VOL. 3 NO. 3, DECEMBER, 2010

FluxLetter

THE NEWSLETTER OF FLUXNET

ISSUE DEVOTED TO SCALING FLUXES

An examination of methods and tools for using remote sensing data and FLUXNET measurements to enable modeled estimates of grid-scale fluxes 'everywhere all the time"

Scaling Carbon and Water Fluxes from Patches to the Globe: A Challenge and an Opportunity for the Future

An Editorial by Dennis Baldocchi, Rodrigo Vargas and Laurie Koteen

Today, a new scientific revolution is emerging among the FLUXNET community where groups of scientists are producing global scale information on carbon and water fluxes. They are doing so by merging of information from networks of flux towers, biophysical models, ecological databases and satellite-based remote sensing to produce a new generation of flux maps on monthly, yearly and decadal intervals. The success of this effort is only possible by the altruistic sharing of data by each and every one of us, and represents a joint effort. This issue of the FLUXLETTER profiles several groups who are leading the global upscaling charge with a combination of statistical and biophysical models. The effort to produce flux information with a global extent grew out of a general desire to solve problems related to perturbations to the Earth's terrestrial carbon and water cycles. A consensus is now emerging that to be most effective, scientists must produce a measurement and modeling system that is 'everywhere, all of the time'. The global network of networks, FLUXNET, is a step in this direction because it produces a system of flux measurement towers that are 'many places, most of the time'. But is this good enough?

An alternative is to model carbon and water fluxes across the globe and to use the flux towers to validate and test the models. But this task requires quantifying a set of coupled and highly non-linear equations that explain biophysical processes that span 14 orders of magnitude in time and space (Osmond 1989, Jarvis 1995).

Looking back, it is interesting to see how our orientation towards scaling approaches has changed. Such an enterprise would not have arisen in the early days of scaling, when there was much resistance to develop models that span more than 3 'levels', (scales), let alone fourteen. In the proceedings of the famous Trebon, Czechoslovakia workshop of 1969, the pioneering modeler C.T. deWit (de Wit 1970) wrote:

> *'now I believe that the simulation in the biological sciences has to be used to fill the gap that exists between specialists at various 'levels' and that we may come to a strategy of model-building in biology when we keep this purpose in mind. To build a model we have to consider and join **two levels of knowledge**. The level with the sort of relaxation times is then the level which provides the explanation or the*

Announcement of Upcoming FLUXNET Workshop......23

In This Issue:

Scaling Carbon and Water Fluxes from Patches to the Globe: A Challenge and an Opportunity for the Future

An Editorial by Dennis Baldocchi, Rodrigo Vargas and Laurie Koteen

explanatory level and the one with the long relaxation times, the level which is to be explained or the explainable level.'

The words of warning by Prof. deWit seem to have been forgotten because there has been rapid growth in the production of global scale products on carbon, water and energy exchange these past few decades. The emergence of new ecological scaling laws and databases on fluxes, climate and plant structure and function are helping us move forward to develop new and defensible global upscaling systems that transcends more than two 'levels'. Using these databases in concert to test, validate and parameterize flux upscaling exercises seems to be an avenue with much scientific potential.

While all the approaches may not be perfect, there remain many reasons to pursue this line of intellectual inquiry because the benefits will eventually outweigh the costs or limitations. First, a diverse 'modeling ecosystem' is warranted in this new era to find the best approach and to find out 'how good is good enough'. Success and improvement will evolve as we continue to learn by comparing the output of different models with one another and with measured flux data, by revising the models and expanding the FLUXNET database with more and newer data. Further, globally-gridded datasets of fluxes that are data-driven will have great utility to set priors for Bayesian models that predict carbon fluxes from the top down approach. We will also be able to produce grid averaged fluxes at the native resolution of climate models. And there is potential to mine the databases to extract information at regional and local scales that is pertinent for policy and management decisions, (eg. What is the potential GPP of an area for biomass fuel production? How much water is lost by a certain watershed? Is my land a carbon source or sink for selling carbon credits?) If anything we'd like to welcome these new gridded data products to the FLUXNET database for all to use, share and examine.

It is fair to note that we have to be conscious of and cautious about the potential pitfalls of any global modeling exercise. We should not be seduced by producing pretty global maps and global integrals because models can suffer from the pitfalls of 'garbage in/garbage out'. On the other hand, the field of global ecology would have never emerged if we had confined our scope to only two 'levels' in our models. Consequently, we would have remained intellectually stagnant if we blindly abide by the advice of dear Dr. deWit.

Contact:

Dennis Baldocchi
(baldocchi@berkeley.edu)

Dept of Environmental Science, Policy and Management
University of California, Berkeley, USA

References

Beer, C., M. Reichstein, E. Tomelleri, P. Ciais, M. Jung, N. Carvalhais, C. Rodenbeck, M. A. Arain, D. Baldocchi, G. B. Bonan, A. Bondeau, A. Cescatti, G. Lasslop, A. Lindroth, M. Lomas, S. Luyssaert, H. Margolis, K. W. Oleson, O. Roupsard, E. Veenendaal, N. Viovy, C. Williams, F. I. Woodward, and D. Papale, (2010) Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate. Science **329**, 834-838.

de Wit, C. T. (1970), Dynamic Concepts in Biology. Pages 17-23 *in* I. Malek, editor. Preditcion and measurement of photosynthetic productivity. Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands.

Jarvis, P. G. (1995) Scaling Processes and Problems. Plant Cell and Environment **18**, 1079-1089.

Jung, M., M. Reichstein, and A. Bondeau, (2009) Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. Biogeosciences **6**, 2001-2013.

Osmond, C. B. 1989. Photosynthesis from the molecule to the biosphere: a challenge for integration. Pages 5-17 *in* W. R. Briggs, editor. Photosynthesis. A. R. Liss, New York, NY.

Scaling Carbon and Water Fluxes from Patches to the Globe: A Challenge and an Opportunity for the Future

An Editorial by Dennis Baldocchi, Rodrigo Vargas and Laurie Koteen

Xiao, J., Q. Z, D. D. Baldocchi, B. E. Law, A. D. Richardson, J. Chen, R. Oren, G. Starr, A. Noormets, S. Ma, S. B. Verma, S. Wharton, S. C. Wofsy, P. V. Bolstad, S. P. Burns, D. R. Cook, P. S. Curtis, B. G. Drake, M. Falk, M. L. Fischer, D. R. Foster, L. Gu, J. L. Hadley, D. Y. Hollinger, G. G. Katul, M. Litvak, T. A. Martin, R. Matamala, S. G. McNulty, T. P. Meyers, R. K. Monson, J. W. Munger, W. C. Oechel, K. T. Paw U, H. P. Schmid, R. L. Scott, G. Sun, A. E. Suyker, and M. S. Torn (2008) Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data. Agricultural and Forest Meteorology **148**,1827-1847.

Yuan, W. P., Y. Q. Luo, A. D. Richardson, R. Oren, S. Luyssaert, I. A. Janssens, R. Ceulemans, X. H. Zhou, T. Grunwald, M. Aubinet, C. Berhofer, D. D. Baldocchi, J. Q. Chen, A. L. Dunn, J. L. Deforest, D. Dragoni, A. H. Goldstein, E. Moors, J. W. Munger, R. K. Monson, A. E. Suyker, G. Star, R. L. Scott, J. Tenhunen, S. B. Verma, T. Vesala, and S. C. Wofsy, (2009) Latitudinal patterns of magnitude and interannual variability in net ecosystem exchange regulated by biological and environmental variables. Global Change Biology **15**, 2905-2920.

FLUXNET - From Point to Globe

M. Reichstein, M. Jung, N. Carvalhais, M. Mahecha, C. Beer, E. Tomelleri

The advent of the eddy covariance method has revolutionized the science of terrestrial biosphere-atmosphere interactions, providing non-destructive measurements of carbon and energy fluxes which span across more than eight orders of magnitude in time. The power of the eddy covariance method has led to a rapid proliferation in world-wide research and organization, with more than 520 sites now registered in the FLUXNET network. In past decades, important ecosystem-physiological and eco-climatological questions were successfully addressed at single sites and also integrated across regional and global networks. However, one fundamental question has remained: Can we infer information on global land carbon and energy cycles from FLUXNET? Here, we illustrate recent advances in estimating global fields of carbon and water fluxes based on integration of FLUXNET with remote sensing and other earth observation data streams.

From one perspective, FLUXNET can be viewed as acupuncture of the biosphere; the global cumulative footprint area of all FLUXNET stations is not more than 0.003 ‰ of the terrestrial land surface. There is a gap of 4-5 orders of magnitude in space between a flux tower footprint and the global land surface (Fig. 1). Statistical sampling theory tells us that a few hundred samples allow for accurate estimates of the population mean, but only for well-behaved distributions, and only if the sample is representative of the population. The FLUXNET network is not fully representative of the terrestrial biosphere, however, as the sites are heavily clumped in areas with high population density and/or good infrastructure and research funding. Hence, deriving globally relevant information on biosphere-atmosphere exchanges remains a challenge. And yet, an opportunity exists to provide global coverage on carbon and energy exchange if generalized relationships between observed fluxes at the site level and corresponding variables with global coverage can be extracted and exploited. The key is in establishing the level of synergy that exists between global satellite remote sensing products, (many of which monitor relevant land surface properties at a spatial resolution compatible with flux tower footprints) and individual flux tower measurements. If synergy exists, the capability arises for bridging the gap between local measurements and globally relevant information (Fig. 1).

There are many different conceivable ways of integrating flux data and remote sensing products with models. For instance, both data sources can be used as constraints for

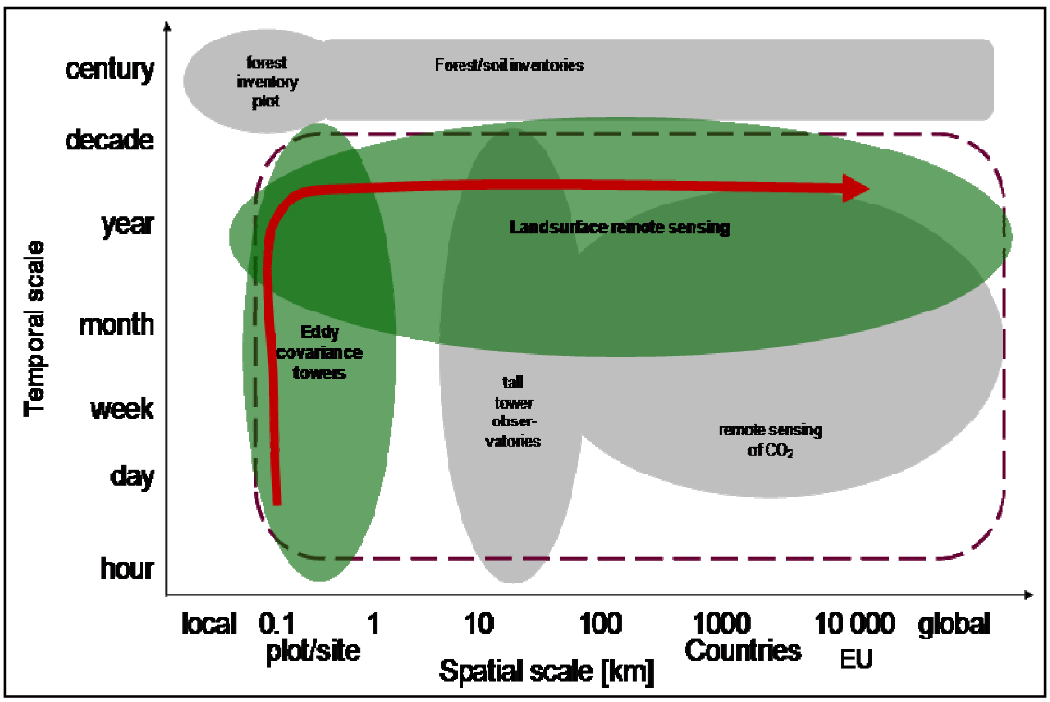

Figure 1: Observation systems spanning across time and space scales and the synergistic role of eddy covariance flux towers and remote sensing

FLUXNET - From Point to Globe

M. Reichstein, M. Jung, N. Carvalhais, M. Mahecha, C. Beer, E. Tomelleri

process-based models in a "data assimilation" approach. Results from this approach contain knowledge, or rather a conceptualization, of land surface processes and represent a compromise between model assumptions, prior parameters and (new) observations. Data-assimilation in this way is an effective tool for minimizing (but also quantifying) simulation uncertainties arising from uncertain parameter values, and as a means for assessing the adequacy of the model structure (Raupach *et al.* 2005; Williams *et al.* 2009).

A conceptually different and somewhat radical methodology is a fully data-oriented or data-adaptive approach: "let only the data speak". In this case, the modeler refrains from injecting strong theoretical assumptions, such as functional relationships between variables. Such an approach relies on machine learning algorithms and non-parametric statistical methods. It has only become possible through recent developments in computational statistics (Mjolsness & DeCoste 2001) and the increasing data richness in Earth Observation itself. Development and application of this methodology for observation of biosphere-atmosphere interactions is still in its infancy, but promising first examples exist at continental (Papale & Valentini 2003; Reichstein *et al.* 2007a; Jung *et al.* 2008; Xiao *et al.* 2008) and global scales (Jung *et al.* 2009; Beer *et al.* 2010; Jung *et al.* 2010). Most recently, we have combined flux tower data with meteorological and remote sensing observations in just such a data-driven up-scaling framework, and this approach has allowed us to generate a new global data stream for Earth system science. Currently, the newly created data stream comprises 27 years of monthly fluxes at 0.5° resolution for biosphere CO_2, H_2O, and energy fluxes, with the capability for yet greater spatial and temporal resolution enhancement.

Can we estimate the carbon balance of the land biosphere accurately from upscaling FLUXNET measurements? The short answer is 'no', or 'not as stand- alone product'. But perhaps in tandem with other earth observation data streams, we can. Currently, one major limitation to carbon estimation is the lack of global data on disturbance. We know that site history and past disturbances largely determine the mean carbon balance, and possibly other fluxes as well, at the local scale. While the FLUXNET network does include recently disturbed sites, upscaling capability is limited by the lack of a "scaling" variable (like fPAR for GPP and ET) which would allow for the diagnosis of disturbance at the global scale. Once a map of disturbance history and intensity (fire, wind throw, harvests) and structural information (e.g. from LiDAR) is available globally, more realistic global estimates of the global carbon balance might be possible. Other problems relate to the precision of the eddy-covariance method. The net global terrestrial carbon uptake inferred from

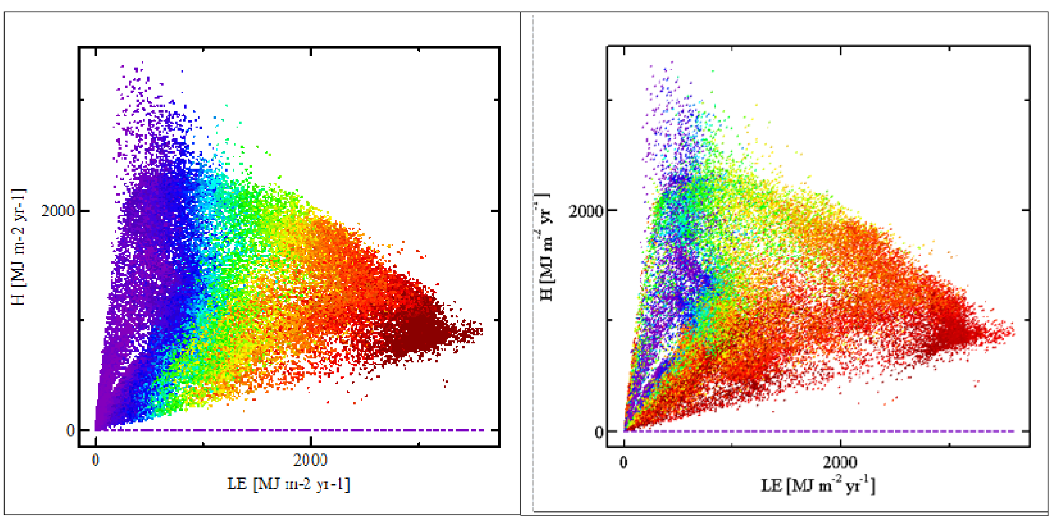

Figure 2: Global patterns of empirically derived annual biosphere-atmosphere flux integrals. Both panels show sensible versus latent heat fluxes, where colors code gross primary productivity and water-use efficiency on the left and right-hand side, respectively (red: high, violet: low). Derived from flux maps in Jung et al. submitted to JGR-Biogeosciences (special issue).

FLUXNET - From Point to Globe

M. Reichstein, M. Jung, N. Carvalhais, M. Mahecha, C. Beer, E. Tomelleri

BOX 1: Principles of empirical up-scaling

The general procedure behind empirical upscaling begins with formation of a “training data set” which contains the variable of interest as a target variable, and an array of potential explanatory variables that must also be available as global grids. Once equipped with this training data set, machine learning algorithms “learn” how to map the explanatory variables to the target variable. The result of this training phase is an empirical model which is then applied at the global scale using the global gridded explanatory variables. The methods of constructing the empirical model depend on the type of learning machine used, e.g. artificial neural networks, support vector machines, regression and model trees. Ensembles of empirical models generally yield more accurate results, and methods of creating ensembles are an important aspect too. The search for a regression model suitable for global upscaling purposes remains a great playground and we should strive for systematic tests to find the most suitable approaches. Box 2 outlines briefly the method called model tree ensembles (MTE) that we have followed.

atmospheric data is around 2 PgC/yr, which equals 15 gC/m^2/yr if we distribute these 2 PgC/yr equally over the global vegetated area. As currently deployed, however, this is likely below the detection limit of the eddy covariance method. Hence, instead of hunting for the global carbon balance, it is more rewarding to extract robust global multivariate patterns of the relationships between the variables that describe biosphere-atmosphere exchange, (*i.e.* Fig. 2). This approach enables insights into eco-climatological problems, biosphere-atmosphere feedbacks and eventually to improved climate-carbon cycle models.

Integration of remote sensing data into this effort is also providing new opportunities for coupled climate-biosphere model evaluation that circumvents some of the conceptual problems with direct model-site comparisons This is especially true when coupled with the machine-learning based derivation of generalized relationships. Advantages of this approach include: 1) Upscaled fields are integrated over the grid-box heterogeneity, hence the point-to -grid-box scale mismatch is largely avoided, 2) the generalization reduces the importance of site particularities which might not be relevant for global models, 3) the representativeness problem of FLUXNET is greatly reduced (e.g. that site based model comparisons are heavily biased towards temperate sites). Still, it is clear that these data-driven up-scaling products also depend on some assumptions and are subject to several conceptual limitations (Box 2). For example, the choice of potentially relevant predictor variables is partly subjective and constrained by the availability and quality of corresponding global data sets. Further, the prediction of carbon fluxes in response to rare or previously unseen conditions (e.g. climate extremes, where no training data are available), remains an extrapolation, and as such, is highly uncertain. In addition, results from a machine learning approach in land areas, vegetation types and climate conditions that are not well sampled by FLUXNET measurements might still be heavily influenced by single sites. Hence, these products should not be taken as truth and used blindly for biosphere model evaluation. Any mismatch between biosphere and data-oriented models should be critically examined for respective weaknesses. Nevertheless, we have gained confidence in the FLUXNET derived global data streams based on various consistency checks that have been performed with independent approaches (Jung *et al.* 2009; Beer *et al.* 2010; Jung *et al.* 2010; Jimenez *et al.* 2010 (in press)).

In summary, the fusion of site level flux data with satellite-based Earth Observation offers great opportunities for increasing the relevance of FLUXNET in Earth System science. Nevertheless, we must keep in mind that the aforementioned global data-driven up-scaling products depend on many assumptions, are subject to several limitations, and should always be critically

FLUXNET - From Point to Globe

M. Reichstein, M. Jung, N. Carvalhais, M. Mahecha, C. Beer, E. Tomelleri

BOX 2: Example: Model Tree Ensembles

Model trees are decision trees that split the training data set hierarchically into smaller units, or leaves, based on a series of decision steps. Within each leaf, a multiple linear regression is used to describe the mapping of explanatory variables to the target variable. The decision steps are determined by finding the split variable (i.e. a variable by which the data set can be stratified), and value for which the sum of squared errors of both sub-domains is minimal. Tree growth continues as long as the Bayesian information criterion decreases. The criterion is calculated from the mean squared error of the model as determined by cross-validation in the leaves, and plus a term penalizing for the number of model parameters given a number of samples. In our group, we created ensembles of model trees using an evolutionary method. This method takes an existing tree, removes part of it, and then grows the tree for subsequent steps based on random split decisions. Model trees have the advantage of inherent stratification. Stratification is helpful because we expect that the relevance of each controlling factor, and its sensitivity to the target variable, changes among different environmental domains (e.g. tropical vs temperate, trees vs grasses, summer vs winter, …).

evaluated against independent approaches. Even so we think that the FLUXNET derived global data stream contains robust patterns which can be used to constrain and reduce the uncertainty of process-based biosphere models. Future directions include integration of information at scales not considered so far(e.g. organism or landscape), other trace gases and isofluxes, hyperspectral and structural remote sensing. Central to this endeavor will be the need to embrace novel developments in the field of machine learning and empirical inference in general (e.g. detection of causality). This effort will be greatly facilitated by a spirit of open data and code sharing across disciplines.

Contact

Markus Reichstein
(markus.reichstein@bgc-jena.mpg.de)

Max Planck Institute for Biogeochemistry
Munich, Germany

References

Beer C., Reichstein M., Tomelleri E., Ciais P., Jung M., Carvalhais N., Rodenbeck C., Arain M.A., Baldocchi D., Bonan G.B., Bondeau A., Cescatti A., Lasslop G., Lindroth A., Lomas M., Luyssaert S., Margolis H., Oleson K.W., Roupsard O., Veenendaal E., Viovy N., Williams C., Woodward F.I. & Papale D. (2010) Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate. *Science*, **329**, 834-838

Jimenez C., Prigent C., Mueller B., Seneviratne S.I., McCabe M.F., Wood E.F., Rossow W.B., Balsamo G., Betts A.K., Dirmeyer P.A., Fisher J.B., Jung M., Kanamitsu M., Reichle R.H., Reichstein M., Rodell M., Sheffield J., Tu K. & Wang K. (2010 (in press)) Global inter-comparison of 12 land surface heat flux estimates. *journal of Geophysical Research - Atmospheres*

Jung M., Reichstein M. & Bondeau A. (2009) Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, **6**, 2001-2013

Jung M., Reichstein M., Ciais P., Seneviratne S.I., Sheffield J., Goulden M.L., Bonan G., Cescatti A., Chen J., de Jeu R., Dolman A.J., Eugster W., Gerten D., Gianelle D., Gobron N., Heinke J., Kimball J., Law B.E., Montagnani L., Mu Q., Mueller B., Oleson K., Papale D., Richardson A.D., Roupsard O., Running S., Tomelleri E., Viovy N., Weber U., Williams C., Wood E., Zaehle S. & Zhang K. (2010) Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, **467**, 951–954

Jung M., Verstraete M., Gobron N., Reichstein M., Papale D., Bondeau A., Robustelli M. & Pinty B. (2008) Diagnostic assessment of European gross primary production. *Global Change Biology*, **14**, 2349-2364

BOX 3: Conceptual Limitations

Many conceptual limitations are inherent to the approach we have outlined. For one, the training dataset is composed of time-series data from many different sites, while information about time and site is generally not explicitly considered in the learning machine. There are several conceptual problems with that: 1) mixing of spatial and temporal scales, 2) time-scale specific relations among variables, 3) time independence and dynamic effects. The 'time for space' substitution approach is inherently used. However, relations in the temporal and spatial domain are likely different from each other, in particular if the spatial and temporal scales in the training data set are decoupled (e.g. half-hourly data of globally distributed sites). We also anticipate that the relationships among variables are time scale specific and that therefore in the training phase on time series that capture several orders of magnitude of temporal scale sensitivities of some variables that are confounded by effects of scale may arise [Mahecha et al., 2010]. The effect of 'mixing scales' may actually be reduced, however, if explanatory variables are provided that allow for stratification of the training data set, which implicitly accounts for such confounding effects in part. For example, if we allow the algorithm to split the training data set along relevant variables that vary in space, but not in time, (e.g. mean annual temperature, mean annual precipitation), or along variables that vary in time, but not in space, (e.g. mean seasonal cycle of meteorological variables), the confounding effects of mixing scales can be reduced. Another conceptual problem stems from the time-independence of upscaled results. Under this scheme, memory effects of, for example, extreme events, are not explicitly accounted for. Some implicit accounting is likely, however, if the memory effects are incorporated by remote sensing indices. Moreover, memory effects may also be captured if we include lagged explanatory variables in the training data set. Despite these ameliorating factors, however, the fact remains that we currently lack systematic means to test the degree to which problems of mismatched scales and memory effects remain. Addressing these issues further, and finding ways to circumvent or integrate them, should be next steps in upscaling exercises.

Mjolsness E. & DeCoste D. (2001) Machine learning for science: State of the art and future prospects. *Science*, **293**, 2051-2055

Papale D. & Valentini R. (2003) A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization. *Global Change Biology*, 9, 525-535

Raupach M.R., Rayner P.J., Barrett D.J., DeFries R.S., Heimann M., Ojima D.S., Quegan S. & Schmullius C.C. (2005) Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. *Global Change Biology*, **11**, 378-397

Reichstein M., Ciais P., Papale D., Valentini R., Running S., Viovy N., Cramer W., Granier A., Ogée J., Allard V., Aubinet M., Bernhofer C., Buchmann N., Carrara A., Grünwald T., Heimann M., Heinesch B., Knohl A., Kutsch W., Loustau D., Manca G., Matteucci G., Miglietta F., Ourcival J.M., Pilegaard K., Pumpanen J., Rambal S., Schaphoff S., Seufert G., Soussana J.-F., Sanz M.-J., Vesala T. & Zhao M. (2007a) Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modeling analysis. *Global Change Biology*, **13**, 634–651

Williams M., Richardson A.D., Reichstein M., Stoy P.C., Peylin P., Verbeeck H., Carvalhais N., Jung M., Hollinger D.Y., Kattge J., Leuning R., Luo Y., Tomelleri E., Trudinger C.M. & Wang Y.-P. (2009) Improving land surface models with FLUXNET data. *Biogeosciences*, **6**, 1341-1359

Xiao J.F., Zhuang Q.L., Baldocchi D.D., Law B.E., Richardson A.D., Chen J.Q., Oren R., Starr G., Noormets A., Ma S.Y., Verma S.B., Wharton S., Wofsy S.C., Bolstad P.V., Burns S.P., Cook D.R., Curtis P.S., Drake B.G., Falk M., Fischer M.L., Foster D.R., Gu L.H., Hadley J.L., Hollinger D.Y., Katul G.G., Litvak M., Martin T.A., Matamala R., McNulty S., Meyers T.P., Monson R.K., Munger J.W., Oechel W.C., U K.T.P., Schmid H.P., Scott R.L., Sun G., Suyker A.E. & Torn M.S. (2008) Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data. *Agricultural and Forest Meteorology*, **148**, 1827-1847

Global Remote Sensing in a PC: Cloud Computing as a New Tool to Scale Land Surface Fluxes from Plot to the Globe

Youngryel Ryu, Jie Li, Catharine van Ingen, Deb Agarwal, Keith Jackson, You-Wei Cheah, Marty Humphrey

When YR arrived at the UC Berkeley in fall of 2006, his adviser offered a PC. That was the only computing resource. YR received the NASA fellowship with a proposal on the global mapping of GPP and ET using MODIS. At that time, YR did not realize how much that project would require computing resources. Two years later, YR learned that a one year global mapping calculation requires 2.6 TB of source data, 20 days for source data download from NASA ftp servers, 2,000 days for reprojection, and 15 days for ET and GPP calculations. How could this possibly be done with a PC before YR gets his PhD degree? The answer is cloud computing.

1. Overview of the MODIS-Azure system

Cloud computing offers the potential for internet-based, on-demand, and highly scalable services. We built the MODIS-Azure service on the Microsoft cloud computing platform, Azure, to process large satellite data (Li et al., 2010). The basic idea is “to download data within the Cloud, to process/ analyze data in the Cloud, and download results to my PC” (Figure 1). The MODIS-Azure web portal allows 1) submitting job requests and 2) monitoring the processing job status in real-time. One can request the number of virtual machines (1-250 VMs currently) depending on the estimated computing needs, which offers highly

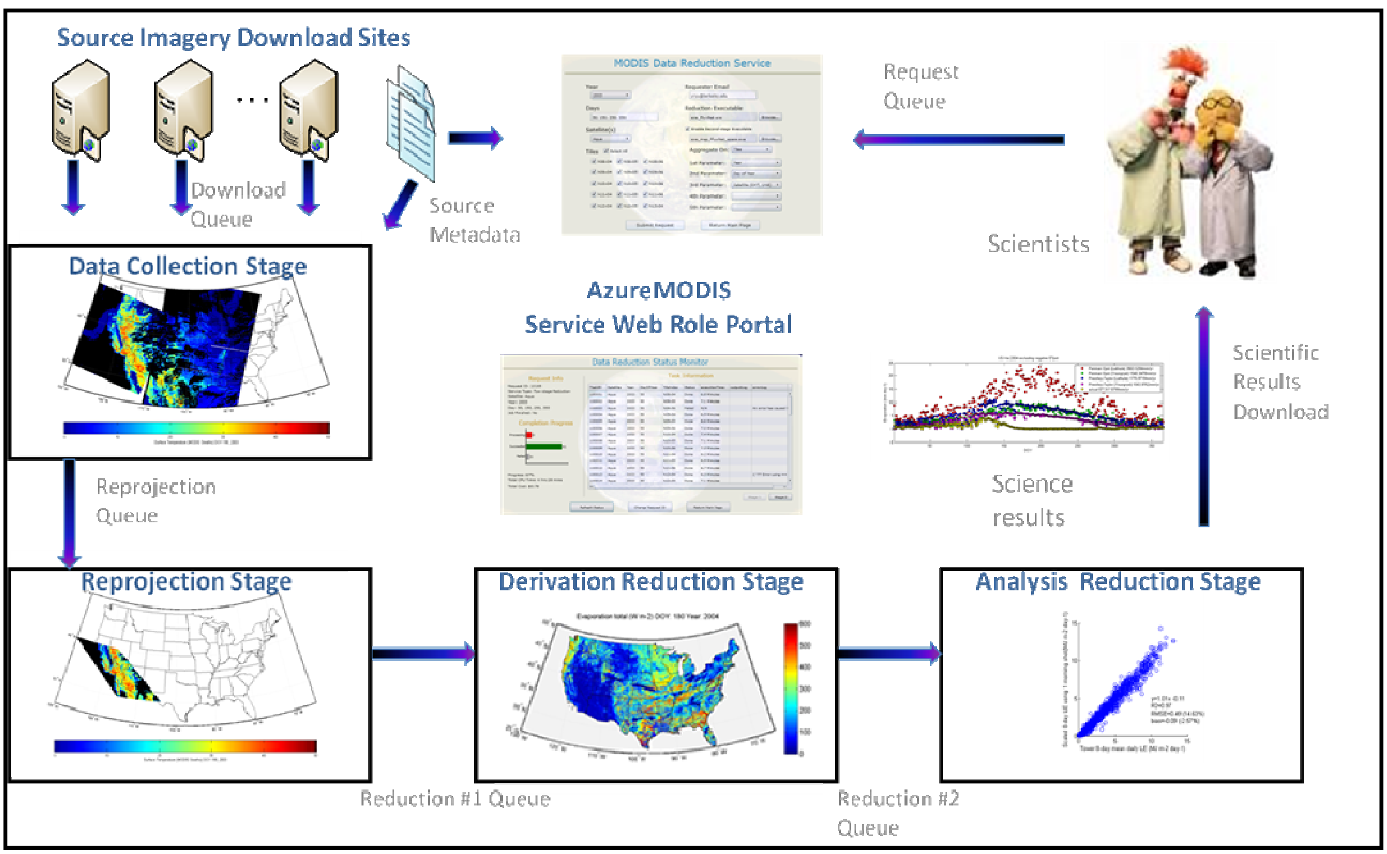

Figure 1: Work flow of MODIS-Azure system (http://research.microsoft.com/en-us/projects/azure/azuremodis.aspx)

Global Remote Sensing in a PC: Cloud Computing as a New Tool to Scale Land Surface Fluxes from Plot to the Globe

Youngryel Ryu, Jie Li, Catharine van Ingen, Deb Agarwal, Keith Jackson, You-Wei Cheah, Marty Humphrey

scalable performance. The service performs data download, reprojection, resampling, analysis, and emails web links, that include final results, to the user. Anyone with authorized credentials and internet access can use the service.

2. Overcoming barriers for global remote sensing study with the MODIS-Azure system

We identified three barriers for the global remote sensing study to overcome. ***First, data download***. NASA supports several web portals that provide various satellite data, but the download processing is really tedious (eg. with numerous clicks) and time-consuming. Our contribution has been to revamp this process. Now, we determine and automatically down-load needed files from the NASA ftp servers, and the data download runs in the background of all the VMs. ***Second, data heterogeneity.*** MODIS provides land, atmosphere, and ocean disciplines, and all have different data structures. There are pressing needs to develop high resolution meteorological forcing variables (e.g. radiation, temperature, and humidity) which are compatible with the resolutions of other key variables like LAI, leaf area index and VI (vegetation index). Several studies have demonstrated that the integration of MODIS atmospheric and land products can provide reliable land surface radiation components (Houborg and Soegaard, 2004; Ryu et al., 2008; Van Laake and Sanchez-Azofeifa, 2004). However, to the best of our knowledge, there has been no study that integrates the MODIS atmospheric and land products at continental or global scales because of the incongruity of the data types. The former uses swath format (following the satellite in latitude, longitude) whereas the latter uses sinusoidal projections. To couple the two disciplines, both should have the same projection and spatial resolution, which requires large computational resources and careful geospatial lookup. With the MODIS-Azure system that we developed, we are able to reproject the MODIS atmospheric products to the sinusoidal projection at a compatible (1 km) resolution. We follow the same tile convention of MODIS land products (h and v). For example, when the user submits a reprojection request for the h08v05 tile for a certain date, the MODIS-Azure system accesses the NASA ftp servers, searches the swaths that cover the tile (Hua et al., 2007), reprojects the swaths to the sinusoidal projection, and finally, generates the h08v05 tile of atmospheric variables (see the left figures in the Fig 1). With a single researcher and PC, the reprojection process alone took the largest computation time (2,000 days for 1 year of global reprojections). ***Third, data analysis***. The co-located land and atmospheric products provided rich forcing data globally. The next barrier we faced was how to analyze the data. To facilitate analysis, we developed a 2-step reduction analysis tool in the MODIS-Azure service. For example, for global tile coverage, in the 1st reduction step, one could integrate all tiles over the globe, and in the 2nd step, one could integrate the global data over the weeks, months, or years. The two stage analysis enables a very large scale computation to be processed in the cloud that produces results that can be readily examined on a local PC.

3. Observing the biosphere with the Breathing Earth Simulator System (BESS)

Here, we briefly introduce a coupled gross primary productivity (GPP) and evapotranspiration (ET) model based on the MODIS cloud computing platform we originated (Ryu et al. In preparation). We developed a two-leaf (sunlit-shade), dual-source (vegetation and soil) model. The model first calculates instantaneous radiation components under all sky conditions at 1 km resolution including incoming shortwave, beam photosynthetically active radiation (PAR) and diffuse PAR (Iwabuchi, 2006; Kobayashi and Iwabuchi, 2008), incoming longwave (Prata, 1996), outgoing longwave (Wan, 2008) and outgoing shortwave radia-

Global Remote Sensing in a PC: Cloud Computing as a New Tool to Scale Land Surface Fluxes from Plot to the Globe

Youngryel Ryu, Jie Li, Catharine van Ingen, Deb Agarwal, Keith Jackson, You-Wei Cheah, Marty Humphrey

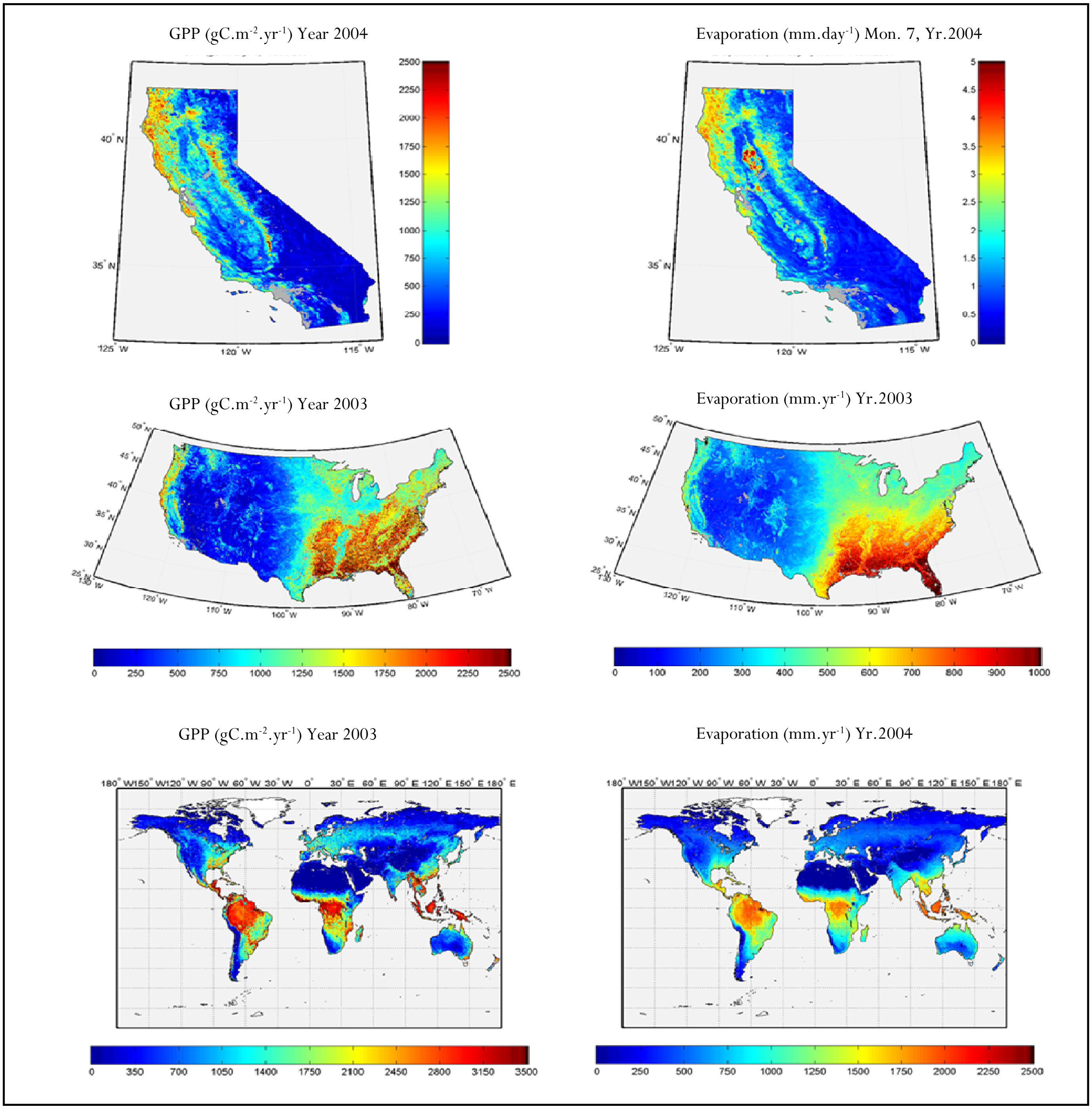

Figure 2: GPP and Evapotranspiration maps from California, conterminous US, and the globe

Global Remote Sensing in a PC: Cloud Computing as a New Tool to Scale Land Surface Fluxes from Plot to the Globe

Youngryel Ryu, Jie Li, Catharine van Ingen, Deb Agarwal, Keith Jackson, You-Wei Cheah, Marty Humphrey

us to calculate instantaneous net radiation (R_n) over all sky conditions. We used MODIS albedo data to estimate percent leaf N (Ollinger et al., 2008), which we converted to an areal based -N estimate using a leaf mass per area data product based on a global dataset of leaf traits (Wright et al., 2004). Then we quantified the maximum carboxylation rate, (Vcmax), at 25°C using the nitrogen use efficiency (i.e. Vcmax/N (area based)) (Kattge et al., 2009). The detailed radiation components and Vcmax information allowed us to apply Farquhar's photosynthesis model to the two leaves separately (dePury and Farquhar, 1997). In this model, the two-leaf photosynthesis and two-leaf energy balance are iteratively solved until roots converge. The Ball-Berry equation calculates the two-leaf canopy conductance (Collatz et al., 1991). To calculate canopy evapotranspiration, we applied the quadratic form of the Penman Monteith equation for the two-leaf types separately (Paw U and Gao, 1988). The soil evaporation was calculated as the equilibrium evapotranspiration constrained by a water stress factor (RH^{VPD}) (Fisher et al., 2008). The modeled GPP and ET are instantaneous values, and we upscaled the "snap-shot" values to the daily integrals using the ratio of snap-shot solar radiation to the daily sum potential solar radiation (Ryu et al. To be submitted). The test against 34 fluxnet sites proved the efficacy of this approach. This state-of-the-art biophysical model coupled with the temporal and spatial upscaling strategy in the MODIS-Azure service enabled us to upscale fluxes across a range of scales. Some examples are included in Fig. 2. First, it shows the wave of green growth that migrates from south to north during the northern hemisphere spring. Second, it identifies areas of intense agricultural activity across the Midwest and along the Mississippi River and reveals that they have lower potential for assimilating carbon compared to forest ecosystems. Third, it shows the hotspots of ET associated with wide scale rice production in the Sacramento Valley and forested mountain ranges like the Black Hills which are surrounded by drier grasslands. Fourth, it identifies the effect of climatic anomalies like wide spread heat spells and droughts across wide regions of forest and agricultural land.

4. Conclusion

Many environmental scientists do not have access to super computers or even small cluster computers. This limit in the computational resources has hampered data-intensive science. We demonstrated that the synergystic collaboration between computer and environmental scientists can advance our ability to process and analyze complex terabyte size data sets. This is the first study that has used the cloud computing system in environmental science, and we believe this novel tool could facilitate data intensive science.

Contact:

Youngryel Ryu
(ryuyr77@gmail.com)

Department of Organismic and Evolutionary Biology, Harvard University
Cambridge, MA, USA

References

Collatz, G.J., Ball, J.T., Grivet, C. and Berry, J.A., (1991) Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. Agric. For. Meteorol., **54**, 107-136.

dePury, D.G.G. and Farquhar, G.D., (1997) Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. Plant Cell and Environment **20** (5): 537-557.

Fisher, J.B., Tu, K.P. and Baldocchi, D.D., (2008) Global estimates of the land-atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. Remote Sens. Environ., 112(3): 901-919

Houborg, R.M. and Soegaard, H., (2004) Regional simulation of ecosystem CO2 and water vapor exchange for agricultural land using NOAA AVHRR and Terra MODIS satellite data. Application to Zealand, Denmark. Remote Sens. Environ., 93: 150-

Global Remote Sensing in a PC: Cloud Computing as a New Tool to Scale Land Surface Fluxes from Plot to the Globe

Youngryel Ryu, Jie Li, Catharine van Ingen, Deb Agarwal, Keith Jackson, You-Wei Cheah, Marty Humphrey

167.

Hua, X.M., Pan, J.F., Ouzounov, D., Lyapustin, A., Wang, Y.J., Tewari, K., Leptoukh, G. and Vollmer, B., (2007) A spatial prescreening technique for earth observation data. IEEE Geoscience and Remote Sensing Letters, **4**(1) 152-156.

Iwabuchi, H., (2006) Efficient Monte Carlo methods for radiative transfer modeling. J. Atmos. Sci., **63**(9) 2324-2339.

Kattge, J., Knorr, W., Raddatz, T. and Wirth, C., (2009) Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. Global Change Biol., **15**, 976 - 991.

Kobayashi, H. and Iwabuchi, H., (2008) A coupled 1-D atmosphere and 3-D canopy radiative transfer model for canopy reflectance, light environment, and photosynthesis simulation in a heterogeneous landscape. Remote Sens. Environ., **112**(1) 173-185.

Li, J., Humphrey, M., Agarwal, D., Jackson, K., van Ingen, C. and Ryu, Y., (2010) eScience in the cloud: A MODIS satellite data reprojection and reduction pipeline in the Windows Azure platform, Parallel & Distributed Processing (IPDPS), (2010) IEEE International Symposium on, pp. 1-10.

Ollinger, S.V., Richardson, A.D., Martin, M.E., Hollinger, D.Y., Frolking, S.E., Reich, P.B., Plourde, L.C., Katul, G.G., Munger, J.W., Oren, R., Smith, M.L., Paw U, K.T., Bolstad, P.V., Cook, B.D., Day, M.C., Martin, T.A., Monson, R.K. and Schmid, H.P., (2008) Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. Proc. Natl. Acad. Sci. U. S. A., **105**(49) 19336-19341.

Paw U, K.T. and Gao, W.G., (1988) Applications of solutions to non-linear energy budget equations. Agric. For. Meteorol., **43** (2) 121-145.

Prata, A.J., (1996) A new long-wave formula for estimating downward clear-sky radiation at the surface. Q. J. R. Meteorol. Soc., **122**(533) 1127-1151.

Ryu, Y., Kang, S., Moon, S.K. and Kim, J., (2008) Evaluation of land surface radiation balance derived from Moderate Resolution Imaging Spectrometer (MODIS) over complex terrain and heterogeneous landscape on clear sky days. Agric. For. Meteorol., **148**(10) 1538-1552.

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y. and Muller, J.-P., (2002) First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sens. Environ.,**83**(1-2) 135-148.

Van Laake, P.E. and Sanchez-Azofeifa, G.A., (2004) Simplified atmospheric radiative transfer modelling for estimating incident PAR using MODIS atmosphere products. Remote Sens. Environ., **91**(1) 98-113.

Wan, Z.M., (2008) New refinements and validation of the MODIS Land-Surface Temperature/Emissivity products. Remote Sens. Environ., **112** (1) 59-74.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J. and Villar, R., (2004) The worldwide leaf economics spectrum. Nature, **428**(6985) 821-827.

Highlight Young Scientist

Rodrigo Vargas

I was born in a "big" town but I always had the opportunity to explore the outdoors. This may not be common for every kid in Mexico City, but having an archeologist and anthropologist for parents made the outdoors the primary activity for weekends and holidays. As an undergraduate I worked on the weekends as a river rafting guide and later on there were always excuses to travel and see the outdoors.

As an undergraduate, I studied biology at the *Universidad Nacional Autónoma de México*. My undergraduate thesis focused on understanding the spatial and temporal variation of nitrogen fixation in a tropical wetland, and inspired me to pursue a career in environmental science.

I was fortunate to be awarded the Fulbright (US government) and CONACyT (Mexican government) scholarships to pursue a PhD at the *University of California, Riverside*. I worked with Michael Allen, as my PhD advisor, and in his lab, I was given the opportunity to explore several projects. During my first year I started a "side project" working with wireless sensors in collaboration with the Center for Embedded Networked Sensing (CENS; http://research.cens.ucla.edu/) at the *University of California, Los Angeles*. The goal of my project was to monitor, in high temporal resolution, root and mycorrhizae dynamics and to couple these observations with those from a large network of soil CO_2 efflux measurements in a temperate forest. We found that mycorrhizae turnover was faster than expected and correlated well with soil CO_2 efflux (Vargas & Allen, 2008b).

Rodrigo Vargas

My primary PhD project brought me to study carbon dynamics in a tropical dry forest in the Yucatan Peninsula in Mexico. I began by combining biometric forest measurements and experimental forest management treatments to demonstrate the large capability of these forests to store carbon (Vargas et al., 2008). Then, just when I thought everything was under control and all the experiments were ready, hurricane Wilma passed over the study site, causing large damage. Thus, I had the opportunity to study the effects of an extreme event on the carbon dynamics of a tropical forest. In the wake of this event, we measured soil CO_2 efflux continuously and reported emissions >3000 gC m^2 $year^{-1}$ (Vargas & Allen, 2008a), which are the highest so far recorded from soils (Bond-Lamberty & Thomson, 2010). Through a radiocarbon analysis we also demonstrated the unexpected capacity of trees to allocate old stored carbon (>10 yrs old) to produce fine roots (and probably sustain mycorrhizal fungi colonization as well) following a hurricane disturbance (Vargas et al., 2009).

After finishing my PhD, I worked with Dennis Baldocchi at the *University of California, Berkeley*, and for three years I had the opportunity to interact with FLUXNET scientist around the world. Many of these collaborations moved from synthesis studies to friendships and future research projects. One unor-

Highlight Young Scientist

Rodrigo Vargas

thodox FLUXNET synthesis activity that Dennis let me pursue was to test a theory developed for the global distribution of mycorrhizal fungi (Read & Perez-Moreno, 2003) with flux tower data. We found that tower data supports and complements the hypothesis that bioclimatic conditions influence the global distribution of mycorrhizae, but there are differences in the major climatic factors controlling ecosystem CO_2 fluxes (Vargas et al., 2010). As flux data continues to grow one can only imagine the opportunities that will arise for the scientific community with this global dataset.

At present, I am an Assistant Research Professor at the *Centro de Investigación Científica y de Educación Superior de Ensenada* (CICESE); a top national research center in Mexico. Currently, we have an eddy covariance tower in chaparral vegetation and we are planning to expand the network to a few more sites along a precipitation gradient in the Baja California peninsula. Furthermore, we are working in collaboration with Mexican scientists to create a regional flux network (MexFlux), and perhaps FLUXNET will soon see a few more dots on the global flux tower map. That said, we always welcome collaborations and exchange of ideas from the community, and we are working hard to make FLUXNET a global network to see the "breathing of the Earth".

Finally, I want to thank all my colleagues who have contributed to FluxLetter. Without your enthusiasm and support, this project could not have been possible. For me, it is time to pass on the editorial responsibilities, but I encourage all scientists and especially those "young scientists" to keep participating to make FluxLetter a powerful information, networking, and communication resource for the community….*muchas gracias*.

Contact

Rodrigo Vargas
(rvargas@cicese.mx)
Centro de Investigación
Científica y de Educación
Superior de Ensenada
Department of Conservation Biology
Tijuana, Mexico

http://nature.berkeley.edu/~rvargas

References

Bond-Lamberty B, Thomson A (2010) Temperature-associated increases in the global soil respiration record. *Nature,* **464**, 579-U132.

Read Dj, Perez-Moreno J (2003) Mycorrhizas and nutrient cycling in ecosystems - a journey towards relevance? *New Phytologist,* **157**, 475-492.

Vargas R, Allen Mf (2008a) Diel patterns of soil respiration in a tropical forest after Hurricane Wilma. *Journal of Geophysical Research-Biogeosciences,* **113**, G03021.

Vargas R, Allen Mf (2008b) Environmental controls and the influence of vegetation type, fine roots and rhizomorphs on diel and seasonal variation in soil respiration. *New Phytologist,* **179**, 460-471.

Vargas R, Allen Mf, Allen Eb (2008) Biomass and carbon accumulation in a fire chronosequence of a seasonally dry tropical forest. *Global Change Biology,* **14**, 109-124.

Vargas R, Baldocchi DD, Querejeta JI *et al.* (2010) Ecosystem CO2 fluxes of arbuscular and ectomycorrhizal dominated vegetation types are differentially influenced by precipitation and temperature. *New Phytologist,* **185**, 226-236.

Vargas R, Trumbore Se, Allen Mf (2009) Evidence of old carbon used to grow new fine roots in a tropical forest. *New Phytologist,* **182**, 710-718.

Upscaling Fluxes from Towers to Regional, Continental and Global Scales Using Data-driven Approaches

Jingfeng Xiao

Eddy covariance flux towers have been providing continuous, high-frequency measurements of whole-ecosystem carbon and water fluxes since the early 1990s (Wofsy et al., 1993; Baldocchi et al. 2001). At present, over 500 flux towers are operating on a long-term and continuous basis over the globe. This global network, FLUXNET, encompasses a large range of climate and biome types (Baldocchi et al., 2001). However, these measurements only represent the fluxes at the scale of the tower footprint. To quantify the net exchange of carbon and water between the terrestrial biosphere and the atmosphere, we need to upscale these flux measurements to regions, continents, and the globe (Running et al. 1999; Davis 2008).

We used a data-driven approach to upscale carbon fluxes from the AmeriFlux network to the continental scale and produced gridded fields of gross primary productivity (GPP) and net ecosystem carbon exchange (NEE) with high spatial (1km) and temporal (8-day) resolutions for temperate North America over the period 2000-2006 (Xiao et al., 2008, 2010a,b). Our GPP and NEE fields were derived from eddy covariance (EC) flux measurements and MODIS data, and are referred to as EC-MOD hereafter. We used our continuous flux fields to assess the magnitude, distribution, and interannual variability of recent U.S. ecosystem carbon exchange (Figure 1). One of the main innovations in our estimates is the use of daily NEE measurements from flux towers. These measurements represent direct samples of net CO_2 exchange from sites encompassing a wide variety of U.S. biomes and climate types, which have not been previously utilized in U.S. carbon budget studies (e.g., Houghton et al., 1999; Caspersen et al., 2000; Schimel et al., 2000; Pacala et al., 2001; SOCCR, 2007). Our analysis provides an alternative, independent, and novel perspective on recent U.S. ecosystem carbon exchange (Xiao et al. 2010a,b). We are currently using this data-driven approach to upscale FLUXNET data to the global scale to examine the patterns and trends of water use efficiency.

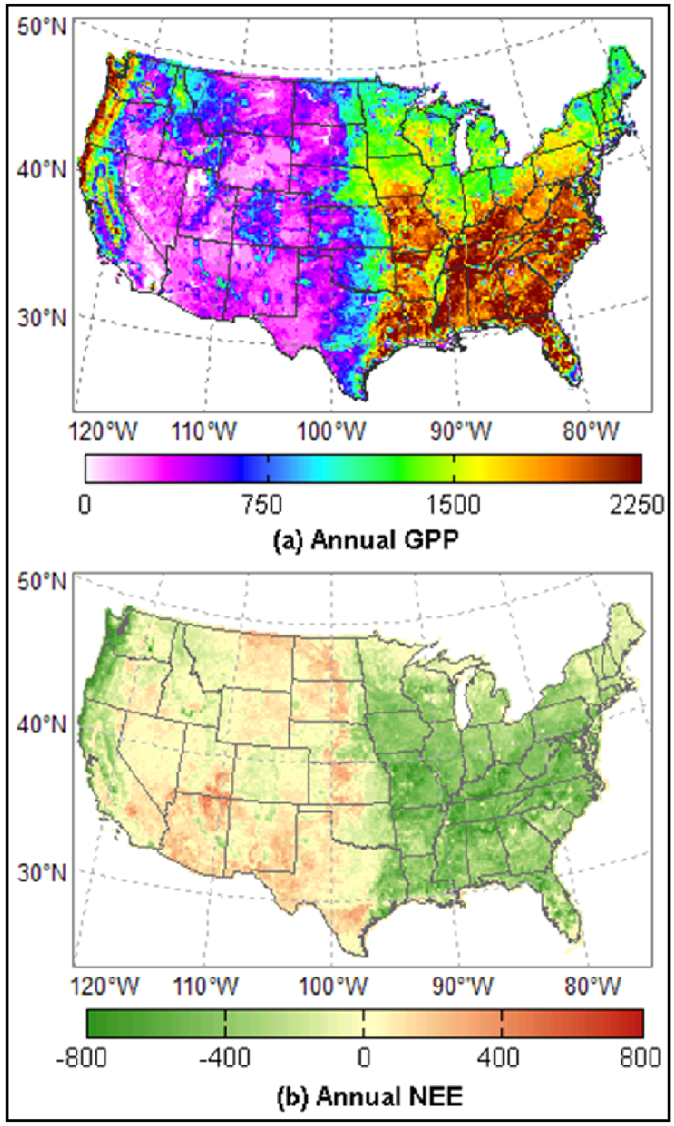

Figure 1: Annual GPP and NEE for the conterminous United States in 2006 (Xiao et al. 2010a,b). The units are

Upscaling Fluxes from Towers to Regional, Continental and Global Scales Using Data-driven Approaches

Jingfeng Xiao

Data-driven Approach

Figure 1. Annual GPP and NEE for the conterminous United States in 2006 (Xiao et al. 2010a,b). The units are g C m^{-2} yr^{-1}.

Our EC-MOD system upscales fluxes from tower footprint to regional, continental or global scales to produce gridded flux fields over these broad regions (Figure 2). The core of our system is a data-driven approach based on a modified regression tree method (Xiao et al. 2008). This approach relies on rule-based models, each of which is a set of conditions associated with a multivariate linear submodel. These rule-based, piecewise regression models allow both numerical (e.g., carbon fluxes, temperature, vegetation index) and categorical variables (e.g., land cover type) as input variables, and account for possible nonlinear relationships between predictive and target variables.

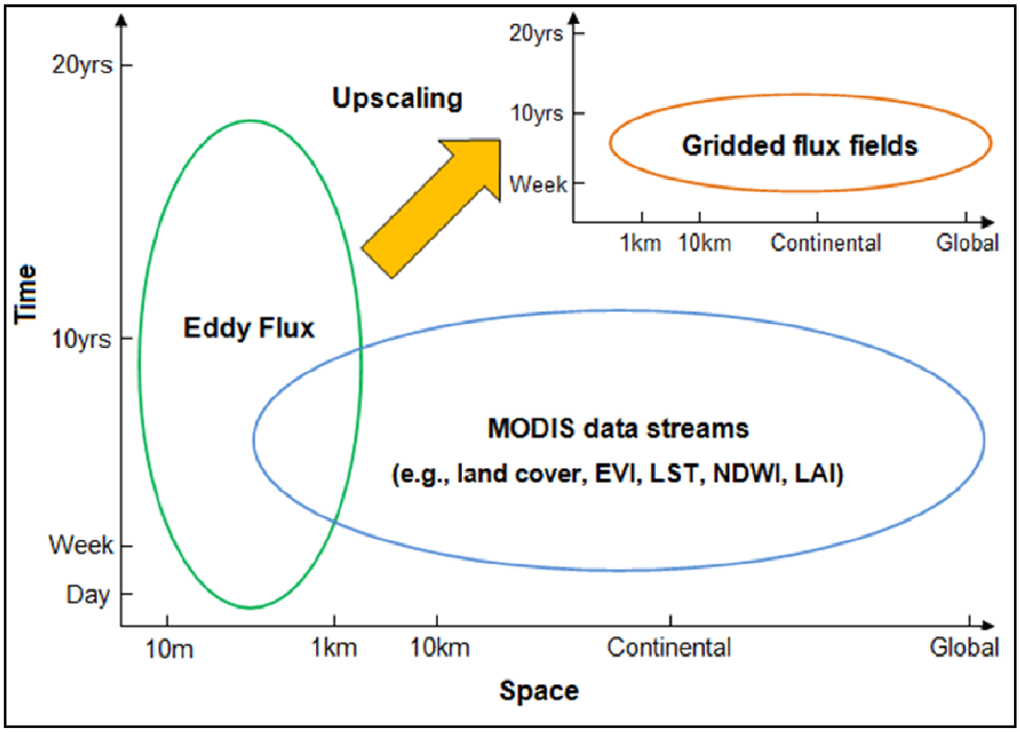

Figure 2: Conceptual framework of data-driven approaches for upscaling fluxes from tower footprint to regional, continental and global scales.

In this approach, the predictive accuracy of a rule-based model can be improved by combining it with an instance-based/ nearest-neighbor model that predicts the target value of a new case using the average predicted values of the *n* most similar cases (RuleQuest, 2008). The use of the composite model can improve the predictive accuracy relative to the rule-based model alone. This approach can also generate committee models made up of several rule-based models, and each member of the committee model predicts the target value for a case (RuleQuest, 2008). The member's predictions are averaged to give a final prediction.

We constructed predictive GPP and NEE models based on AmeriFlux and MODIS data. The predictive variables we used include a variety of MODIS data streams, such as vegetation type, enhanced vegetation index (EVI), land surface temperature (LST), normalized difference water index (NDWI), and leaf area index (LAI). Given the diversity in ecosystem types, age structures, fire and insect disturbances, and management practices, the performance of our approach for estimating ecosystem carbon fluxes is encouraging (Xiao et al. 2008, 2010a).

Upscaling Fluxes from Towers to Regional, Continental and Global Scales Using Data-driven Approaches

Jingfeng Xiao

Our data-driven approach has several advantages over ecosystem models. Our approach can make use of continuous flux measurements from a large number of towers and various satellite data streams. Moreover, most ecosystem models (e.g., Terrestrial Ecosystem Model, TEM; Xiao et al. 2009) are dependent on site-level parameterizations that are used as default parameters for model simulations over large areas. Our predictive models, however, are highly constrained by flux data from towers encompassing a range of ecosystem and climate types, and can lead to model parameters that are more representative of the full spectrum of vegetation and climate types. Therefore, they are more likely to produce more accurate estimates of carbon fluxes at broad scales. Our approach can also substantially reduce computational complexity compared to most ecosystem models. On the other hand, our data-driven approach is an empirical method, and does not incorporate ecosystem processes such as photosynthesis and nitrogen cycling. In addition, our approach does not explicitly consider some important factors influencing ecosystem carbon exchange such as the sizes of carbon pools and nitrogen availability that may be explicitly simulated in ecosystem models at stand scales (e.g., Ollinger et al. 2002). All these differences can lead to discrepancies in flux estimates between our data-driven approach and ecosystem models.

Challenges

Several challenges remain for more accurate quantification of carbon and water fluxes using data-driven approaches. First, the representativeness of the FLUXNET network will presumably affect flux estimates. Although FLUXNET is generally representative of the major climate and ecosystem types over the globe, some geographical regions, ecoregions, and biome types are still underrepresented. For instance, no open savanna (tree cover 10-30%) towers are affiliated with the AmeriFlux network. Therefore, in our upscaling studies, we treated open savannas as woody savannas (tree cover 30-60%), which can lead to significant biases to the flux estimates for open savannas, as woody savannas typically have larger biomass and higher productivity than open savannas. Gridded flux estimates will exhibit larger uncertainties in regions under-represented by towers. Upscaling studies will benefit from the availability of flux data from more towers so that the differentiation of woody savannas from open savannas, open shrublands from closed shrublands, and C_3 from C_4 plants can be made.

Second, our data-driven approach does not sufficiently account for the factors influencing heterotrophic respiration. Heterotrophic respiration is influenced by substrate availability, soil temperature, and soil moisture. LST and NDWI are used to account for soil moisture and temperature, and surface reflectance also partly accounts for non-photosynthetic material (e.g., litter). However, satellite data streams cannot account for the sizes of soil organic carbon pools. The inability of our model to sufficiently account for transient carbon pools may lead to significant uncertainties in NEE fields.

Third, data-driven approaches currently cannot sufficiently account for the effects of disturbance on carbon fluxes. Disturbances can substantially alter ecosystem carbon fluxes and regional carbon budgets (Law et al. 2004). More recently, Amiro et al (2010) showed that North American forests provided carbon sources in the years immediately following stand-replacing disturbances, and became carbon sinks by 10-20 years. To date, however, no spatially explicit information on disturbance and stand age is available at continental to global scales. Due to the lack of spatially explicit information on disturbance/stand age, our upscaling approach can only partly account for the ef-

Upscaling Fluxes from Towers to Regional, Continental and Global Scales Using Data-driven Approaches

Jingfeng Xiao

fects of disturbances on ecosystem carbon exchange by using MODIS observations that provided real-time observations of vegetation before and after disturbances.

The Future

Upscaling studies are expected to advance at least in the following four areas: (1) developing and improving methods for upscaling fluxes from towers to broader regions, (2) quantifying uncertainties associated with gridded flux estimates and regional carbon and water balances, (3) better constraining carbon sinks and sources over regions, continents, or the globe, and (4) assessing the interannual variability of carbon and water fluxes, particularly the impacts of extreme climate events and disturbance. These results will improve our understanding of the variability in carbon and water fluxes over broad regions.

Future research should include additional explanatory variables to better account for live and dead vegetation carbon pools and other factors that influence decomposition of woody detritus and soil respiration. In particular, future upscaling studies should explicitly consider the impacts of disturbance on ecosystem carbon exchange. Over North America, forested flux sites encompass a large range of disturbance history and stand age (Figure 3). Notably, 27 out of the 70 sites are younger than 25 yrs of age. In the meanwhile, spatially explicit information on disturbance and stand age is expected to be available at continental scales in the near future. The integration of tower-scale disturbance history and continental-scale disturbance products is expected to improve the estimation of carbon fluxes.

Upscaling studies have rarely, if ever, conducted comprehensive, quantitative analyses of uncertainties associated with the resulting flux fields. There are several sources of uncertainty associated with flux estimates: uncertainties in eddy flux measurements, uncertainties in input data (e.g., land cover), model structural uncertainty, and uncertainties arising from the representativeness of the flux networks. Future upscaling efforts, however, should gauge the uncertainty in flux estimates by consider-

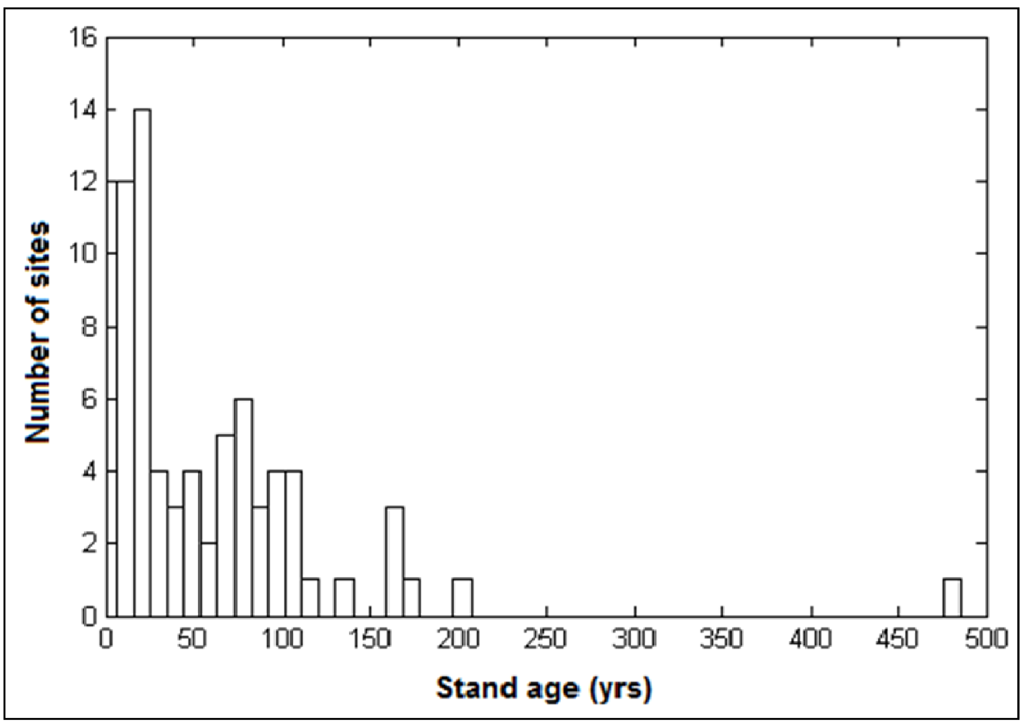

Figure 3: The distribution of stand age of forested flux sites across North America.

Upscaling Fluxes from Towers to Regional, Continental and Global Scales Using Data-driven Approaches

Jingfeng Xiao

ing uncertainties of eddy flux measurements and other input data (e.g., land-cover maps), propagating the probability distributions of parameters through the models, and comparing changes in fluxes caused by systematically removing individual flux tower data from the development of models (Xiao et al. 2010b). State-of-the-art uncertainty analyses will provide robust uncertainty estimates and error bounds to annual fluxes and to the estimates of regional to continental carbon sinks/sources.

Upscaling efforts will also benefit from the intercomparison of multiple upscaling methods (data-driven and data assimilation approaches) and of the resulting flux fields. The juxtaposition of flux estimates resulting from these approaches can provide complementary information for the diagnostics of ecosystem carbon exchange at regional, continental and global scales and valuable information for improvement of these approaches.

Upscaling studies provide alternative and independent fluxes estimates from conventional ecosystem modeling and inventory approaches. The resulting gridded flux estimates have been used to examine terrestrial carbon budgets and evaluate ecosystem models and inversion estimates. For instance, our EC-MOD flux fields have been used to examine the U.S. carbon sink (Xiao et al. 2010b), validate atmospheric inversions over temperate North America (Deng et al. under review), assess North American carbon dynamics through the North American Carbon Program (NACP) Interim Synthesis, and evaluate watershed-scale carbon fluxes derived from a water-centric model (Sun et al. under review). Future gridded flux fields derived from upscaling efforts will likely be more widely used for carbon and water analyses, model evaluation, and model intercomparison studies, and provide valuable databases for developing climate change mitigation and adaptation strategies.

Contact:

Jingfeng Xiao
(j.xiao@unh.edu)

Complex Systems Research Center,
Univ. of New Hampshire
Durham, New Hampshire, USA

References:

Amiro, B.D., Barr, A.G., Barr, J.G., Black, T.A., Bracho, R., Brown, M., Chen, J., Clark, K.L., Davis, K.J., Desai, A.R., Dore, S., Engel, V., Fuentes, J.D., Goldstein, A.H., Goulden, M.L., Kolb, T.E., Lavigne, M.B., Law, B.E., Margolis, H.A., Martin, T., McCaughey, J.H., Misson, L., Montes-Helu, M., Noormets, A., Randerson, J.T., Starr, G., and Xiao, J. (2010) Ecosystem carbon dioxide fluxes after disturbance in forests of North America. *J. Geophys. Res.*, **115**, G00K02,

Baldocchi, D.D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K.T., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S. (2001) FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide,

Upscaling Fluxes from Towers to Regional, Continental and Global Scales Using Data-driven Approaches

Jingfeng Xiao

water vapor, and energy flux densities. *Bull. Amer. Meteor. Soc.,* **82**, 2415-2434.

Caspersen, J.P., Pacala, S.W., Jenkins, J.C., Hurtt, G.C., Moorcroft, P.R., and Birdsey, R.A. (2000) Contributions of land-use history to carbon accumulation in U.S. forests. *Science*, **290**, 1148-1151.

Davis, K. J. (2008) Integrating Field Measurements with Flux Tower and Remote Sensing Data. In Field Measurements for Landscape-Scale Forest Carbon Monitoring, Hoover, Coeli M (Ed.) 2008, XVIII, 242 p. 20

Deng, F., Chen, J.M., Pan, Y., Peters, W., Birdsey, R., McCullough, K., and Xiao, J. (2010) Forest stand age information improves an inverse North American carbon flux estimate. *Global Biogeochem. Cycles*, under review.

Houghton, R.A., Hackler, J.L., and Lawrence, K.T. (1999) The U.S. carbon budget: contributions from land-use change. *Science*, **285**, 574-578.

Law, B.E., Turner, D., Campbell, J., Sun, O.J., Tuyl, S.V., Ritts, W.D., and Cohen, W.B. (2004) Disturbances and climate effects on carbon stocks and fluxes across Western Oregon USA. *Global Change Biol.*, **10**, 1429-1444.

Ollinger, S.V., Aber, J.D., Reich, P.B., and Freuder, R. (2002) Interactive effects of tropospheric ozone nitrogen deposition, elevated CO and land use history on the carbon dynamics of northern hardwood forests. *Global Change Biol.*, **8**, 545-562.

Pacala, S.W., Hurtt, G.C., Baker, D., Peylin, P., Houghton, R.A., Birdsey, R.A., Heath, L., Sundquist, E.T., Stallard, R.F., Ciais, P., Moorcroft, P., Caspersen, J.P., Shevliakova, E., Moore, B., Kohlmaier, G., Holland, E., Gloor, M., Harmon, M.E., Fan, S.M., Sarmiento, J.L., Goodale, C.L., Schimel, D., and Field, C.B. (2001) Consistent land- and atmosphere-based U.S. carbon sink estimates. *Science*, **292**, 2316-2320.

RuleQuest (2008) http://www.Rule quest .com. Visited on 10/18/2007.

Running, S.W., Baldocchi, D.D., Turner, D.P., Gower, S.T., Bakwin, P.S., and Hibbard, K.A. (1999) A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. *Remote Sens. Environ.*, **70**, 108-127.

Schimel, D., Melillo, J., Tian, H., McGuire, A.D., Kicklighter, D., Kittel, T., Rosenbloom, N., Running, S., Thornton, P., Ojima, D., Parton, W., Kelly, R., Sykes, M., Neilson, R., and Rizzo, B. (2000) Contribution of increasing CO_2 and climate to carbon storage by ecosystems in the United States. *Science*, **287**, 2004-2006.

SOCCR (2007) *The First State of the Carbon Cycle Report (SOCCR): the North American Carbon Budget and Implications for the Global Carbon Cycle,* eds King AW, Dilling L, Zimmerman GP, Fairman DM, Houghton RA, Marland GA, Rose AZ, Wilbanks TJ (US Climate Change Science Program, Washington, DC), pp. 19.

Sun, G., Caldwell, P., Noormets, A., Cohen, E., McNulty, S.G., Treasure, E., Domec, J.-C., Mu, Q., Xiao, J., John, R., and Chen, J. (2010) Upscaling key ecosystem functions across the conterminous United States by a water-centric ecosystem model. *J. Geophys. Res.*, under review.

Wofsy, S.C., Goulden, M.L., Munger, J.W., Fan, S.-M., Bakwin, P.S., Daube, B.C., Bassow, S.L., and Bazzaz, F.A. (1993) Net exchange of CO_2 in a mid-latitude forest. *Science*, **260**, 1314-1317.

Xiao, J., Zhuang, Q., Baldocchi, D.D., Law, B.E., Richardson, A.D., Chen, J., Oren, R., Starr, G., Noormets, A., Ma, S., Verma, S.B., Wharton, S., Wofsy, S.C., Bolstad, P.V., Burns, S.P., Cook, D.R., Curtis, P.S., Drake, B.G., Falk, M., Fischer, M.L., Foster, D.R., Gu, L., Hadley, J.L., Hollinger, D.Y., Katul, G.G.,

Upscaling Fluxes from Towers to Regional, Continental and Global Scales Using Data-driven Approaches

Jingfeng Xiao

Litvak, M., Martin, T.A., Matamala, R., McNulty, S., Meyers, T.P., Monson, R.K., Munger, J.W., Oechel, W.C., Paw U, K.T., Schmid, H.P., Scott, R.L., Sun, G., Suyker, A.E., and Torn, M.S. (2008) Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data. *Agric. For. Meteorol.*, **148**, 1827-1847.

Xiao, J., Zhuang, Q., Liang, E., McGuire, A.D., Moody, A., Kicklighter, D.W., Shao, X., and Melillo, J.M. (2009) Twentieth-century droughts and their impacts on terrestrial carbon cycling in China. *Earth Interactions,* **13**, Paper No. 10, 1-31.

Xiao, J., Zhuang, Q., Law, B.E., Chen, J., Baldocchi, D.D., Cook, D.R., Oren, R., Richardson, A.D., Wharton, S., Ma, S., Martin, T.A., Verma, S.B., Suyker, A.E., Scott, R.L., Monson, R.K., Litvak, M., Hollinger, D.Y., Sun, G., Davis, K.J., Bolstad, P.V., Burns, S.P., Curtis, P.S., Drake, B.G., Falk, M., Fischer, M.L., Foster, D.R., Gu, L., Hadley, J.L., Katul, G.G., Matamala, R., McNulty, S., Meyers, T.P., Munger, J.W., Noormets, A., Oechel, W.C., Paw U, K.T., Schmid, H.P., Starr, G., Torn, M.S., and Wofsy, S.C. (2010a) A continuous measure of gross primary production for the conterminous U.S. derived from MODIS and AmeriFlux data. *Remote Sens. Environ.*, **114**, 576-591,

Xiao, J., Zhuang, Q., Law, B.E., Baldocchi, D.D., Chen, J., Richardson, A.D., Melillo, J.M., Davis, K.J., Hollinger, D.Y., Wharton, S., Oren, R., Noormets, A., Fischer, M.L., Verma, S.B., Cook, D.R., Sun, G., McNulty, S., Wofsy, S.C., Bolstad, P.V., Burns, S.P., Curtis, P.S., Drake, B.G., Falk, M., Foster, D.R., Gu, L., Hadley, J.L., Katul, G.G., Litvak, M., Ma, S., Martin, T.A., Matamala, R., Meyers, T.P., Monson, R.K., Munger, J.W., Oechel, W.C., Paw U, K.T., Schmid, H.P., Scott, R.L., Starr, G., Suyker, A.E., and Torn, M.S. (2010b) Assessing Net Ecosystem Carbon Exchange of U.S. Terrestrial Ecosystems by Integrating Eddy Covariance Flux Measurements and Satellite Observations. *Agric. For. Meteorol.*, in press, doi:10.1016/j.agrformet.2010.09.002.

And Announcing…

FLUXNET and Remote Sensing Open Workshop: Towards Up-scaling Flux Information from Towers to the Globe

June 7-9, 2011
Berkeley, California, USA

To register, visit:
http://nature.berkeley.edu/biometlab/fluxnet2011/fluxwkshp.html

Registration is limited to 120 participants.

Workshop registration is free, and there are limited travel funds available for graduate students, postdocs and workshop speakers.

FluxLetter
The Newsletter of FLUXNET
VOL. 3 NO. 3 DECEMBER, 2010

FluxLetter is produced quarterly at the FLUXNET Office with support from the National Science Foundation.

THIS ISSUE OF FLUXLETTER WAS EDITED, DESIGNED AND PRODUCED BY:

LAURIE KOTEEN
DENNIS BALDOCCHI

FLUXNET OFFICE:
137 MULFORD HALL
DEPT. OF ESPM
UNIVERSITY OF CALIFORNIA,
BERKELEY, CA 94720
PH: 1-(510)-642-2421
FAX: 1-(510)-643-5098
LKOTEEN@BERKELEY.EDU

We hope to make the FLUXNET newsletter a powerful information, networking, and communication resource for the community. If you want to contribute to any section or propose a new one please contact the FLUXNET Office. THANKS!!