Ensemble Kalman Filter data assimilation in global MPAS/DART

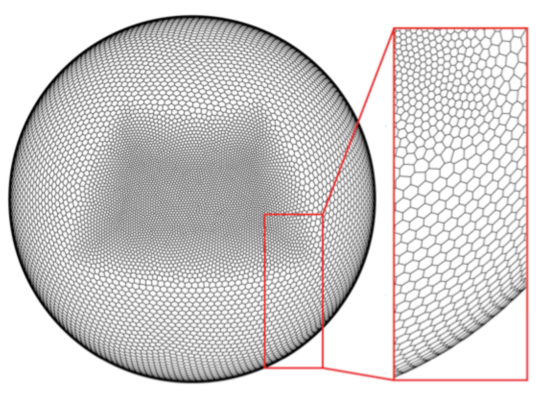
So-Young Ha and Chris Snyder

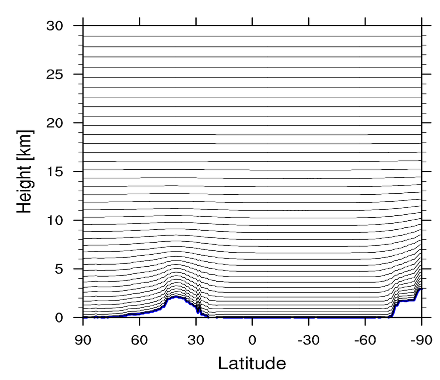
National Center for Atmospheric Research, Boulder, Colorado, U.S.A

Sponsored by the National Science Foundation

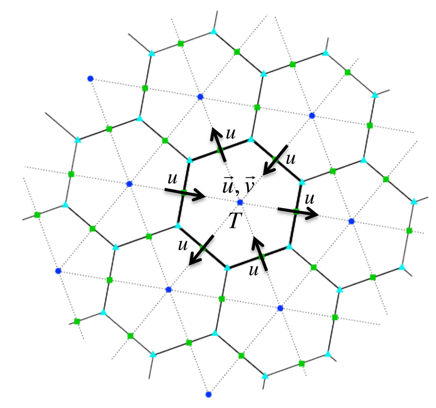
The Model for Prediction Across Scales (MPAS)

The atmospheric component of MPAS (MPAS-A) has been developed in the *Mesoscale and Microscale Meteorology (MMM)* division at NCAR since 2008. The MPAS uses unstructured grid meshes based on spherical centroidal Voronoi tessellation (SCVT) which allows variable resolution meshes with smooth transition from coarse to fine grids. The MPAS hydrostatic and non-hydrostatic solvers use a terrain-following height coordinate and a C-grid staggering of the prognostic variables.





Cross-section of the vertical coordinate along 110 W.



The dual mesh of a Voronoi tessellation is a Delaunay triangulation, where the corners of the triangles are the generating points of the Voronoi tessellation.

- All scalar fields and reconstructed winds are defined at "cell" locations (blue circles)
- Normal wind (u) is defined at "edge" locations (green squares)
- "Vertex" locations (cyan triangles) are used in the searching algorithm for an observation point in the observation operator

Ensemble Kalman filter Data Assimilation

The nonhydrostatic MPAS solver is now coupled to the Data Assimilation Research Testbed (DART) system with a full capability of ensemble Kalman filter data assimilation for various observations.

- All the prognostic variables in MPAS – potential temperature (θ), dry density (ρ), vertical velocity (w), normal wind speed (u), hydrometeors (qv, qc, qi, qs, …) – can be used as state vectors in DART.
- In MPAS, the normal wind speed on the edges (u) is the only prognostic variable for wind components. Meteorological u-, and v-wind ($\vec{u}, \vec{v}$) are reconstructed winds at cell centers and are diagnostic variables in the model.
- In DART, reconstructed winds at cell centers ($\vec{u}, \vec{v}$) can be used as state vectors instead of normal wind speed (u).

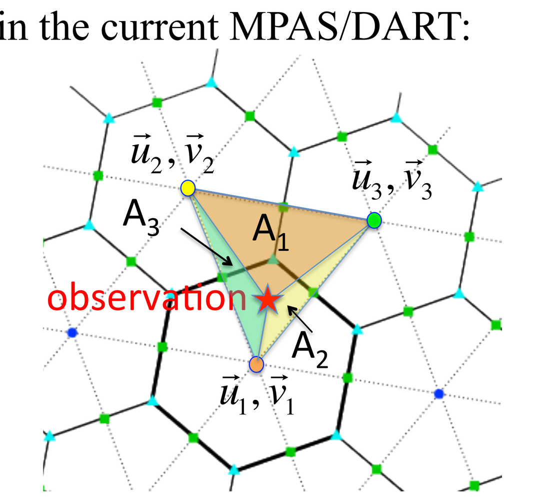
Observation operators use dual mesh at cartesian coordinate:

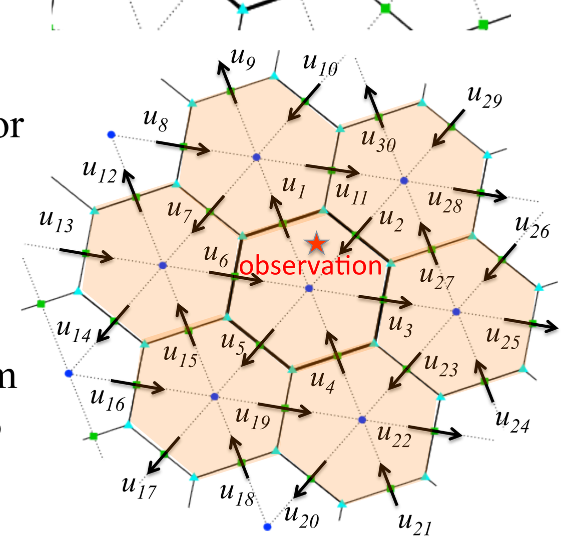
- No interpolation issues at the poles
- For all scalar variables, we find a triangle with the closest cell center to a given (observation) point, then do the barycentric interpolation in the triangle.

$(A_1x_1 + A_2x_2 + A_3x_3)/(A_1 + A_2 + A_3)$ *where x can be reconstructed* $\vec{u}, \vec{v}$, *or any scalar variable*

- For horizontal winds, three different options are available in the current MPAS/DART:

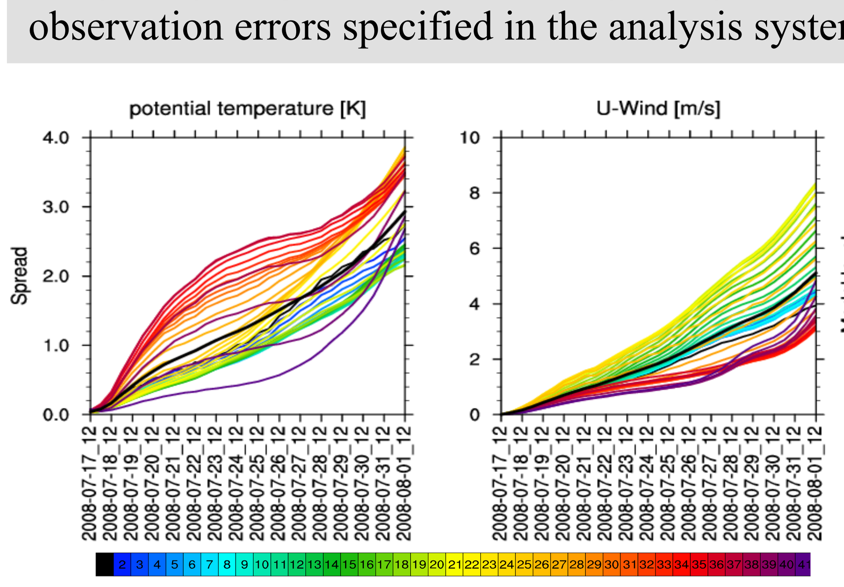
1. Use of reconstructed u-, v-wind at cell centers both in the forward operator and the filter update; then, update the prognostic normal wind (u) by increments of u-, v-winds at cell centers. Same as the scalar variables => currently being used.
2. Use of reconstructed u-, v-wind at cell centers in the forward operator but update the normal wind (u) directly in the filter (i.e., as a state vector) => more memory required.
3. Use of the normal wind (u) both in the forward operator (w/ radial basis functions; RBFs) and the filter update. => more computational resources required; discontinuity at edge boundaries or smoothing effect depending on the number of edges in use for the interpolation. Default behavior uses normal winds from all the edges of the cells surrounding the cell closest to an observation location (e.g., $u_1, u_2, \ldots, u_{30}$).



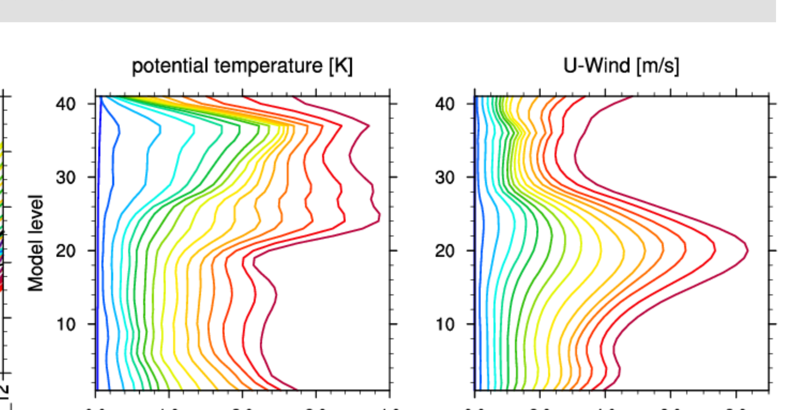


Ensemble spread in the free ensemble forecast

An initial 80-member ensemble was generated by adding tiny random perturbations to the GFS analysis data and running ensemble forecast for 15 days. An initial ensemble for the analysis cycling was constructed by 7-day forecast of which spread is comparable to observation errors specified in the analysis system.



Domain-averaged ensemble spread at different model levels. Each color shows each level.

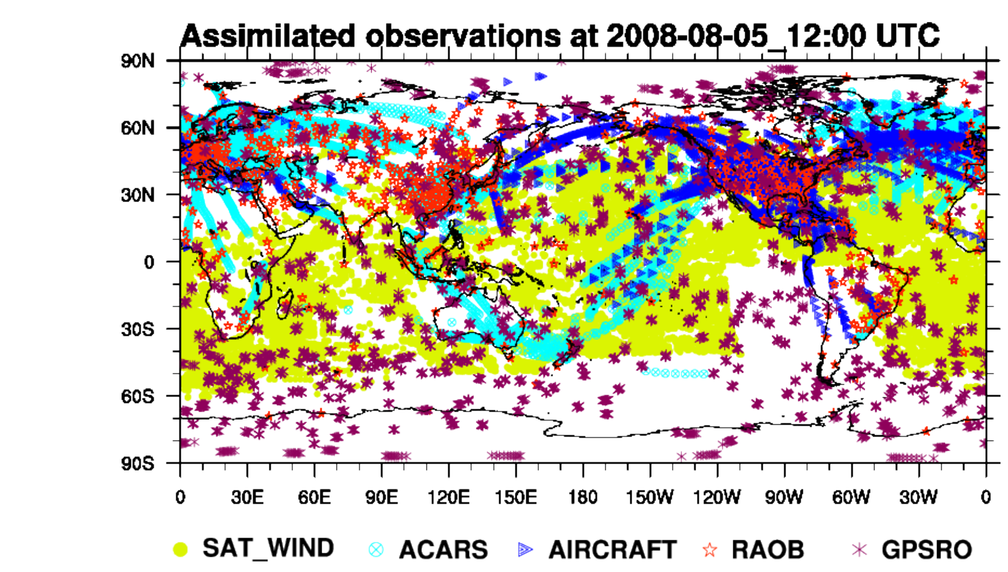


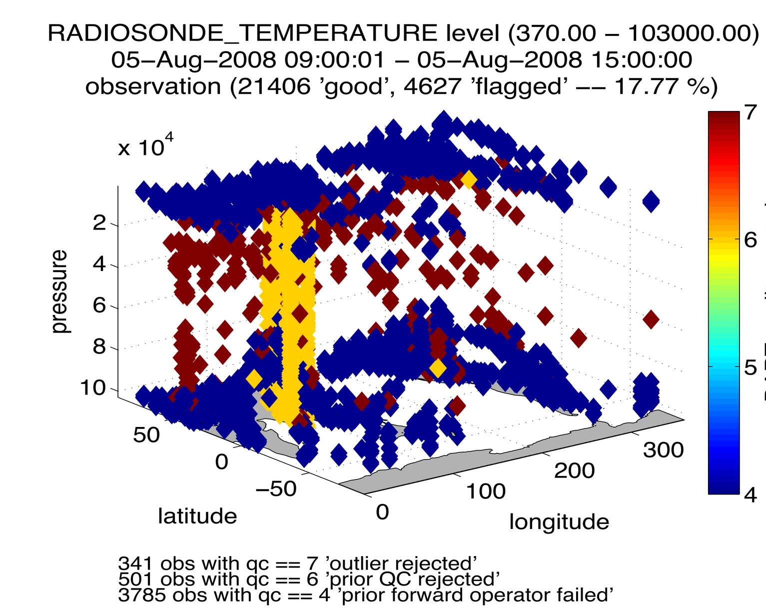
Domain-averaged ensemble spread at different cycles. Each color shows each day from 2008-07-17_12Z.

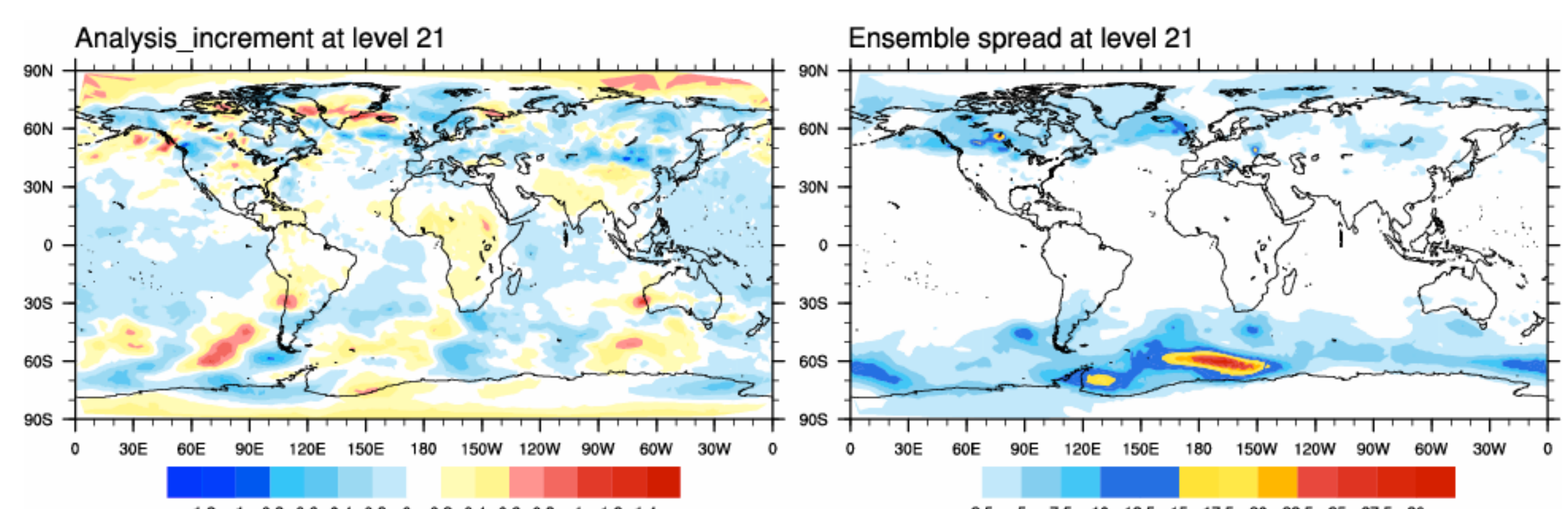
The interface between MPAS and DART was developed in collaboration with Jeff Anderson, Nancy Collins, and Tim Hoar, and the CAM/DART was run by Kevin Raeder in IMAGe/UCAR.

Real data assimilation

- Model configuration: 80-member ensemble at ~2-degree resolution, 41 vertical levels w/ the model top at 30-km
- Filter design: localization with half-width radius of ~1200-km (horizontally) and 4-km (vertically), adaptive inflation in prior state, 6-hrly cycling for one month of August 2008.
- Physics: WSM6 microphysics, YSU PBL, NOAH LSM, Tiedtke cumulus parameterization, CAM SW/LW radiation schemes



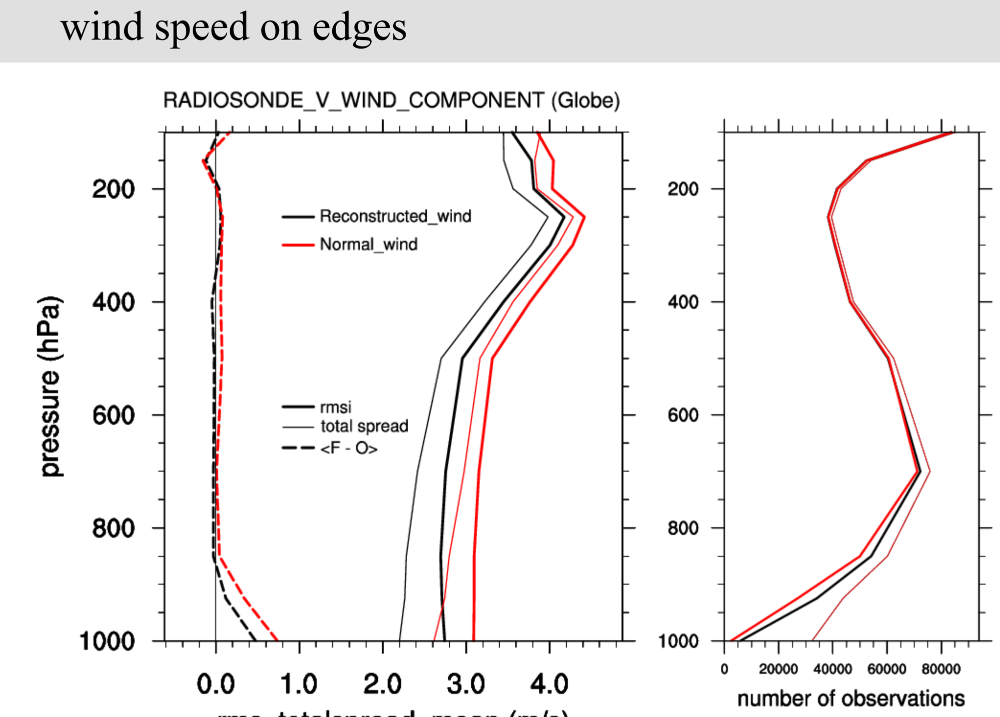




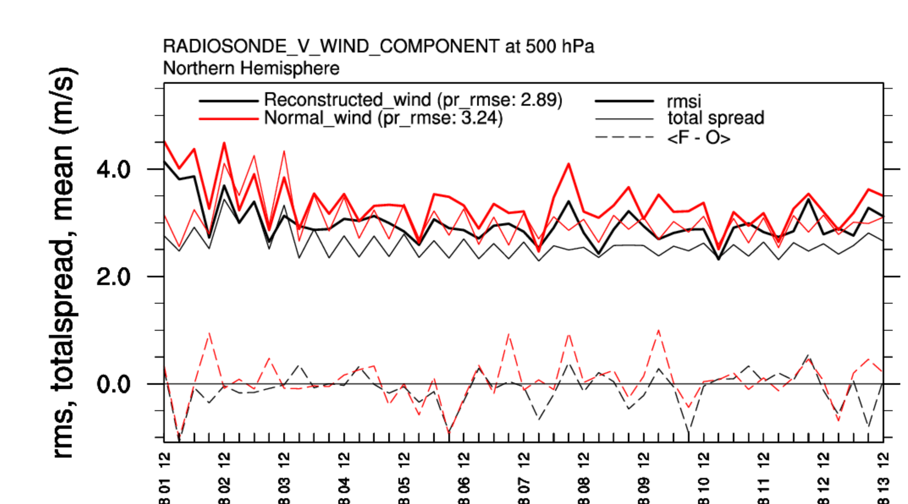
The analysis increment (left) and ensemble spread (right) at level 21 (~ 200 mb) in potential temperature [K] averaged over the last 10 days of August 2008.

Wind data assimilation strategy

- Comparison of different observation operators for wind assimilation in the real data assimilation context
- Method #1with the use of reconstructed winds at cell centers vs. method #3 with RBF functions on the normal wind speed on edges



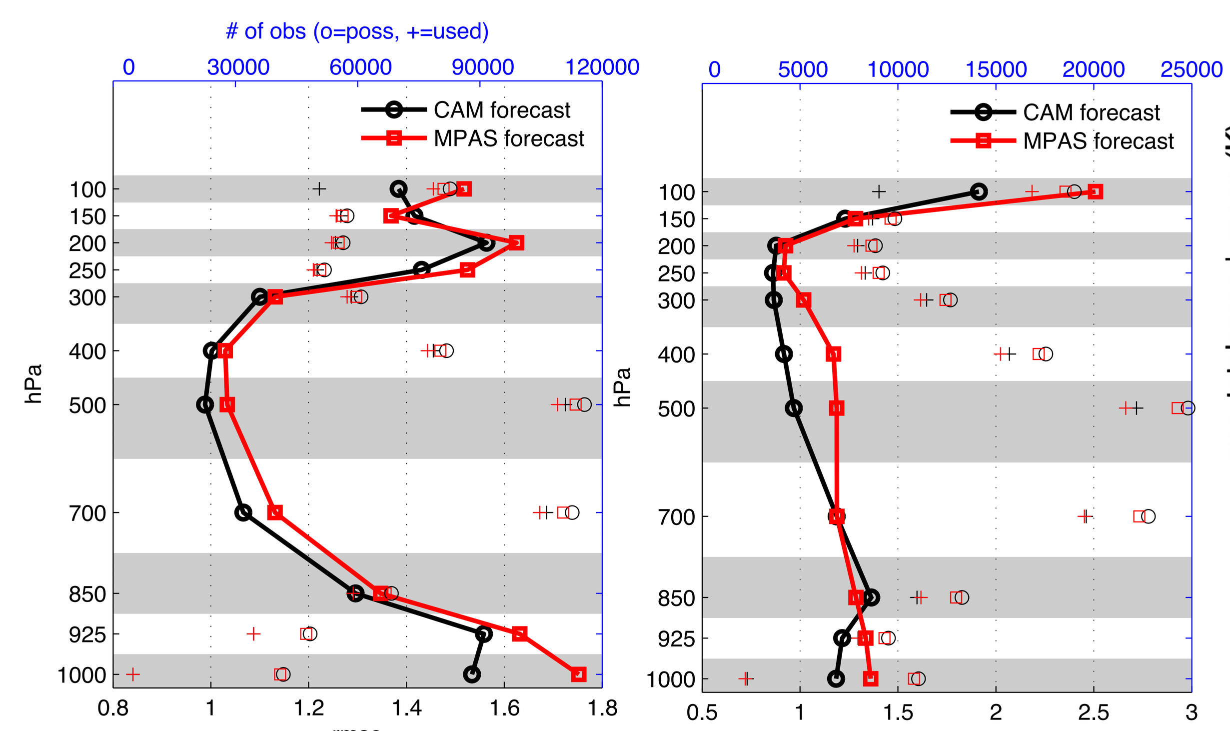
RMS errors of 6-hr ensemble mean verified against sounding temperature in the whole globe (left) and the numbers of observations available (thick lines) and used (thin lines) in the verification (right). Vertical profiles are averaged from 08-06-12_00:00 to 08-13-12:00:00 UTC 2008.



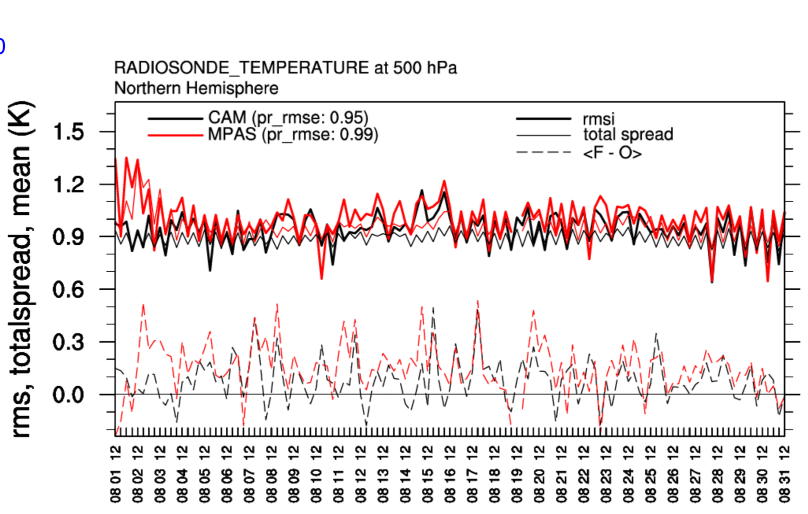
A time series of rms errors (thick solid lines), bias errors (dashed lines) and total ensemble spread (thin solid lines) of 6-hr ensemble mean forecast verified against sounding temperature at 500 mb in northern hemisphere. The numbers in the parenthesis next to the model names are the time-mean rms errors.

Comparison to CAM/DART

- CAM available in CCSM4 at 2-degree resolution
- Same filter design and same observations used for the same period



RMS errors of 6-hr ensemble mean forecast in CAM (black) and MPAS (red) verified against sounding temperature in northern hemisphere (left) and tropics (right). Vertical profiles are averaged for the last three weeks of 2008. The numbers of observations available (marked as "o") and used ("+") in the verification are also drawn with the x-axis on the top.



A time series of rms errors (thick solid lines), bias errors (dashed lines) and total ensemble spread (thin solid lines) of 6-hr ensemble mean forecast in CAM (black) and MPAS (red) verified against sounding temperature at 500 mb in northern hemisphere.

Summary and future plans

Ensemble Kalman filter data assimilation was successfully implemented for MPAS using the DART system. The analysis cycle with real data assimilation is reliable over the whole globe and the performance is comparable to CAM/DART.

We plan to test the same analysis/forecast cycling on the higher resolution or variable grids. For the better performance skills, various efforts on importing global physics to MPAS are ongoing.