

## ***Interactive comment on “Evaluating terrestrial CO<sub>2</sub> flux diagnoses and uncertainties from a simple land surface model and its residuals” by T. W. Hilton et al.***

### **Anonymous Referee #2**

Received and published: 17 October 2013

Given that I am the 2nd of 2 referees to upload comments I will limit myself to things not covered by R1 (I note that I am in complete agreement with his comments).

This is an excellent paper overall and a worthy contribution to the corpus of upscaling FLUXNET literature. After minor technical fixes this is ready for publication. I have a few higher-order comments:

I would prefer a more comprehensive review of upscaling to date. The authors cite and detail a few studies but there are several that are left off. As examples: the Jung Nature paper that upscales ET (not strictly a 1:1 correspondence with C fluxes as here but very much a game changer for FLXUNET-inspired upscaling wrt visibility). Schwalm et al

C5889

(2010,2011ab) have done FLUXNET-based global upscaling of changes in NEP, GPP, and TER that are solely attributable to hydrological intensification and drought. Yuan et al 2010 detail the derivation of the EC-LUE model. Yang et al 2007 use FLUXNET sites in conjunction with MODIS and SVM to get at GPP. There are others. These are all worthy contributions that have advanced this field. While a summary as per the original is certainly too verbose perhaps a summary sentence (or two) that showcases the depth and breadth of upscaling approaches would be useful.

I am intrigued why RMSE was chosen as the metric? I am curious if you looked at how your results would change if you scored fit differently? Note that I am not suggesting submitting another set of parallel results. But, how sensitive are the conclusions of your study to the skill metric chosen?

As R1 I am intrigued by the 27 vs. 65 split and was very curious as to selection criteria. But R1 has discussed this already and I have nothing to add. But I would emphasize the issue's importance. Consider that you show different maps based on different parameter sets. How about different maps based on different site splits?

Did you consider downscaling your driver data, e.g., 16-day MODIS data? You use 3h data to drive VPRM and there are canned routines for downscaling (for “imposing” the diurnal cycle).

Wrt Eq. [5], this has minimal skill ( $r^2 = c. 0.3$ ). I'm not sure I have a high degree of comfort with any map generated based on this equation. I would like some words on why, given the clear lack of skill of Eq. [5], it has any value wrt uncertainties as discussed in the original.

Use NEP and NOT NEE. NEE is the integrated vertical exchange of CO<sub>2</sub>. FLUXNET does not measure this (so you can't upscale this C term either). FLUXNET gets at CO<sub>2</sub> exchange as the disequilibrium between GPP and TER only. The processes that are part of NEE (e.g., aquatic evasion, disturbance emission [fLUC or fire flux as examples] and product decay) are not “seen” by FLUXNET at all. Put another way, you cannot

C5890

compare your NEE values to something that comes from an inversion framework. So I find this misleading, NEP is what FLUXNET does well, not NEE.

Figures: What does black represent? This is not detailed/explained.

Pg 13770: wrt "Once again, instead of concluding that respiration is causing the mixed forests of the southeastern USA to release on the order of  $150 \text{ g Cm}^{-2} \text{ yr}^{-1}$  to the atmosphere, the explanations discussed in Sect. 3.3 seem more plausible." Could you parenthetically include what Sect 3.3 stated? This would help the reader.

---

Interactive comment on Biogeosciences Discuss., 10, 13753, 2013.