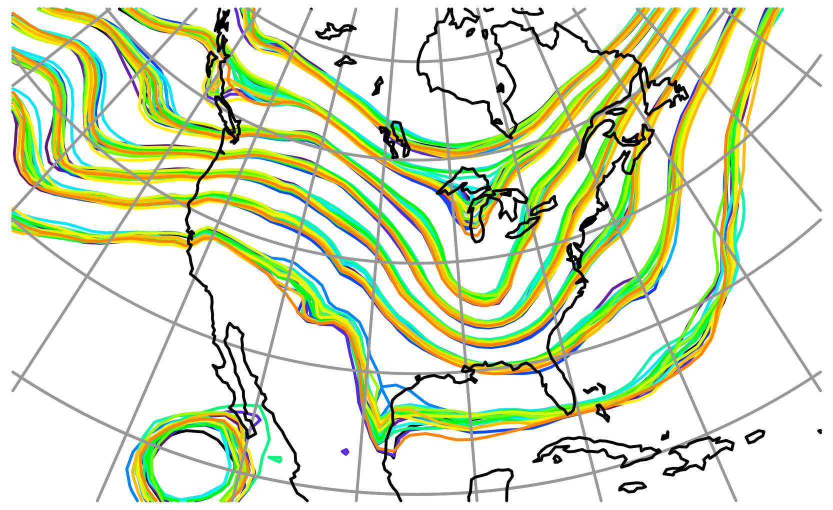


New Developments in the Data Assimilation Research Testbed

K. Raeder, J. Anderson, T. Hoar, N. Collins G. Romine, H. Liu,
A. Karspeck, T. Matsuo, J. Berner, M. Vertenstein, A. Craig, J. Edwards

National Center for Atmospheric Research
Boulder, Colorado, USA

dart@ucar.edu

Institute for Mathematics Applied to Geosciences

1. DART is ...

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 continues to expand in breadth of:

+ models, which now include MPAS Ocean and CESM-CLM,
+ observation sets available, which now include ground water, flux tower, COSMOS, microwave brightness temperature and radar observations such as graupel, hail, etc.,
+ and applications of its products to geophysical research.

With a relatively small investment of effort, it provides both state-of-the-art ensemble data assimilation capabilities and an interactive educational platform to researchers and students. This poster highlights a few recent DART developments.

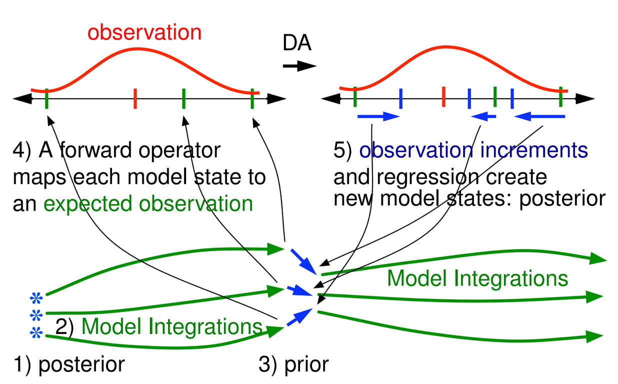


Figure 1: *Illustration for a toy ensemble size of 3.*

2. DART and Multi-Instance CESM

DART can now be used in the "multi-instance" ensemble environment of the Community Earth System Model (CESM; NCAR's global climate model), which facilitates assimilation with any CESM component(s). Multi-instance means that an ensemble of forecasts can be made simultaneously, starting from varying initial conditions and/or using different model parameters.

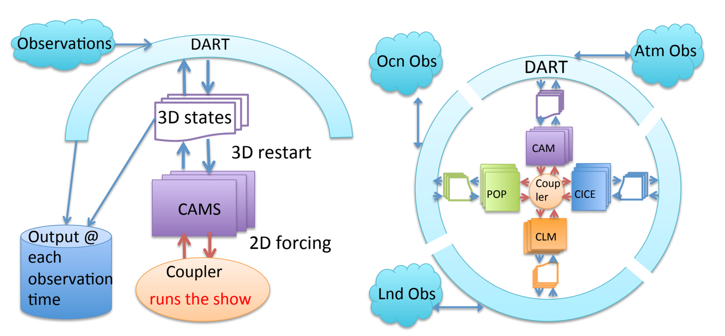
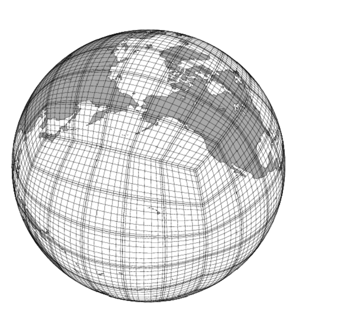


Figure 2: *In the new paradigm, the CESM coupler runs the assimilation, advancing CAM when necessary, and stopping for DART to run when there are observations to assimilate.*

This leverages the software development work of the CESM to make data assimilation available in a cutting edge climate model. It also helps with CESM development, by challenging the model with actual observations.



Using DART within CESM enables assimilation using new CESM components shortly after they become available, e.g. the Spectral-Element dynamical core version of CAM, which uses a non-rectangular grid.

Adding DART capabilities to CESM will facilitate "cross-component" data assimilation in a fully coupled model (active atmosphere (CAM), ocean (POP), and land surface (CLM)). Observations of one component of the earth system will potentially influence all components of the coupled earth system model. For example, observations of cloud cover could directly influence not only the modeled cloud cover, but the modeled ocean and land temperatures as well.

3. Representing Model Error by a Stochastic Kinetic-Energy Backscatter Scheme (SKEBS)

Introduction The development of methods to estimate flow-dependent uncertainty in climate predictions has become an important addendum to the development of climate models. Some model uncertainty is caused by the truncation of the underlying, continuous, differential equations onto a finite grid. Truncation uncertainties can be represented by stochastic-dynamic parameterization techniques, such as SKEBS (Berner et al. 2009), in which energy associated with subgrid processes, instead of being dissipated within the gridbox, is injected back into the resolved scales using a stochastic pattern generator. This method has been successfully used for operational and research forecasts across scales ranging from short-range weather forecasting to annual predictions. Here, we show preliminary results of its CAM implementation.

Results

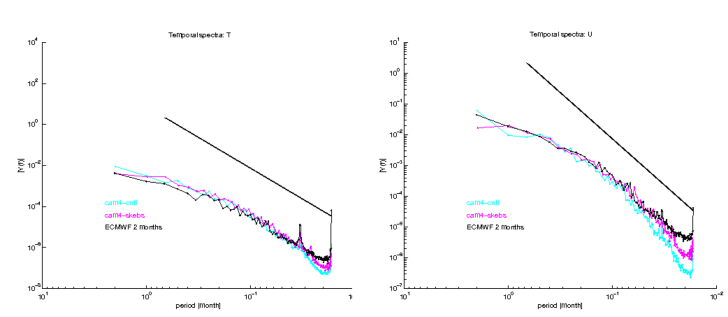


Figure 3: *Temporal power spectra of T (left, K^2) and U (right, m^2/s^2) in CAM4 (cyan), CAM4 with SKEBS (magenta), and of ERA-interim (black). SKEBS introduces power in the sub-monthly scales so that the spectra of CAM-SKEBS are closer to those of the ERA-interim analysis. The straight line denotes a power spectrum $\propto k^{-3}$.*

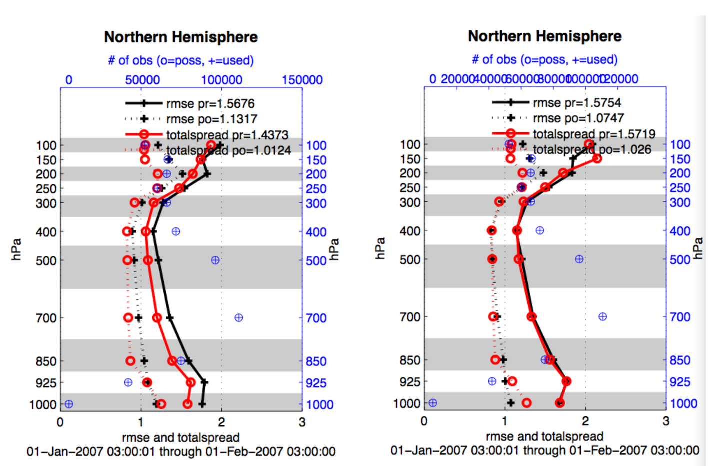


Figure 4: *Total spread (includes observational error variance) and RMS error of posterior in data assimilation experiments utilizing the Ensemble Kalman filter in CAM-DART. The ensemble simulations without SKEBS are underdispersive; the ensemble total spread is smaller than the RMS error of the ensemble (left). SKEBS increases the spread which leads to a better agreement between spread and RMS error (right).*

Conclusions

- CAM4 underestimates the power in the spectra of temperature and wind on timescales associated with sub-synoptic scales when compared to atmospheric ERA-Interim reanalyses.
- SKEBS is able to introduce some of the missing variability in these scales without injecting too much energy at the larger scales (Fig. 3).
- When used within DART, SKEBS leads to a better agreement between spread and error (Fig. 4), but currently produces an increase in the RMSE of the prior.

References

- Berner, J., T. Jung, and T. N. Palmer, 2011b: Systematic model error: The impact of increased horizontal resolution versus improved stochastic and deterministic parameterizations. *J. Climate, in press*, http://dx.doi.org/10.1175/JCLI-D-11-00297.1 .
- Berner, J., G. Shutts, M. Leutbecher, and T. Palmer, 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.*, **66**, 603–626.

4. Global Ocean Assimilation

The ocean assimilation system

- 48 member ensemble adjustment Kalman filter implemented using DART.
- Each of the 48 ocean members forced by a unique member of a CAM4/EaKF atmospheric analysis. Prescribed sea-ice concentrations.
- Assimilation state consists of the prognostic ocean-model state (T,S,U,V,SLH) on the POP2 $1^\circ \times 1^\circ$, 60 vertical level grid.
- Horizontal localization of 11°. No localization in the vertical, allowing the observations at any depth to impact the entire water column.
- No covariance inflation. Spread in the ensemble maintained primarily through the use of an ensemble of atmospheric forcing states.
- Daily assimilation of temperature and salinity data from the 2009 World Ocean Database from Jan 1998-Dec 2005. No satellite SST or altimetry assimilated.

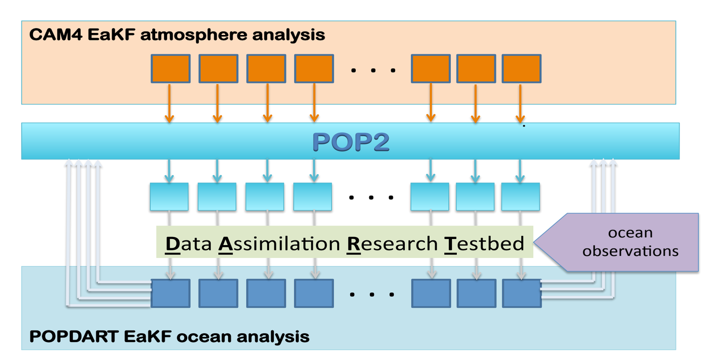
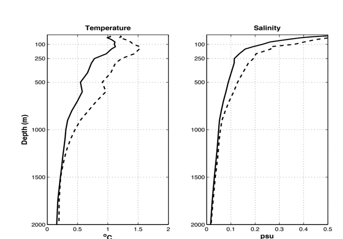


Figure 5: *Chart of the POP+DART data flow. Sets of small boxes are ensembles of model states.*

Comparison to identically forced ocean without assimilation

To understand the role that data plays in constraining the ocean, we compare the ensemble mean analysis with assimilation (hereafter "Assim") to the ensemble mean of an identically atmospheric-forced collection of 48 ocean simulations ("NoAssim"). Generally, the assimilation of data brings the model solution into better agreement with the WOD09 observations above 1000m. Below 1000m, there is no consistent improvement due to assimilation.



This is due in part to a lack of ensemble spread in the deep ocean, limiting the impact of Argo floats from 1000-2000m depth that become available after 2003.

Figure 6: *Average absolute error (misfit) between the daily observations in the WOD09 and the value of the model monthly average interpolated to the geographic location of the observation. Averages from 2000-2006.*

Most of this improvement stems from the correction of systematic time-mean bias in the model and is most notable in strong frontal regions such as the western and eastern boundary currents and the Antarctic Circumpolar Current (Fig. 7).

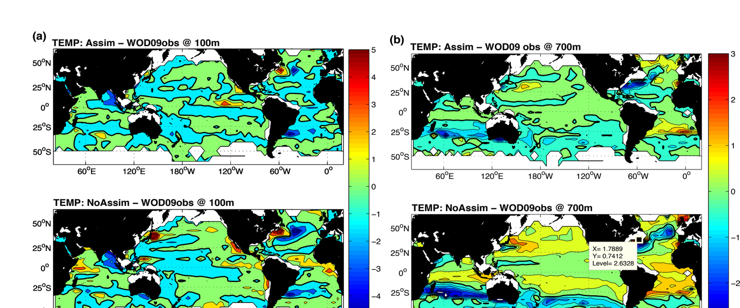


Figure 7: *For Assim and NoAssim, the time-mean model-data bias at 100m and 700m depth. Averages from 2000-2006 are taken over all WOD09 temperature observations in 5° bins at each depth level.*

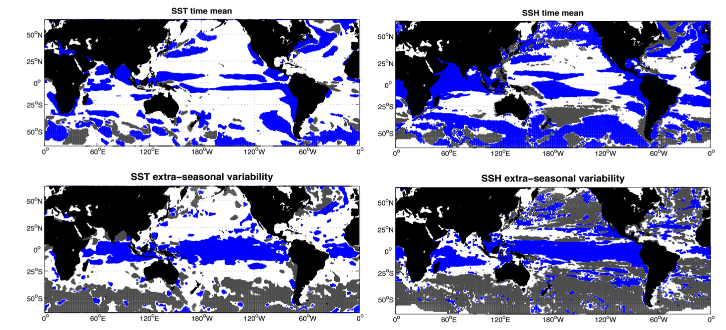


Figure 8: *Blue: assimilation reduces the time-mean (top) and variability (bottom) misfit for satellite based SST and SSH. Gray: degradation of the solution. White: no statistically significant difference.*

References:

- Hurrell et al., 2008: A New SST and sea ice boundary dataset for the CAM. J. Climate, 21, 5145-5153
- Karspeck et al., 2012: An Ensemble Adjustment Kalman Filter for the CCSM4 ocean component. J. Climate submitted.
- Johnson et al., 2009: World Ocean Database 2009 Documentation. NODC Internal Rep. 20.

5. WRF Real-time Ensemble Data Assimilation Supporting Explicit Convective Forecasts

NCAR has engaged in real time, Spring season, forecast exercises since 2003 to highlight capabilities in the Weather Research and Forecast (WRF) model. Since 2011 this has included a data assimilation component, which uses the WRF-DART system to provide initial conditions for convection permitting forecasts. This approach enables crucial testing and evaluation of both the WRF model and the data assimilation system, allowing for significant gains in forecast skill by leveraging the diagnostics provided by the WRF-DART system. Changes in the WRF model physics suite to minimize mean analysis bias improved initial conditions and the skill of the convection-permitting forecast model over previously favored model configuration. Efforts will expand in Spring 2013 to include ensemble forecasts in real time.

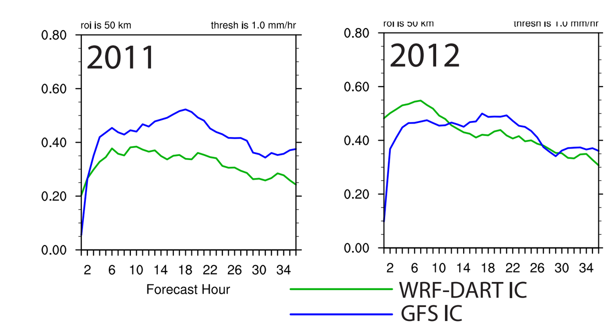


Figure 9: *Fractions skill score (FSS) for precipitation exceeding 1 mm/h, as a function of lead time from the 2011 and 2012 real-time forecasts. The forecasts use 3-km resolution and begin from the WRF/DART analysis (green) or NCEP's GFS analysis (blue). At each forecast grid point and for a specified influence radius (here 50 km), the FSS compares the fraction of points within that radius exceeding the threshold (1 mm/h) against the fraction of observed points in the same area exceeding the threshold. A value of 1 means perfect agreement at all grid points between the forecast and observed fractions. Forecasts with no skill have FSS = 0.*

During the Spring 2012 real time exercise several high impact severe weather events were well forecast, including the June 29 2012 derecho. We are currently investigating initial condition sensitivity drawn from the EnKF analysis to provide useful forecast guidance for events such as this.

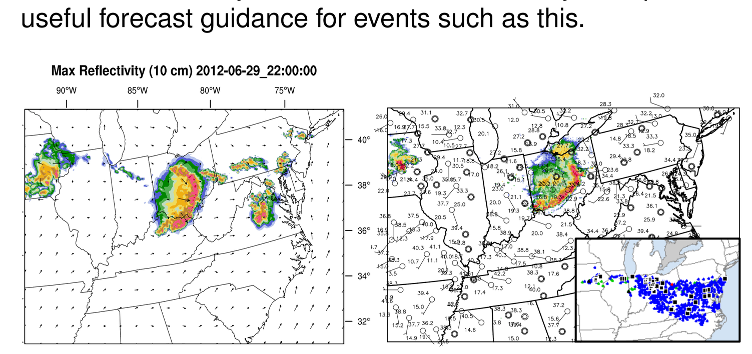


Figure 10: *Simulated reflectivity from 10 h WRF forecast (top), observed radar reflectivity composite and surface observations (bottom) both valid at 22:00 UTC on June 29, 2012, as well as preliminary storm reports from the 29 June 2012 derecho (inset, bottom panel).*

6. Thermosphere-Ionosphere Electrodynamics

TIEGCM+DART is ensemble Kalman filtering applied to NCAR's Thermosphere Ionosphere Electrodynamics General Circulation Model. TIEGCM uses self-consistent solutions for the coupled nonlinear equations of hydrodynamics, neutral and ion chemistry, and electrodynamics. It incorporates the feedback between plasma and neutral variables in both the analysis and forecast steps of filtering so that thermospheric parameters can be inferred from ionospheric observations and vice versa.

The experiments in Fig. 11 assimilate FORMOSAT-3/COSMIC (F3/C) GPS Occultation Experiment observations. The F3/C electron density profiles (EDPs) are uniformly distributed around the globe, which provide an excellent opportunity to monitor the ionospheric electron density structure. The three experiments assimilate:

1. Only electron density, "NE".
2. NE and thermospheric temperature ("TN") and winds ("UN" and "VN").
3. Exp 2, plus the atomic and molecular oxygen mixing ratios ("O" and "$O2$").

This is a "perfect model" study, where the truth is a free model run, and observations are created from the truth, with realistic observation errors added to them.

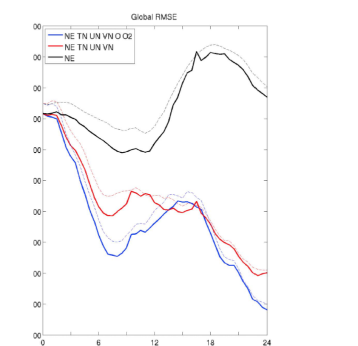


Figure 11: *The RMSE of the electron density analysis (the posterior mean) relative to the truth, summed over the globe, is shown by solid lines for Exp 1 (black), Exp 2 (red), and Exp 3 (blue). Dashed lines represent the RMSE of the prior mean.*

While adding thermospheric variables to the assimilation (Exp 2) has a dramatic effect on the RMSE, the impact of adding compositions (O and $O2$) to the assimilation (Exp 3), may seem small, but it grows over time. Considering that the thermosphere-ionosphere coupling is fully taken into account in the forecast step in all three experiments, it indicates that self-consistent treatment of the thermosphere and ionosphere in data assimilation schemes significantly improves a global ionospheric specification.

Figure 12 illustrates adjustments to the altitude and magnitude of the electron density distribution near the F3/C EDP locations due to assimilating F3/C EDP. Both the peak density and height of the F region are changed. The equatorial ionization anomaly (EIA) features become more prominent in the Northern Hemisphere except for 11:00, 15:00 and 19:00 LT. Moreover, the two EIA crests shift closer to the geomagnetic equator from the 11:00 LT slice, and acquires a sharper poleward edge in the Southern Hemisphere at 14:00 and 15:00. The electron density is enhanced above 250 km over the geomagnetic equator for 22:00, 23:00 and 01:00 LT.

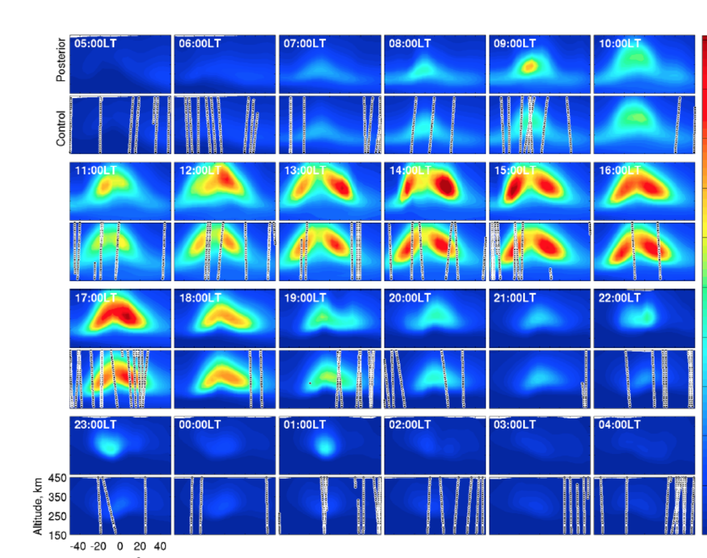


Figure 12: *Latitude/height slices along $75^\circ E$ from 2008/04/12 05:00 UT (10:00 LT) to 2008/04/13 04:00 UT (23:00 LT) before and after assimilating the FORMOSAT-3/COSMIC electron density profiles. The upper and lower row of each panel displays the posterior and control, respectively. The black dots indicate the observation profiles located between 60 and $90^\circ E$ longitude within 30 min of a given assimilation time.*

Figure 13 is taken from a forth-coming study comparing of the effect on modeled neutral mass density (NMD) of assimilating CHAMP neutral mass density or F3/C electron mass density (EMD) in a "perfect model" context. The global NMD RMSD is significantly reduced by assimilation of F3/C electron density profiles, and the degree of the reduction is far greater than assimilation of the CHAMP NMD (CHAMP results not shown; see Matsuo, et al., 2013, below).

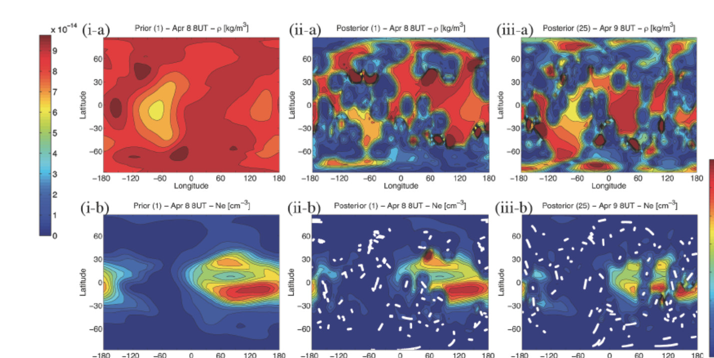


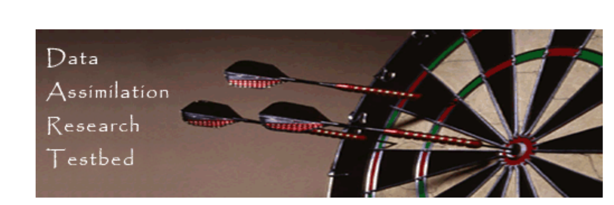
Figure 13: *The root-mean-square deviation (RMSD) of the neutral mass density analysis from the truth for Apr 8 2008, 8 UT (i- and ii-a) and 24 hours later (iii-a) is displayed in the top 3 frames. The number in parentheses on the top of each frame indicates the assimilation cycle. The color scale is at the top left. The units are kgm^{-3}. The contour interval is 0.75×10^{-14}. The bottom 3 frames are similar, but for the RMSD of the electron density analysis. The color bar is at the bottom right. The units are cm^{-3}. The contour interval is 0.25×10^{4}. White dots indicate the location of COSMIC/FORMOSAT-3 profiles.*

TIEGCM figure references:

- Lee, I. T., T. Matsuo, A. D. Richmond, J. Y. Liu, W. Wang, C. H. Lin, J. L. Anderson, and M. Q. Chen (2012), Assimilation of FORMOSAT-3/COSMIC electron density profiles into a coupled thermosphere/ionosphere model using ensemble Kalman filtering, J. Geophys. Res., 117, A10318, doi:10.1029/2012JA017700.
- Matsuo, T., and E. A. Araujo-Pradere (2011), Role of thermosphere-ionosphere coupling in a global ionospheric specification, Radio Sci., 46, RS0D23, doi:10.1029/2010RS004576.
- Matsuo, T., I-Te Lee, and J. L. Anderson, Thermospheric Mass Density Specification using an Ensemble Kalman Filter. J. Geophys. Res. (in review for minor revisions, expected 2013).

7. Further Information

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



8. References

J. Anderson, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Arellano, 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.

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