correctly classified as “exponential” by the AIC test. This suggests that either 1) the spatial sampling process has significant impacts on the tests; 2) the estimated spatial parameters (range, sill, and nugget) have a very large uncertainty or simply not reliable. This problem is clearly illustrated by Fig. 5 (see below).**

As noted by Reviewer 4, the pseudodata experiment results reported in figure 5 correctly identified the underlying exponential covariance structure in only 7% of the realizations tested. The realizations were generated using the median covariance range identified by our analysis. In contrast, the same analysis method identified the exponential model as fitting the real VPRM-observed NEE residuals better than the pure nugget structure more than 36% of the time. It is certainly true that the spatial sampling process employed will influence the results of the analysis. We are, however, for the time being, restricted to the single realization of North American eddy covariance towers that actually exists. Moreover, we believe that there is a third plausible explanation for the difference in the detection rate between the exponential covariance structure pseudodata and the actual residuals: the true underlying covariance structure may be

C8572

stronger than the median results reported here, but the realization of spatial distribution available to us (Fig 1) is inadequate to detect it. This is also consistent with our finding that spurious exponential covariance structures were identified in only 2.5% of realizations, far below the detection rate of 36% for the real residuals.

We will add points (1) and (2) raised by Reviewer 4 to the discussion of the pseudodata experiments in section 3.2.

We thank Reviewer 4 for carefully considering the results in section 3.2.

**3) Fig. 5. Because all the pseudo-data are generated with parameters estimated from the VPRM-NEE residuals, one would expect that the Monte-Carlo simulated samples well encompass the original VPRM-NEE residuals. However, this is not the case shown here (in the right panel). The seemingly agreement in the left-panel can be misleading because it is a statistical principle not to compare data with subjectively selected data.**

Similarly, we believe that another plausible explanation for the detection rate of exponential covariance structure for the actual residuals exceeds that in the pseudodata is that the true residual covariance structure is more coherent than the median of detected structures.

We certainly agree with Reviewer 4 that it is unwise to compare data with subjectively-selected data. We believe it is useful to compare the subsets of pseudodata and real residuals in which the exponential covariance structure fit better than the pure nugget structure because it demonstrates that the spread of detected range values among pseudodata is similar to the (admittedly large) spread detected from the real residuals. The left panel of figure 5 attempts to illustrate the similarities in spread. A box-and-whisker plot in a separate figure could make that point equally well and perhaps could better separate the points we wish to make with in figure 5. We will make this change in the manuscript.

C8573

**4) Table 3, Fig. 6-7. The estimated values of “range” vary dramatically at interannual basis. For instance, the first row in Table 3 shows that the range is only 4km in 2003 but is over 3500km and 1500km in 2002 and 2004, respectively. Similar large interannual variability is found in many other rows as well. Such variability can hardly be interpreted in a physically meaningful way. In addition, Fig. 6-7 are also rather confusing.**

Points (3) and (4) are related: As reviewer 4 notes, the interannual variability in table 3 is certainly large. Similarly large is the overall spread within the exponential range values presented in table 3. However, the spread among exponential range values from the real residuals (table 3) is quite similar to the spread within pseudodata exponential range values (figure 5, left panel). We therefore believe that the interannual variability exhibited in table 3 is consistent with what one might expect from this group of realizations of annual covariance structures from 65 locations within a domain of 6500 km per side.

**5) Fig. 4. It is hard to compare these plots because the residuals have different standard deviations. For instance, the “all sites” plots have higher semivariance because they have fewer parameters and thus less “fit” with the data. It may be better to normalize the plots by their own standard deviations.**

Thank you for this point – we will change figure 4 to show semivariance normalized by standard deviation. The main point we wish to make with figure 4 is the similarity the length scale of spatial covariance across different VPRM parameterizations. This feature of the semivariograms will remain after normalization.