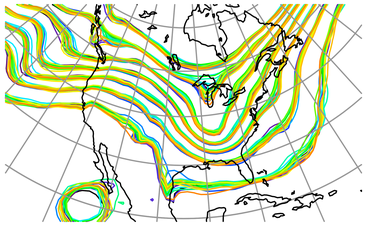


# The Data Assimilation Research Testbed (DART): Ensemble Data Assimilation for NCAR Community Earth System Models


**J. Anderson[1], N. Collins[1], M. El Gharamti[1], A. Fox[2], S. Ha[1], J. Hendricks[1], T. Hoar[1], J. Liang[3], J. McCreight[1], A. Mizzi[5], N. Pedatella[1], K. Raeder[1], A. RafieeiNasab[1], Y. Richter[1], C. Riedel[4], G. Romine[1], and J. Tribbia[1]**

1 National Center for Atmospheric Research, Boulder, CO; 2 University of Arizona, School of Natural Resources and the Environment, Tucson, AZ; 3 University of Texas at Austin, Jackson School of Geosciences, Austin, TX; 4 University of Oklahoma, School of Meteorology, Norman, OK; 5 Colorado Department of Public Health and Environment, Air Pollution Control Division, Denver, CO

dart@ucar.edu


CISL Computational & Information Systems Lab

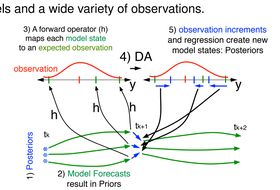
Data Assimilation Research Testbed

## 1. Overview

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 provides data assimilation capabilities for nearly all NCAR community earth system models. The ensemble data assimilation tools provided by DART allow NCAR models to produce ensemble forecasts. The data assimilation process involves combining short model forecasts with observations to produce ensemble analyses that can be used for subsequent forecasts of any length. This process of confronting the model with observations facilitates model evaluation and improvement. The ensemble analyses and forecasts from DART enable analysis and understanding of the earth system.

## 2. DART

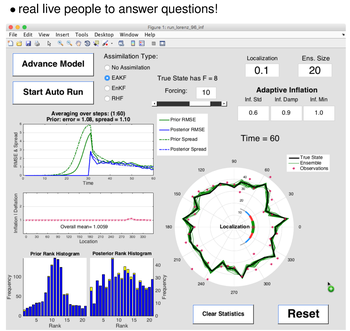
DART has been free and publically available for more than 10 years. Building an interface between DART and a new model does not require an adjoint and **generally requires no modifications to the model code**. DART works with dozens of models and a wide variety of observations.




**www.image.ucar.edu/DAReS/DART** has information about how to download DART, the DART educational materials, and how to contact us.

DART provides both state-of-the-art ensemble data assimilation capabilities and an interactive educational platform to researchers and students. DART contains a variety of instructional material to appeal to different types of learning:

- a tutorial directory with 23 self-paced modules,
- a MATLAB® tutorial with point-and-click GUI examples,
- a user Application Program Interface (API),
- a web site dedicated to explaining how to use DART, and
- real live people to answer questions!




## 3. Regional MPAS *S. Ha, C. Riedel*

A limited-area version of Model for Prediction Across Scales (MPAS) has been recently developed using the same unstructured spherical centroidal Voronoi meshes as in its global configuration (Figure 1). Lateral boundary conditions (LBCs) and the mesoscale physics used in the rectangular grid WRF model (https://www.mmm.ucar.edu/weather-research-and-forecasting-model) are adapted to the Voronoi mesh discretization.

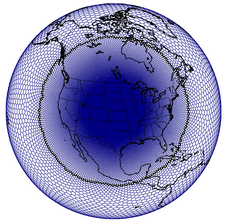


**Figure 1:** *Illustration of the MPAS mesh showing the refinement region between the global and regional configurations.*

When regional domains are expanded to include a coarser outer mesh, the variable-resolution configurations recover most of the error reduction compared to a uniform high resolution at much-reduced cost. New features include the detection of domain boundaries in the unstructured meshes, the update of LBCs in the relaxation zone (Figure 2), and observation processing to reject observations and/or adjust observation error variances near the lateral boundaries (Figure 3).

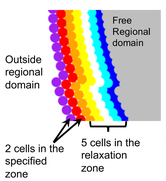



**Figure 2:** *In updating LBCs, the regional analysis is blended with driving data in the relaxation zone while the specified zone is fixed with driving data. The boundary zone is composed of the relaxation zone and the specified zone.*

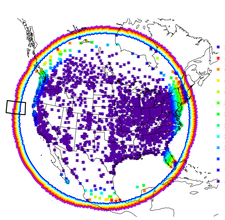



**Figure 3:** *The land surface altimeter observation error variance (marked as a colored star) is gradually increased near the lateral boundary. The boundary zone is depicted as colored dots. The small box is the area shown in Figure 2.*


W. C. Skamarock, M. G. Duda, S. Ha, and S.-H. Park, 2018: *Limited-Area Atmospheric Modeling Using an Unstructured Mesh. Mon. Wea. Rev.*, **146** doi.org/10.1175/MWR-D-18-0155.1


## 4. CAM6 Reanalysis *K. Raeder, Y. Richter, J. Tribbia*

An 80-member DART reanalysis using the latest version of the Community Atmospheric Model (CAM6-FV) is in initial production. In addition to documenting CAMs capabilities as a forecast model, ensembles of atmospheric forcing are being produced that can be used to generate DART reanalyses with other CESM components like CLM and CICE. Results for 2017 will be available soon to facilitate the use of modern remote sensing platforms. Results for 2010-present should be available within the next year.

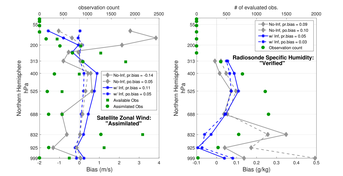



**Figure 4:** *A representative example of the quality of the reanalysis. Left: The monthly average profile of the bias of satellite zonal wind observations for September 2010, which were assimilated. The 'No-Inf' experiment is one without prior inflation, 'w/Inf' uses prior inflation from Gharamti 2018. Right: Radiosonde specific humidity observations were not assimilated and used for verification.*

## 5. WACCM-X *N. Pedatella*

DART and the Whole Atmosphere Community Climate Model with Thermosphere and Ionosphere Extension (WACCM-X) has been used to investigate connections between the tropics and northern latitudes, where the Arctic and North Atlantic oscillations (AO/NAO), and sudden stratospheric warmings (SSW) occur.

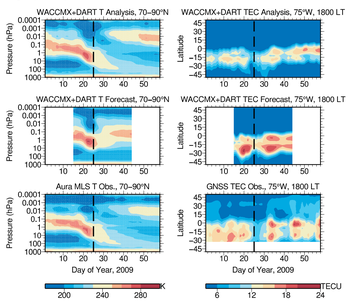



**Figure 5:** *WACCMX+DART analysis fields (top), forecast initialized on January 15, 2009 (middle), and observations (bottom) for the high-latitude zonal mean temperature (left) and total electron content (right). The results show that the model can generally capture the middle and upper atmosphere variability during the 2009 SSW, and that it can be forecast ≈10 days in advance.*


Pedatella N.M., et al. 2018: *Analysis and Hindcast Experiments of the 2009 Sudden Stratospheric Warming in WACCMX+DART. J. Geophys. Res.*, **123**, doi:10.1002/2017JA025107


## 6. CLM *A. Fox*

Our objective is to develop an integrated modeling capability to assimilate observations from JPLs ECOSTRESS and Orbiting Carbon Observatory 3 (OCO-3) data, UMDs Global Ecosystem Dynamics Investigation Lidar (GEDI) data and the Japanese Hyperspectral Imager Suite (HISUI) data using the Community Land Model (CLM) and DART. This system is capable of producing a reanalysis of the land surface with well-characterized uncertainties, with an emphasis on a tight integration between carbon stocks and fluxes.

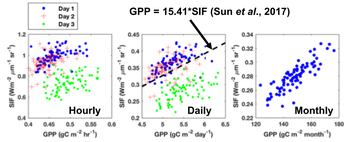



**Figure 6:** *The relationship between Solar Induced Fluorescence (SIF) and Gross Primary Productivity (GPP) in a CLM grid cell across an 80 member ensemble. The OCO2 SIF vs. flux tower GPP relationship is also plotted.*

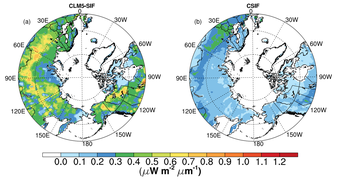



**Figure 7:** *The mean SIF during June, July and August (JJA) of 2012-2014. (a) from CLM5, (b) from CSIF - an OCO-2 SIF observation-based product (Zhang et al. 2018).*

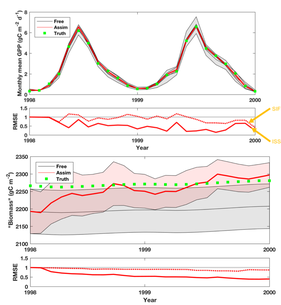
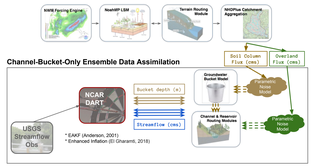



**Figure 8:** *The results from a single site Perfect Model experiment assimilating SIF, Biomass, Leaf Nitrogen and Evapotranspiration. The assimilation of monthly SIF observations improved the model's ability to simulate SIF. However, the greatest improvements in GPP error came from asssimilating all four data streams.*


Fox, A. M., et al. 2018: *Evaluation of a data assimilation system for land surface models using CLM4.5. Journal of Advances in Modeling Earth Systems*, **10**, doi.org/10.1002/2018MS001362


## 7. WRF-Hydro *J. McCreight, A. RafieeiNasab, M. El Gharamti*

The Weather Research and Forecasting Hydrologic model (WRF-Hydro) is a community modeling system and framework for hydrologic modeling and model coupling. WRF-Hydro is configured to use the Noah-MP Land Surface Model to simulate land surface processes. Combined with DART, the facility is called HydroDART.




In this initial configuration, there is little expectation that the posterior states will result in an improved forecast as the system is dominated by components unaffected by the assimilation.

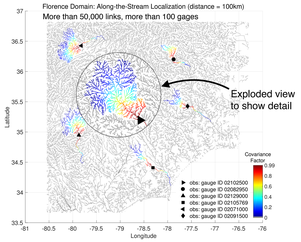



**Figure 9:** *The test domain 'Florence' for WRF-Hydro and DART. The domain is approximately 100,000 km$^2$ on the Carolina coast. The symbols are at stream gage locations and the coloring depicts the localization of an observation. A localization distance of 100km is used for visualization only. The LSM runs at ≈ 1 km resolution.*

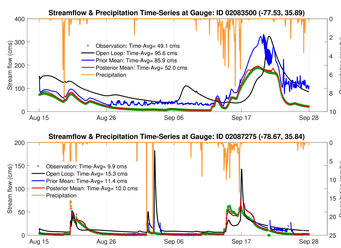



**Figure 10:** *A summary of the performance of the hourly assimilation with 80 ensemble members for two gages. The open loop run is clearly biased high and the timing of the streamflow does not match the gage observations. The assimilation prior is generally much closer to the observed streamflow, and the posterior is better still.*

## 8. WRF-Chem *A. Mizzi*

WRF-Chem/DART is a community resource for realtime chemical weather forecasting/data assimilation research. It couples the Weather Research and Forecasting model (WRF) with online chemistry (WRF-Chem) and DART. WRF-Chem/DART assimilates MOPITT and IASI total and partial column CO, IASI total and partial column $O_3$ (under testing), OMI total column $NO_2$, MODIS AOD retrievals, AirNOW *in situ* observations, and Retrieval profiles as raw retrievals (RETRs) or "compact phase space retrievals" (CPSRs).

WRF-Chem is a state-of-the-art numerical forecast model that simulates the emission, transport, mixing, and transformation of trace gases and aerosols simultaneously with meteorology. WRF-Chem development is a collaborative effort between: NOAA/ESRL, DOE/PNNL, NCAR/ACOM, and various universities. It is well documented and contains an educational platform.

*Nanjing University - Prof. Tijian Wang and Mr. Chaoqun Ma*
Prof. Wang and Mr. Ma are using WRF-Chem/DART to assimilate *in situ* observations of atmospheric chemistry together with MODIS AOD to update emissions and study the analysis and prediction of air quality over eastern China.

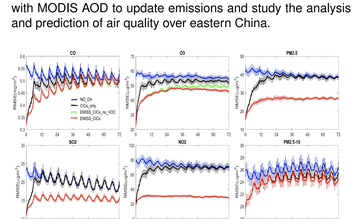



**Figure 11:** *Forecast RMSE ($\mu g/m^3$) at different forecast hours verified against the in situ air chemistry observations.*

## 9. NOAH-MP *J. Liang*

The Noah-multiparameterization land surface model (Noah-MP) and DART are being used to explore the assimilation of total water storage (TWS) observations from the Gravity Recovery And Climate Experiment (GRACE) satellites and MODIS. A single model state was used for 40 ensemble members and then advanced using an ensemble of CAM forcing for 3 years before the experiments.

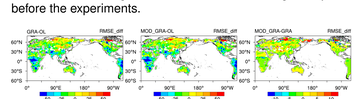



**Figure 12:** *Differences in the RMSE of TWS for 2003. From the left to right each figure indicates GRA (GRACE-only DA) minus OL (Open Loop), MOD_GRA (GRACE and MODIS DA) minus OL, and MOD_GRA minus GRA.*


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

We would like to acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation.