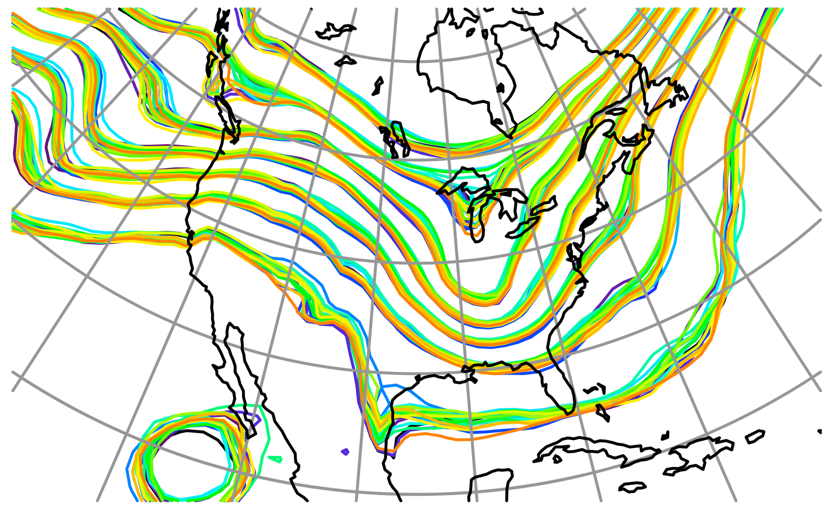


Open-Source Data Assimilation for Land Models and Multiscale Observations.

T. Hoar[1], A. Fox[2], Y. Zhang[3], R. Rosolem[4], A. Toure[5], B. Evans[6], J. McCreight[1]

[1]National Center for Atmospheric Research, Boulder, Colorado, USA [2]National Ecological Observatory Network, Boulder, Colorado, USA
[3]University of Texas at Austin, Austin, Texas, USA [4]University of Bristol, Bristol, England
[5]NASA GSFC, SSAI, Greenbelt, Maryland, USA [6]Macquarie University, Sydney, NSW, Australia

dart@ucar.edu

Institute for Mathematics Applied to Geosciences

1. Introduction

The Data Assimilation Research Testbed (DART) is an open source community software facility for ensemble data assimilation developed at the National Center for Atmospheric Research (NCAR). DART works with a wide variety of models and observations. Building an interface between DART and a new model does not require an adjoint and generally requires no modifications to the model code.

DART works with several land models, including:

- the Community Land Model – CLM,
- the NOAH-LSM,
- the Weather Research and Forecasting Model Hydrological modeling extension package – WRF-Hydro, and
- the Community Atmosphere-Biosphere-Land Exchange (CABLE) model.

DART assimilates dozens of observation types from a variety of sources. Some of the observations of interest for land assimilation are:

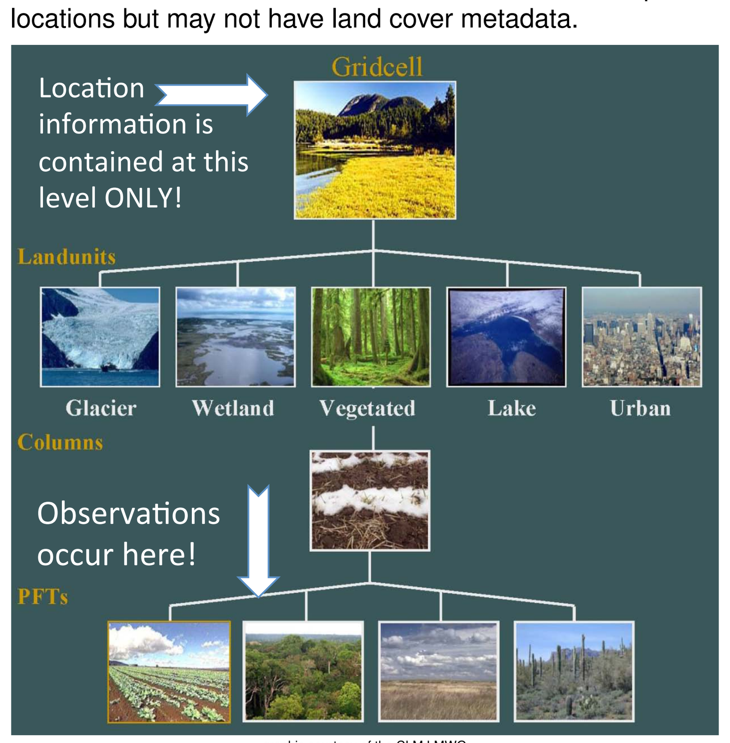
- in-situ measurements of soil moisture, temperature,
- tower fluxes,
- leaf area index,
- total water storage anomalies (i.e. GRACE),
- cosmic ray neutron intensities,
- and microwave brightness temperatures (Tb).

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

http://www.image.ucar.edu/DAReS/DART has information about how to download and install DART, a full DART tutorial (included with the distribution), and how to contact us.

1.1 Land Model Structure Complications

Many land models divide gridcells into proportional units based on land cover characteristics. This is a challenge for data assimilation as the land units generally have no unique location information of their own. Observations have specific locations but may not have land cover metadata.



graphic courtesy of the CLM LMWG

1.2 Observation Metadata

All ensemble data assimilation systems require the ability to calculate the expected value of the observation given a model state. The accurate application of this calculation (the observation operator) may require:

- knowledge of what land cover unit(s) or PFT to use for the calculation,
- soil properties, and
- instrument-specific parameters

Some of these could come from a lookup table based on PFT or location, but the lookup table generally must be recomputed to match the model resolution. Some (like Tb polarizations and frequencies) must be part of each observation.

2. Ecological State Estimation

Andrew Fox *National Ecological Observatory Network*

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 $\approx$ 520,000 synthetic observations over a 3 month period.

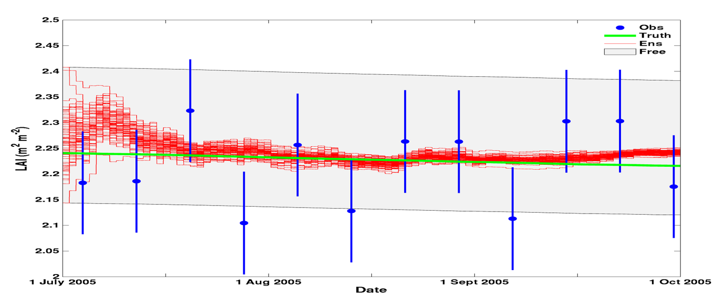


Figure 1: *This "sawtooth" plot shows LAI simulated by all 80 ensemble members in a grid cell with observations. The increments (updates) calculated by the filter move the ensemble towards the observations and result in a reduction in uncertainty (spread) around the truth. In this case, uncertainty is reduced too much and the result is slightly biased.*

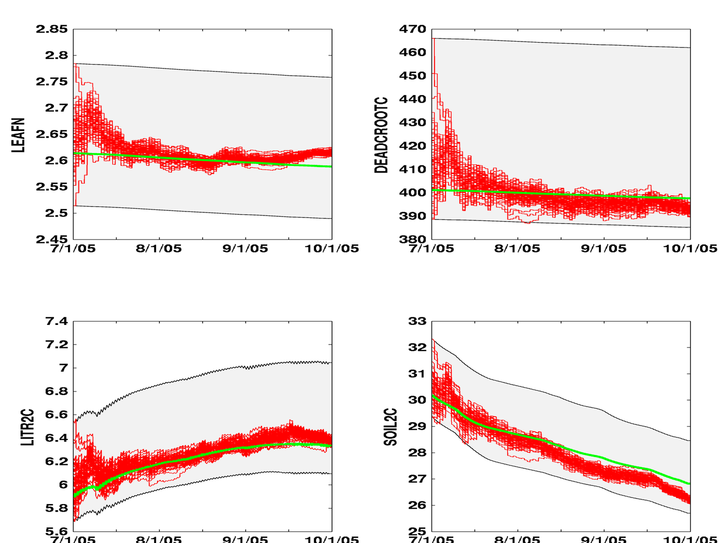


Figure 2: *The DART state vector contains more than 20 variables, including all the large carbon and nitrogen pools. These can all be updated by the filter through their covariance with observed variables. The allocation algorithms in CLM mean observations provide a strong constraint on many unobserved variables.*

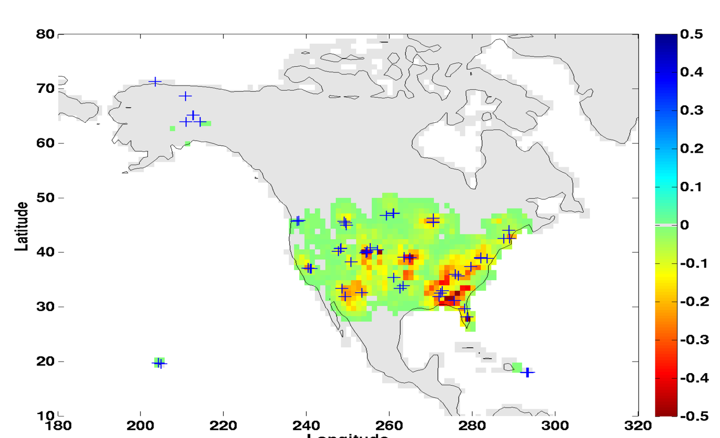


Figure 3: *Change in LAI spread in posterior ensemble, 31 July 2005. The largest innovations are near the observations, but not necessarily in the exact grid cell. Carbon pools from all grid cells are in the DART state vector and information can propagate from sites to regions. A cutoff value limits the distance over which this can occur.*

3. Multisensor Assimilation

Yongfei Zhang, *University of Texas at Austin.*

The DART algorithms can assimilate observations with uncorrelated observation errors in any order (Anderson, 2003). This allows one to simultaneously assimlate MODIS/Terra snow cover fraction (SCF – with little information about snow amount) and Gravity Recovery and Climate Experiment (GRACE) estimates of total water storage anomalies.

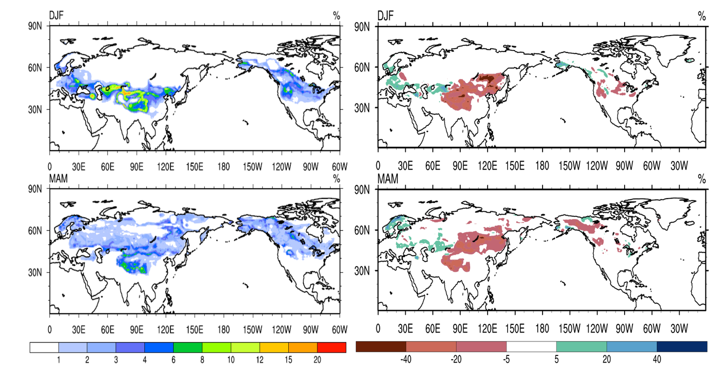


Figure 4: *Left: Ensemble spread of SCF for (top) DJF and (bottom) MAM in 2002-2003. Ensemble spread is calculated as the standard deviation of SCF among 40 ensemble members. Right: The difference of SCF between the data assimilation case and the open loop case averaged for (top) DJF and (bottom) MAM.*

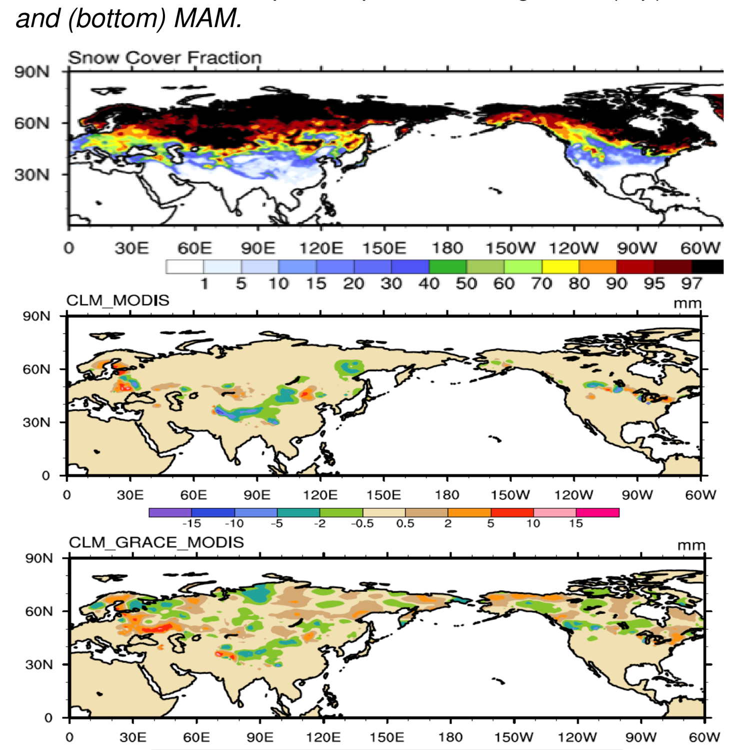


Figure 5: *Top: Mean SCF for Dec 2002 through Feb 2003. Middle: Impact (assimilation minus open loop) on snow water equivalent for 15 March 2003 for MODIS-only assimilation. Bottom: Impact on SCF for an assimilation with both MODIS and GRACE observations. GRACE is clearly providing additional information in regions where the SCF is saturated.*

Zhang, Y.-F., et al., 2014 *Assimilation of MODIS snow cover through the Data Assimilation Research Testbed and the Community Land Model version 4*
DOI: 10.1002/2013JD021329

4. Brightness Temperature Observations

Ally M. Toure, *NASA GSFC, USRA.*

The objective is to assess the performance of the brightness temperature (Tb) prediction in the Community Land surface Model version 4 (CLM4) coupled with a snow Radiative Transfer Model (RTM).

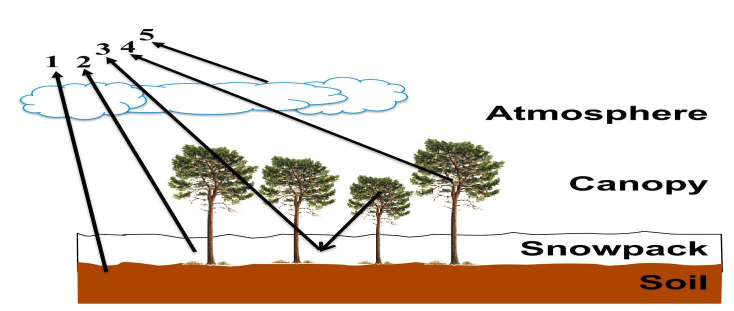


Figure 6: *The main contributions to the microwave emission measured by a spaceborne radiometer: 1) upward emitted soil emission, 2) snowpack, 3) combined canopy and snowpack, 4) canopy, and 5) the atmosphere.*

The RTM used to predict Tb using the CLM4 output is the Microwave Emission Model of Layered Snowpacks (MEMLS). It simulates Tbs for multi-layer snowpack and is valid for the frequency range of 5 GHz to 100 GHz. Typical inputs to the model are: Tb : the snowpack brightness temperature, S_0 : the ground-snow interface reflectivity, T_0 : the ground temperature, S_j : the interface reflectivity on top of each snow layer j, d_j : layer thickness, T_j : layer temperature, r_j : layer internal reflectivity, e_j : layer emissivity, t_j : transmissivity of each layer, and T_{sky} : the downwelling (sky) radiation.

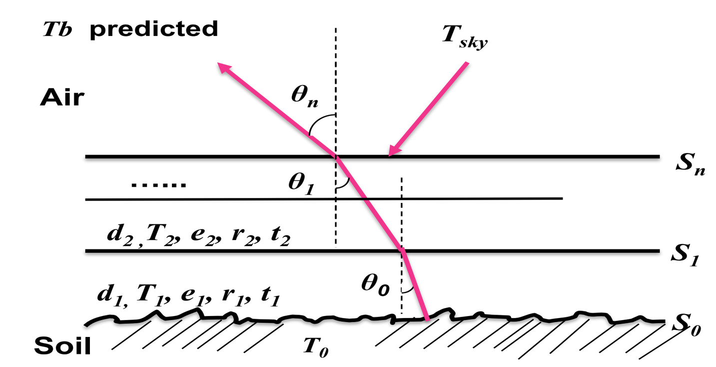


Figure 7: *Schematic of the MEMLS snow RTM (Wiesmann and Mätzler, 1999).*

RTM's also have parameters that must be estimated and are usally spatially varying. (See, for example, De Lannoy et. al, 2013: Global Calibration of the GEOS-5 L-Band Microwave Radiative Transfer Model over Nonfrozen Land Using SMOS Observations. J. Hydrometeor, 14, 765-785. doi: http://dx.doi.org/10.1175/JHM-D-12-092.1)

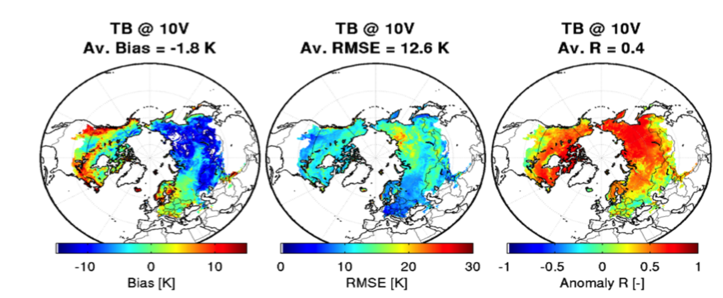


Figure 8: *Annual mean bias, RMSE and anomaly correlation coefficient (R) between predicted Tb and AMSR-E observations during 2002-2010. This is just 1 of 6 frequencies, and 1 polarization (10.7 Ghz, V pol).*

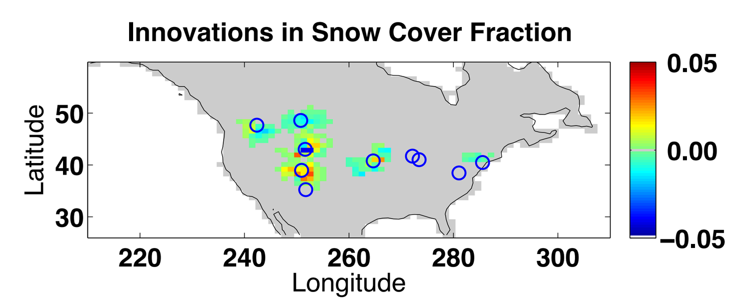


Figure 9: *This figure shows the amount of change induced in the snow cover fraction from assimilating synthetic brightness temperatures at 10 locations (shown as ◯) for 31 January 2001.*

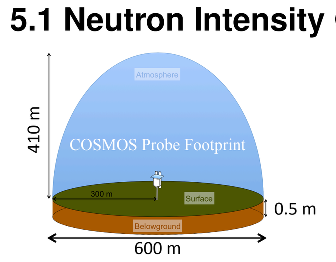
Tb assimilation presents other problems for DA. If you were to use all 6 AMSR-E frequencies at both polarizations for both ascending and descending swaths, you would be assimilating more than 6,000,000 observations *per day* in the Northern Hemisphere alone!

5. Soil Moisture Observations

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 NOAH model state.

5.1 Neutron Intensity Observations



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

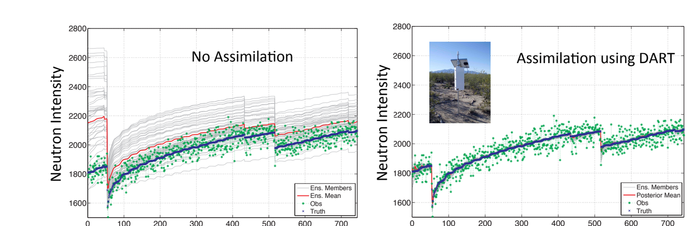


Figure 10: *An early result for the Santa Rita site.*

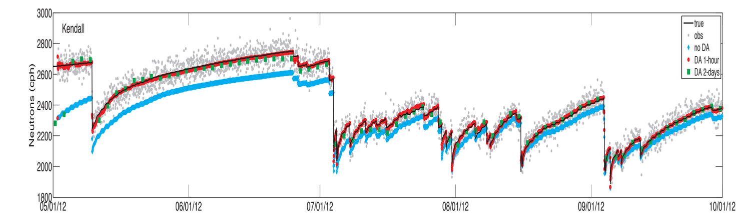


Figure 11: *Expected neutron intensity (counts per hour – cph) from a set of 3 experiments. The (known) true state is a solid black line, realistic observations (i.e. truth plus noise) are indicated by the gray dots, blue diamonds are a free run (i.e. no DA), red dots are assimilation every hour, green squares are assimilation every 2 days.*

5.2 Real Observations

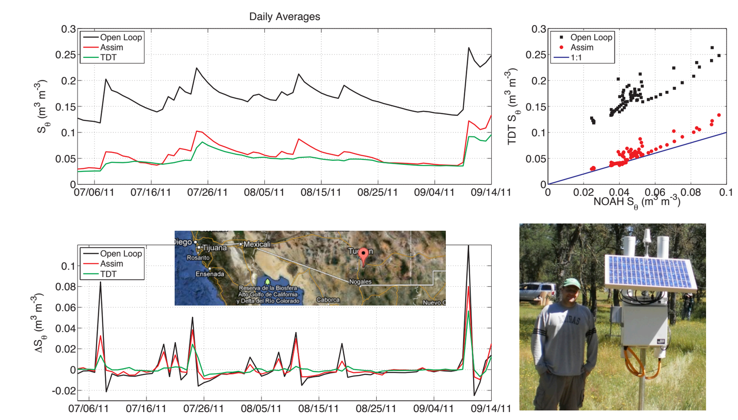


Figure 12: *These graphics assess the performance of the assimilation of neutron intensity observations on soil moisture to withheld traditional soil moisture observations. The posterior mean is plotted in red.*

Rosolem, R., et al., 2014 *Translating above-ground cosmic-ray neutron intensity to high-frequency soil moisture profiles at sub-kilometer scale.* *HESS* **18**, pp. 4363-4379

6. Hydrologic Assimilation

James McCreight, *NCAR.*

The Weather Research and Forecasting Model Hydrological modeling extension package (WRF-Hydro) is a community-based model coupling framework designed to link multi-scale process models of the atmosphere and terrestrial hydrology. Research with DART and WRF-Hydro will enable: 1) improved forecasts by reducing error in initial conditions, 2) a high-quality reanalysis, 3) diagnosis of model or observation errors, and 4) exploration of targeted observations.

6.1 Streamflow Assimilation

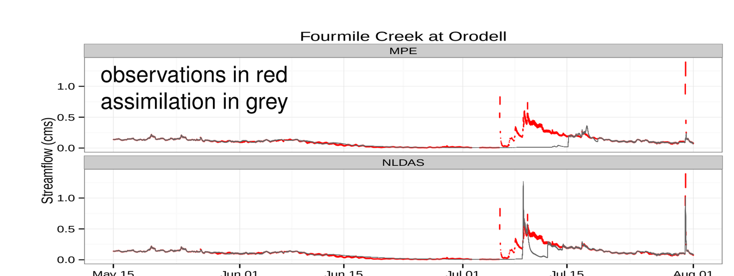


Figure 13: *Simulation before including several parameters into the DA for the two sets of precipitation forcing data. Top: precipitation from NOAA's Multisensor Precipitation Estimate (MPE), Bottom: NLDAS precipitation. The assimilation result runoff (grey) is highly dependent on the forcing.*

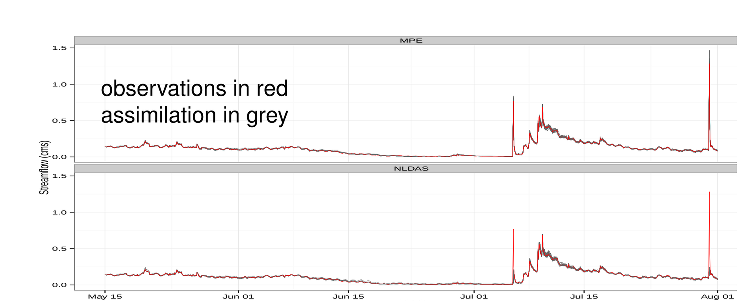
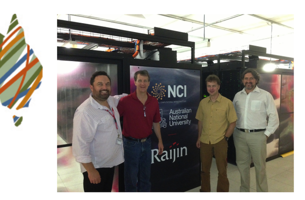


Figure 14: *Simulation after including several parameters into the DA for the two sets of forcing data.*

7. Land Surface Data Assimilation – CABLE

Brad Evans, *TERN/eMAST* **Luigi Renzullo**, *CSIRO*



Researchers from CSIRO, Macquarie University and the National Computing Infrastructure (NCI) teamed up with US collaborators to install and run DART on NCI's supercomputer (Raijin) and coupled it to Australia's Community Atmosphere Biosphere Land Exchange (CABLE) land surface model. The endeavour marks significant progress toward the vision of the Ecosystem Modelling and Scaling Infrastructure (eMAST) facility under the Terrestrial Ecosystem Research Network (TERN) to develop Australia's first modelling and data integration system for ecosystem science and monitoring at unparalleled scales in space and time. The system brings together a range of disparate ecological observations from ground- and space-based sensing networks into CABLE's modelling framework.

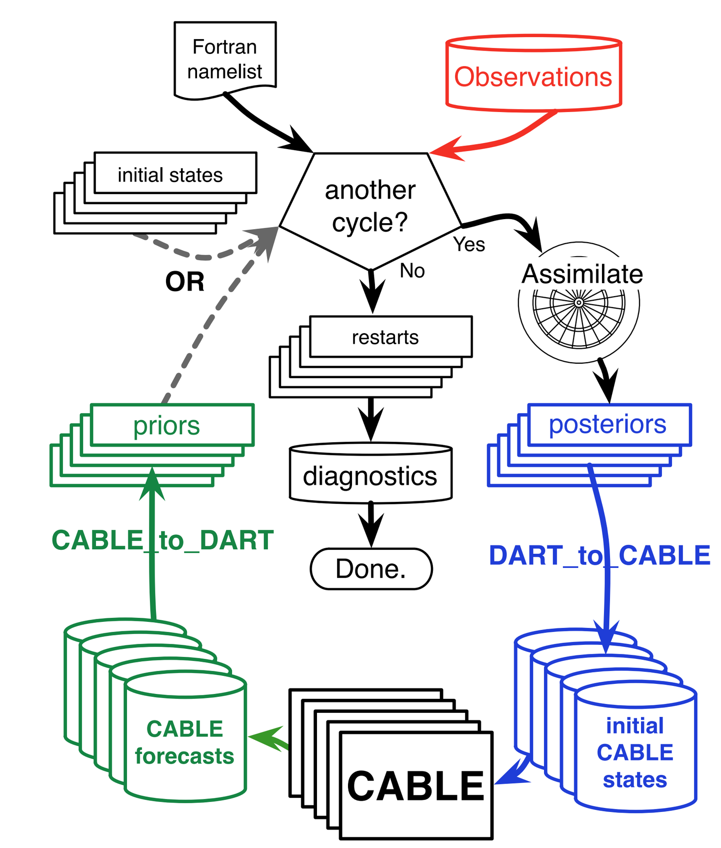


Figure 15: *A schematic of the assimilation system with DART and CABLE. Starting at the top: DART reads in an initial ensemble, the observations, the run-time control information and performs an assimilation to create posterior estimates of the CABLE variables.* DART_TO_CABLE *conveys the posteriors to a set of CABLE restart files which are advanced by CABLE to the time of the next observation.* CABLE_TO_DART *then extracts the prognostic state variables of interest and converts them to a DART-compliant format.*

Some of the instruments providing the observations that can be assimilated in the CABLE/DART system. Left-to-right: Eddy Covariance (Cape Tribulation), OzFlux (Scott Farm), CosmOz (Tullochgorum).

NCRIS National Research Infrastructure for Australia An Australian Government Initiative

The Terrestrial Ecosystem Research Network Ecosystem Modelling and Scaling Infrastructure is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS).

Anderson, J.-L., 2003
A local least squares framework for ensemble filtering.
MWR **131**, pp. 634-642

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