

Interactive
Comment

Interactive comment on “Using atmospheric observations to evaluate the spatiotemporal variability of CO₂ fluxes simulated by terrestrial biospheric models” by Y. Fang et al.

Anonymous Referee #1

Received and published: 16 July 2014

General Comments: (Overall Quality of Paper)

The purpose of this manuscript is twofold: first it tests a methodological approach that compares continuous CO₂ observations against simulated NEE fluxes, and second applies this methodological approach to evaluate 4 models from the Regional Continental Interim Synthesis (RCIS) across 35 towers within North American during 2008.

The strength of this paper is a test of the methodology, in part established by Gourdjii and others and combined with the Bayesian Information Criterion, for selection of TBMs that best match variation in atmospheric CO₂. The authors rigorously test the design of experiment with pseudo-data to evaluate the effectiveness of the method for idealized

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

situations. The authors find that only a subset of biomes (e.g. grassland, conifer and deciduous) provide a large enough biogenic signal to make atmospheric CO₂ a useful diagnostic for surface fluxes. This is not surprising considering the method is incapable of detecting skillful model performance in poorly sampled regions or biomes with minimal variation in seasonal NEE (e.g. tundra).

When applied to real model output, the authors found that two EK models (SIB3, ORCHIDEE) perform better than two LUE models (CASA-GFED, VEGAS2), which the manuscript states is most likely the result of the time step and number of PFTs included in these models. In general, the models perform better during the growing season and poorer during the transition seasons.

Overall, the evaluation of real model data is brief, with limited interpretation, and largely reiterates findings of previous manuscripts within the NACP synthesis. This is the main weakness of the paper. This reviewer recommends including more years and models to improve the scope of the paper, with more focus upon the attribution of model performance. This reviewer also recommends comparing the TBM model performance against those findings in Schwalm et al. 2010 (site runs) and Raczka et al. 2013 (regional runs) as part of the NACP Synthesis. This should better identify consistencies/inconsistencies across the different model validation approaches.

Specific Comments: (Individual Scientific Questions/Issues)

It was unclear to what extent this method is a test of the terrestrial biosphere model vs. the spatial coverage of the observation network. A part of the methods identifies biome-months that were not well constrained by atmospheric concentrations and eliminates them from further consideration, nevertheless, the approach is influenced by the observation network. Furthermore, the method attempts to tease out the behavior of biogenic fluxes, when the variation of real data is due to boundary conditions, fossil fuel emissions and function of model transport. The method 'subtracts out' the influence of boundary conditions and fossil fuel emissions, but still the strength of the biogenic sig-

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

nal is weakened because of these confounding influences. The method requires fitting of the TBM NEE to a statistical model, and it is this statistical model, which is tested for how well it can simulate the observed atm. CO₂. This seems like a challenging undertaking to connect the model surface fluxes to network of atm. CO₂ observations.

This reviewer remained skeptical how a method that emphasizes the ability to assess fine-scale processes, can use atmospheric CO₂ observations, a state variable that is inherently dependent upon an integral of influence across a large spatial area, as an effective way to accomplish this. Even though the TBMs provide fluxes at relatively fine spatial/temporal resolution, the measured flux tower CO₂ can be the result of many confounding, fine scale processes within the footprint.

It is unclear to this reviewer whether this approach offers any additional insights into model processes other than what has already been demonstrated from previous RCIS publications. The most important findings that the models perform better during the growing season, rather during the transition seasons is already established by works such as Richardson et al. 2012. Although the author attempts to make a distinction between 2 types of models (enzyme kinetic, light-use efficiency), there is limited attribution (e.g. model time step, # pfts) to why the performance is different.

In general, TBMs perform the best at simulating diurnal and seasonal variation, which it seems this methodology is sensitive to. Models do not simulate magnitude of integrated carbon fluxes well, nor do they simulate interannual variation in flux well. This is something this method/paper does not address. It is unclear to this reviewer whether this approach distinguishes between models that have large biases in the carbon flux, or only evaluates model skill through variation in model flux.

The author provides cursory explanation of the difference in model performance between growing season and transition season. The author does not make a clear distinction between phenology (e.g. leaf timing, leaf activity) and model representation of photosynthetic/respiration processes influenced by environmental variation. They can

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

be viewed as distinct entities. In other words, do the models perform more poorly during the transition seasons because simply the leaf timing (phenology submodels) are wrong, or because other physiological processes are not well represented during this time?

It was difficult for this reviewer to reconcile why SIB3 and ORCHIDEE (EK models) performed better than CASA-GFED2 VEGAS (LUE models). Given that the models tended to perform worse in the transition seasons likely from the poor phenology. Presumably LUE models (remote sensing of phenology) would do better than EK models (internal prediction of phenology) during transition seasons.

This reviewer was left with the question of how do we improve the models? How does this method, in particular, provide an upgrade in the diagnosis of model behavior from the classical approaches of tower-site evaluation vs. aggregated region analysis between bottom up, inversion and inventory approaches? I think the method is interesting and has potential, but the payoff was not clear at least for this application.

Technical Corrections:

Page 9216, Line 21: "...differences among the examined models are at least partially attributable to their internal structures". The RCIS from which the TBM output was taken from, was not based upon a controlled protocol. It is conceivable models of identical internal structure could provide different results just from differences in driver data.

Page 9217, Line 28: "...novel path..." Assessing the spatial/temporal variability of carbon fluxes to evaluate TBMs is not a 'novel' approach. It has been done quite often. This particular inversion methodology seems to be the 'novel' part.

Page 9227, Line1: Seems contradictory to say that the tundra biome has a weak biospheric signal, but it plays an important role in the carbon cycle in climate. I think you need to be clear that it is expected to play a more important role in future.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper



Page 9231, Lines 17-24: Not sure what this means. What specifically are the models not simulating correctly?

Page 9234, Lines 7-9: This is vague. What are the seasonal differences in performance attributable to?

Page 9234, Lines 16-19: It would be more worthwhile to perform this analysis as part of the MsTMIP (Huntzinger) experiment, which includes a controlled protocol.

Figure 1: The stars denoting the location of the CO₂ observations are hard to read, especially in dark-colored biome regions.

Figure 2: Shouldn't 3rd row be labeled as 'correlated flux residuals' not just 'flux residuals'?

Figure 3 caption: Unclear what the 'grey' threshold (average number of months) is for map figures, which were determined to not be well constrained by atmospheric data.

Interactive comment on Biogeosciences Discuss., 11, 9215, 2014.

BGD

11, C3548–C3552, 2014

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

