Ensemble Data Assimilation for Soil-Vegetation-Atmosphere Systems

T. Hoar (NCAR), Y. Zhang (University of Texas), A. Fox (NEON), R. Rosolem (University of Bristol)

National Center for Atmospheric Research
Boulder, Colorado, USA
dart@ucar.edu

1. Motivation

There are many open research questions in the relatively new field of data assimilation for land surface models. This poster will focus on a method that directly informs and updates the model state with the information content of the observations.
The Data Assimilation Research Testbed (DART) is a community facility for ensemble data assimilation developed and maintained at the National Center for Atmospheric Research (NCAR). DART is a software environment that makes it easy to explore a variety of data assimilation methods and observations with different numerical models and is designed to facilitate the combination of assimilation algorithms, models, and real (as well as synthetic) observations to allow increased understanding of all three.

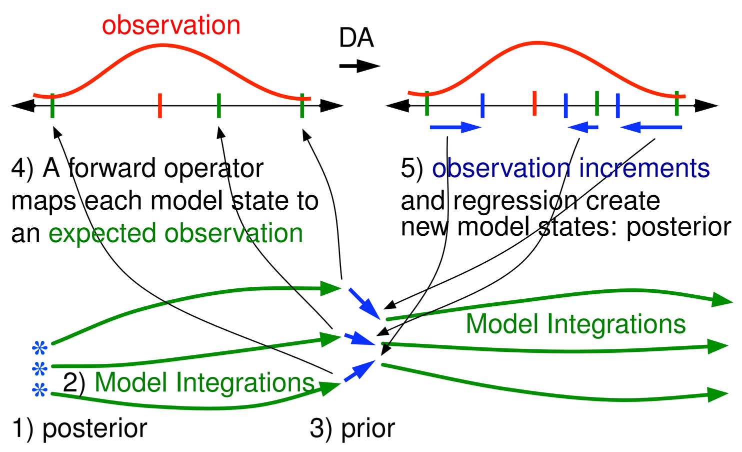


Figure 1: *Illustration for a toy ensemble size of 3.*

DART directly compares the forecast model state to the observations with virtually no additional expense.

2. NOAH LSM

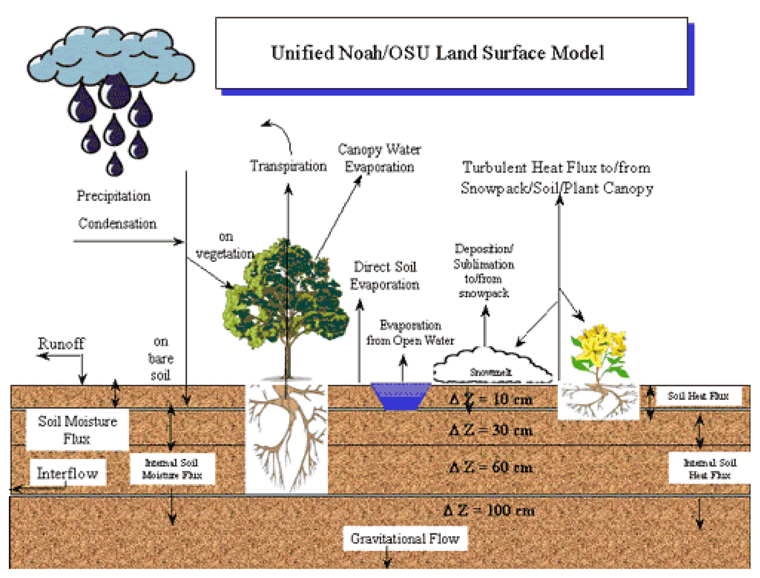


Figure 2: *www.ral.ucar.edu/research/land/technology/lsm.php*

3. DART and CESM/CLM

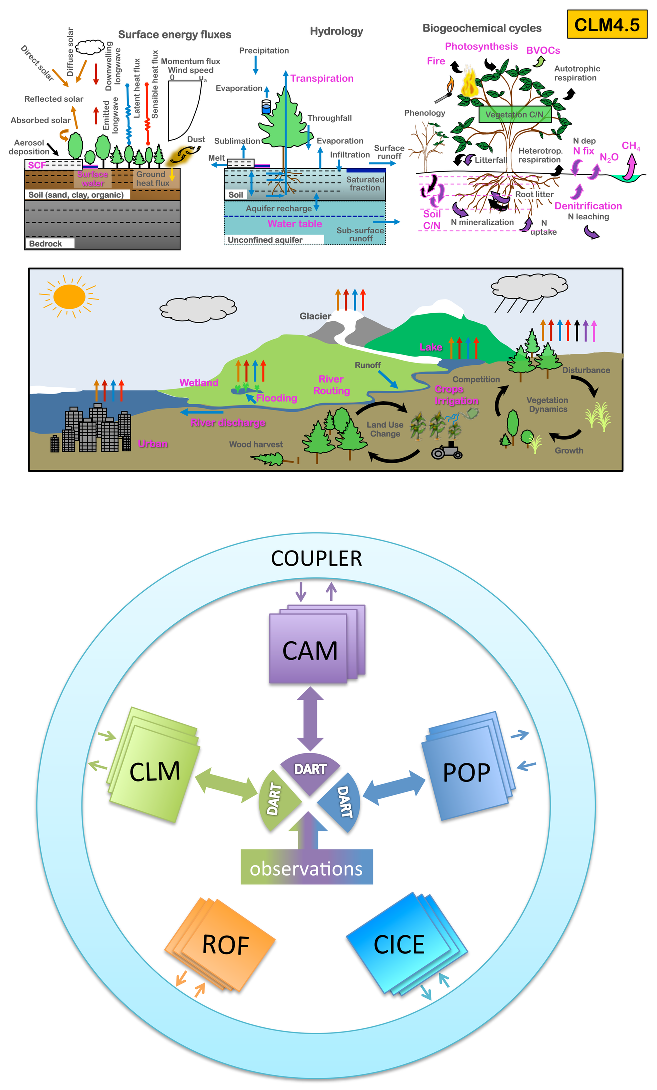


Figure 4: *CESM can advance multiple instances of one or more model components simultaneously, which enables it to use DART to assimilate observations.*

4. Carbon/Nitrogen

This is a collaboration with **Andrew Fox**, *National Ecological Observatory Network* (NEON).

4.1 Flux Tower Observations

A single CLM4.5 model state was spun up for 1500 years with site-specific information for Niwot Ridge, CO to 1 January 2000. This single model state was replicated 64 times and each instance was forced with a different member of an ensemble of 2°-resolution atmospheric reanalyses generated in a previous DA exercise with the Community Atmosphere Model (CAM).

- observations of tower fluxes of latent heat (LE), sensible heat (H), and net ecosystem production (NEP)
- variables influenced by the assimilation: LIVEROOTC, LIVESTEMC, DEADSTEMC, LEAFC, LITR1C, LITR2C, SOIL1C, SOIL2C, SOILLIQ, all of these are unobserved
- when used, the variance inflation has a constant value of 1.1

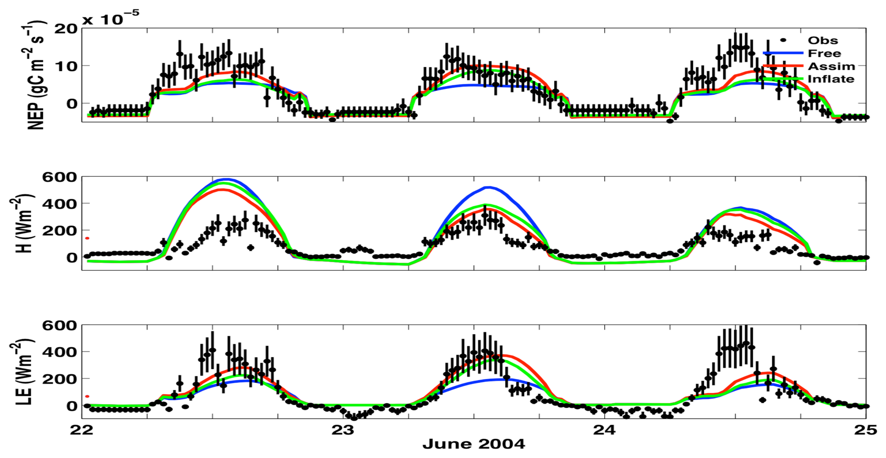


Figure 5: *An examination of the effect of assimilation for a representative 3-day period. The assimilated model state is more consistent with the observations. The solid lines represent the ensemble means. The observation error bars are $\pm$ 1 S.D.*

4.2 Leaf Area Index Observations

In an observation system simulation experiment (OSSE) we treat one ensemble member as "truth" and sample with appropriate noise at 60 NEON site locations to observe Leaf Area Index (LAI) every 8 days, Leaf Nitrogen every 12 days, and Net Ecosystem Productivity and Evapotranspiration every 0.5 hours. We then investigate the impacts of assimilating these $\approx$ 520,000 synthetic observations over a three-month period.

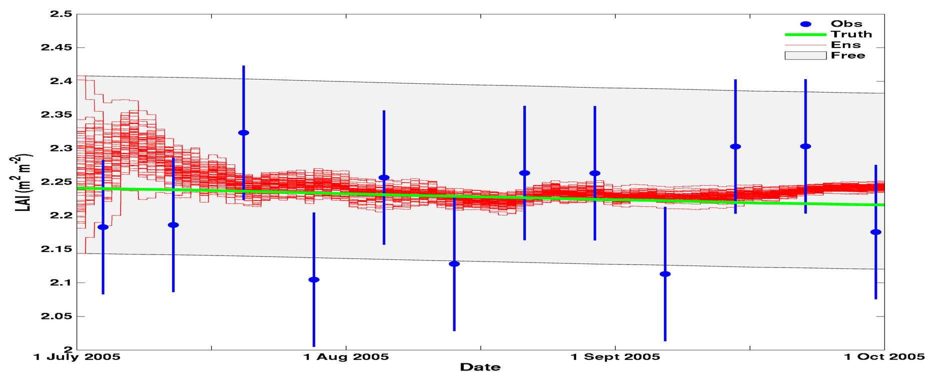


Figure 6: *This "sawtooth" plot shows LAI simulated by all 80 ensemble members in a grid cell with observations. The increments (updates) calculated by the filter move the ensemble towards the observations and results in a reduction in uncertainty (spread) around the truth. In this case, uncertainty is reduced too much and the result is slightly biased.*

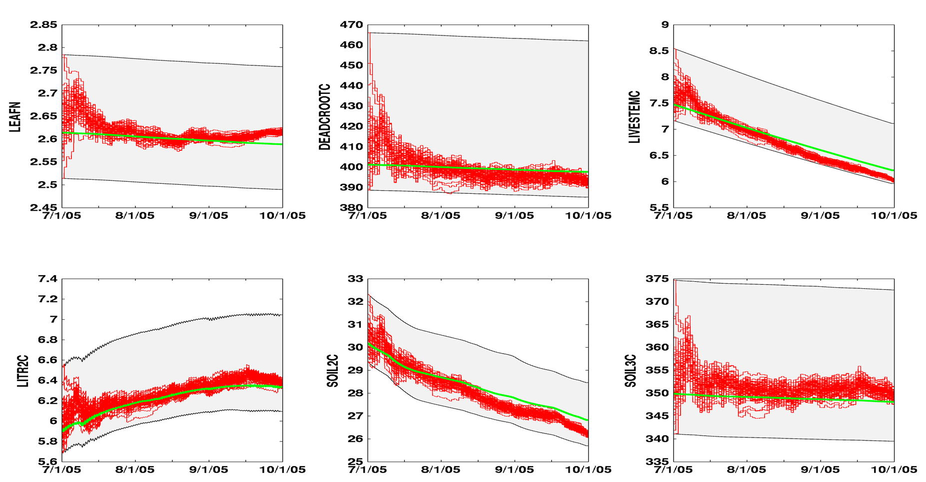


Figure 7: *The DART state vector contains more than 20 variables, including all the large carbon and nitrogen pools. These can all be updated by the filter through their covariance with observed variables. The allocation algorithms in CLM mean observations provide a strong constraint on many unobserved variables.*

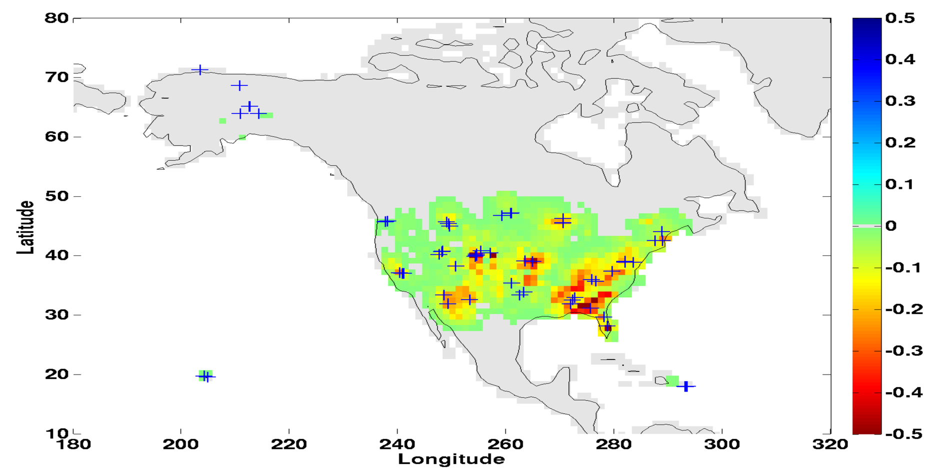


Figure 8: *Change in LAI spread in posterior ensemble, 31 July 2005. The largest innovations are near the observations, but not necessarily in the exact grid cell. Carbon pools from all grid cells are in the DART state vector, and are updated through the covariance matrix, propagating information from sites to regions. A cutoff value limits the distance over which this can occur.*

5. Snow Cover

This is a collaboration with **Yongfei Zhang**, *Jackson School of Geosciences, University of Texas at Austin.*

5.1 Precipitation Forcing Datasets

The snow representation in CLM is strongly determined by the precipitation from forcing datasets. The ensemble of forcing data comes from an offline CAM4/DART assimilation (See Raeder 2012). Here is a brief overview of that product compared to the Global Precipitation Climatology Project (GPCP). CAM4/DART produces more precipitation over Canada, western America and central Siberia; less precipitation over eastern America and western Eurasia.

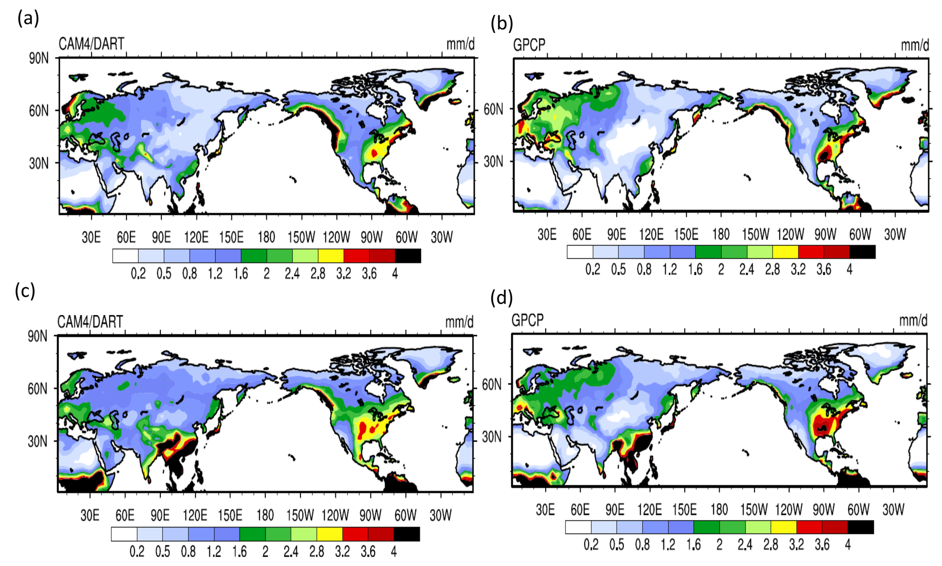


Figure 9: *10-year (1999-2008) DJF mean precipitation for (a) CAM4/DART and (b) GPCP, and MAM mean precipitation for (c) CAM4/DART and (d) GPCP.*

5.2 MODIS Snow Cover Fraction Observations

The MODIS/Terra daily snow cover dataset (MOD10C2; 0.05° resolution, northern hemisphere) was pre-processed to 0.9° x 1.25° "Level 4" data following Rodell and Houser (2004). Pixels with lower than 20% of clear sky were discarded.

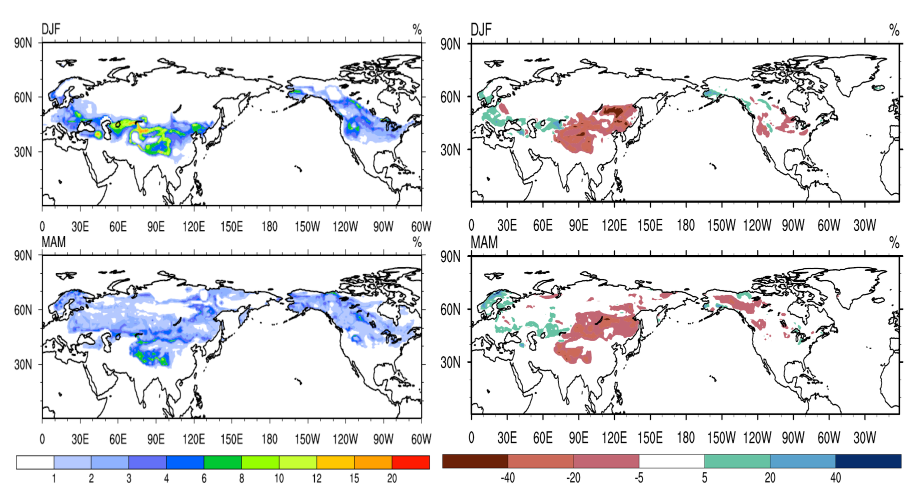


Figure 10: *Left: Ensemble spread of SCF for (top) DJF and (bottom) MAM in 2002-2003. Ensemble spread is calculated as the standard deviation of SCF among 40 ensemble members. SCF values are averaged for two seasons before calculating the ensemble spread. Right: The difference of SCF between the data assimilation case and the open loop case averaged for (top) DJF and (bottom) MAM.*

Raeder, K.D., et al., 2012: *DART/CAM: An ensemble data assimilation system for CESM atmospheric models.* DOI 10.1175/JCLI-D-11-00395.1

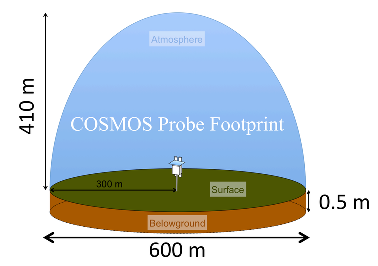
Rodell, M., et al., 2004: *The global land data assimilation system.* DOI 10.1175/BAMS-85-3-381

6. Soil Moisture

This is a collaboration with **Rafael Rosolem**, *University of Bristol.*

DART has been coupled to the NOAH Land Surface Model (HRLDAS-V3.3) and provides an operator to return neutron intensity "observations" given a soil moisture profile. This can be used to update the unobserved variables in the NOAH state.

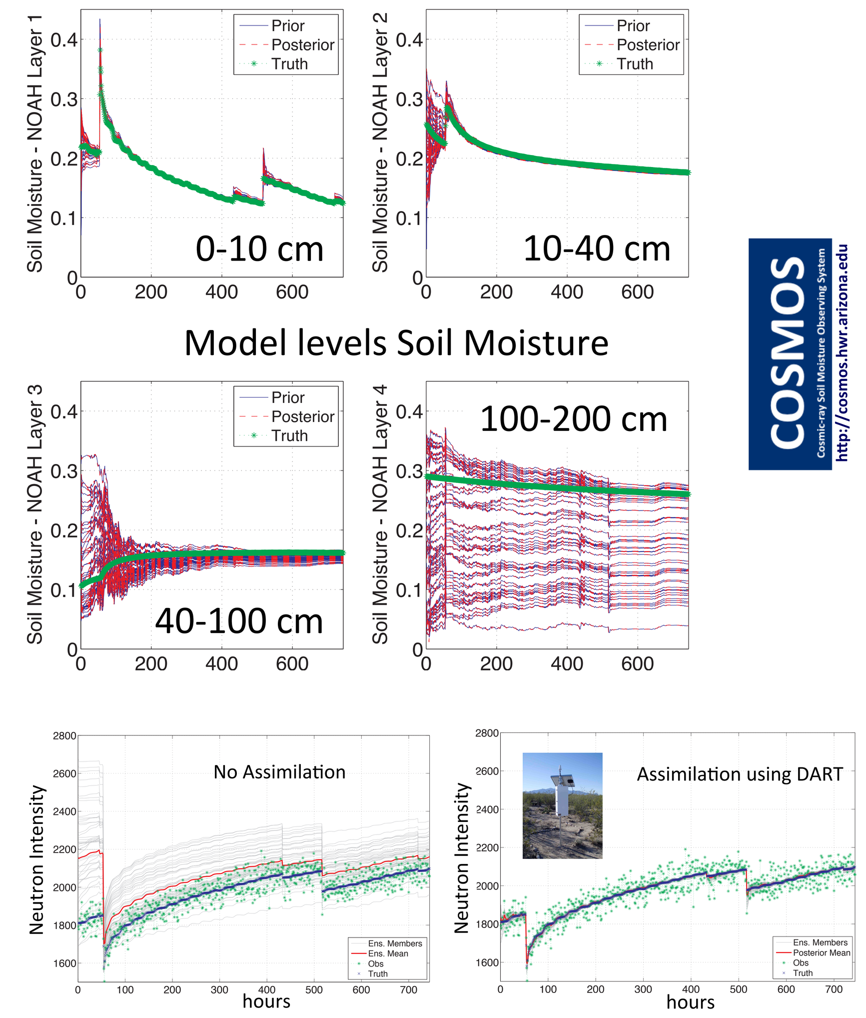
6.1 Neutron Intensity Observations



The COSMOS probe measures the neutron intensity for a given volume. The neutron intensity is related to the amount of Hydrogen present. The COSMIC model relates the neutron intensity to total soil moisture.

6.2 Synthetic Observations

This is for the Santa Rita site (about 20 miles SE of Tucson) for July 2009 using perturbed atmospheric forcing for a 54 member ensemble assimilating hourly observations.



6.3 Real Observations

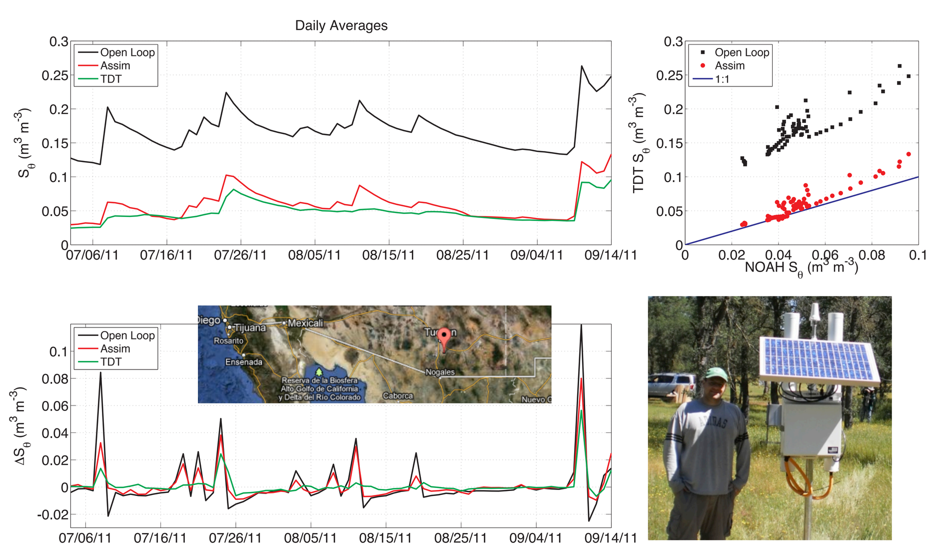


Figure 11: *These graphics assesses the performance of the assimilation of neutron intensity observations on soil moisture to withheld TDT soil moisture observations. The posterior mean is plotted in red.*

7. Further Information and References

http://www.image.ucar.edu/DAReS/DART has information about downloading DART (a full DART tutorial is included with the distribution) and contacting us.

Anderson, J. et. al., 2009: *The Data Assimilation Research Testbed: A Community Data Assimilation Facility. BAMS* **90** No. 9 pp. 1283–1296

The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Some of the computational resources were provided by the Computational and Information Systems Laboratory at NCAR.