

# Modeling and Data Assimilation systems of the future

Jeff Whitaker

NOAA/ESRL

# Trends in NWP modeling

↑ Polyhedral global grids (cube, icosahedron) for scalability

↓ Spectral and lat/lon dycores

↑ Global (variable resolution) models

↓ Regional (nested) models

↑ "Cloud-permitting" non-hydrostatic global models (still with parameterized shallow convection).

↓ Hydrostatic models with parameterized deep convection

↑ Ensembles with stochastic physics.

↓ Deterministic forecast systems.

# Trends in Data Assimilation

↑ EnKF, “weak-constraint” 4DVar, hybrid “incremental” 4DVar/EnKF (freely evolving **B**).

↓ “Incremental” 4DVar, 3DVar (depending on specified **B**).

↑ Direct assimilation of satellite radiances.

↓ Assimilation of retrievals.

↑ Online, adaptive bias correction.

↓ Offline bias correction.

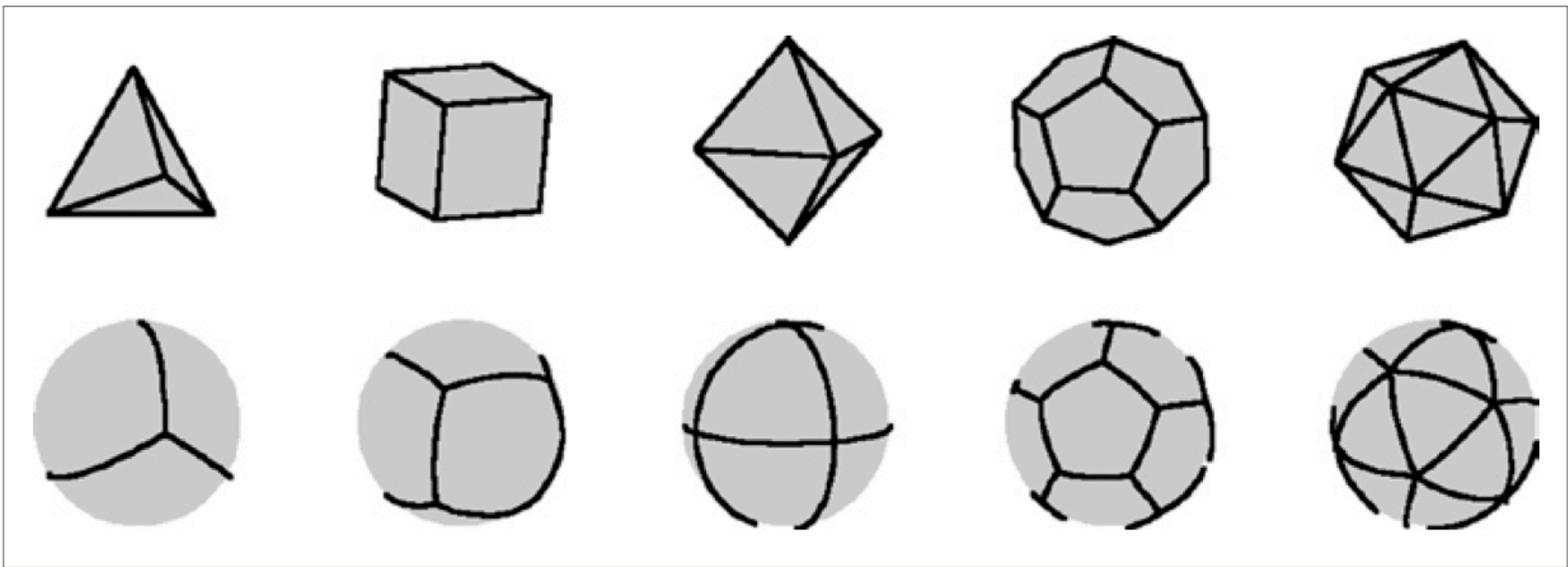
↑ Accounting for model error in **B**.

↑ Assimilation of observations influenced by clouds/precip

↑ Non-gaussian observation errors and priors.

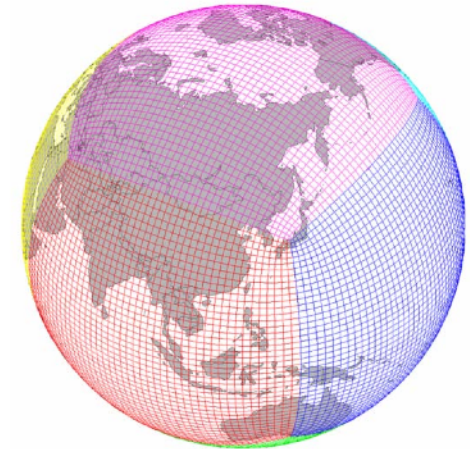
# Spherical Polyhedral Grids

- Spectral models can't scale to 100,000's of cores.
- lat/lon grids have singularity at pole, filtering kills scalability.
- Solution: grids based on platonic solids (cube, icosahedron) inscribed on the unit sphere.



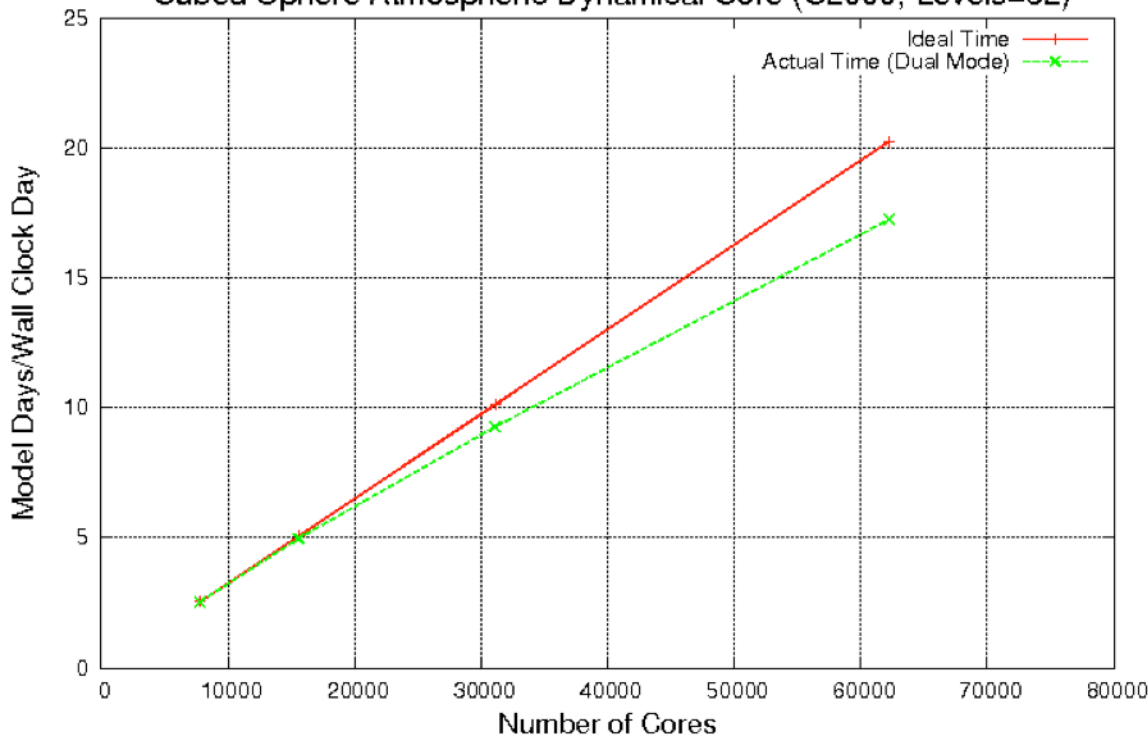
**Figure 2.** Planar and spherical versions of the five platonic solids: the tetrahedron, hexahedron (cube), octahedron, dodecahedron, and icosahedron.

# Cubed Sphere (GFDL, NASA, NCAR)



GFDL C2000L32 (~4.5 km) on IBM Blue Gene

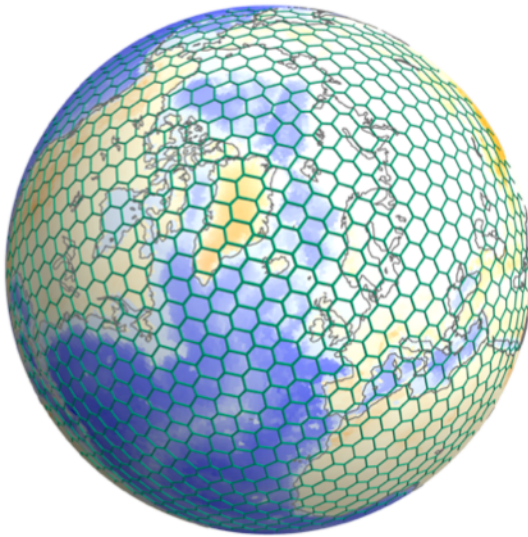
Performance on IBM:BG/P of Held-Suarez Test Case with Non-Hydrostatic Cubed-Sphere Atmospheric Dynamical Core (C2000, Levels=32)



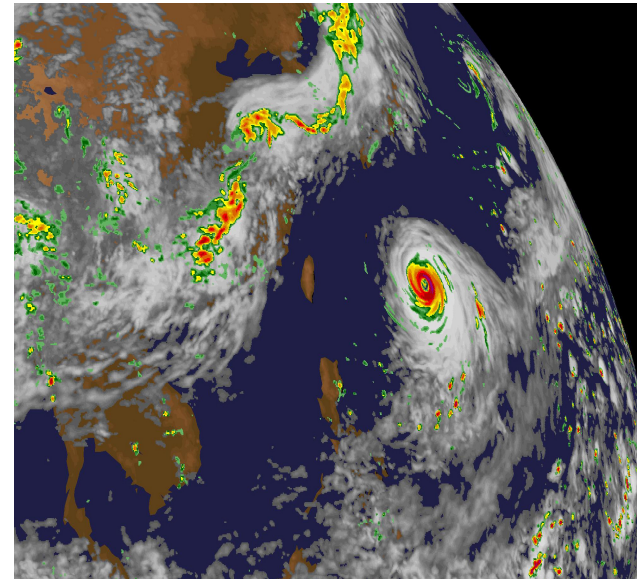
- Super linear scaling below 10,000 cores
- Maximum scalability with pure MPI: 1.5 million cores
- Scalability can be even higher with MPI-OpenMP hybrid programming

# Icosahedral Grids (NCAR MPAS, ESRL FIM/NIM, CSU, Japan)

12 “problem” points instead of 8 on cubed sphere, 1 in lat/lon (good, since singularities are weaker). However, grid is completely unstructured (cubed sphere still has cartesian structure).

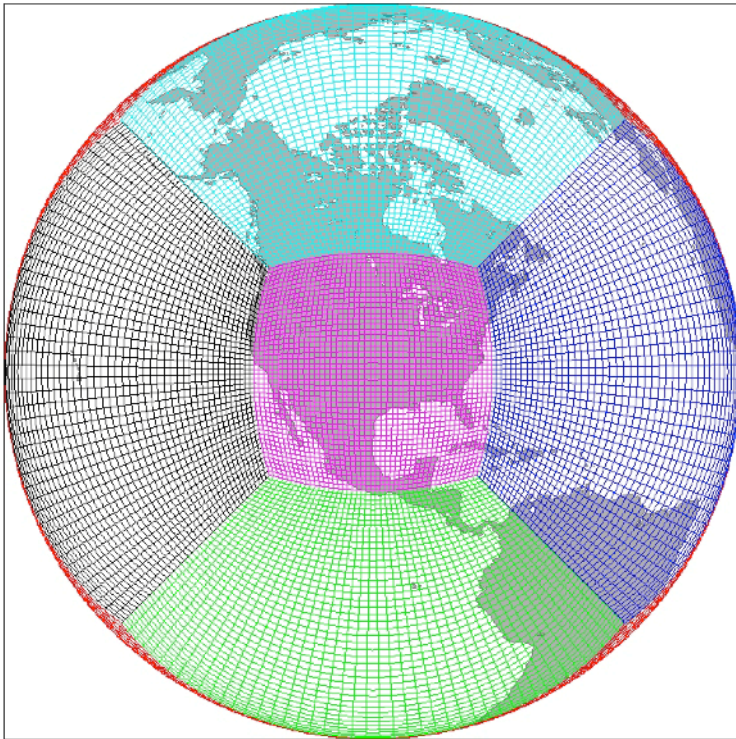


NICAM 7km hurricane simulation

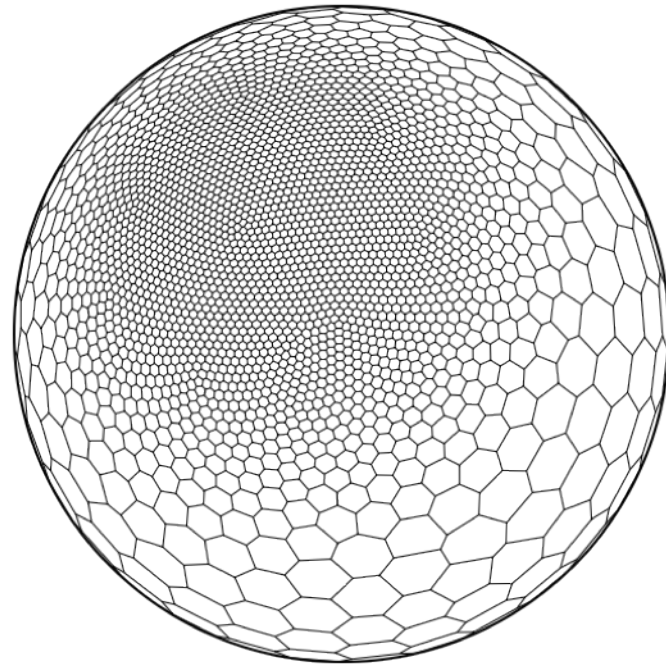


# Global models with variable resolution *avoid boundary conditions (nesting)*

Cubed Sphere with conformal transformation  
(GFDL model)

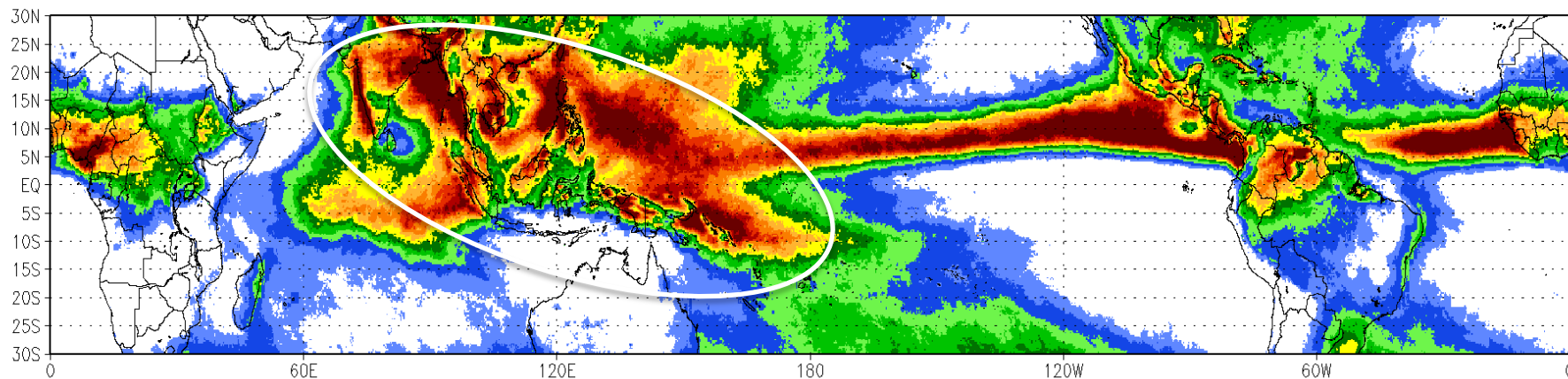


NCAR MPAS w/nonuniform Voronoi tessellation

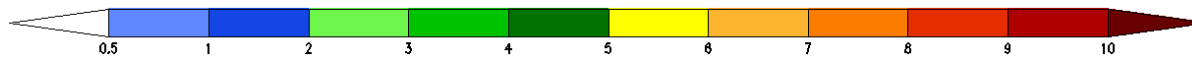
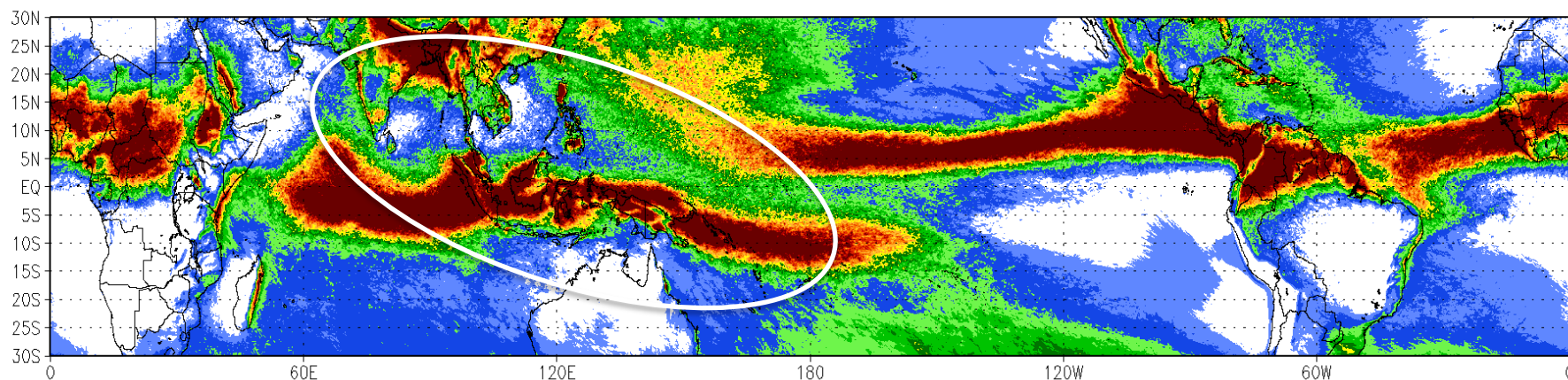


# Total JJA Precipitation (2000-2008)

TRMM JJA Total Precipitation (mm/day)



NICAM JJA Total Precipitation (mm/day)



# Importance of Flow-Dependent Background Errors

## Hurricane Fred 00Z 9 Sep

Single ob increments for  
850 hPa u ob 1 m/s  
different than background.

*Analysis “knows” where  
the hurricane is.*

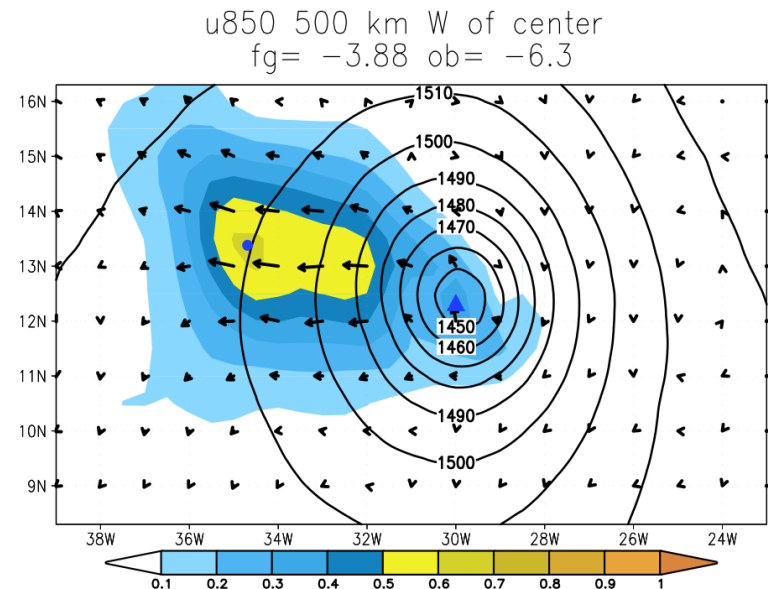
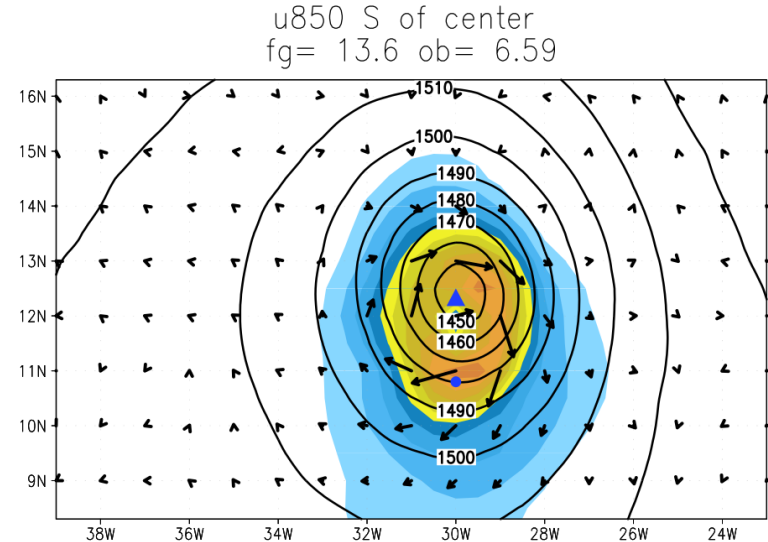
**Solid contours:** 850 hPa  
background geopotential height.

**Colors:** wind speed increment

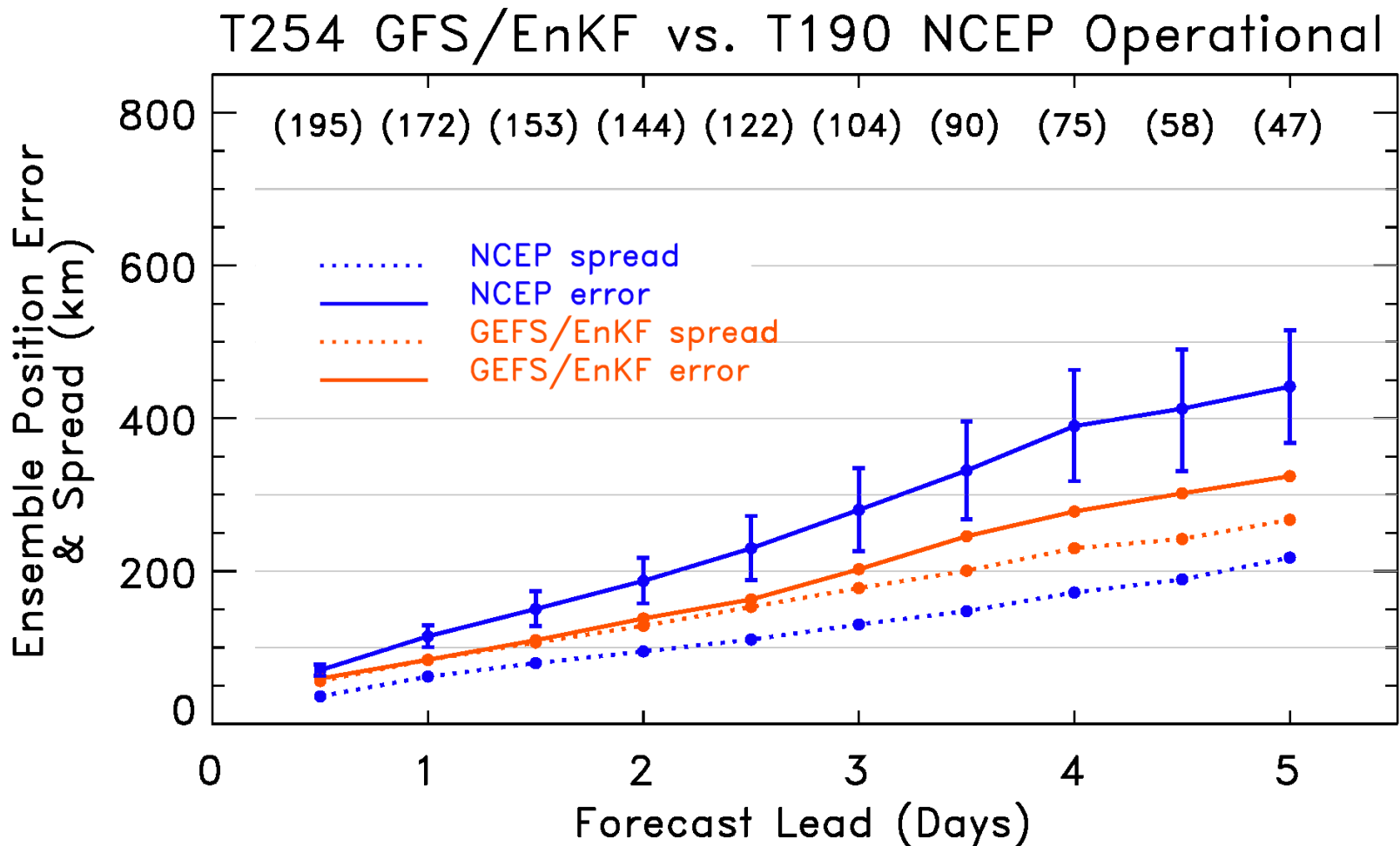
**Arrows:** vector wind increment

**Blue triangle:** hurrican center.

**Blue circle:** location of ob.



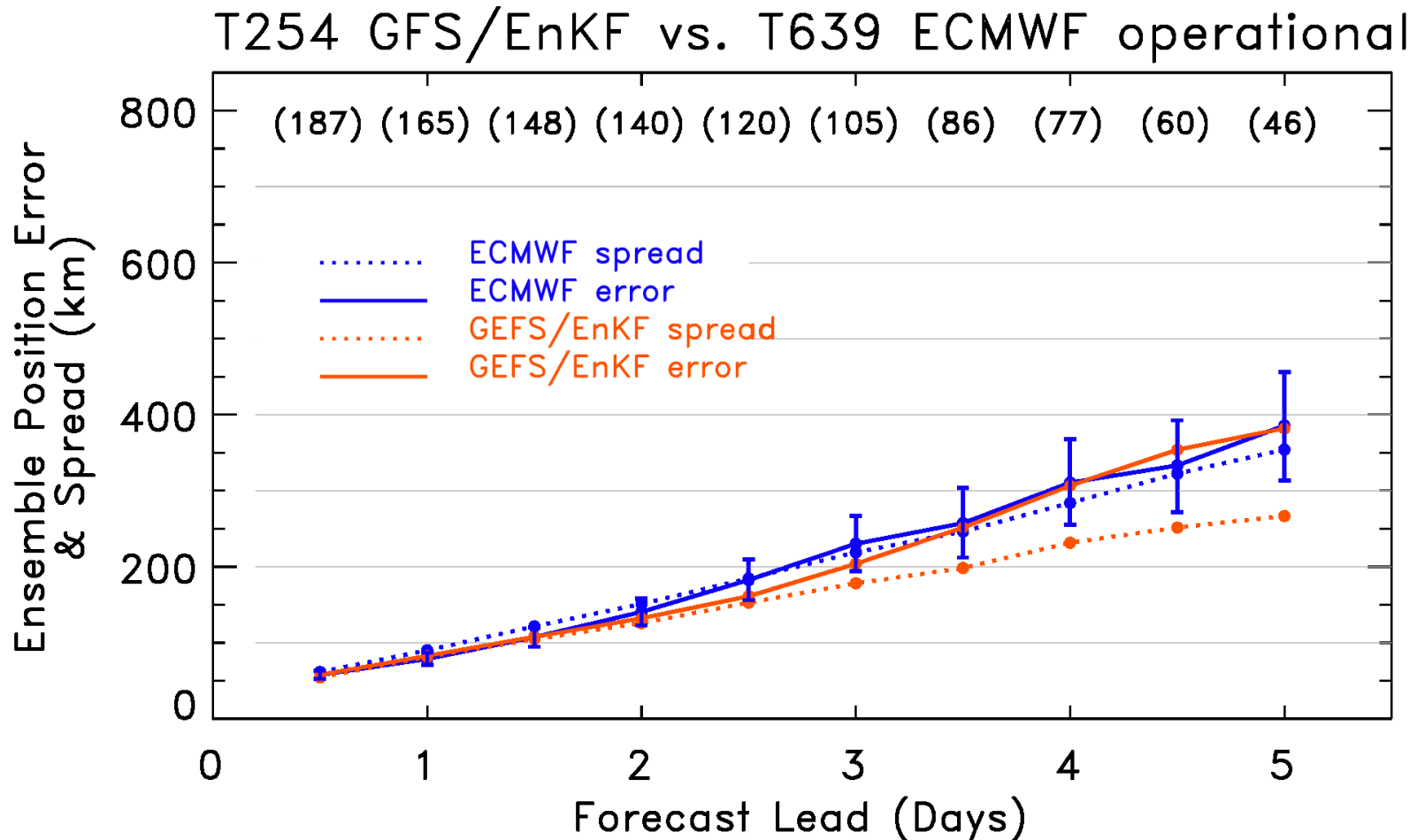
# Global statistics, GFS/EnKF vs. NCEP



At most every lead, GFS/EnKF is statistically significantly better than NCEP operational ens., which uses (a) older GFS model, lower resolution; (b) ETR perturbations around GSI control, and (c) vortex relocation.

# Global statistics, GFS/EnKF vs. ECMWF

(ensemble statistics, 5 June to 21 Sep 2010; all basins together)



GFS/EnKF competitive despite lower resolution (T254 vs. ECMWF's T639)

# EnKF

vs

# 4DVar

- evolves covariances continuously in time, in a low dim. space.

Localization required to increase rank of sample estimate.

- cov. evolution with full NL model.

- Full (potentially nonlinear)  $\mathbf{H}$  applied to each ensemble member.

- Model error accounted for in ensemble (inflation, stochastic physics).

- evolves covariances from an initial  $\mathbf{P}^b$  over a 6-12 h window in a high-dim. space (TLM).

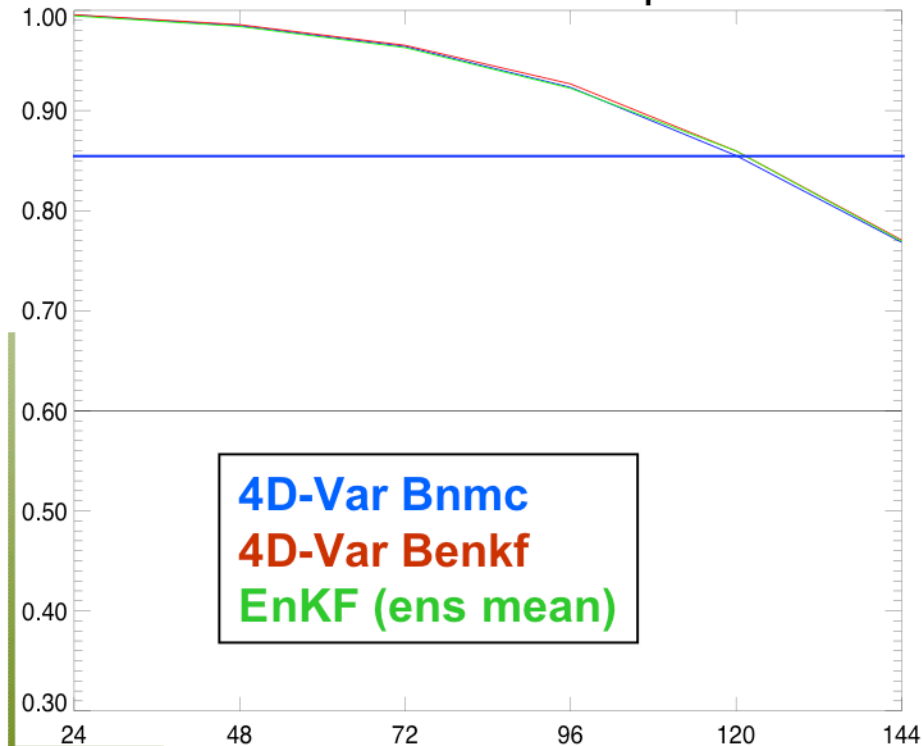
- cov. evolution with simplified linear pert. model.

- Linearized  $\mathbf{H}$  used in inner loop iteration, full nonlinear  $\mathbf{H}$  only applied to nonlinear control trajectory.

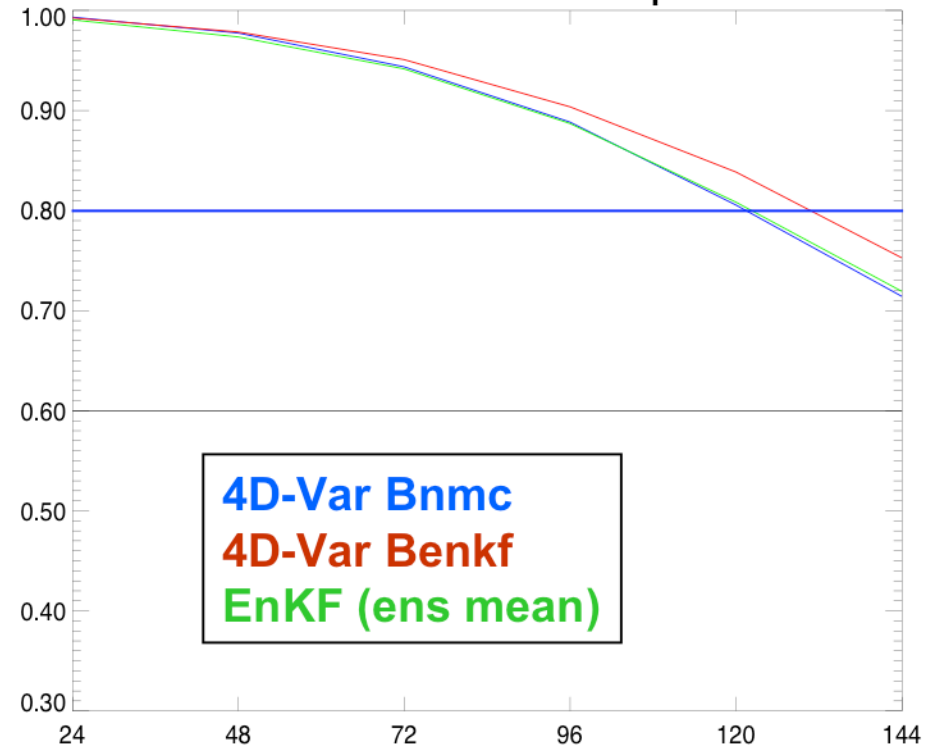
- model error for accounted for in initial  $\mathbf{P}^b$  (strong constraint) or model error covariance matrix (weak constraint).

# Hybrid Var/EnKF (*Mark Buehner Env Canada*)

Northern extra-tropics



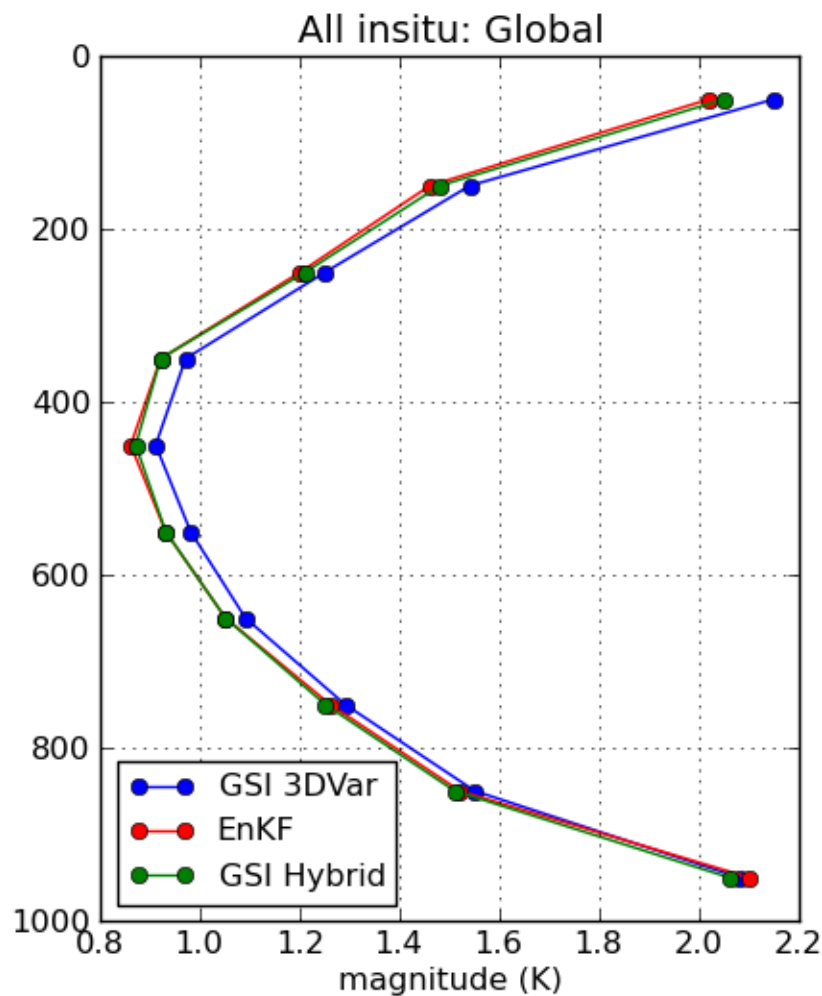
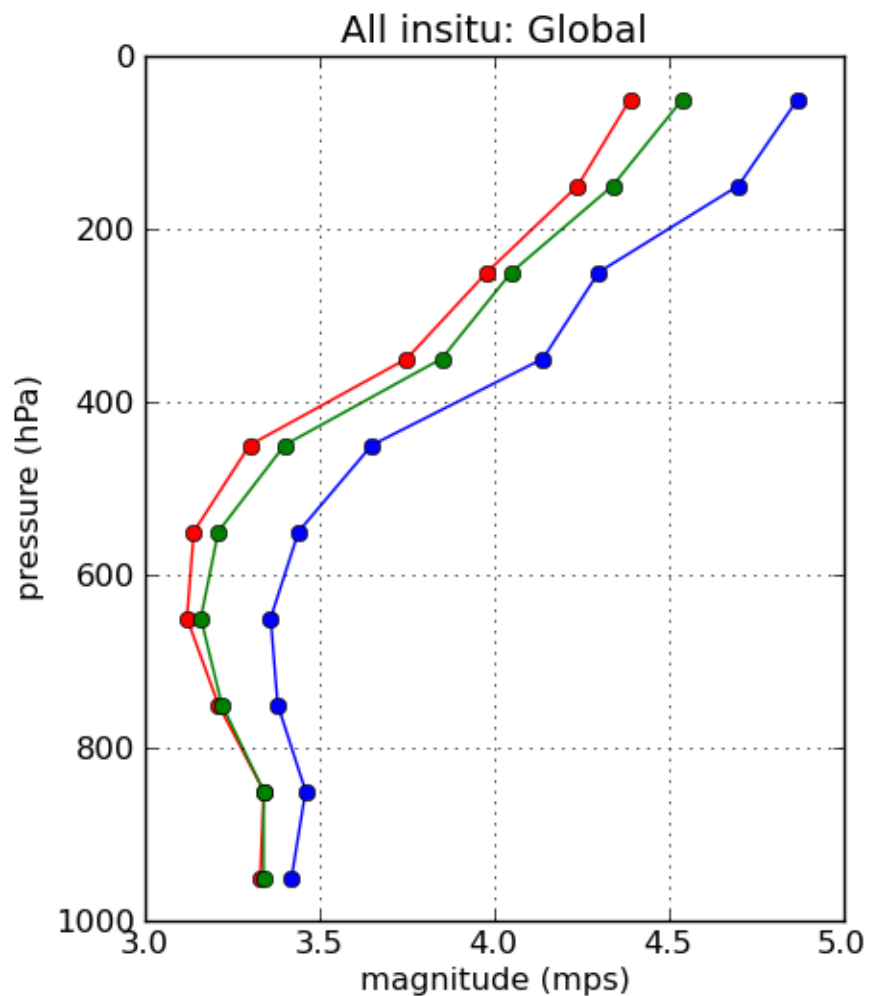
Southern extra-tropics



*EnKF performance nearly identical to operational 4DVar (but using EnKF  $\mathbf{P}^b$  benefits 4DVar in SH).*

# GSI 3DVar vs 3DVar Hybrid vs EnKF

Vector Wind (left) and Temp (right) O-F (2009123012-2010013012)



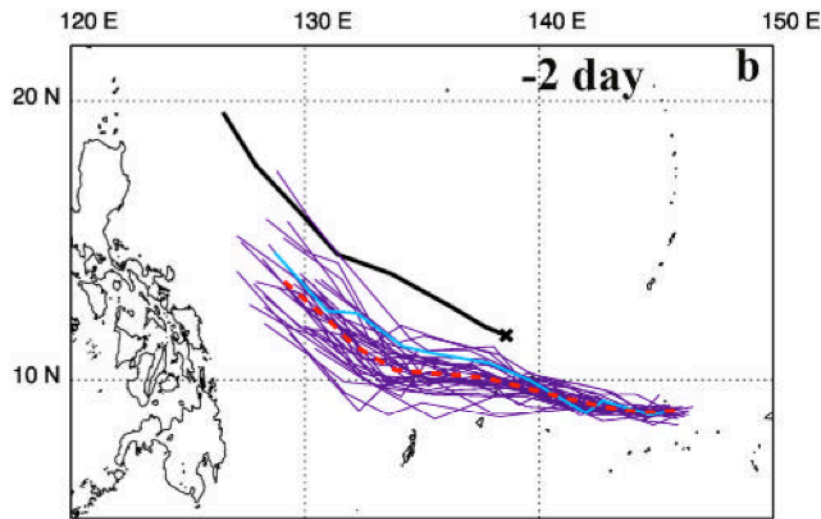
# “Weak-constraint” 4DVar

- Allows for longer windows by relaxing assumption that model is perfect (strong constraint).
  - Allows assimilation to “forget” initial static **B**.
  - Ensemble not needed for flow dependence.
  - BUT model error covariance must be specified.
- Being developed at ECMWF (alongside ensemble/VAR hybrid).

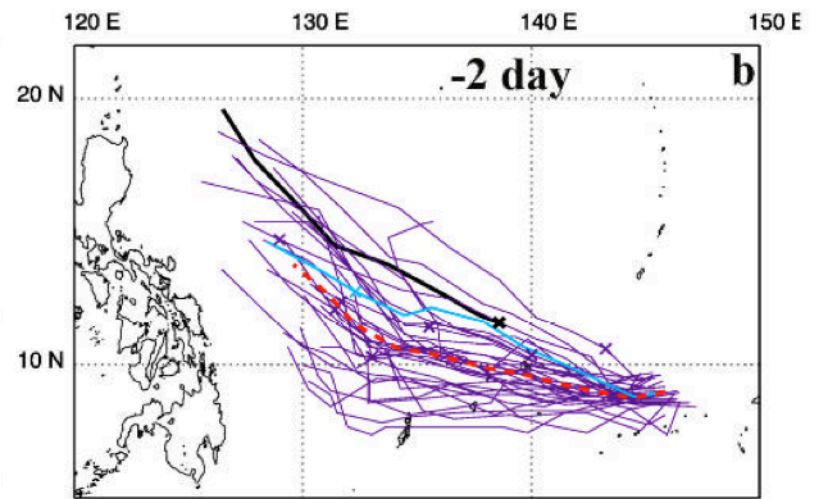
# Stochastic Physics

## *impact on ensemble TC forecasts*

w/out stochastic convection



With stochastic convection



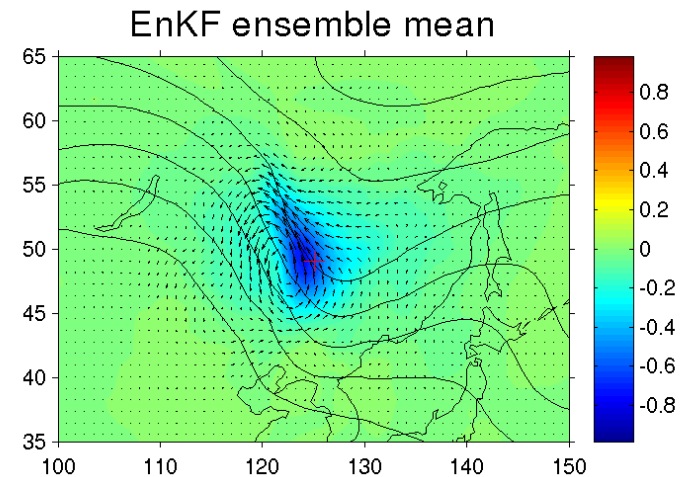
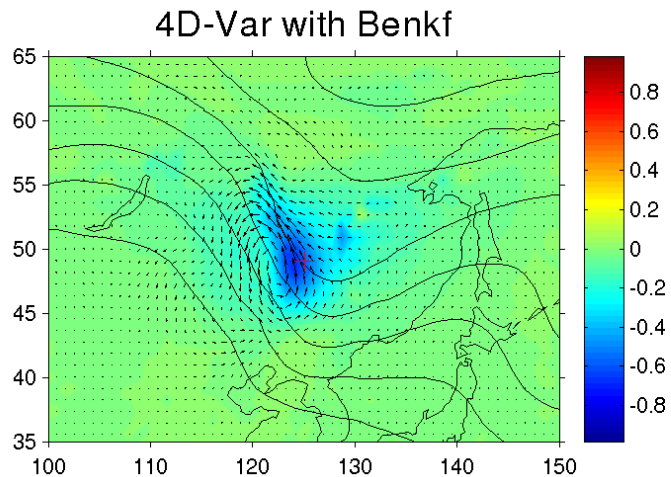
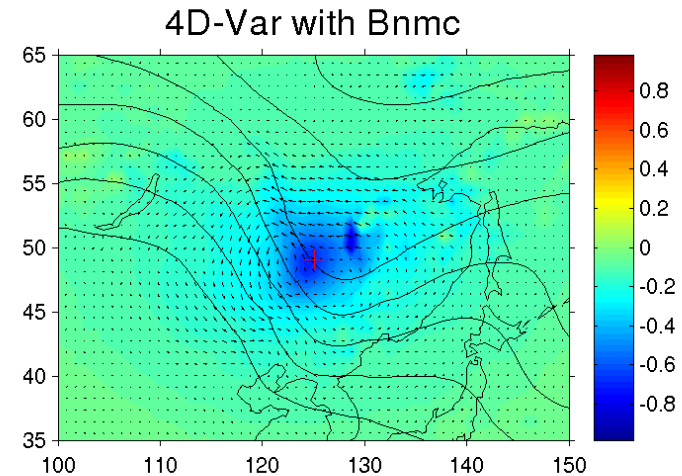
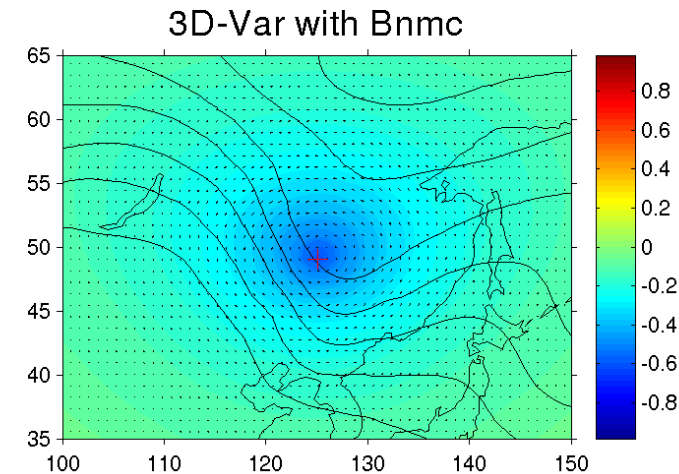
From Snyder, Pu and Reynolds (2011) MWR (early online release)

# Summary

- Global cloud permitting, variable resolution models on spherical polyhedral grids are coming.
- Data assimilation systems with fully flow-dependent background error covariances are here.
- Important challenges:
  - Reducing model errors (esp. in convection dominated regimes).
  - Representation of model error (in model and DA).
  - Dealing with non-gaussian backgrounds (clouds, hydrometeors, precip).

# Hybrid 4DVar/EnkF

*Env. Canada (Buehner et al MWR 2010)*



*EnKF increment more similar to 4DVar than 3DVar, especially when EnKF  $\mathbf{P}^b$  used at beginning of 4DVar window.*