A MODEL-DATA FUSION APPROACH TO INTEGRATE NATIONAL ECOLOGICAL OBSERVATORY NETWORK OBSERVATIONS INTO AN EARTH SYSTEM MODEL

Andrew M. Fox[1,2], Timothy J. Hoar[2], William J. Sacks[2], David J. P. Moore[3], Steve Berukoff[1], David S. Schimel[1]
[1] National Ecological Observatory Network; [2] National Center for Atmospheric Research; [3] University of Arizona

National Ecological Observatory Network

1. Background

The National Ecological Observatory Network (NEON) is a continental scale facility that will collect biogeochemical and biogeophysical data from 60 sites in the US over 30 years. NEON and other ecological observatories (ICOS, LTER, FLUXNET) provide data that can be utilized for continental-scale ecological forecasting, but new statistical and analytical tools are required that are capable of synthesizing this "data deluge".

NEON data products include gridded datasets of carbon and water fluxes across the continent. We are developing a model-data fusion framework in which NEON data will be combined with the Community Land Model (CLM) to produce optimal solutions for model parameter values, states and fluxes. This framework will enable spatial extrapolation of observations and ecological forecasting.

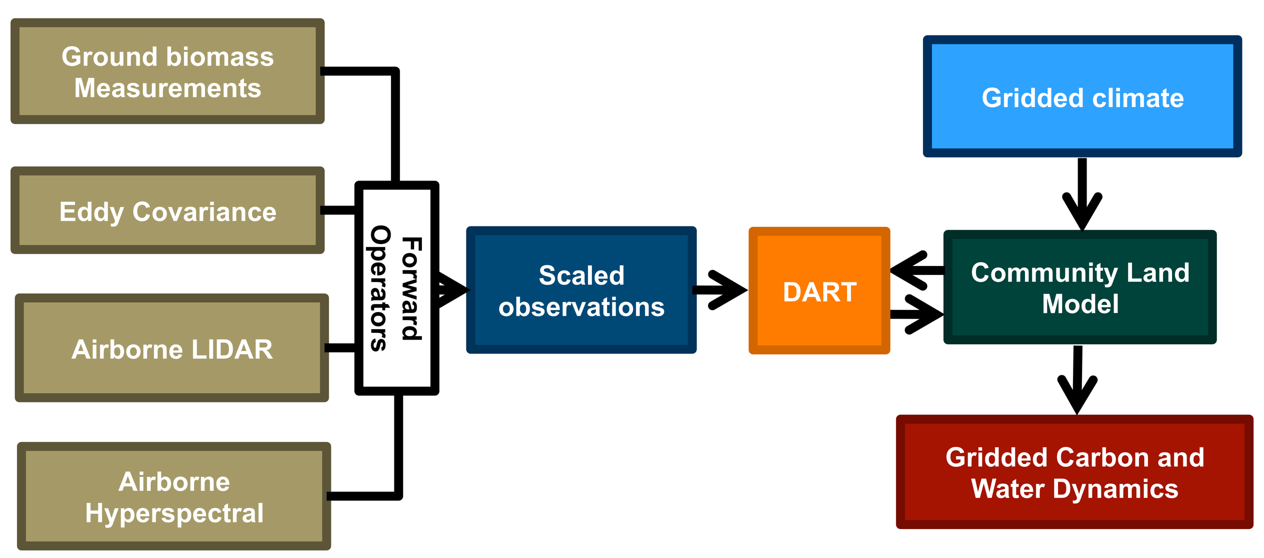

Figure 1. CLM/NEON Model-data fusion framework

2. Community Land Model (CLM)

The Community Land Model (CLM) is used as the land model in the Community Earth System Model (CESM). It simulates terrestrial ecosystem processes including the cycling of energy, water, carbon and nitrogen. CLM is driven by a limited set of climate variables, while the sensitivity of ecosystem processes to climate is controlled by the initial states and parameter sets of the model.

The performance of CLM is being evaluated at a number of existing flux tower sites with long data records, many of which are NEON candidate core sites, including Harvard Forest, and Niwot Ridge. Typically, we find energy fluxes to be more accurately simulated than carbon fluxes (Figure 2).

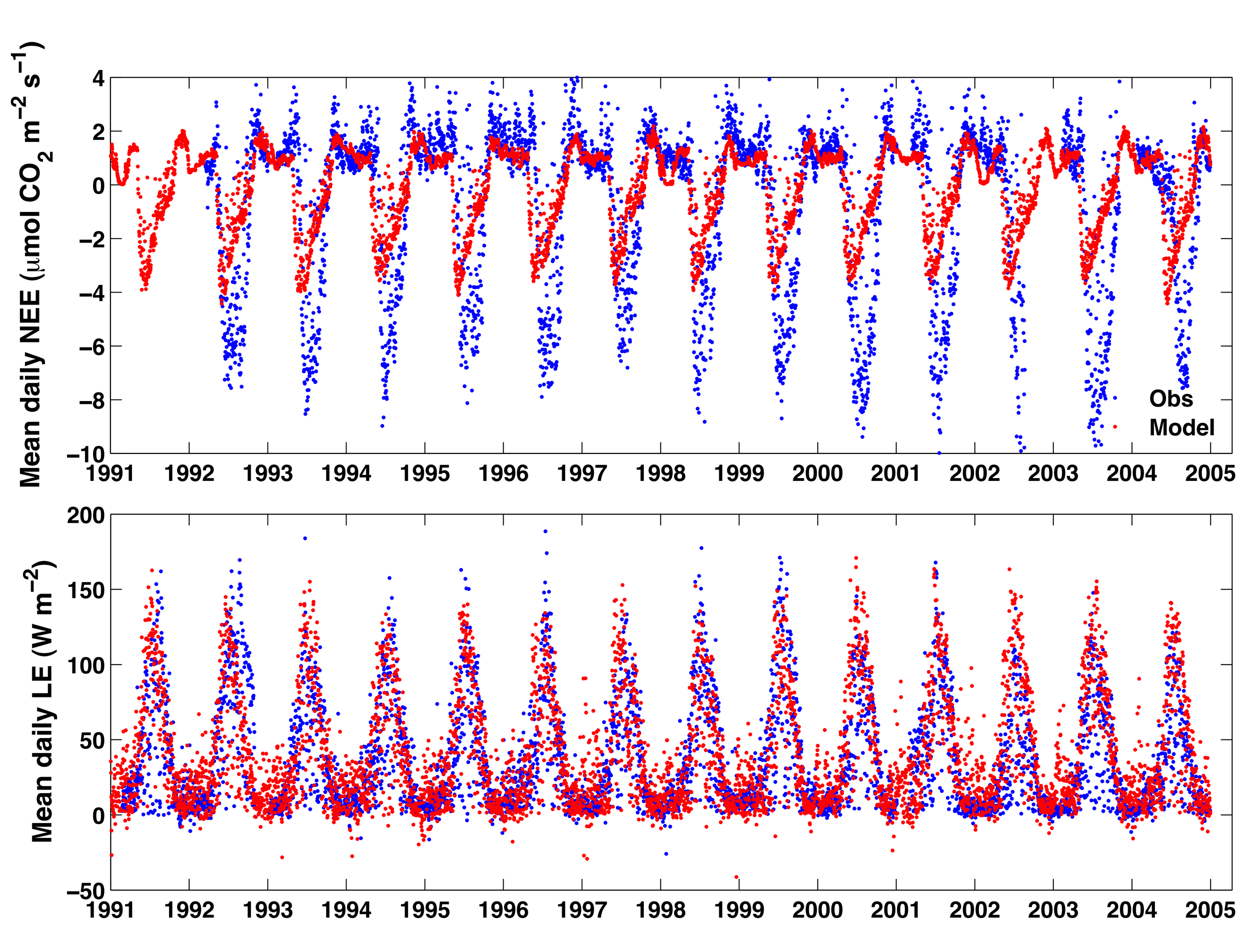

Figure 2. Observed (blue) and modeled (red) carbon and water fluxes at Harvard Forest FLUXNET site

3. CLM-DART coupling

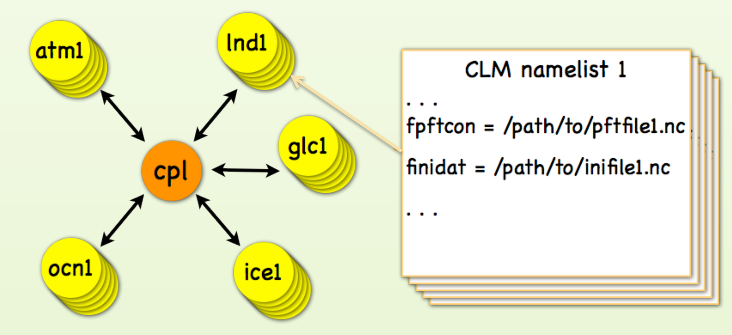

Figure 3. Multi-Instance CESM

A multi-instance version of CESM has been developed that more easily facilitates ensemble-based model-data fusion techniques. For example, multiple land models can be driven by multiple data-atmospheres in a single executable. This capability will be available in the next CESM public release.

We are applying expertise developed in numerical weather prediction to carbon cycle science by coupling CLM to the Data Assimilation Research Testbed (DART). An important difference is our goal is also to estimate parameter values, which give insight into process-level responses to environmental variation.

DART is a community facility for ensemble data assimilation developed and maintained at the National Center for Atmospheric Research (NCAR) that provides a number of enhancements to basic filtering algorithms (Figure 4).

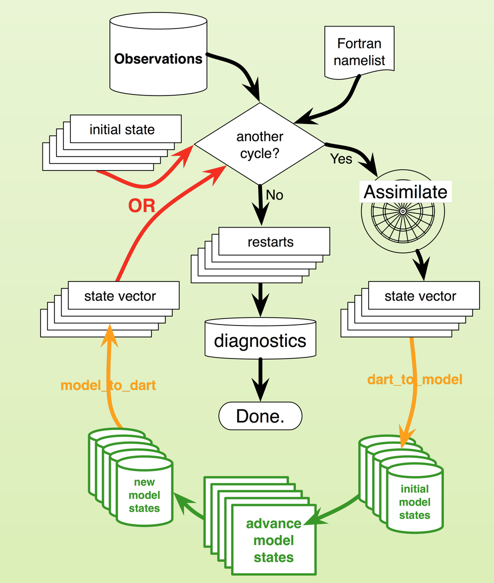

Figure 4. DART schematic

4. Observing System Simulation Experiment

Starting from a 'spun up' global CLM case for year 2000 we advance a 40-member ensemble of the global model for 120 days. Each instance of CLM is coupled with an individual from an ensemble of data-atmospheres. A single instance of CLM from the ensemble is then assigned to be the 'truth' and run forward. Daily 'observations' of leaf carbon (*leafc*) are collected at sites (grid cells) across the globe by adding Gaussian noise to the actual *leafc* state at these locations. These 'observations' are then assimilated by DART as the whole 40-member ensemble is run forward.

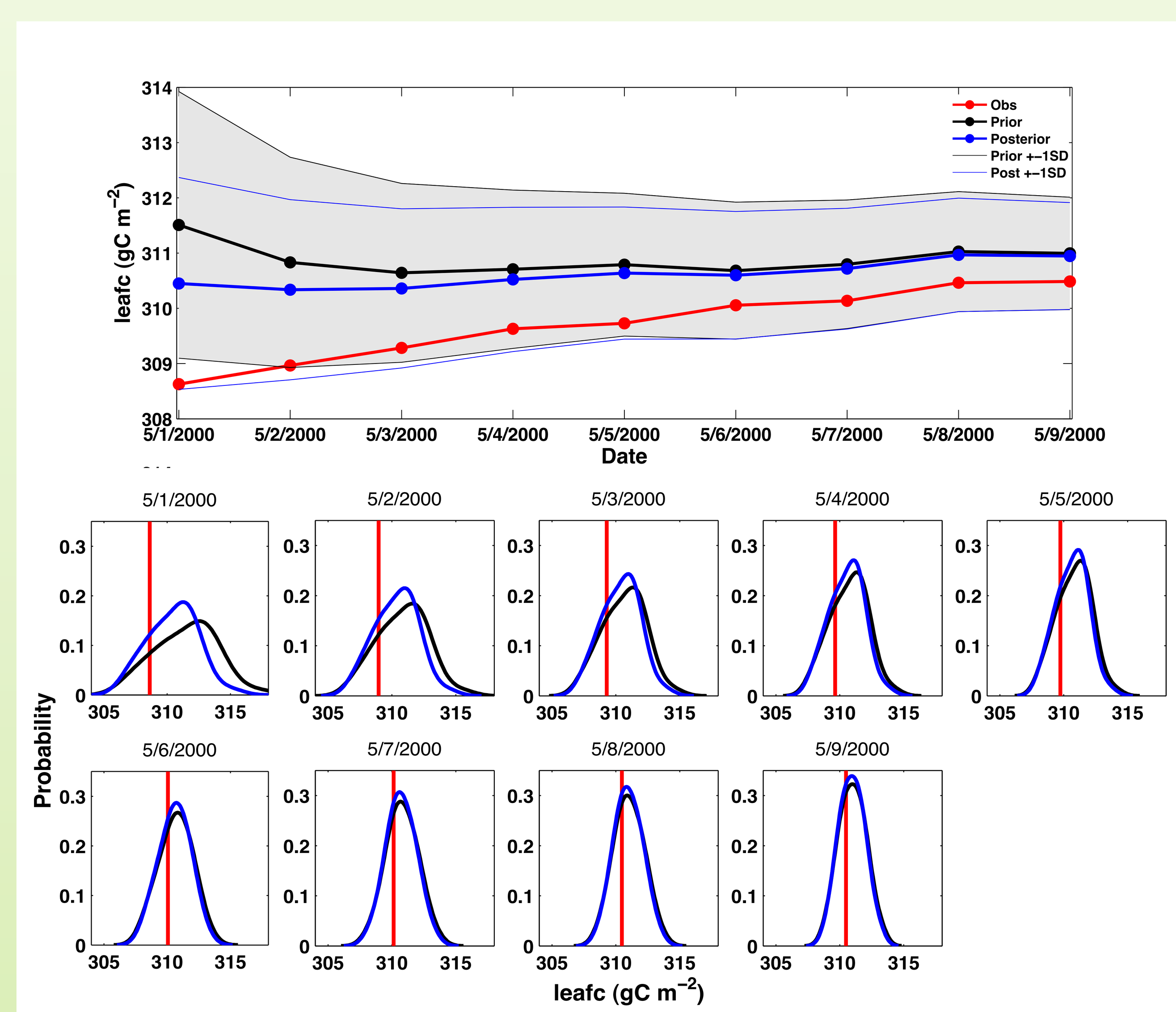

Figure 5. Prior and posterior probability distributions of leaf carbon in a single grid cell at 60°W, 4°S for nine days of assimilation

5. Parameter sensitivity analysis

Prior to estimating parameters through model-data fusion we are undertaking parameter sensitivity analyses to identify parameters to modify through data assimilation. At first we are assessing parameters contained in the PFT physiology parameter namelist file. Here we show results from perturbing four parameters: (i) leaf carbon to nitrogen ratio (leafcn); (ii) fraction of leaf N in rubisco (flnr); (iii) specific leaf area at top of canopy (slatop); and (iv) allocation ratio between fine root carbon and leaf carbon (froot_leaf).

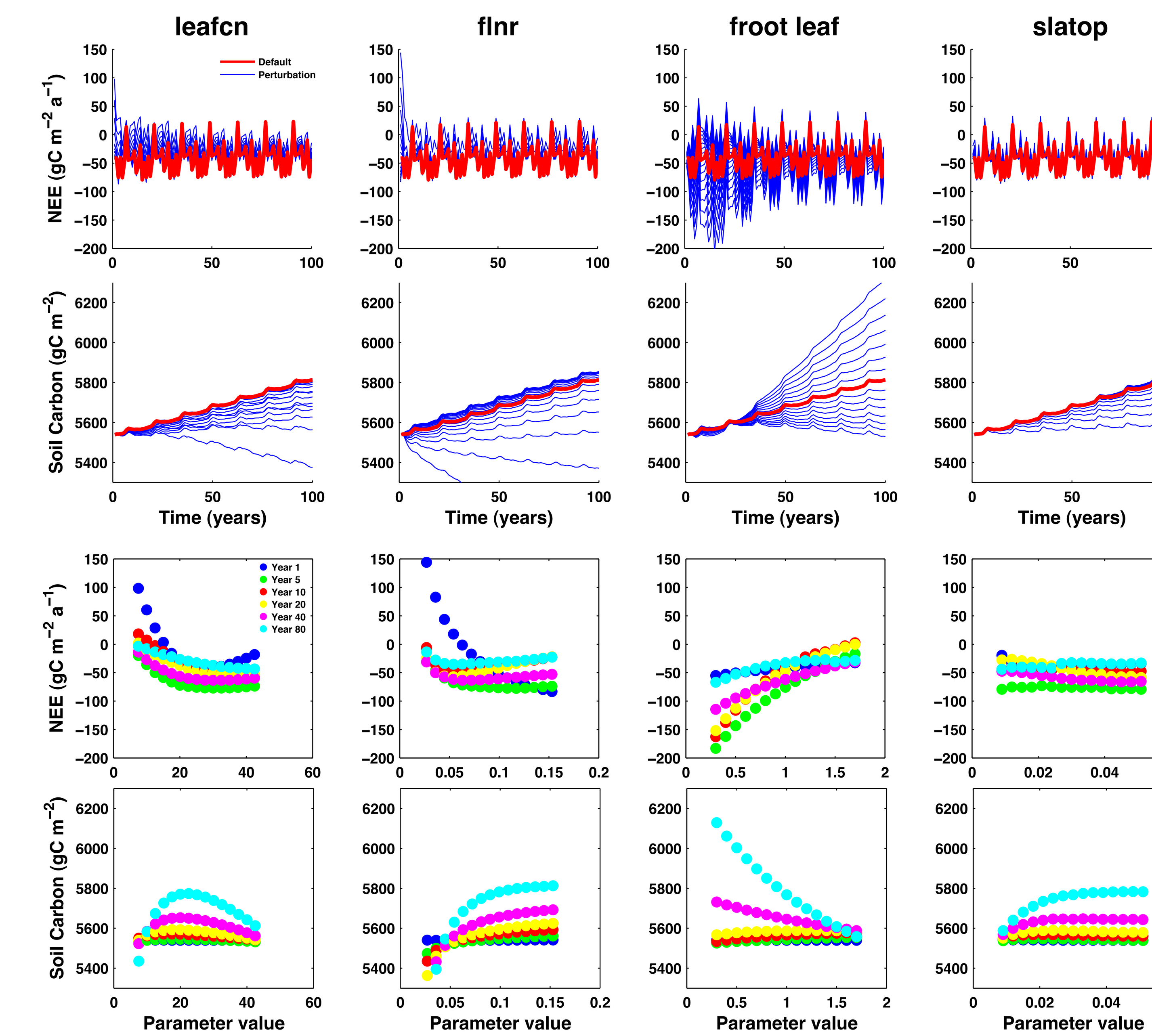

Figure 6. Annual NEE and soil carbon calculated using default parameter values (red) and 14 parameter value perturbations (blue) and scatter plots of annual NEE and soil carbon against parameter value at six time intervals

6. Future Work

For details as to how this work fits into the larger picture of model-data fusion facilitated by NEON see **B11D-0514** Observatory enabled modeling of the Global Carbon Cycle, Schimel *et al*, in this session.

We will continue undertaking a series of observing system simulation experiments to test the ability of the filter to update a number of CLM states. We will use a suite of synthetic observations of carbon and water fluxes and pools available at different temporal frequencies based on those that will be made at NEON sites in the future.

We will then extend the DART state vector to include CLM parameters and test how we can us this ensemble approach in parameter optimization. Will we investigate how this approach works with parameters which control processes operating at very different timescales.

We will investigate how to assimilate data into a spun-up model if the true initial state of the system at the time when observations become available is unknown.

Contact Information: afox@neoninc.org

The National Ecological Observatory Network is a project sponsored by the National Science Foundation and managed under cooperative agreement by NEON, Inc. This material is based in part upon work supported by the National Science Foundation under the following grants: EF-1029808, EF-1138160, EF-1150319 and DBI-0752017. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.