Improving the Representation of Land Surface Processes using the Data Assimilation Research Testbed (DART)

Brett Raczka[1], Xueli Huo[2], Andrew M. Fox[3], Moha Gharamti[1], Rolf Reichle[3], Daniel Hagan[4], Anthony Holmes[5], Ying Sun[5], Lewis Kunik[2], John Lin[2], Jeffrey Anderson[1]

[1]National Center for Atmospheric Research, Boulder, CO, [2]University of Utah, Salt Lake City, UT, [3]NASA Goddard Space Flight Center, [4]Nanjing University of Information Science & Technology, Nanjing, China, [5]Cornell University, Ithaca, NY

Contact: bmraczka@ucar.edu

Data Assimilation Research Testbed (DART)

DART is an open-source ensemble DA software system developed by the National Center for Atmospheric Research (Anderson et al., 2009). The DART system has been previously coupled with many geophysical and theoretical models. Here we focus on applications of the ensemble Kalman Filter within DART to the Community Land Model (CLM; Lawrence et al., 2019).

To generate ensemble spread CLM-DART relies on atmospheric reanalysis data (Raeder et al., 2012, Fig. 1, 2). DART includes important features such as a time and space varying 'inflation' which adjusted ensemble spread to account for sampling uncertainty and systemic biases. It also includes space and state based 'localization' to isolate the influence of observations upon the most correlated portions of the model state. **See https://dart.ucar.edu/ for more information.**

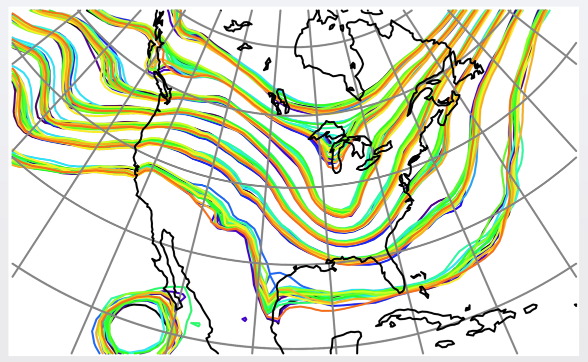


Figure 1. Example of an atmospheric reanalysis (CAM-DART) that samples uncertainty in atmospheric state. This serves as a boundary condition for CLM-DART.

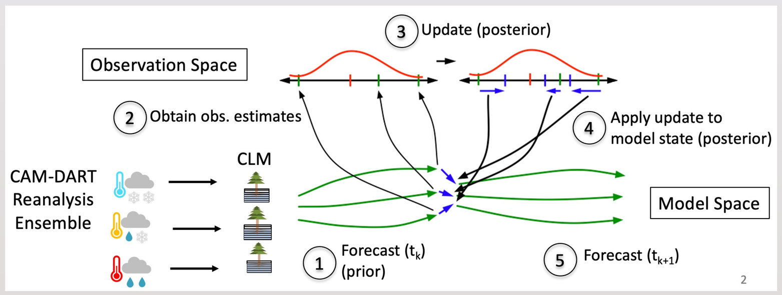


Figure 2. A schematic of a DART assimilation time step which adjusts the ensemble model prior (green) to better match observations (red) to generate a posterior (blue).

CLM-DART Applications: Vulnerable Ecosystems, Drought Monitoring, Seasonal Forecasting, Forest Restoration

CLM-DART can be used to better understand carbon cycling, albedo and canopy height in the Arctic region (Fig. 3); improve drought monitoring (Fig. 4), seasonal forecasting (Fig. 5) and forest restoration verification systems (Fig. 6). These applications assimilate observations including leaf area, soil moisture, terrestrial water storage, and solar-induced fluorescence (SIF).

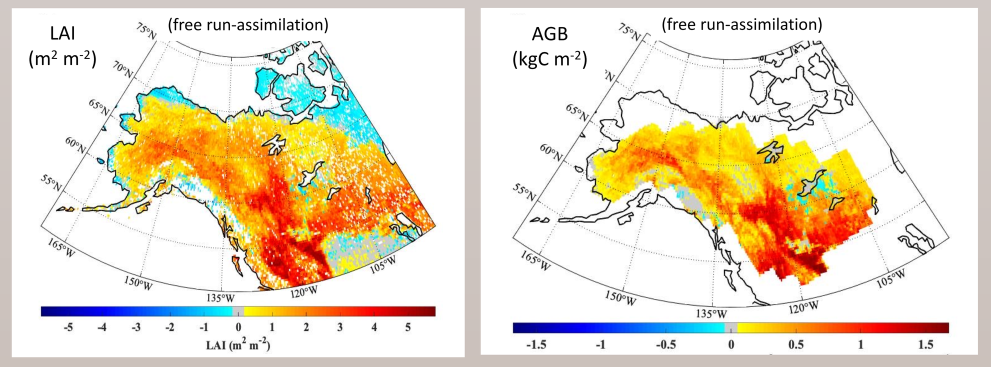


Figure 3. A simulation of CLM5 across the Arctic-Boreal domain (2011-2019) overpredicts observed leaf area leaf area (MODIS) and above-ground biomass (Wang et al., 2021).

Huo et al., (in prep)

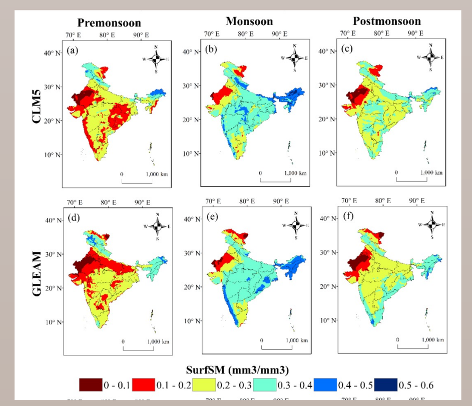


Figure 4. Simulation of CLM-SP across India (1981-2020) compared against the validation data set of GLEAM surface soil moisture across the monsoon season.

D.C. Naik
Danya C.T.

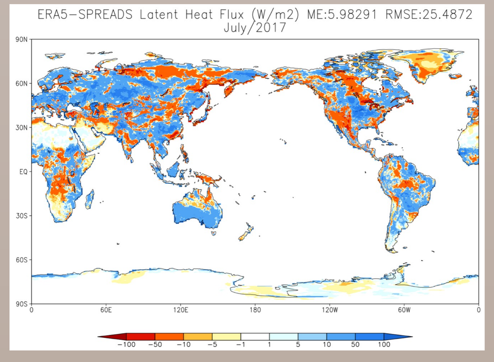


Figure 5. Global simulation of CLM5 during July 2017. Assimilation of GLASS LAI observations using LDAS-SPREAD dramatically influenced latent heat flux with implications for land-atmospheric energy/water coupling.

Luis Gustavo Goncalves

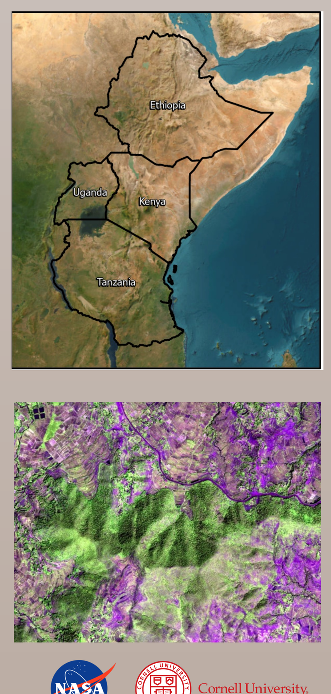


Figure 6. CLM-DART can be used as a verification system for forest restoration projects (East Africa, top). Restored forest is shown in green for a region in Ethiopia (bottom).

Ying Sun, Anthony Holmes.

Generation of Ensemble Spread for LSM EnKF Applications

A proven method for ensemble generation in CLM-DART is to couple CLM with the CAM Reanalysis (Fig. 1), thereby propagating atmospheric uncertainty directly to the CLM states. There are limitations, however, given the latency of the CAM Reanalysis that limit applications for near real-time forecasting. Furthermore the coarse spatial resolution (1x1 degree) of the CAM6 Reanalysis creates a representation mismatch for site level applications. Potential solutions to these challenges include an alternative AR-1 perturbation approach for single instance meteorological forcing (Fig. 7) and an ensemble bias correction approach for site level simulations (Fig. 8).

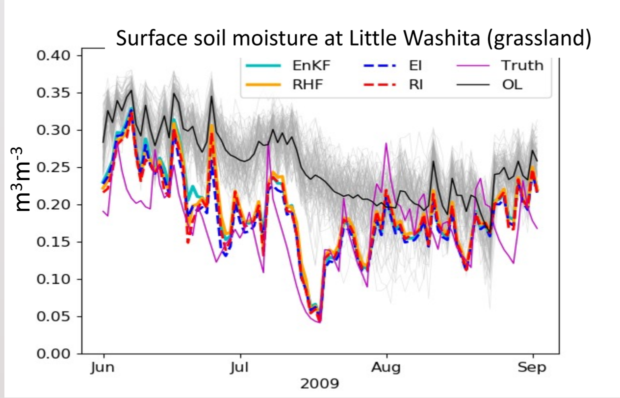


Figure 7. An AR-1 perturbation approach applied to site level meteorological forcing to generate spread in the NASA Catchment model. The OSSE simulations are open loop (OL), rank histogram (RHF), EnKF w/ inflation (EI), rank histogram w/ inflation (RI).

Dibia et al., (2023), (Journal of Hydrometeorology)

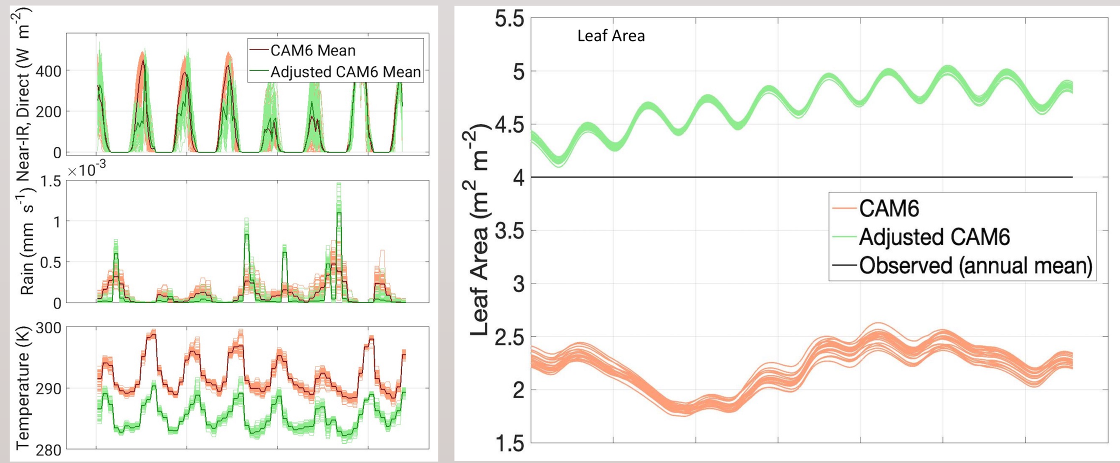


Figure 8. Demonstration of the site level correction applied to the CAM6 reanalysis for Niwot Ridge (US-NR1). The correction approach applies spread to the site level met forcing providing a cooler/wetter simulation and more accurate simulated LAI.

DART Tools: Localization & Adaptive Inflation

Vertical localization is applied to a CLM4.5-SP soil moisture assimilation (Fig. 9) which restricts filter updates (Gaspari-Cohn) to ~100 cm of depth. This removes spurious correlations at deeper levels and promotes model stability. Spatially and temporal adaptive inflation is important for maintaining ensemble spread. Here, a varying inflation damping value was imposed to account for varying seasonal observation density across the Arctic domain for LAI and biomass (Fig. 10).

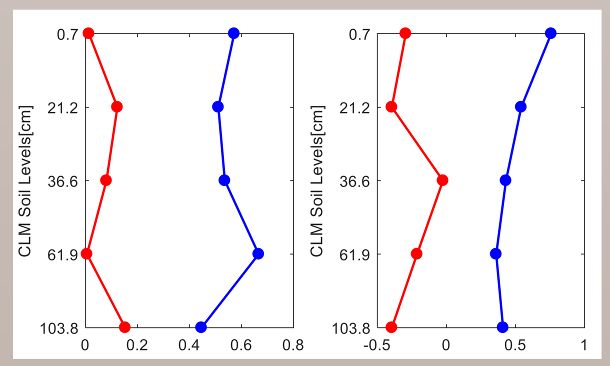


Figure 9. Correlation comparison between the CLM4.5-SP (red) and CLM4.5-SP w/ assimilation (blue) soil moisture with site level soil moisture observations. Observations shown between 0-100 cm. in depth.

Daniel Hagan

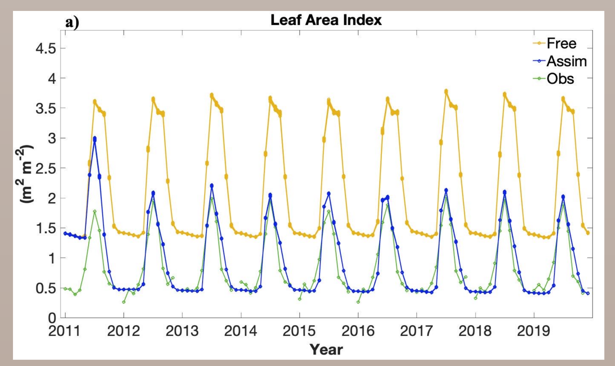


Figure 10. To maintain proper ensemble spread the inflation damping value is set to promote spread during the summer season (high observation density) and to inhibit spread during other seasons (lower observation density).

Huo et al., (in prep)

Development of DART Tools to Support Challenging Observations

DART software tools have been developed to support the assimilation of snow water equivalent (SWE, Fig. 11); carbon and energy flux observations (Fig. 12) and solar-induced fluorescence (Fig.13).

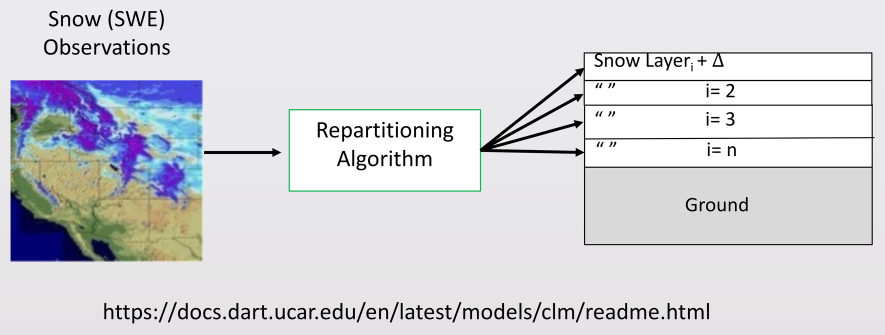
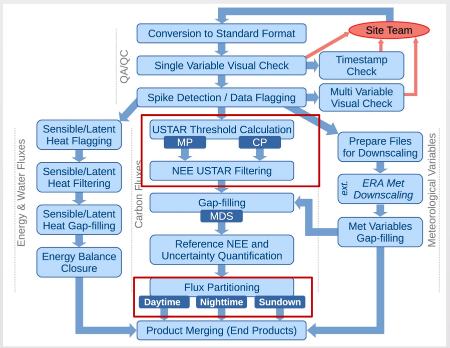


Figure 11. DART-CLM includes a snow layer repartitioning algorithm that allows for a more accurate update based on SWE observations.



Pastorello et al., (2020)

Figure 12. The Fluxnet 'Oneflux' processing database allows for portioned fluxes at multiple temporal resolutions. includeing estimates of random, ustar and partitioning uncertainty.

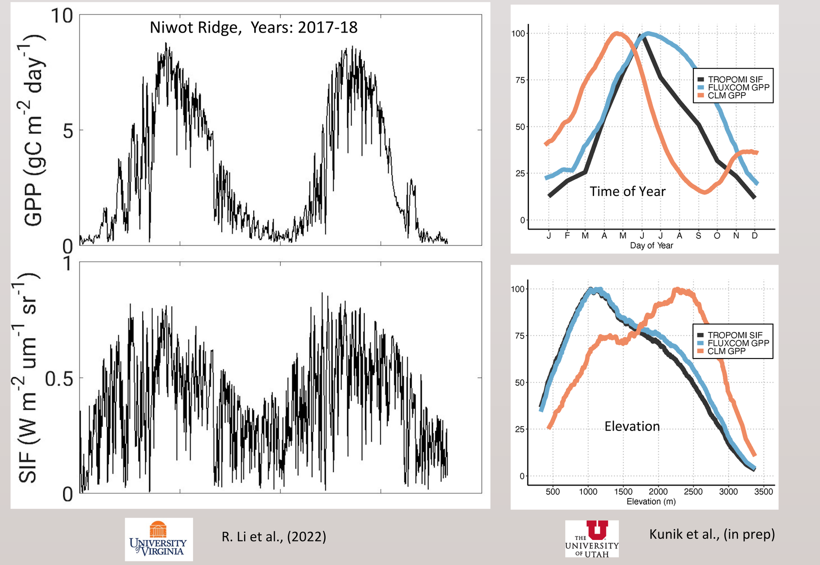


Figure 13. Benchmarking of the canopy SIF forward operator in CLM5 to allow for site/regional level constraint on carbon cycle processes (left panel). SIF provides a potentially strong constraint for the Southern Sierras (right panels) where there is strong mismatch in GPP both in elevation and seasonal timing.

The Quantile Conserving Filter: Addressing Bounded and Non-linear Relationships

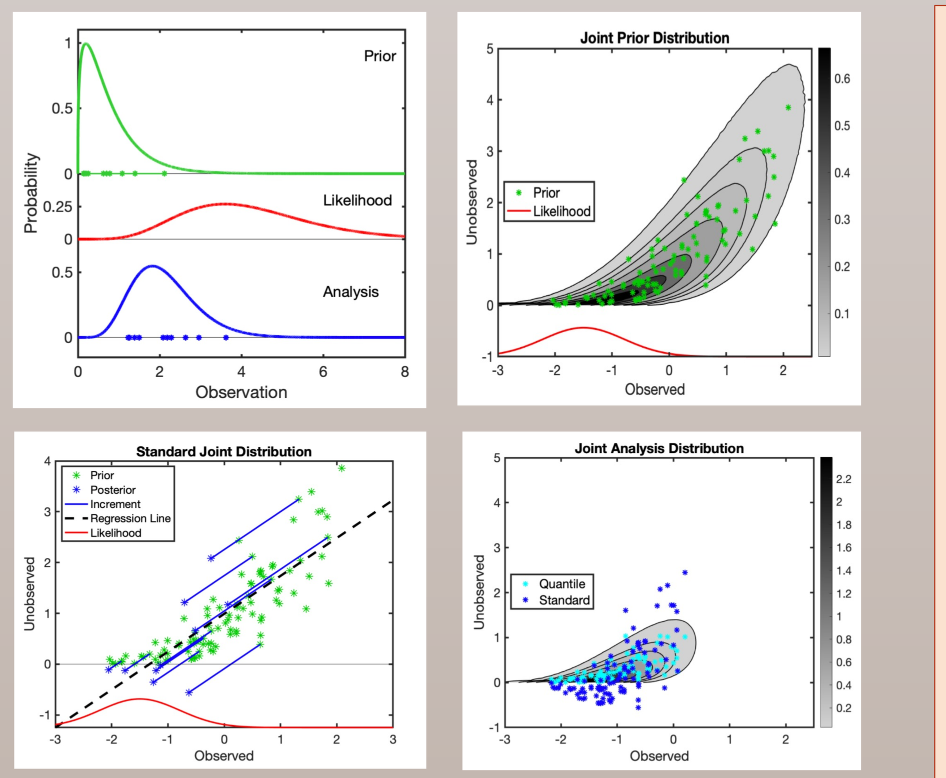


Figure 14. Example of a 0-bounded gamma distributed state confronted with an observation (top left, right). Linear updates with a non-linear joint distribution do not respect either the bounds or the shape of joint posterior distribution (grey, bottom right).

Standard linear filter updates struggle with bounded quantities and non-linear joint distributions. A probit-transformed quantile regression filter, however, respects bounded quantities and better accounts for the curvature in non-linear joint distributions (Fig. 14). Note how the quantile update approach (light blue) better approximates the true joint posterior distribution (grey isolines) as compared to the standard update (dark blue) in the bottom right panel.

CITATIONS

Dibia, E. C., R. H. Reichle, J. L. Anderson, and X.-Z. Liang. 2023. "Non-Gaussian Ensemble Filtering and Adaptive Inflation for Soil Moisture Data Assimilation." *Journal of Hydrometeorology*, **24 (6):** 1039-1053 [10.1175/jhm-d-22-0046.1]

Li et al., (2022) Representation of Leaf-to-Canopy Radiative Transfer Processes Improves Simulation of Far-Red Solar-Induced Chlorophyll Fluorescence in the Community Land Model Version 5; JGR Biogeosciences