Using Remotely-sensed Data Sets for Model Evaluation and Benchmarking

Hoffman, Forrest M.[1,2]
Randerson, James T.[2]
Mao, Jiafu[1]

[1]Oak Ridge National Laboratory, Climate Change Science Institute, Oak Ridge, TN 37831-6301
[2]University of California - Irvine, Department of Earth System Science, Irvine, CA 92697

Bulleted list of abstract highlights:

- While observations and data products derived from satellite remote sensing are important for validation and evaluation of Earth System Models, a variety of technical issues pose challenges to the production of remote sensing data sets useful for this purpose.
- Since models predict properties not directly measurable as reflected radiances, these variables must be estimated from remote sensing measurements using models.
- Because of uncertainties in the measurements and the models that generate remote sensing products, analysis of spatial correlations and comparisons of temporal phase may offer better metrics for model evaluation.
- Presented are results from the Carbon-Land Model Intercomparison Project (C-LAMP) in which comparisons with MODIS data products and *in situ* observations were used to score model performance.
- Results from this study suggest that models can be improved using the current body of remote sensing data.
- Community-accepted standards are needed for model benchmarks using remote sensing data.

Abstract:

Observations and data products derived from satellite remote sensing are important for validation and evaluation of Earth System Models (ESMs) from site to global scales. The wide spatial coverage and frequent repeat observations afforded by satellite remote sensing make such observational data complementary to intensive ground-based *in situ* measurements and less frequent aerial survey data. However, a variety of technical issues pose significant challenges to the production of remote sensing data sets useful for comparison with model results, including spatial mismatch between sensors and models, orbital constraints on observational timing and frequency, difficulties with atmospheric correction, and obscuring by clouds and snow cover. Moreover, because models predict properties not directly measurable as reflected radiances—such as leaf area index, gross and net primary production, and soil temperature and moisture profiles—these variables must be estimated from remote sensing data using processing algorithms that are themselves rudimentary ecosystem models. Uncertainties in the measurements and the models that generate the more-applicable remote sensing products limit the utility of these data for constraining ESMs, suggesting the need for carefully designed metrics when applying them to model evaluation. Spatial and temporal averaging of both model results and remote sensing data are usually required.

Because of measurement uncertainties, it may be more practical to analyze spatial correlations and compare temporal phase of model results with remote sensing data.

A recent model-data comparison called the Carbon-Land Model Intercomparison Project (C-LAMP) used remote sensing and *in situ* observations to evaluate the performance of two terrestrial biogeochemical modules, CASA′ (Carnegie-Ames-Stanford Approach Prime model) and CN (Carbon-Nitrogen model), coupled to the Community Land Model (CLM3.5) and running within the Community Climate System Model (CCSM3). In this study, model estimates of leaf area index (LAI) and net primary production (NPP) were compared with observations from the MODerate Resolution Imaging Spectroradiometer (MODIS). Three aspects of observed leaf area—a key prognostic variable of climate-carbon models that couples biophysics, hydrology, and biogeochemistry—were used to assess model performance: the timing or phase of maximum LAI (seasonality), maximum monthly LAI, and annual mean LAI. MODIS estimates were taken from the MOD15A2 collection 4 LAI product (*Myneni et al.*, 2002) with additional adjustments to interpolate across periods of cloud contamination as described by *Zhao et al.* (2005). Biases may exist in the satellite-derived estimates of mean and maximum LAI due to errors in radiative transfer models used in the retrieval, but the seasonality metric is probably less uncertain because it should be less sensitive to these biases. As a result, more weight was given to the comparison of LAI phase in the C-LAMP scoring system than to comparisons of maximum or mean. Figure 1 shows the comparison of month of maximum LAI for MODIS, CASA′, and CN. Similarly, comparisons were made between estimates of NPP from these models and MODIS using the MOD17A3 collection 4.5 product (*Heinsch et al.*, 2003). However, biases could exist because of errors in underlying algorithms that convert satellite radiances to the fraction of absorbed photosynthetically active radiation (fAPAR) or that convert APAR to NPP using a light use efficiency model. In an attempt to avoid these biases, the model evaluation metrics relied upon the square of the Pearson correlation coefficient (r^2) between MODIS NPP and the model results for all model grid cells and, separately, for latitudinal zonal means. In the C-LAMP scoring system, comparisons with MODIS NPP were weighted less than the comparisons with MODIS LAI.

While the C-LAMP study also included many comparisons with *in situ* measurements, results suggest that models can be improved using the current body of remote sensing data. For example, comparison with MODIS LAI showed that both the CASA′ and CN models lagged the observed timing of maximum leaf area by 1–2 months (Figure 1). Nevertheless, advances in remote sensing technology and processing algorithms are needed to further reduce uncertainties in model predictions. Satellite mission continuity and development of new combined multi-sensor techniques are critically important for improving Earth observations. In addition, the modeling community can make better use of remote sensing observations by developing satellite platform simulators and by producing community-accepted standards for model benchmarks using remote sensing data. Satellite simulators are embedded within ESMs and compute the measurements that would be made by a satellite platform under the model-projected biogeophysical conditions. The International Land Model Benchmarking (ILAMB) project (`http://www.ilamb.org/`) is a model-data intercomparison and integration effort designed to improve the performance of land models while simultaneously improving the design of new measurement campaigns to reduce uncertainties associated with land surface processes. ILAMB promotes the development of model-data comparison benchmarks that are agreed upon by the international research community, including those based on remote sensing observations. If successful, this project will produce a model assessment toolkit that employs an evolving collection of satellite and ground-based observations to systematically evaluate the fidelity of ESMs.

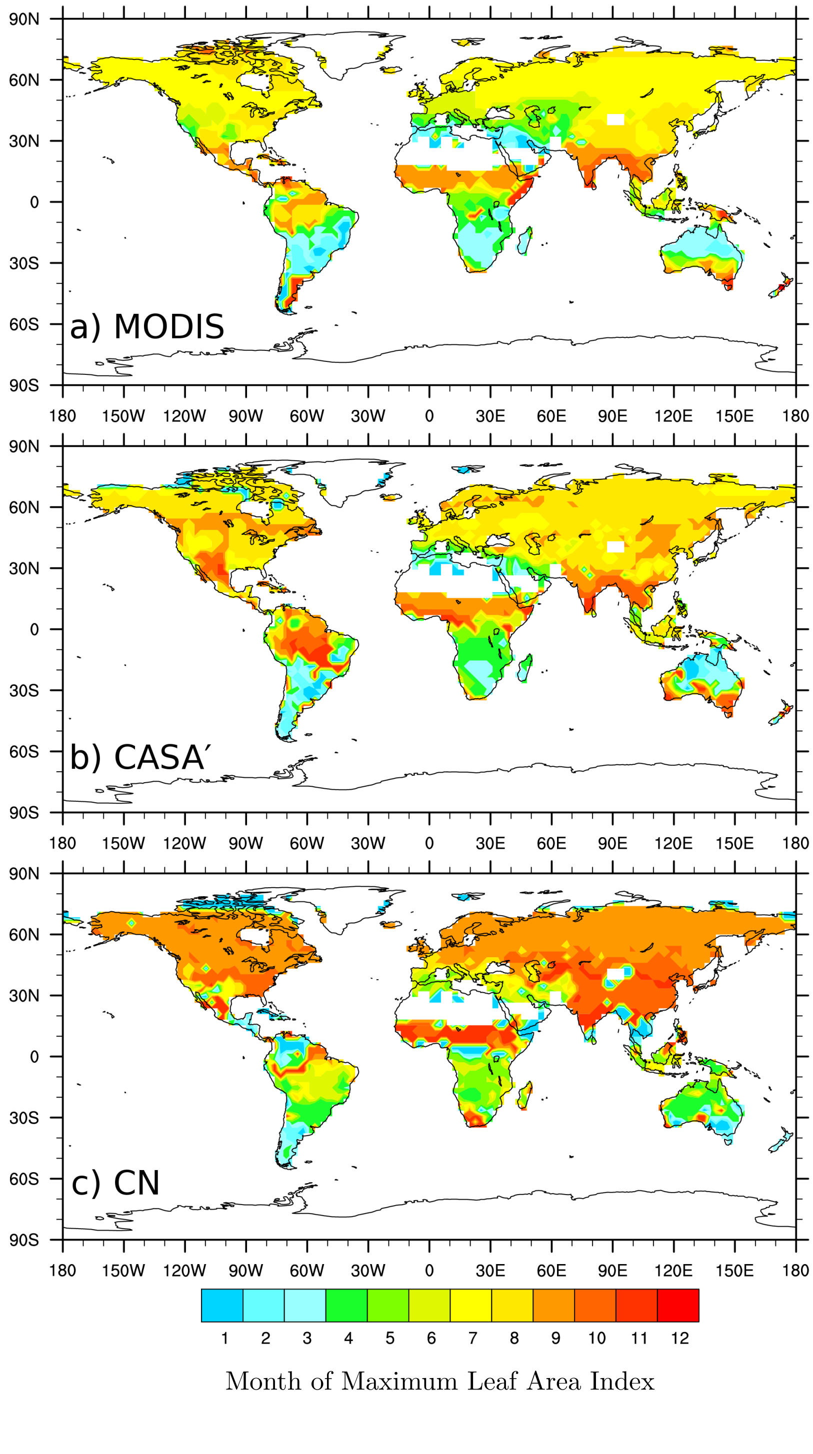

Figure 1: Month of maximum leaf area index from (a) MODIS, (b) CASA′, and (c) CN. The observations are from the MOD15A2 collection 4 LAI product from MODIS (*Myneni et al.*, 2002) with additional adjustments to interpolate across periods of cloud contamination as described by *Zhao et al.* (2005). Figure from *Randerson et al.* (2009).

References:

Heinsch, F. A., M. C. Reeves, P. Votava, S. Kang, C. Milesi, M. Zhao, J. Glassy, W. M. Jolly, R. Loehman, C. F. Bowker, J. S. Kimball, R. R. Nemani, and S. W. Running (2003), User's guide: GPP and NPP (MOD17A2/A3) products, NASA MODIS land algorithm, *Tech. rep.*, University of Montana.

Myneni, R. B., S. Hoffman, Y. Knyazikhin, J. L. Privette, J. Glassy, Y. Tian, Y. Wang, X. Song, Y. Zhang, G. R. Smith, A. Lotsch, M. Friedl, J. T. Morisette, P. Votava, R. R. Nemani, and S. W. Running (2002), Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, *Remote Sens. Environ.*, *83*(1–2), 214–231, doi:10.1016/S0034-4257(02)00074-3.

Randerson, J. T., F. M. Hoffman, P. E. Thornton, N. M. Mahowald, K. Lindsay, Y.-H. Lee, C. D. Nevison, S. C. Doney, G. Bonan, R. Stöckli, C. Covey, S. W. Running, and I. Y. Fung (2009), Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, *Global Change Biol.*, *15*(9), 2462–2484, doi:10.1111/j.1365-2486.2009.01912.x.

Zhao, M., F. A. Heinsch, R. R. Nemani, and S. W. Running (2005), Improvements of the MODIS terrestrial gross and net primary production global data set, *Remote Sens. Environ.*, *95*(2), 164–176, doi:10.1016/j.rse.2004.12.011.