Comparing estimates of North American biospheric carbon flux from process-based models: New tools for understanding differences in predictions

D.N. Huntzinger[1] (dnhuntzi@umich.edu), K. Mueller[1], S. Gourdji[1], A.M. Michalak[1,2]

[1]Department of Civil & Environmental Engineering, University of Michigan, Ann Arbor, MI 48109-2125
[2]Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, MI 48109-2143

Acknowledgements:

Funding for this work was provided by NASA through the ROSES A.6 NACP program, Grant No. NNX06AE84G. Data suppliers include NASA MODIS for NACP, NOAA GLDAS, GEOS-DAS, International Geosphere Biosphere Porgram (IGBP) land cover classification, the World Wildlife Fund land cover classification, Transcom 3 Continuous, CarbonTracker, and Ian Baker.

Goal and Scope

A number of biospheric or process-based models have been developed to estimate the magnitude of carbon sources and sinks across regional and continental scales. These models vary in complexity and tend to include a diverse array of processes that operate on widely different spatial and temporal scales. Both the complexity of the system being modeled, as well as the inherent differences among the various modeling approaches have made comparisons of model-derived net ecosystem eschange (NEE) difficult. However, with the lack of direct observations of NEE at the scales of model estimation, there is no way to validate model-derived NEE; thus model inter-comparisions are a necessary means to assesss model performance, as well as evaluate our understanding of the terrestrial carbon cycle system.

The objective ot this study was to apply a set of aggregation, spatial analysis, and regression tools in an inter-comparison of biospheric model estimates of NEE. The goal was to develop an approach for comparing model results in light of the inherent differences among the models in terms of their formulation and driving variables. The subset of biospheric models examined in this analysis (SiB 3.0, CASA, and CASA GFEDv2) were chosen because (1) they are widely cited and used to study land-atmosphere carbon exchange and (2) these models are commonly incorprated as *a priori* information in Bayesian inverse modeling set-ups used to estimate surface fluxes. There is growing awareness of the strong influence of explicit prior flux estimates on the inferred magnitude and spatial distribution of surface fluxes in inversions.

Approach

The approach is fourfold:

(1) Highlight regions where models exhibit significant differences in the magnitude and seasonal cycle of NEE estimates by spatially aggregating modeled NEE to the entire North American domain, as well as six dominant ecoregions.

(2) Quantify the degree of overall spatial variability of modeled monthly NEE across North America within a geostatistical framework.

(3) Using variable selection methods, attribute the variability in model estimated NEE to those underlying environmental variables that are most significant in explaining the spatial trend of fluxes at regional scales.

(4) Quantify the influence of these environmental variables on modeled NEE through the use of universal kriging, a geostatistical form of multi-linear regression that accounts for the spatial correlation in the portion of the flux not explained by the environmental variables.

Models Examined:

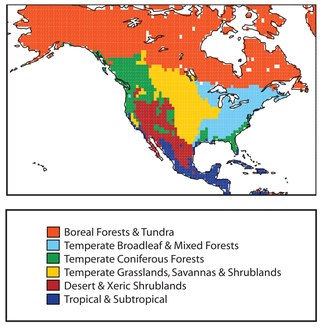
SiB 3.0 (Sellers et al. 1996, Baker et al. 2008); CASA, (Potter et al. 2003, Randerson et al. 1997) and CASA GFEDv2 (van der Werf et al. 2004, 2006) with (NECB) and without (NEE) fire emissions.

Study period: 2002-2003
Domain: 10N to 70N, 50W to 170 W
Spatial resolution: 1 degree X 1 degree
Temporal resolution: monthly

(-) refers to a removal of CO_2 from the atmosphere
(+) refers to a release of CO_2 to the atmosphere

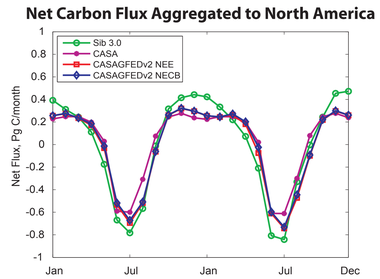
Biome Classification

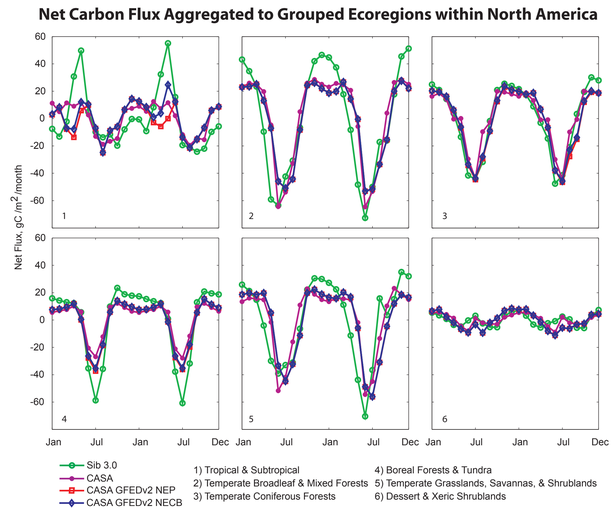
The spatial aggregation of fluxes to spatially contiguous regons with similar land cover or biome types and climatic conditions allows for the examination of regional differences among the models by applying a common mask to the modeled output.

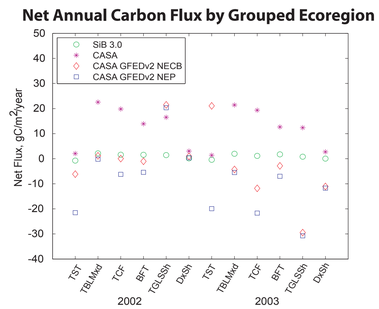


Grouped biome classification based on Olson et al. 2001 and the World Wildlife Fund: \\SASQUATCH\DATA\ecoregions\version_2\new_version\wwf_terr_ecos.shp

Net Flux in North America and Grouped Ecoregions

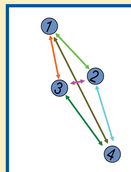






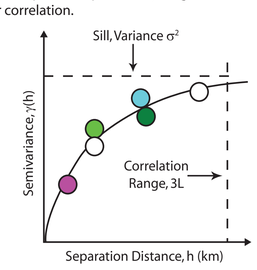
By comparing model results annually and seasonally over different ecosystem types, regions or biomes can be highlighted where differences in the model formulation and assumptions result in large variations in their estimates of NEE. These differences cannot be seen when fluxes are only examined at the scale of North America, as differences between models become masked by aggregation to larger regions. Therefore, examining modeled NEE over smaller regions, linked to ecosystem type and climatic conditions, can be a first step in identifying factors in the model that may be driving these differences.

Quantifying Spatial Variability using Variograms



Variogram

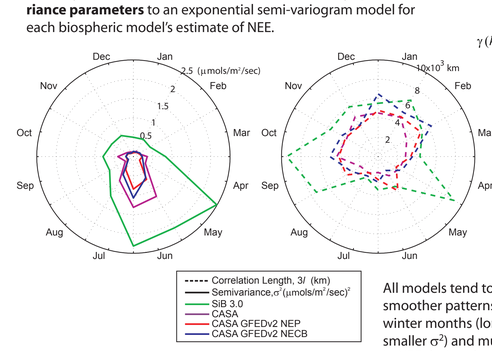
Measure of variability between two points as a function of separation distance. The closer in space two points are, the greater their correlation.



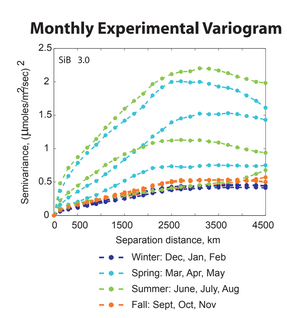
Ordinary Least Squares (OLS) was used to fit **monthly spatial covariance parameters** to an exponential semi-variogram model for each biospheric model's estimate of NEE.

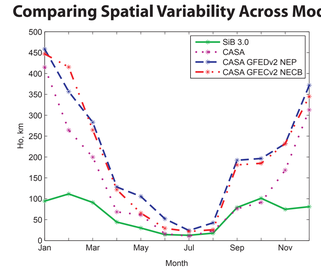
Exponential Variogram

$$\gamma(h) = \sigma^2\left(1-\exp\left(-\frac{h}{l}\right)\right)$$



All models tend to estimate much smoother patterns of NEE in the winter months (longer *l* and and smaller σ^2) and much more variability in NEE during the growing season.





where $h_0 = -L\ \ln\left(1-\frac{V_{max}}{2\sigma^2}\right)$

h_0 provides a common metric with which to compare the overall spatial variability of the models (Alkahed et al. 2008). Both a higher regional variance $(\sigma)^2$ and a shorter correlation range parameter (l) will result in a shorter h_0 indicating a greater degree of spatial variability of modeled NEE over smaller spatial scales.

The use of variogram analysis allows for the quantification and comparison of overall spatial variability among NEE predicted by the different models. In addition, it is a valuable tool for highlighting differences in the spatial pattern of modeled NEE that may not be evident from qualitative or visual assessments.

Variable Selection and Regression

The **Variance Ratio Test (VRT)** (Kitanidis et al 1997) was used to assess the correlation of modeled NEE to a set of environmental variables commonly associated with flux or used as driving variables in forward models.

Environmental Parameter	Source
Enhanced vegetation index, EVI	MODIS
Normalized difference vegetation index, NDVI	MODIS
Fraction absorbed photosynthetically active radiation, fPAR	MODIS
Leaf area index, LAI	MODIS
Net short-wave radiation, SWR, W/m^2	NOAA
fPAR*netSWR, W/m2	MODIS
Evapotranspiration, Evap, g/m^2	NOAA
Precipitation, mm/day	GEOS-DAS
Near surface air temperature, Temp, C	GEOS-DAS
Soil moisture, SM, kg/m^2	NOAA
Q10	calculated
Percent crop land, %CL	IGBP
Precent evergreen needleleaf, %EVNL	IGBP
Precent grassland, %GL	IGBP
Percent mixed deciduous and broadleaf forests, %MXDBL	IGBP
Precent shrubland, %SHL	IGBP

The VRT tests the statistical significance of correlations and therefore avoids indentifying variables with only a spurious relationship to flux. The meaure of fit or correlation between modeled NEE (**y**) and the environmental variables (**X**) is defined by the Weighted Sum of Squares (WSS), where **Q** is defined using the optimized covariance parameters defined by variogram analysis:

$$WSS = \mathbf{y}^T\left(\mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{X}\left(\mathbf{X}^T\mathbf{Q}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Q}^{-1}\right)\mathbf{y} \qquad Q_{ij}\left(h_{ij}\middle|\sigma^2, l\right) = \sigma^2\left(1-\exp\left(-\frac{h_{i,j}}{l}\right)\right) \qquad v = \left(\frac{\left(WSS_0 - WSS_1\right)/q}{WSS_1/\left(n-p-q\right)}\right)$$

where *n* is the number of biospheric model estimates of NEE, *p* is the number of component variables in $\mathbf{X_0}$ and *q* is the number of additional component variables in $\mathbf{X_1}$. The *v* statistic follows an F distribution with *q* and *n-p-q* degrees of freedom. The VRT was conducted in similar manner as in Gourdji et al. 2008.

Environmental Variable	Order of Selection in Variance Ratio Test: SiB 3.0	CASA	CASA GFEDv2 NEP	CASA GFEDv2 NECB
EVI				
NDVI	9			
fPAR	4	9	10	11
LAI	11	8	12	
netSWR	10		11	9
fPAR*netSWR	5		9	10
Evap	2	1	2	3
Precip		10	8	8
Temp	7	7	7	7
SM				
Q10	8	5		
CPL		4	4	4
EVNL	6	3	3	2
GL		6	5	5
MXDF	1	2	1	1
SHL	3		6	6

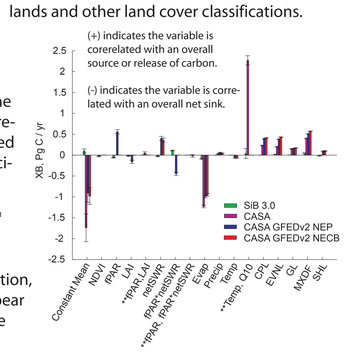
Monthly averaged data were used to assess the annual relationship of NEE to the envrionmental variables in the above table. The order in which a given environmental variable is selected during the VRT in an indication of the strength of its correlation with model estimates of carbon flux. Evapotranspiration and mixed deciduous, broadleaf forests were consistently selected in the first rounds of variable selection, followed by croplands and other land cover classifications.

Geostatistical Regression Analysis

The magnitude and spatial distribution of the different model estimates of NEE (y) are represented as a the linear combination of selected environmental variables (Xβ) and their associated weights (β) and uncertainty (V_β):

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \qquad \beta = \left(\mathbf{X}^T\mathbf{Q}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Q}^{-1}\mathbf{y} \qquad V_\beta = \left(\mathbf{X}^T\mathbf{Q}^{-1}\mathbf{X}\right)^{-1}$$

The annual net contribution (Xβ) to NEE in PgC/yr is shown on the right. Evapotranspiration, land cover, and net short-wave radiation appear to have a strong corrletion to NEE among the models examined.



Conclusions

- Analysis shows that even when models produce similar estimated seasonal cycles of NEE when aggregated spatially to North America, they do not agree on the overall net contribution (source or sink) of North America (or ecoregions) to the overall carbon budget. In addition, neutral biosphere assumptions (e.g., SiB 3.0) make it difficul to compare net sources/sinks.
- Although there are similarities in the general trend or pattern of NEE variability in space, the magnitude of modeled NEE variability over smaller spatial scales varies, and these differences are seem more strongly in the dormant versus the growing season.
- The regression methods applied here indicate low linear correlation of NEE to common climatic and environmental factors (e.g., precipitation, temperature, soil moisture) typically thought to control NEE. This does not mean these factors are not dominant drivers in the models examined, but rather other controls within the models introduce substantial spatial variability.
- The approach presented here allows for the better identification of spatial dissagreements among model estimates of NEE and a more rigorous means of assessing the apparent correlation of multiple environmental factors on modeled NEE.

REFERENCES:

Alkahed, A.A. et al. (2008). A global evaluation of the regional spatial variability of column integrated CO_2 distributions, *Journal of Geophysical Research-Atmospheres, 113*(D20).

Baker, I. et al. (2008). Seasonal drought stress in the Amazon: reconciling models and observations, *Journal of Geophysical Research-Biogeosciences,* 113.

Gourdji, S.M. et al. (2008). Global monthly averaged CO_2 fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data, *Journal of Geophysical Research-Atmospheres,* 113(D21).

Kitanidis, P.K. et al. (1997). A variance-ratio test for supporting a variable mean in kriging, *Mathematical Geology,* 29(3), pp. 335-348.

Olson, D.M. et al. (2001). Terrestrial ecoregions of the World: A new map of life on Earth, *BioScience* 51, pp. 933-938.

Potter, C.S. et al. (1993). Terrestrial ecosystem production: a process model based on global satellite and surface data, *Global Biogeochemical Cycles,* 7(4), pp. 811-841.

Randerson, J.T. et al. (1997). The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, *Global Biogeochemical Cycles,* 11(4), pp. 535-560.

Sellers, P.J. et al. (1996). A revised land surface parameterization (SiB2) for atmospheric GCMs. Part- I: model formulation, *Journal of Climate*, 9, pp. 676-705.

van der Werf, G. R. et al. (2004). Continental-scale partitioning of fire emissions during the 1997 to 2001 El Nino/La Nina period, *Science*, 303 (5654), pp. 73-76.

van der Werf, G.R. et al. (2006). Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmospheric Chemistry and Physics,* 6, pp. 3423-3441.