The Data Assimilation Research Testbed: New Algorithms and Applications

J. L. Anderson, G. Romine, K. Raeder, H. Liu, N. Collins, T. Hoar

National Center for Atmospheric Research – Institute for Mathematics Applied to Geosciences

Boulder, Colorado, USA

DART: The Data Assimilation Research Testbed

NCAR's Data Assimilation Research Testbed (DART) provides a cutting-edge facility to perform ensemble data assimilation with support for a wide range of models, conventional and novel observations, algorithms, and diagnostic tools. Several research projects are presented that demonstrate the effectiveness and flexibility of the DART system to produce high-quality products. Our email address is **dart@ucar.edu**, the DART site is **www.image.ucar.edu/DAReS/DART**

Rank Histogram filtering algorithm

DART continues to add tools to improve ensemble filter performance.

- An inexpensive sampling error correction algorithm that automatically localizes observation impacts and can reduce error in large geophysical applications [JLABergen]

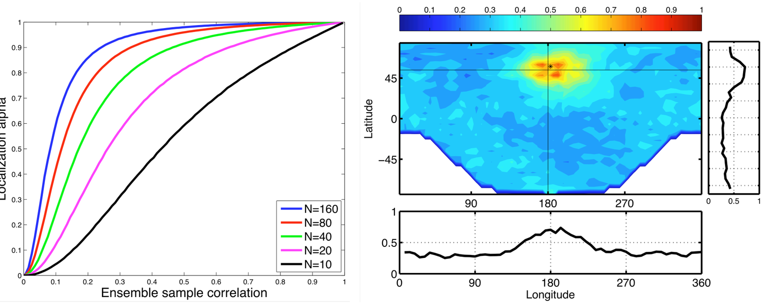


Figure 1: *Left: Localization α based solely on ensemble size (N) and sample correlation. Right: the resulting α for mid-level V obs on the U state variable of a 'perfect model' experiment using 80 members of a dry dynamical core.*

- A non-gaussian observation space update that works with arbitrary prior and observational error distributions [JLA2010]

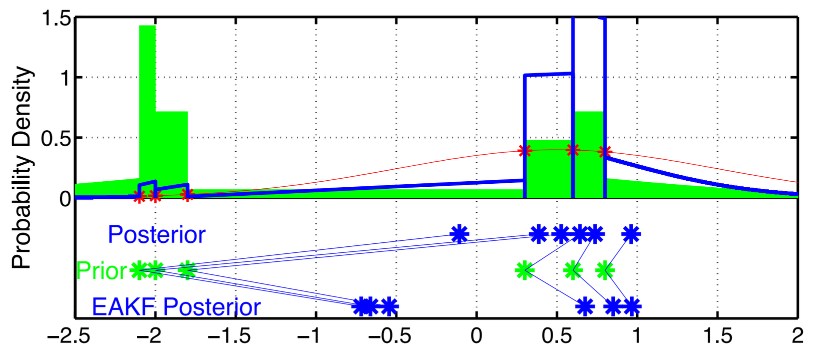


Figure 2: *Non-gaussian ensemble update algorithm: A continuous prior distribution (green histogram) and an observation likelihood (red curve) are combined to give a continuous posterior (blue histogram). Green asterisks mark the prior ensemble while blue asterisks mark the posterior from the new filter (above prior) and an ensemble adjustment filter (EAKF; below). The new filter eliminates the spurious ensemble members near -0.5 in the EAKF.*

Severe Convective Storms on the High Plains

Considerable challenges remain in understanding and predicting the initiation and evolution of high impact convective weather events, particularly in the vicinity of complex terrain. Significant opportunities for improving guidance from storm-scale ensemble forecasts should exist where atmospheric conditions are rapidly evolving such as during convective initiation. A DART/WRF (Weather Research and Forecasting model) ensemble assimilation system has been developed with both meso- and storm-scale probabilistic analyses and forecasts to demonstrate current and future capabilities for storm-scale prediction.

The period of 4-17 June 2009 is examined on a CONUS domain which provides initial and boundary conditions for regional storm-scale analyses centered near the Colorado Front Range (Figure 3).

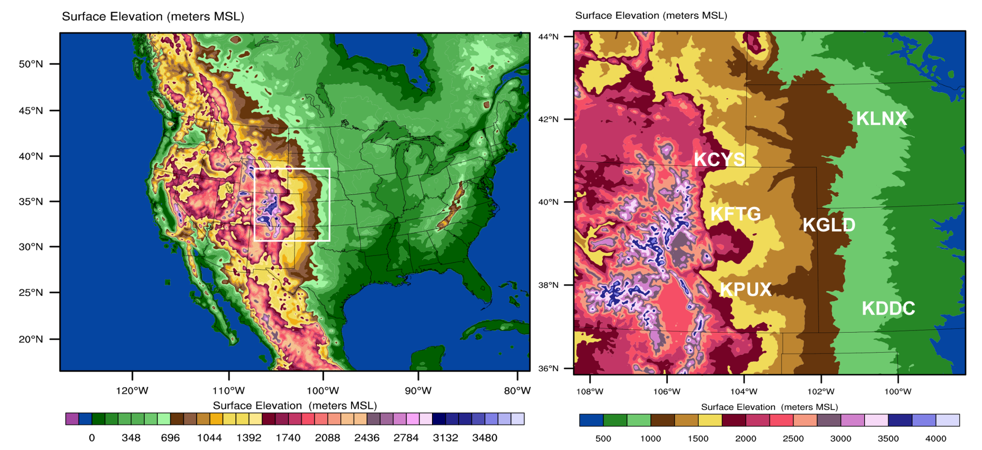


Figure 3: *Surface elevation (shading; m MSL) in the WRF model in the (a) CONUS domain ($\triangle x$ = 15 km) and (b) Front Range domain ($\triangle x$ = 3 km) along with locations of radars where observations were drawn for radar assimilation experiments.*

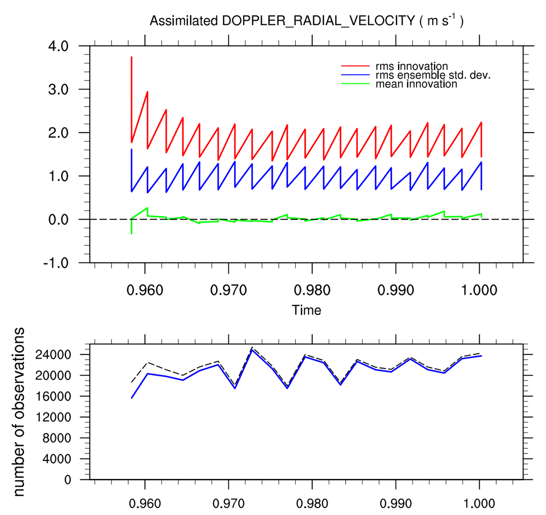
Details about the meso-scale assimilation:

- 3-hourly analysis
- 50 ensemble members
- error statistics comparable to other state-of-the-art mesoscale forecast systems

Details about the storm-scale assimilation:

- 6-hour forecasts every three hours from 15-00 UTC
- continuous assimilation of conventional Doppler radar observations for one hour precedes forecasts

Right: Prior and posterior observation-space diagnostics [m/s; root-mean square of the innovations, ensemble spread, and mean of forecast/analysis minus observations shown in red, blue, and green, respectively] for Doppler velocity during a 1-h assimilation window starting at 23 UTC 11 June, 2009. Times (fractional day) are indicated on the bottom. Assimilated radial velocity observations fit the bred, down-scaled mesoscale background with low RMS and bias.



Control deterministic forecasts, drawn from the meso-scale analysis member with a 'best fit' to the ensemble mean at the start of the forecast period, are also spawned during each ensemble forecast period. In initial tests, radar assimilation leads to improved ensemble storm-scale forecasts (Figure 4) of weather hazards and better quantitative precipitation forecasts.

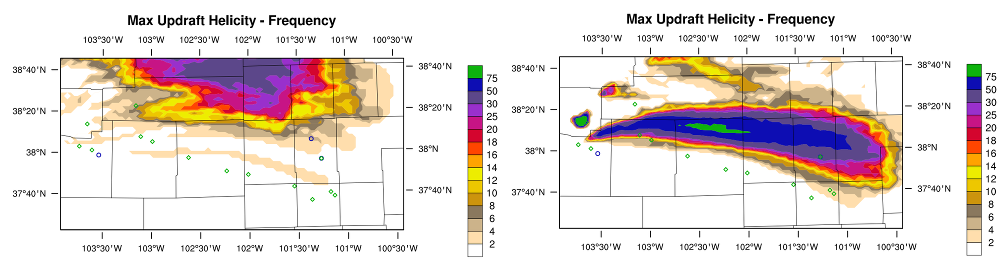


Figure 4: *Ensemble probability (%; shading) of updraft helicity exceeding 75 m^2/s^2 during the forecast period 00$\rightarrow$06 UTC 12 June, 2009 for a portion of the storm-scale domain. SPC preliminary storm reports (hail in green, strong winds in blue) are also shown. Left - control forecast (no radar assimilation), Right - forecast following 21 cycles of Doppler velocity and 'clear air' reflectivity observation assimilation.*

Atmosphere/Ocean Data Assimilation

Decadal forecasts of the ocean are being explored with a loosely coupled data assimilation system in which there is no communication from the ocean assimilation to the atmosphere. A fully-coupled system is planned.

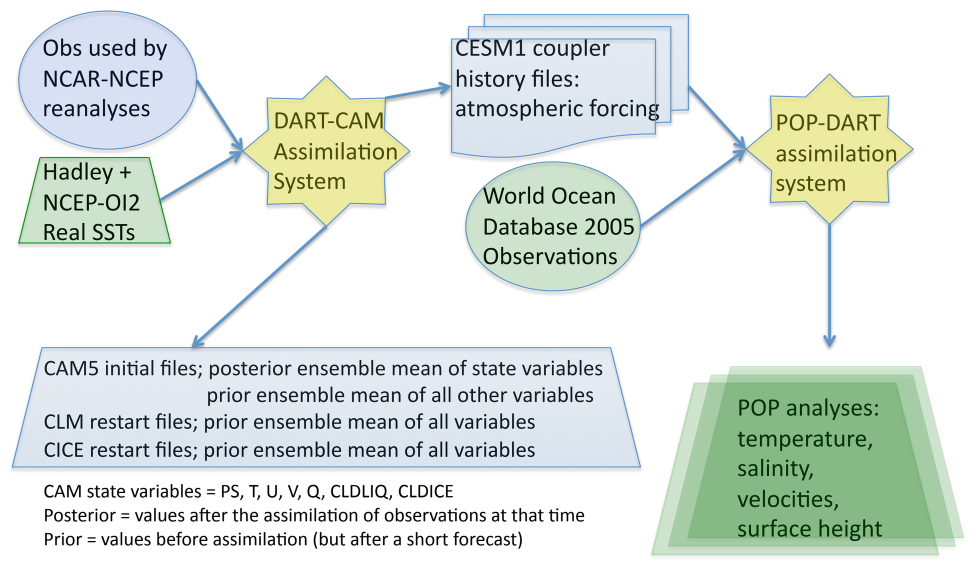


Figure 5: *A schematic describing the coupling between the atmospheric and oceanic data assimilations. DART/CAM is run independently of the ocean assimilation and provides an ensemble of equally likely atmospheric forcings for POP.*

A pair of coupled (ocean-atmosphere-land) forecasts are started at 2000 and will be integrated to 2030 for the IPCC AR5 using the Community Climate System Model version 4 (CCSM4). Initial conditions come from the ends of a forced ocean-ice hindcast, and a two year ocean reanalysis from DART/POP. The initial conditions from DART substantially reduce model error relative to the true state of the ocean, as represented by DART/POP analyses, for up to 2 years, and provide information about the rate of relaxation to the model's climatology.

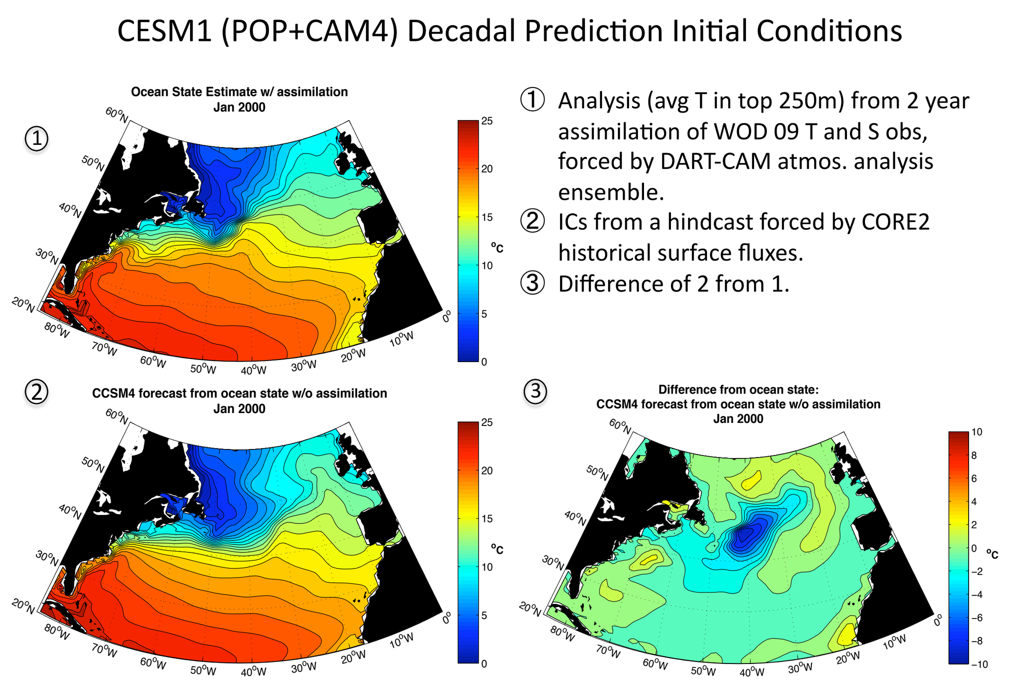


Figure 6: *Average upper ocean temperature (0-250m) in the North Atlantic (1). These represent the initial conditions for the DART/POP forecast. The hindcast initial conditions (2) exhibit a large cold bias, relative to the analyses, where the Gulf Stream fails to turn northward near 45W and a moderate warm bias in the eastern and northern parts of the domain.*

The forecast from hindcast ICs maintains the errors seen in Figure 6, while the forecast from DART/POP ICs (Figure 7; upper right) has much smaller errors (lower panels). Note that the DART/POP forecast is starting to degrade in the area of highest temperature gradient near $50°W$ and $40°N$.

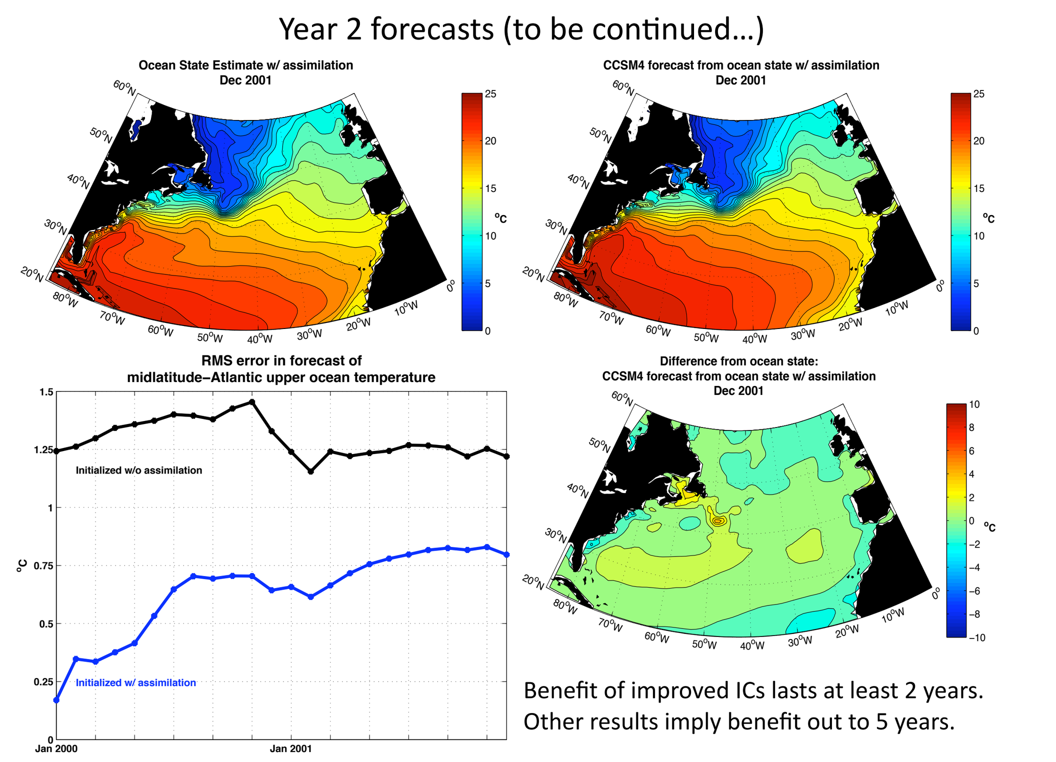


Figure 7: *The ocean state as represented by DART/POP analysis for December 2001 (upper left) is the benchmark for the 2 forecasts. The time series (lower left) show the area weighted RMS error of the average upper 250m temperature for the ocean-ice hindcast simulation (black curve) and the DART/POP IC simulation (blue curve).*

Satellite Soundings and Tropical Storm Forecasts

The DART/WRF (Weather Research and Forecasting model) system has been used to quantify the impact of AIRS water vapor observations on the forecast of intensity and track of Super Typhoon Sinlaku. Sinlaku formed at 06 UTC 8 Sep 2008 over the western Pacific and became a Super typhoon-4 at 18 UTC 10 Sep. We examine the hypothesis that assimilating AIRS Q data for 2 days prior to the genesis can improve the analyses and forecasts of the initial intensification. We use the AIRS Q data as processed by CIMSS – the Cooperative Institute for Meteorological Satellite Studies at the University of Wisconsin-Madison.

These simulations use a standard set of WRF parameterizations and settings:

- 45 km resolution, 45 levels
- 32 ensemble members
- DART adaptive inflation during assimilation
- Initial ensemble mean conditions from NCEP 1 degree global analysis; initial ensemble generated with 3DVar perturbations
- Cycling analysis every 2 hours from 00 UTC 6 Sep to 12UTC 9 Sep
- *FCST run*: Ensemble forecasts from the initial conditions; no observations
- *Only-Q run*: Assimilation of only AIRS Q soundings

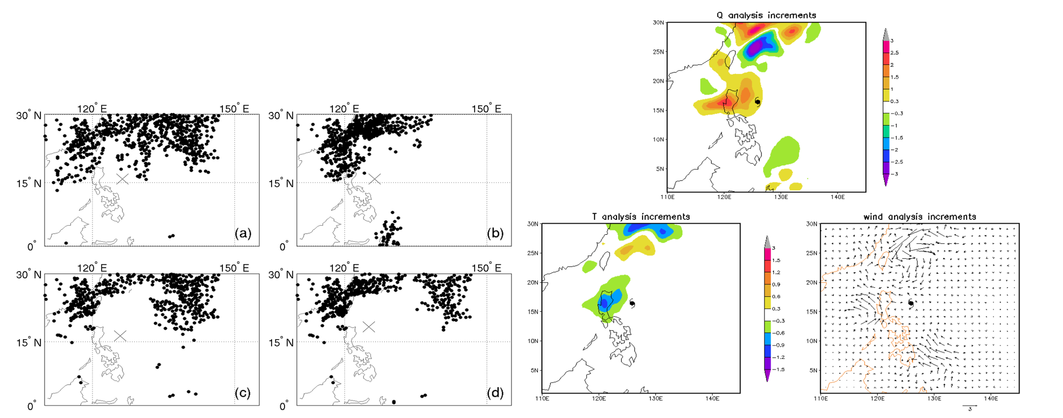


Figure 8: *Left: daily AIRS Q coverage for Sep 6-9, 2008 (a-d, respectively). The need is to spread the information to the unobserved position of the cyclone (the 'X'). Right: The daily analysis increments for 7 Sep 2008. The AIRS Q observations clearly provide information about Q, temperature, and winds.*

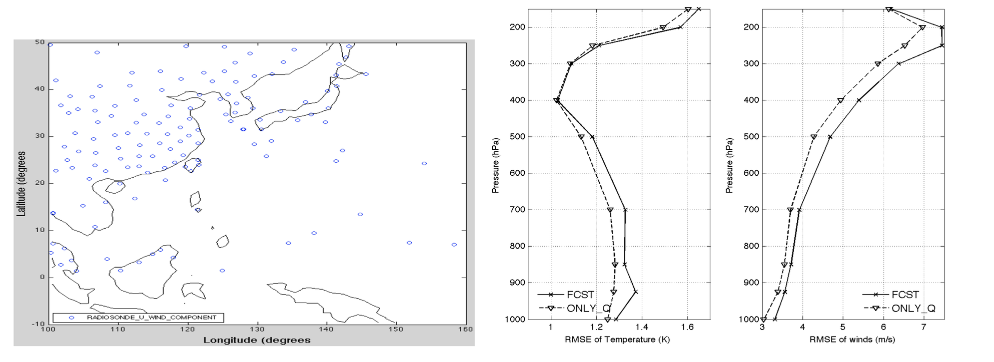
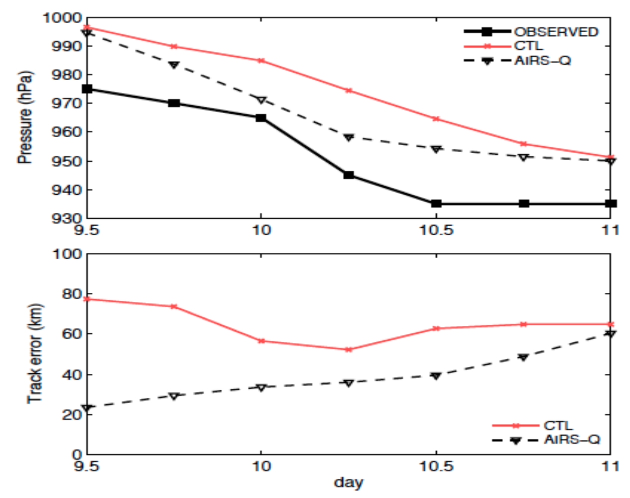


Figure 9: *Left: Locations of the radiosondes withheld for verification. Right: 2-hour forecast fits to the radiosonde observations for 6–9, Sep 2008. The solid lines are for the RMSE of the "no assimilation" experiment, the dashed lines demonstrate the effect of assimilating the AIRS Q soundings .*

A more rigorous test is to compare to an assimilation that uses other observations.

- *CTL run*: Assimilate radiosonde, cloud winds, aircraft data, surface pressure data
- *AIRS-Q run*: Same as CTL plus AIRS Q soundings
- **No** artificial cyclone vortex bogus data is used.
- Each run started at 00 UTC 6 Sep
- cycle every two hours till 12 UTC Sep 12
- launch a forecast and evaluate every 6 hours



Left: Mean of the ensemble forecasts from 12 UTC 9 Sep. Top: The intensity of Sinlaku. The dashed line represents the effect of using the AIRS Q, an improvement over the experiment that did not use the AIRS data. Bottom: The AIRS Q observations clearly improved the track forecast.

References

[DART 09] J. Anderson, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Arellano, 2009: The Data Assimilation Research Testbed: A Community Facility. *BAMS* **90** No.9 pp. 1283–1296

[JLA2010] Anderson, J. L., 2010: A non-gaussian ensemble filter update for data assimilation. *Mon. Wea. Rev.* **138**, pp. 4186-4198

[JLABergen] Anderson, J. L., 2010: Reducing the Impact of Sampling Errors in Ensemble Filters. Presented at The 5th International Workshop on Ensemble Kalman Filter for Model Updating, Bergen, Norway **www.image.ucar.edu/pub/DART/2010/2010 Bergen JLA talk.pdf**

[WOD98] Levitus, S., T. Boyer, M. Conkwright, D. Johnson, T. OBrien, J. Antonov, C. Stephens, and R. Gelfeld, 1998: Introduction, vol. 1. *World Ocean Database 1998*, **NOAA Atlas NESDIS 18**, 346 pp.

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