Evaluation of a hybrid ensemble-variational data assimilation scheme [using an OSSE]

Daryl T. Kleist

Kayo Ide, Dave Parrish, John Derber, Jeff Whitaker

The Ninth International Workshop on Adjoint Model Applications in Dynamic Meteorology
10-14 October 2011; Cefalù, Italy

Outline

- Motivation
  - Hybrid overview
  - Impact from real observation experiments
- Joint OSSE introduction
- Experimental results from assimilation of simulated observations (3DVAR v. Hybrid)
- Future Work

Hybrid Variational-Ensemble

- Incorporate ensemble perturbations directly into variational cost function through extended control variable
  - Lorenc (2003), Buehner (2005)

$$J\left(\mathbf{x}_{\mathrm{f}}^{'},\alpha\right)=\beta_{\mathrm{f}}\frac{1}{2}\left(\mathbf{x}_{\mathrm{f}}^{'}\right)^{\mathrm{T}}\mathbf{B}^{-1}\left(\mathbf{x}_{\mathrm{f}}^{'}\right)+\beta_{\mathrm{e}}\frac{1}{2}\left(\alpha\right)^{\mathrm{T}}\mathbf{L}^{-1}\left(\alpha\right)+\frac{1}{2}\left(\mathbf{y}_{\mathrm{o}}^{'}-\mathbf{H}\mathbf{x}_{\mathrm{t}}^{'}\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{y}_{\mathrm{o}}^{'}-\mathbf{H}\mathbf{x}_{\mathrm{t}}^{'}\right)$$

$$\mathbf{x}_{\mathrm{t}}^{'}=\mathbf{x}_{\mathrm{f}}^{'}+\sum_{k=1}^{K}\left(\alpha_{k}\circ\mathbf{x}_{k}^{\mathrm{e}}\right)$$

$$\frac{1}{\beta_{\mathrm{f}}}+\frac{1}{\beta_{\mathrm{e}}}=1$$

β_{f} & β_{e}: weighting coefficients for fixed and ensemble covariance respectively

$\mathbf{x}_{\mathrm{t}}$: (total increment) sum of increment from fixed/static **B** ($\mathbf{x}_{\mathrm{f}}$) and ensemble **B**

α_{k}: extended control variable; $\mathbf{x}_{k}^{\mathrm{e}}$:ensemble perturbation

L: correlation matrix [localization on ensemble perturbations]

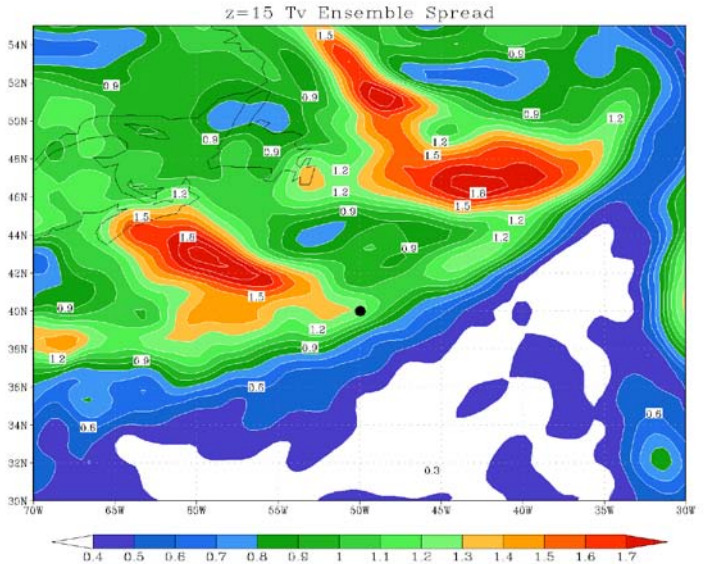
Hybrid with (global) GSI

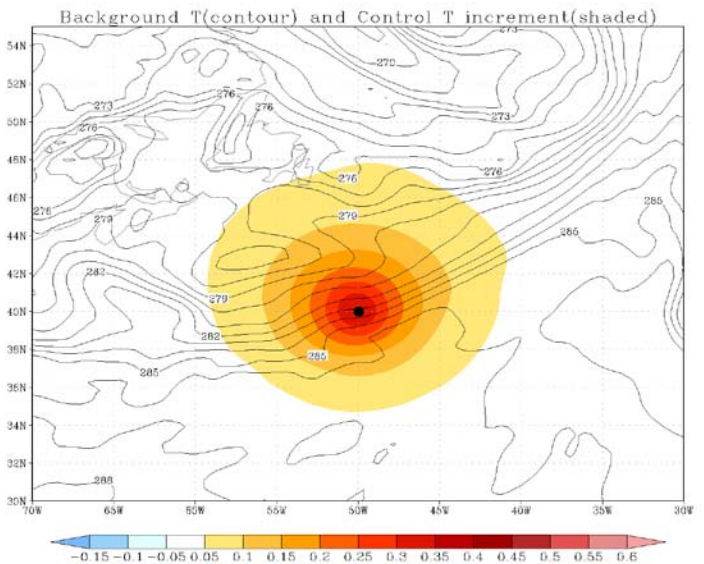
- Control variable has been implemented into GSI 3DVAR*
  - Full **B** preconditioning
    - Working on extensions to $\mathbf{B}^{1/2}$ preconditioned minimization options
      - Collaboration with GMAO
  - Spectral filter for horizontal part of **L**
    - Eventually replace with (anisotropic) recursive filters
  - Recursive filter used for vertical
  - Dual resolution capability
  - Various localization options for **L**
    - Grid units or scale height
    - Level dependent
  - Option to apply TLNMC (Kleist et al. 2009) to analysis increment

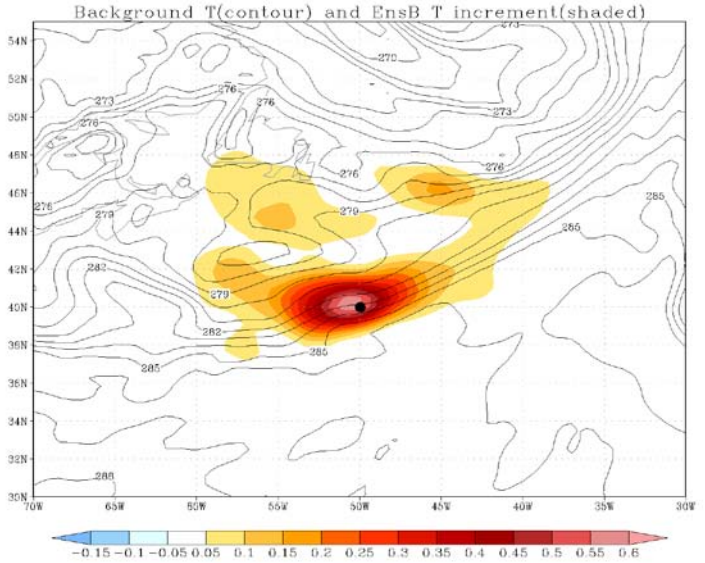
$$\mathbf{x}^{'} = \mathbf{C}\left[\mathbf{x}_{\mathrm{f}}^{'} + \sum_{k=1}^{K}\left(\alpha_k \circ \mathbf{x}_k^e\right)\right]$$

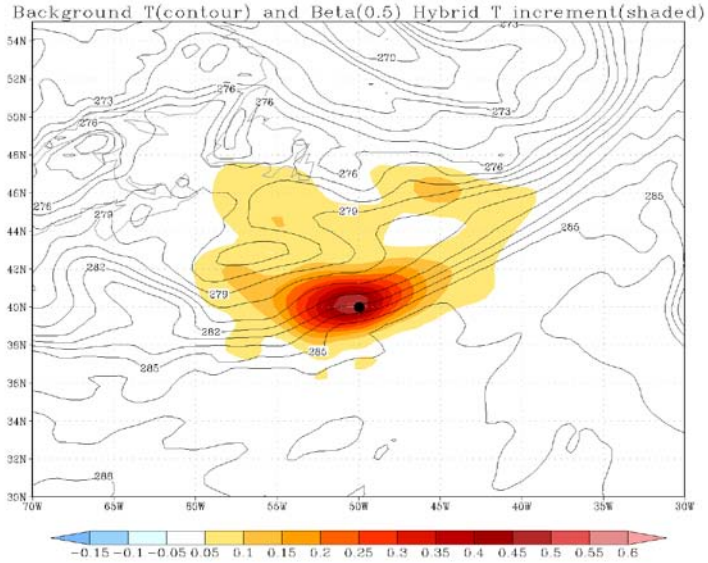
Acknowledgement: Dave Parrish for original implementation of extended control variable

Single Observation



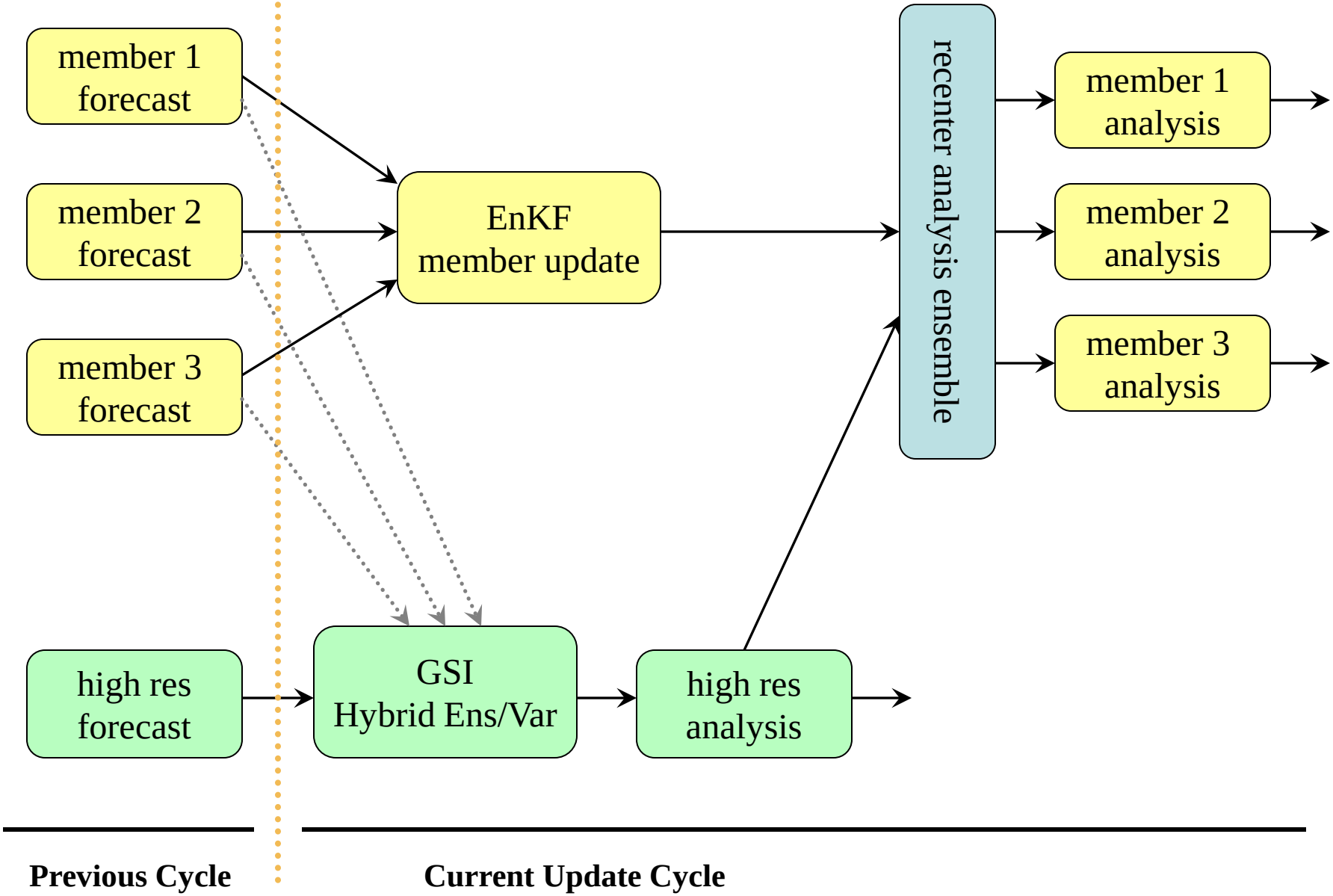






Single 850mb Tv observation (1K O-F, 1K error)

Dual-Res Coupled Hybrid



Hybrid Var-EnKF GFS experiment

- **Model**
  - GFS deterministic (T574L64; post July 2010 version – v9.0.0)
  - GFS ensemble (T254L64)
    - 80 ensemble members, EnKF update, GSI for observation operators

- **Observations**
  - All operationally available observations (including radiances)
  - Includes early (GFS) and late (GDAS/cycled) cycles as in production

- **Dual-resolution/Coupled**
    - High resolution control/deterministic component
    - Ensemble is recentered every cycle about hybrid analysis
      - Discard ensemble mean analysis

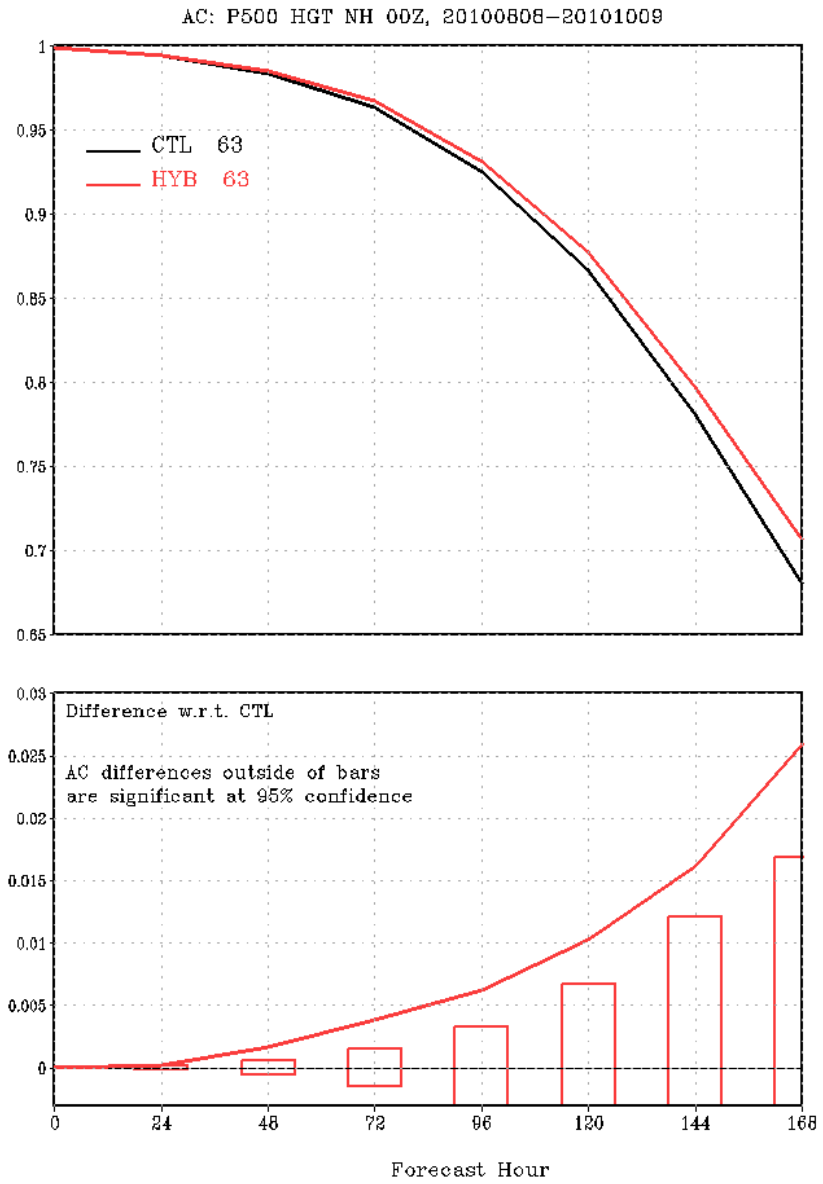
- **Satellite bias corrections**
  - Coefficients come from GSI/VAR

- **Parameter settings**
    - 1/3 static **B**, 2/3 ensemble
    - Fixed localization: 800km & 1.5 scale heights

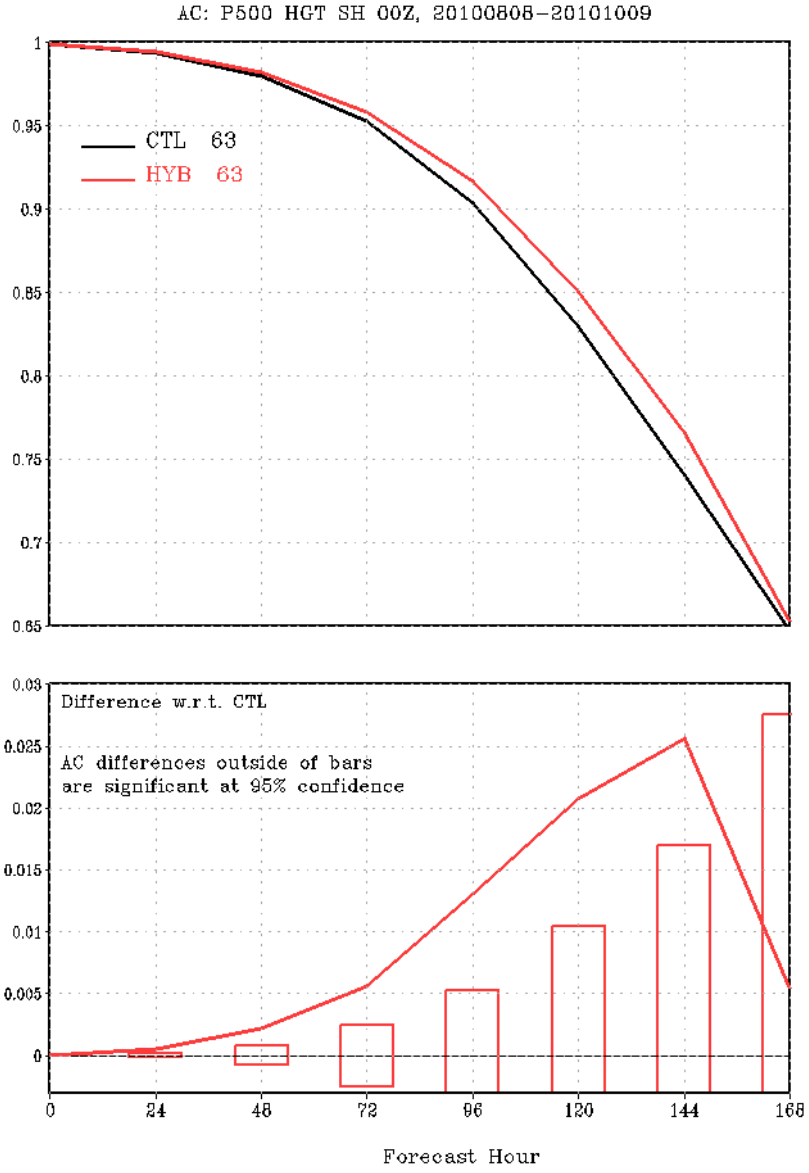
- **Test Period**
  - 15 July 2010 – 15 October 2010 (first two weeks ignored for “spin-up”)

500 hPa Anom.Corr.

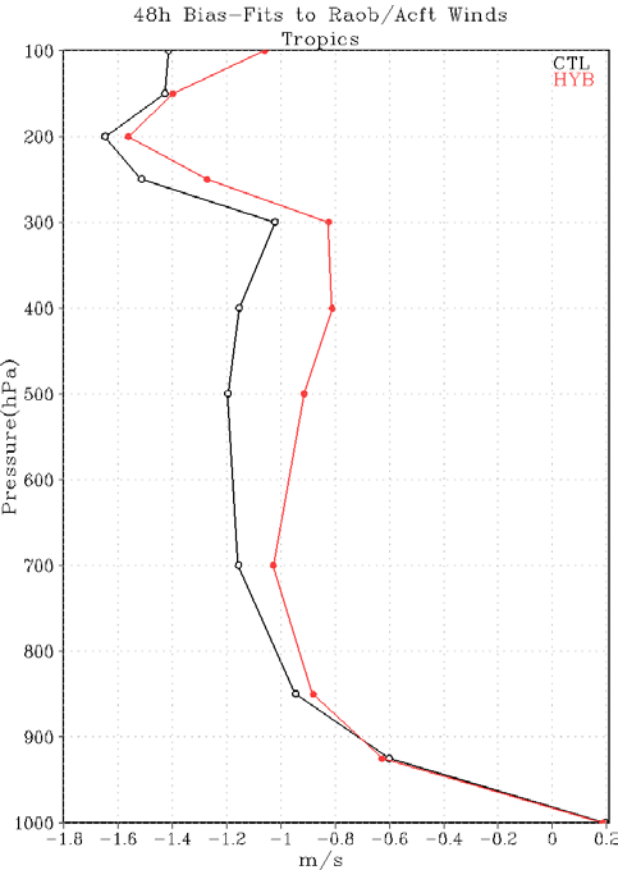
Northern Hemisphere

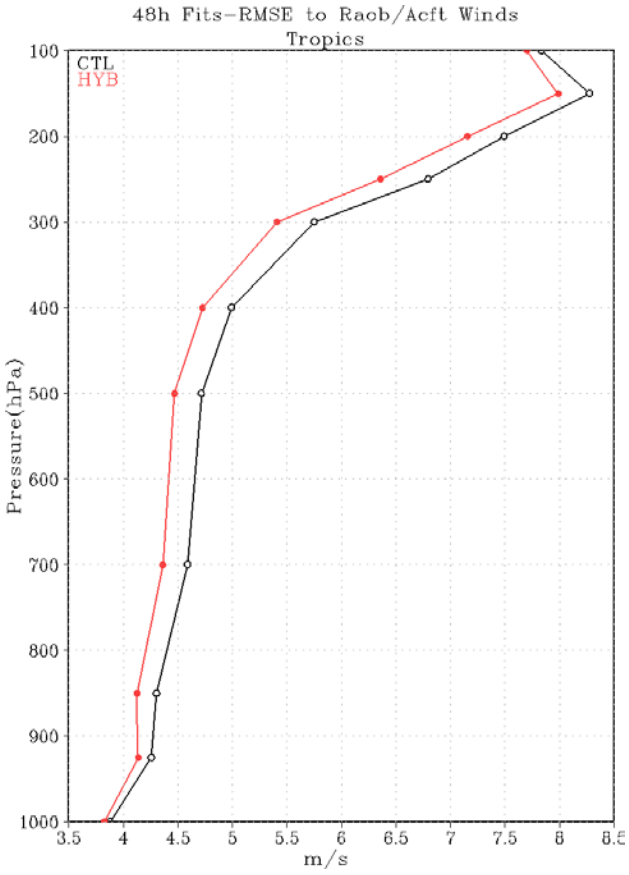


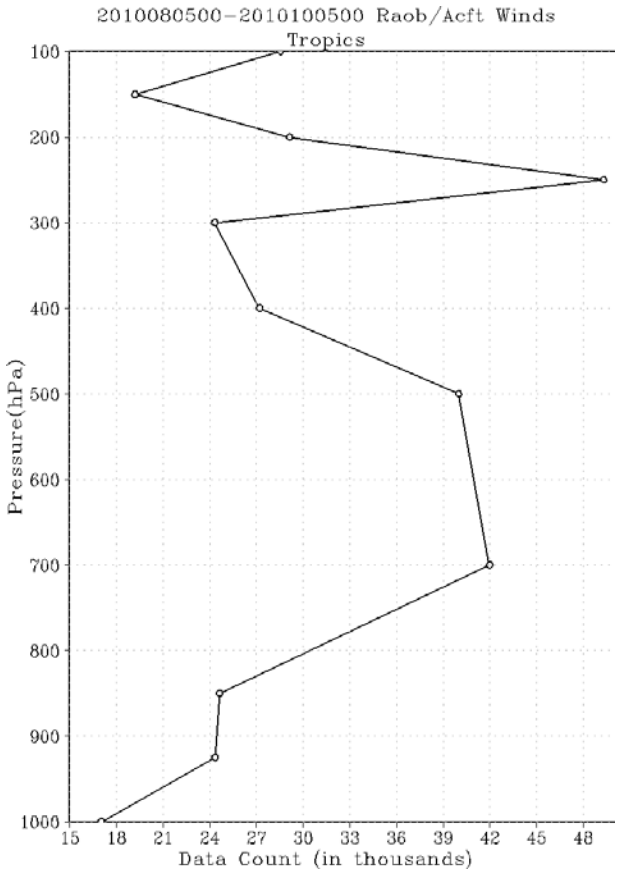
Southern Hemisphere



Forecast Fits to Obs (Tropical Winds)

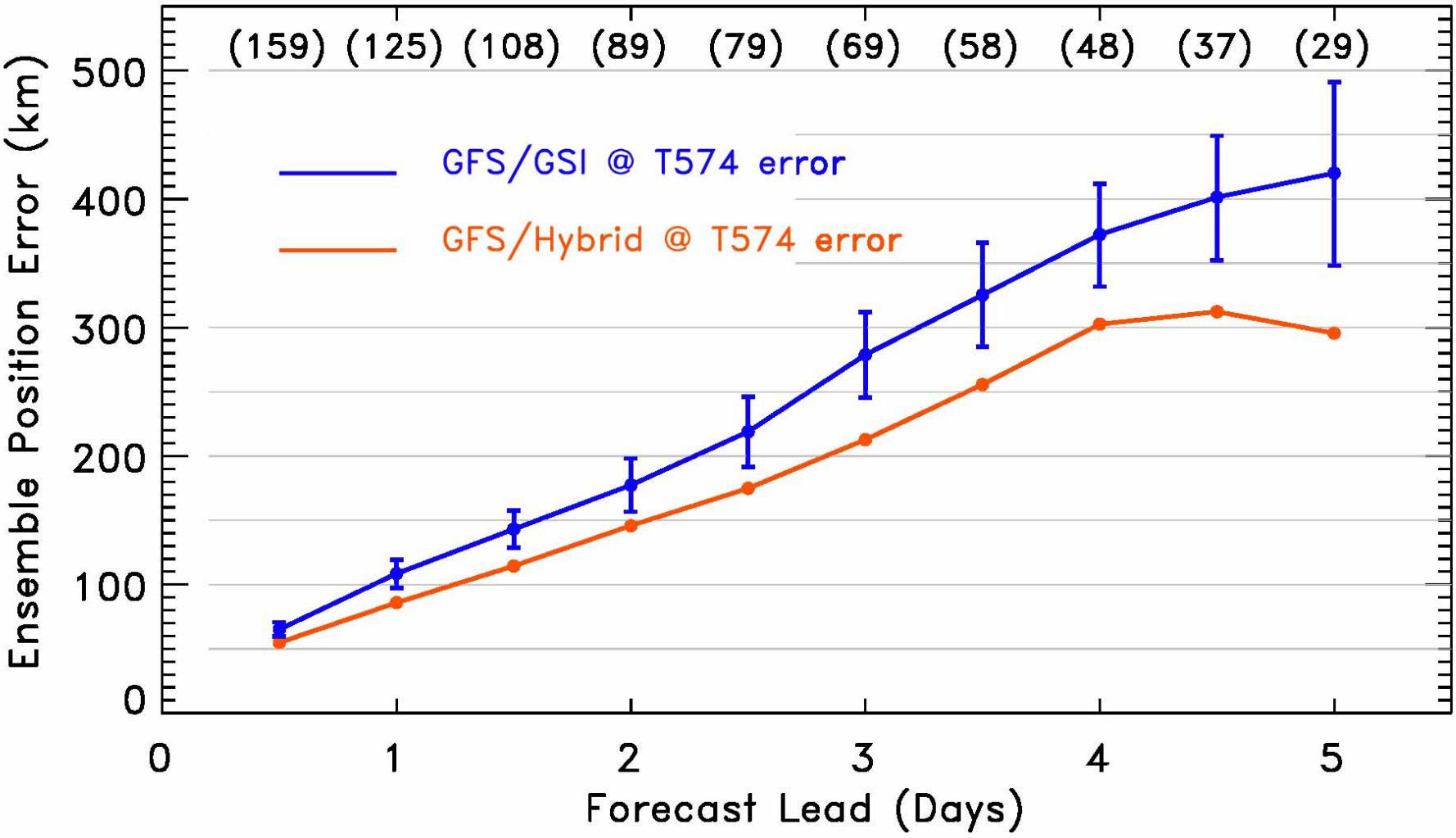






Forecasts from hybrid analyses fit observation much better.

GSI/EnKF Hybrid vs GSI opnl track errors



Hybrid has significantly lower track errors than operational GSI (using static covariance)

Hybrid (3D) for GDAS/GFS

- Prototype dual-resolution, two-way coupled hybrid Var/EnKF system outperforms standard 3DVAR in GFS experiments

- Plan underway to implement into GDAS/GFS operationally
  - Target: Spring 2012 (subject to many potential issues)
    - Porting of codes/scripts back to IBM P6 (**complete**)
    - Cost analysis (**complete-ish**)
    - More thorough (pre-implementation) testing and evaluation (**underway**)
      - More test periods (including NH winter)
      - Other/more verification metrics
    - Testing for implementation now underway

- **Extend hybrid research to realistic OSSE framework**

OSSEs

- Typically used to evaluate impact of future observing systems
- **Useful for evaluating present/proposed data assimilation techniques since 'truth' is known**
- Joint OSSE
  - International, collaborative effort between ECMWF, NASA/GMAO, NOAA (NCEP/EMC, NESDIS, JCSDA), NOAA/ESRL, others
  - ECMWF-generated nature run (c31r1)
    - T511L91, 13 month free run, prescribed SST, snow, ice
  - Shared simulated observations

Synthetic Observations

- Observations from (operational) 2005/2006 observing system developed
  - NCEP: ‘conventional’, sbuv ozone retrievals, GOES sounder radiances
  - NASA/GMAO: all other radiances (AMSUA/B, HIRS, AIRS, MSU)

- Simulated observation errors developed by Ron Errico
  - Horizontally correlated errors for radiances
  - Vertically correlated errors for conventional soundings

- Synthetic observations used in this study were calibrated by Nikki Prive
  - Attempt to match impact of various observation types with results from data denial experiments (OSE)

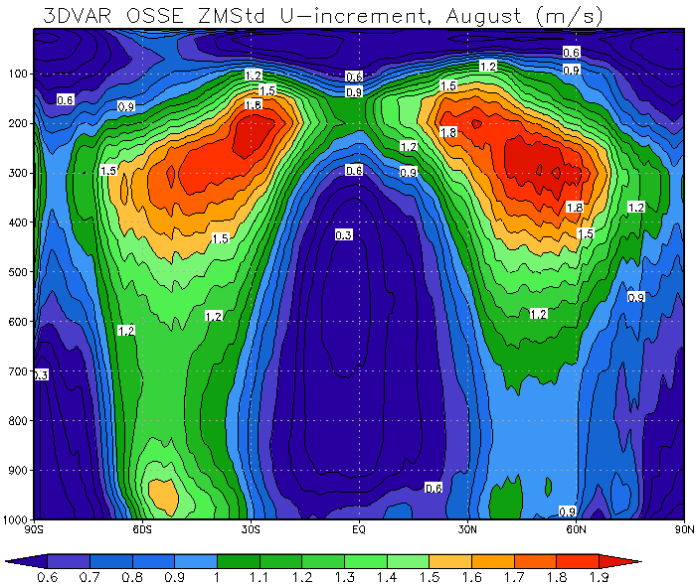
**Thanks to Ron and Nikki for all of their help in getting/using the calibrated synthetic observations.*

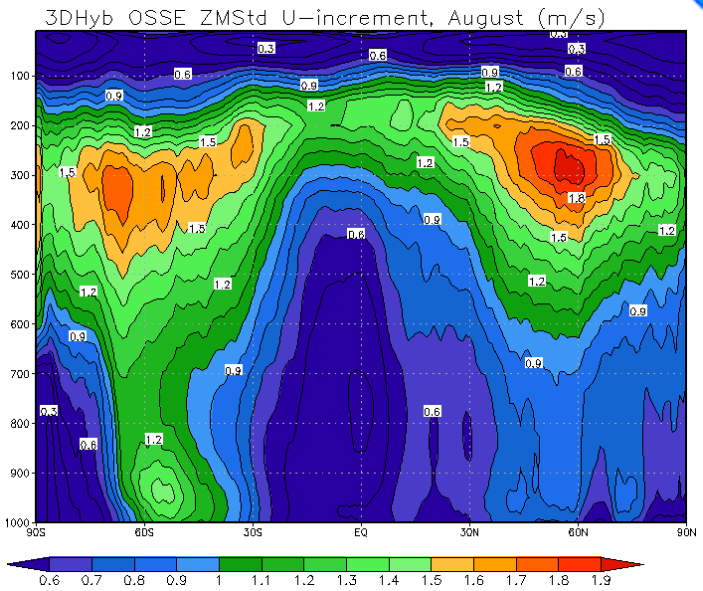
Experimental Design

- **Model**
  - GFS deterministic (T382L64; post May 2011 version – v9.0.1)
  - GFS ensemble (T190L64)
    - 80 ensemble members, EnKF update, GSI for observation operators
- **Observations**
  - Synthetic observations from 2005 observing system (courtesy Ron Errico/Nikki Privi)
- **Dual-resolution/Coupled**
    - High resolution control/deterministic component
    - Ensemble is recentered every cycle about hybrid analysis
      - Discard ensemble mean analysis
- **Satellite bias corrections**
  - Coefficients come from GSI/VAR
- **Parameter settings**
    - 1/4 static **B**, 3/4 ensemble
    - Level-dependent localization
- **Test Period**
  - 01 July 2005-31 August 2005 (3 weeks ignored for spin-up)

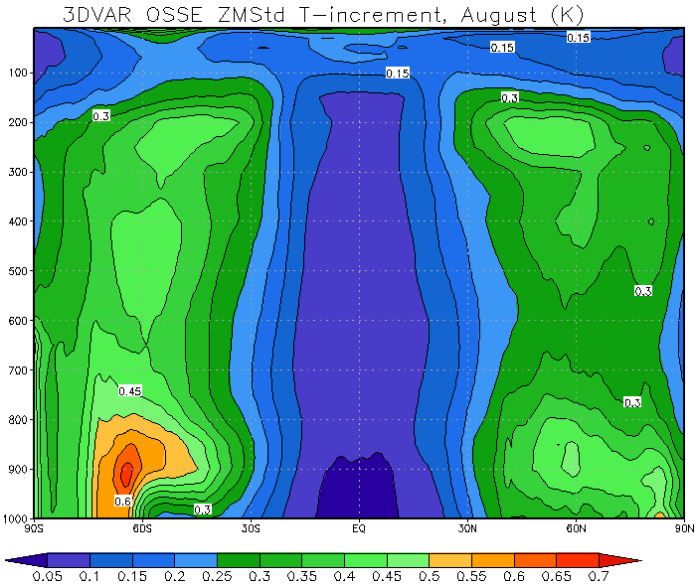
Time mean increment

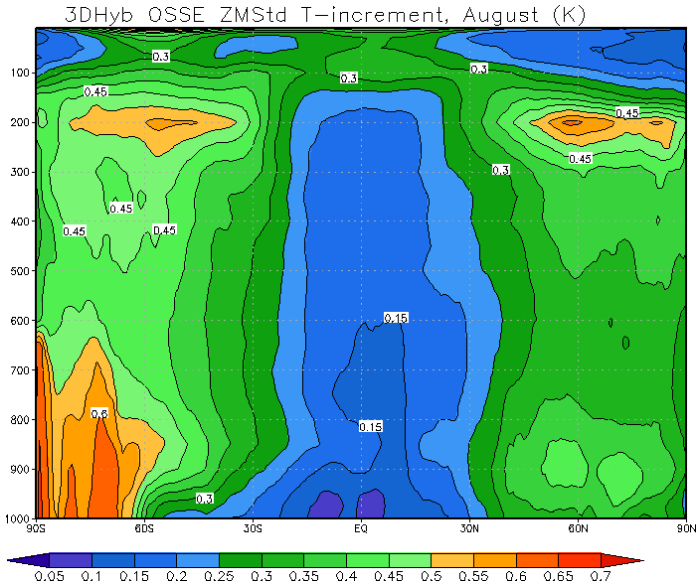
U





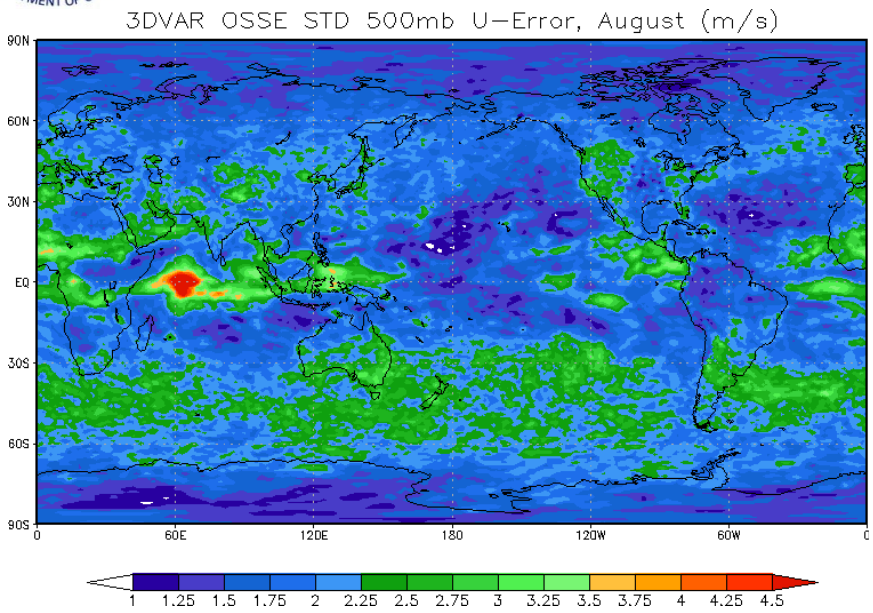
T

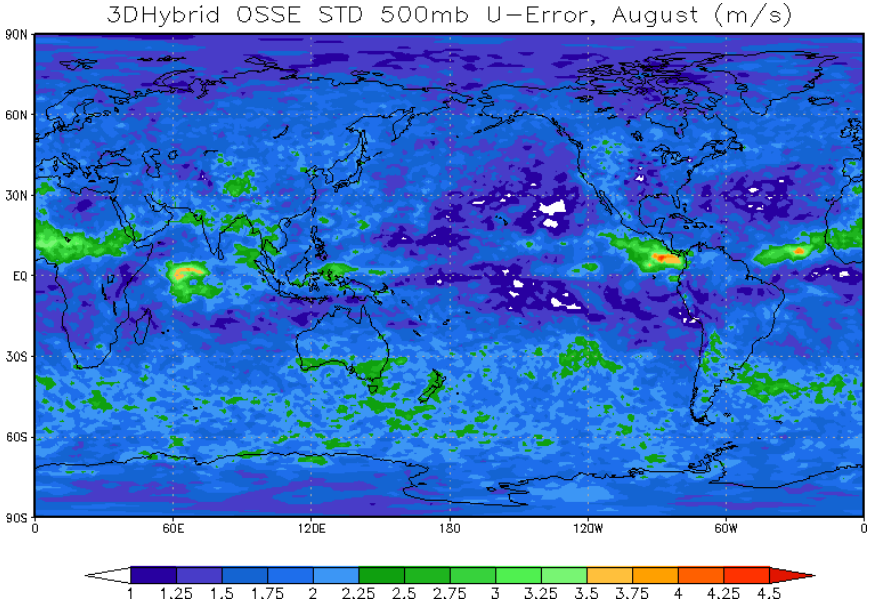


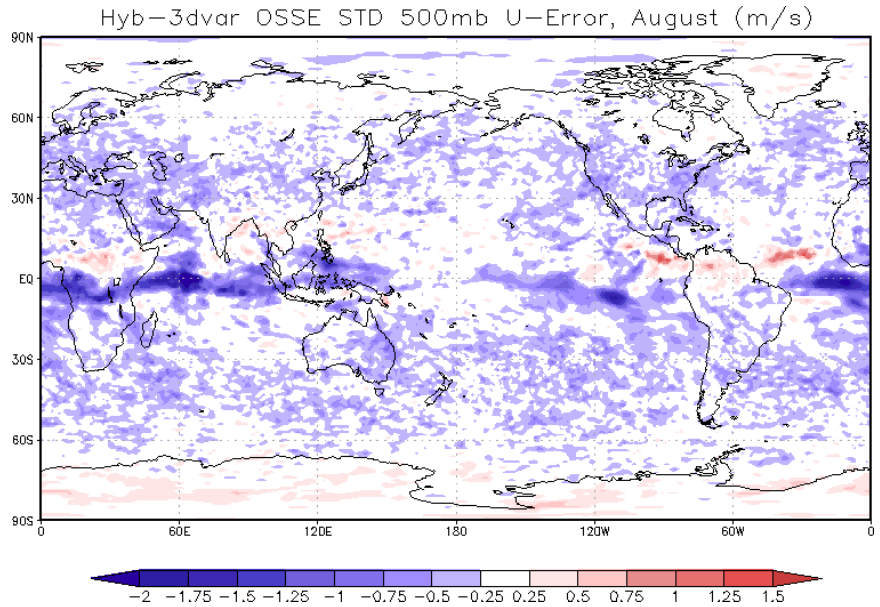


3DVar ***3DHyb***

Analysis Error

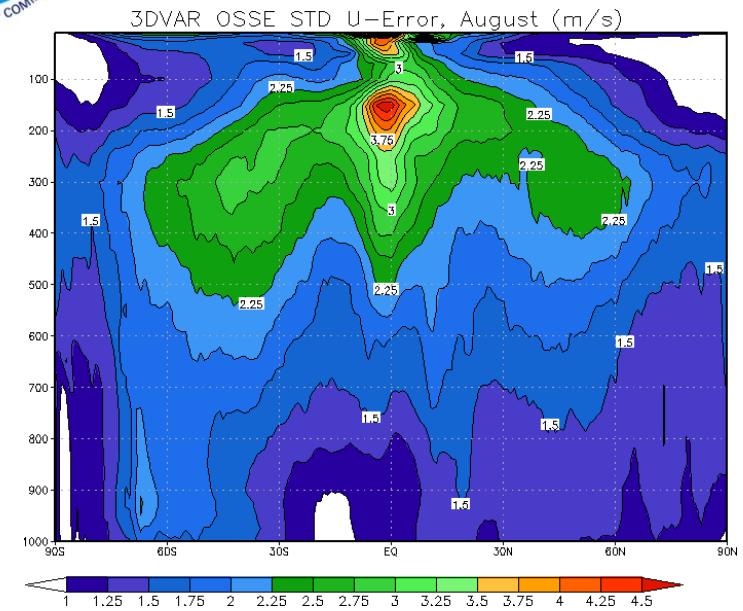


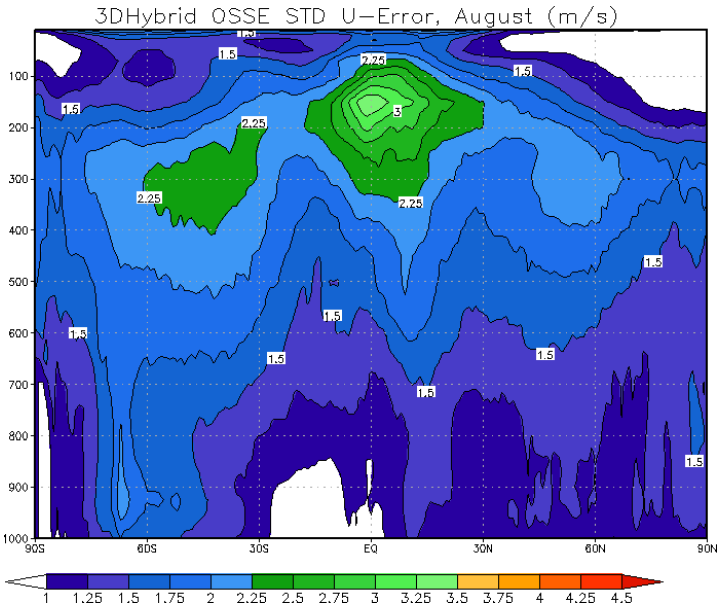


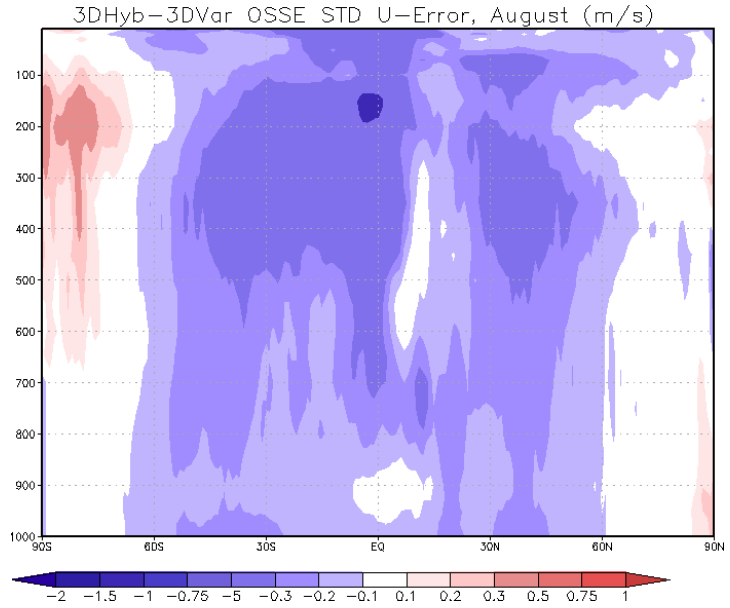


500 hPa zonal wind analysis RMSE for 3DVAR (upper left) and 3D-Hyb(upper right) and difference (Hybrid-3DVAR lower left) for August, all cycles.

Analysis Error Cross Section

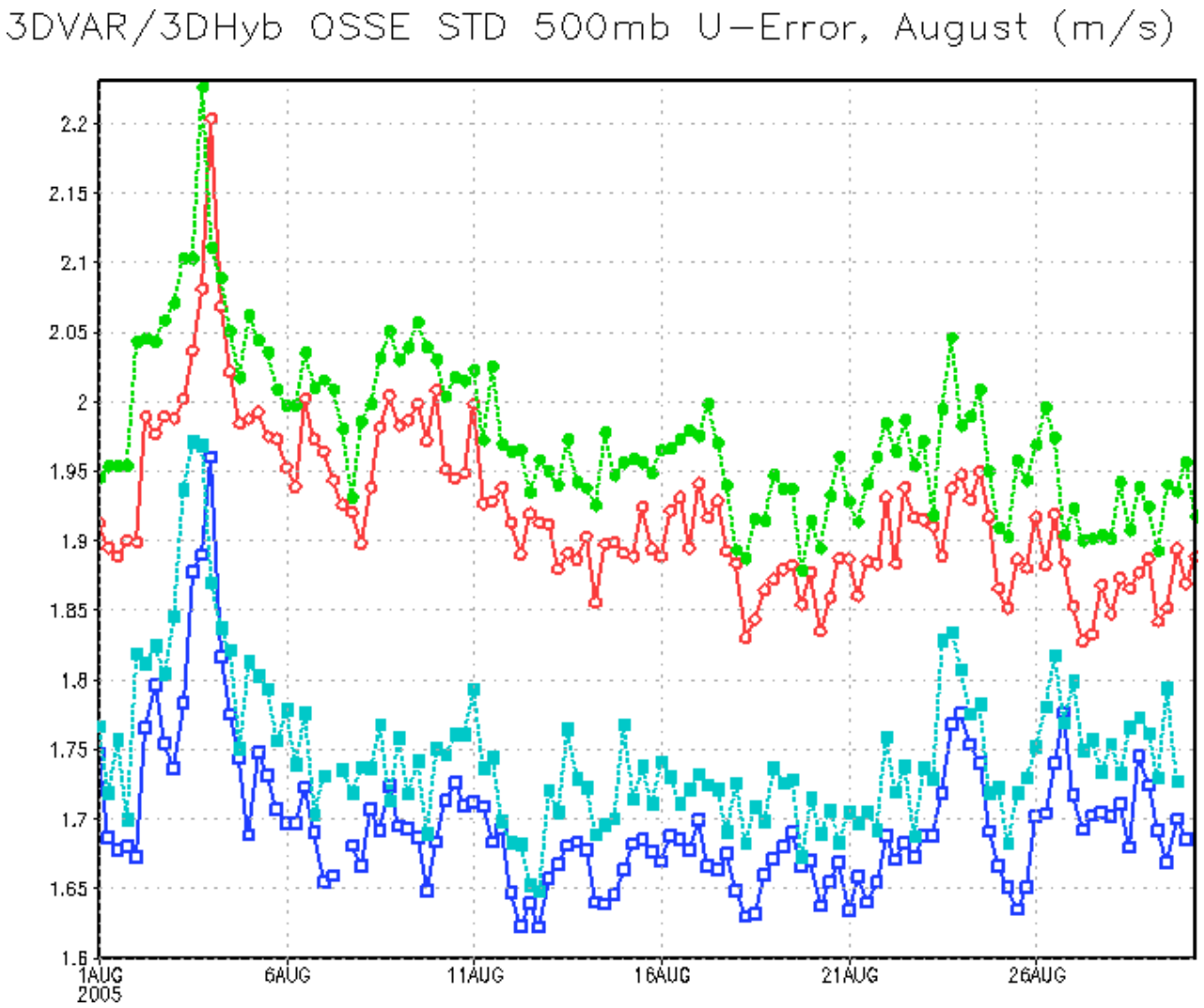






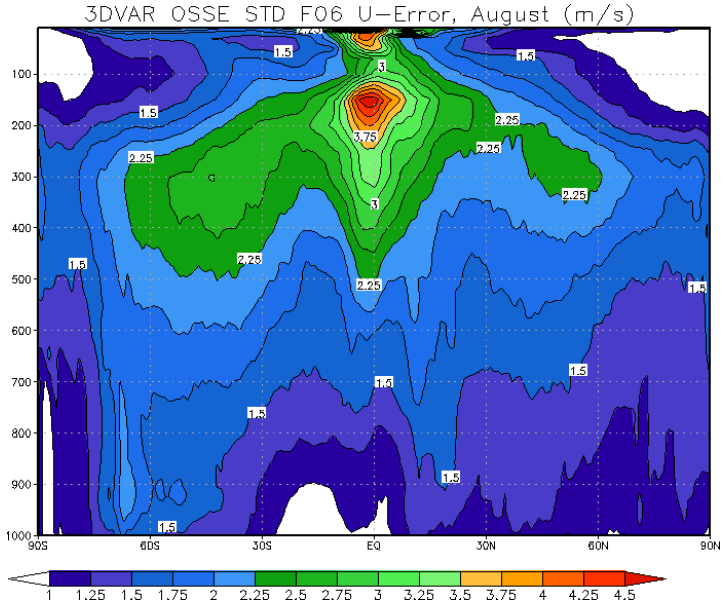
Zonal wind analysis RMSE for 3DVAR (upper left) and 3D-Hyb (upper right) and difference (Hybrid-3DVAR lower left) for August, all cycles.

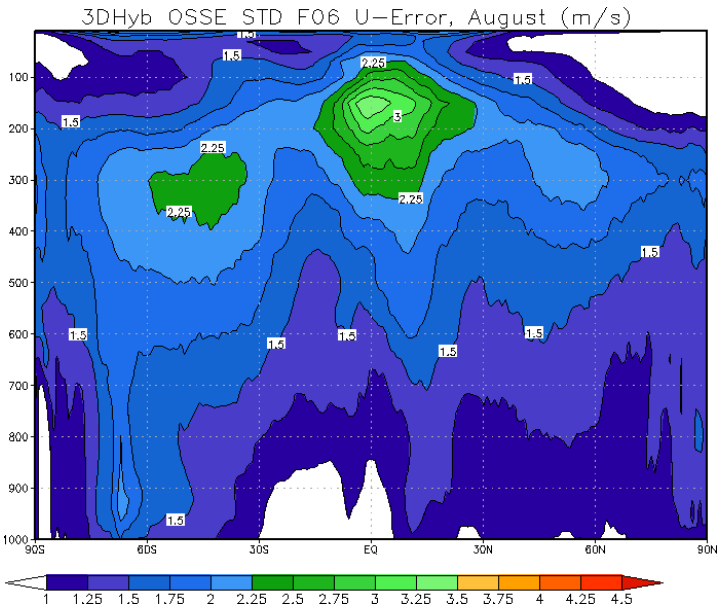
Zonal Wind Error (guess/analysis)

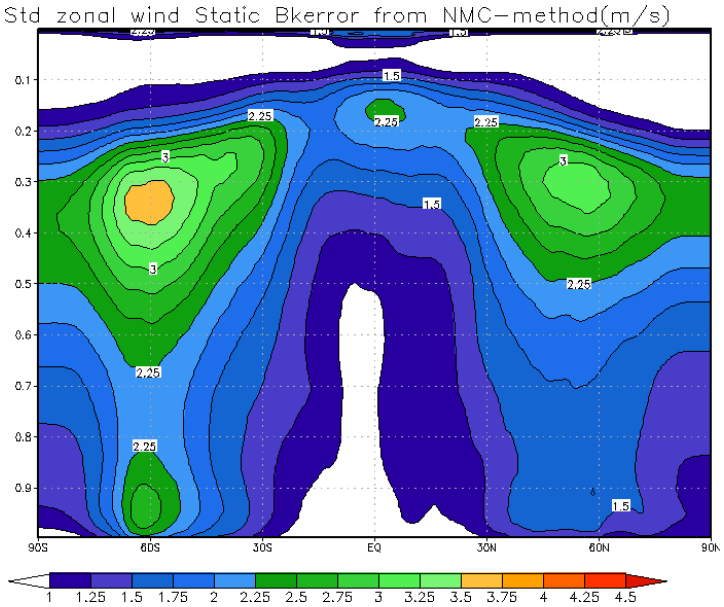


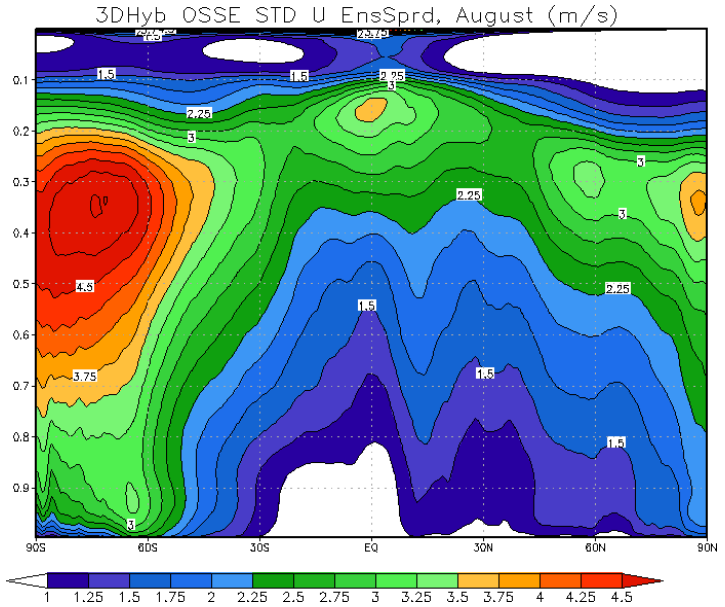
500 hPa zonal wind analysis RMSE for 3DVAR guess (green dashed) and analysis (red solid) and 3D-Hyb guess (aqua dashed) and analysis (blue solid) for August.

Background Error



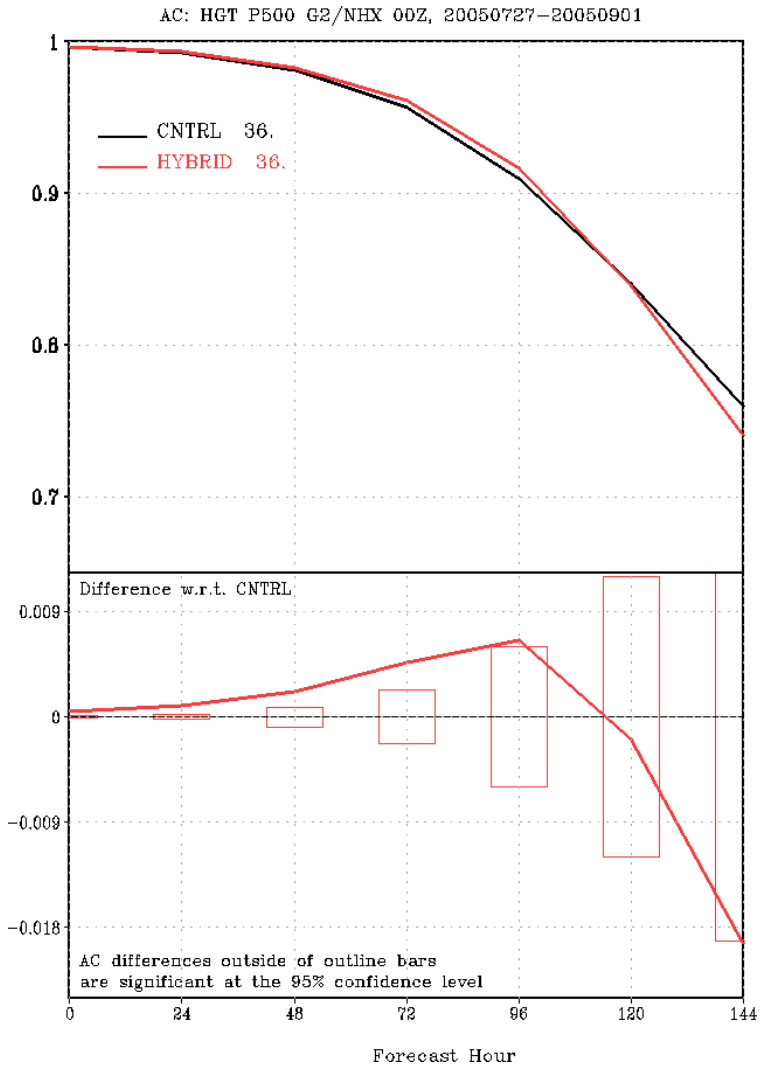




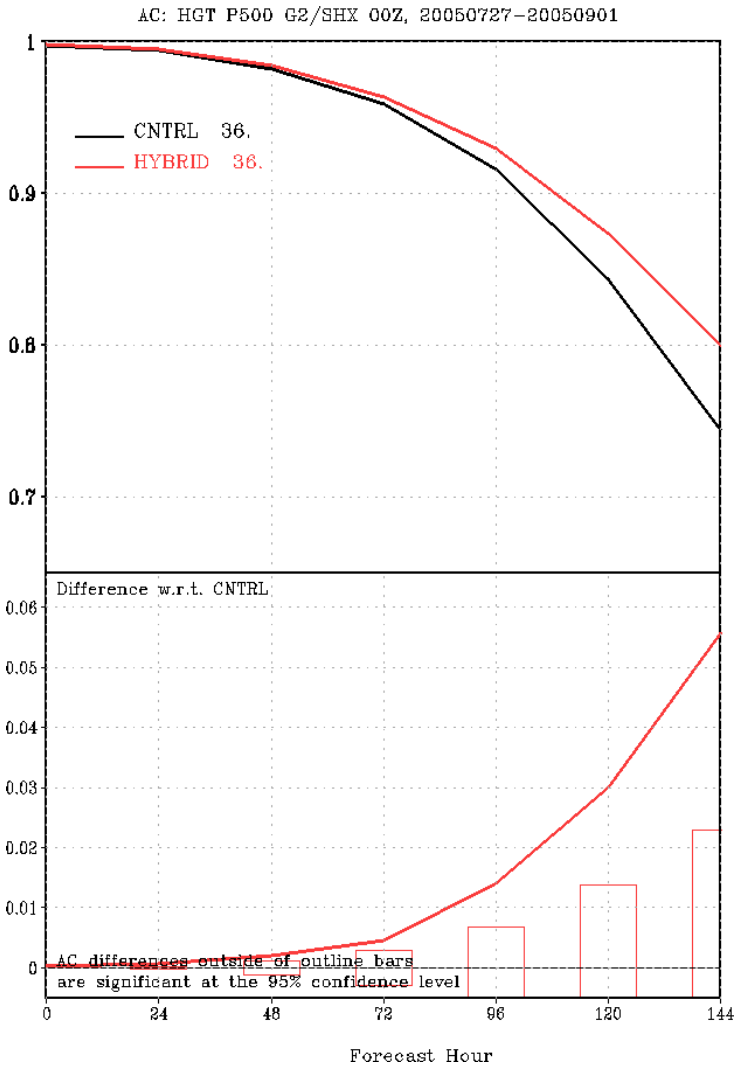


Forecast Impact:
500 hPa Anomaly Correlation

Northern Hemisphere

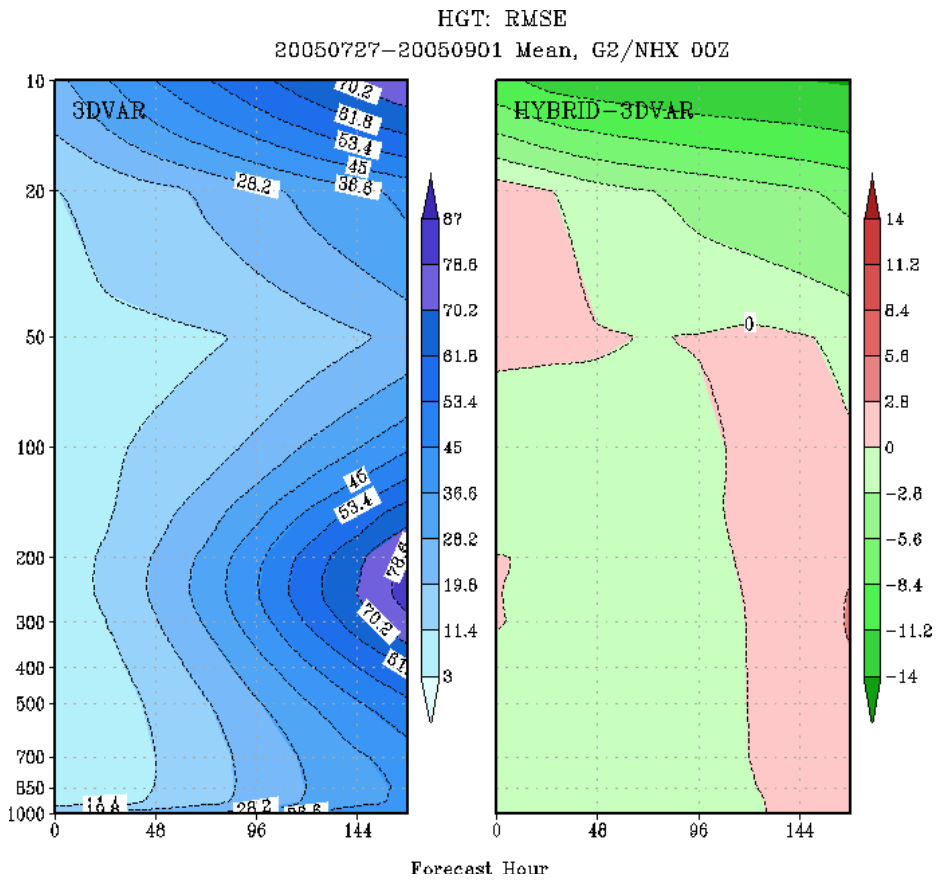


Southern Hemisphere

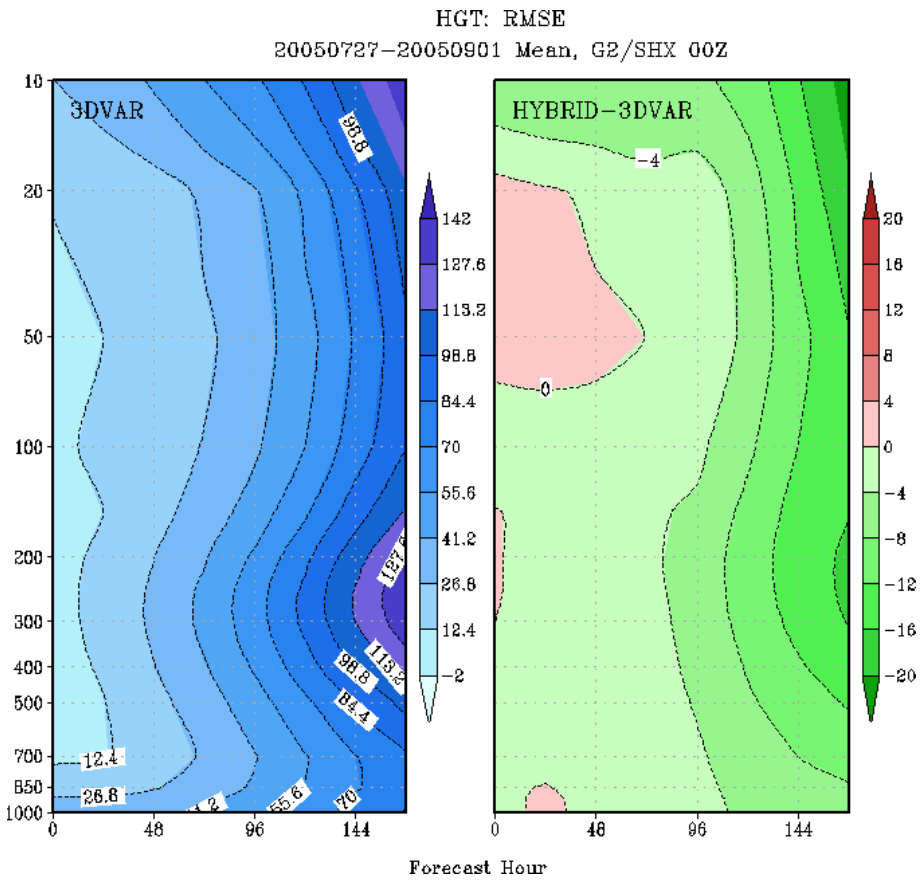


Forecast Impact: Geopotential Height RMSE

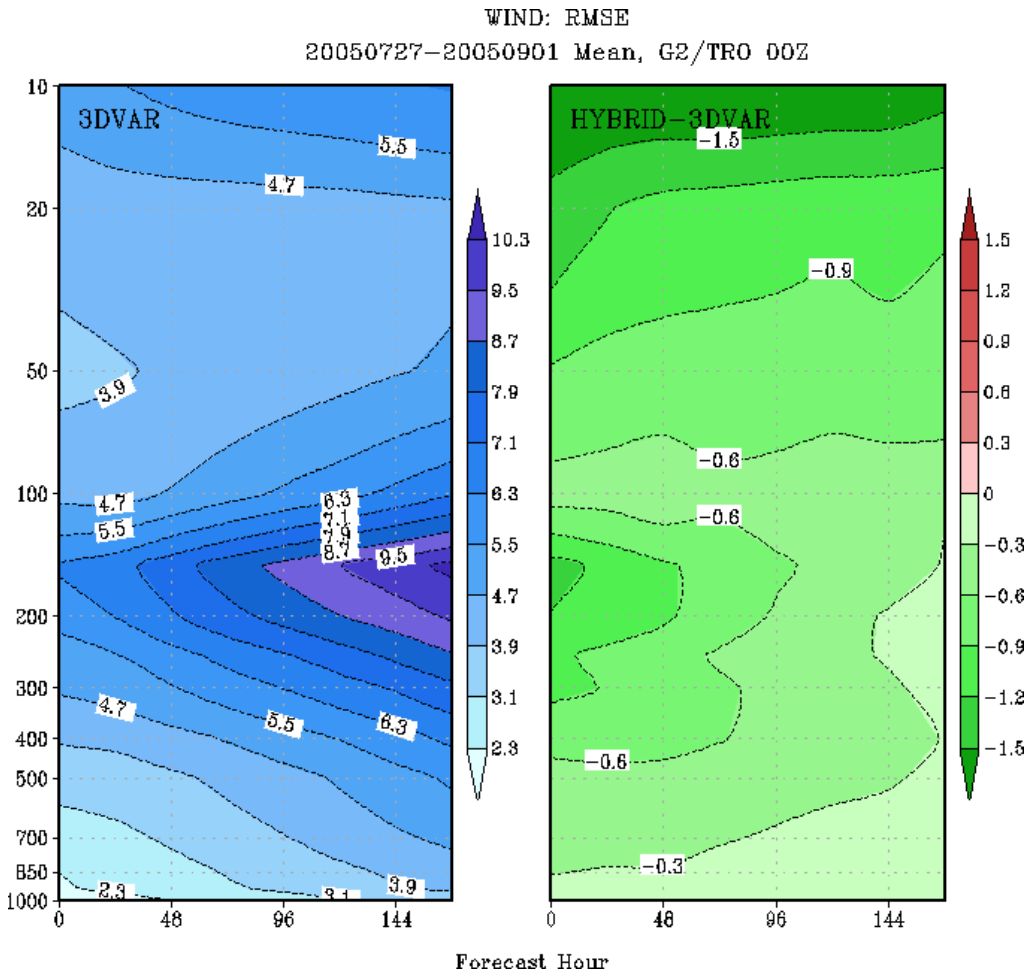
Northern Hemisphere



Southern Hemisphere



Forecast Impact:
Vector Wind RMSE (Tropics)



OSSE Summary

- Control (3DVAR) experiment using synthetic observations performed comparably to real system
- Hybrid experiment yielded improvements over the control similar to real-data experiments
  - Improved forecast skill
  - Impact on winds and stratosphere significant
- Analysis errors much smaller for hybrid than 3DVAR
  - For some variables/levels, the background error in hybrid experiment smaller than analysis error in 3DVAR control

Future Work

- What else can be learned from OSSE runs?
  - Parameter tuning
  - How well does ensemble represent error?
  - Balance/multivariate aspects of ensemble

- Scale-dependent weighting between static and ensemble estimates

- Extensions to “4d ensemble-var”
  - As well as “hybridizing” with 3D-FGAT
  - Impact of quasi-outer loop and full outer loop

$$J\left(\mathbf{x}'_{3D},\alpha\right)=\frac{1}{2}\beta_{3D}\left(\mathbf{x}'_{3D}\right)^{T}\mathbf{B}_{3D}^{-1}\left(\mathbf{x}'_{3D}\right)+\frac{1}{2}\beta_{e}\left(\alpha\right)^{T}\mathbf{L}^{-1}\left(\alpha\right)+\frac{1}{2}\sum_{k=1}^{K}\left[\left(\left(\mathbf{y}_{o}\right)'_{k}-\mathbf{H}\mathbf{x}'_{k}\right)^{T}\mathbf{R}^{-1}\left(\left(\mathbf{y}_{o}\right)'_{k}-\mathbf{H}\mathbf{x}'_{k}\right)\right]$$

$$\mathbf{x}'_{k}=\mathbf{x}'_{3D}+\sum_{n=1}^{N}\left(\alpha_{k}^{n}\circ\left(\mathbf{x}_{e}\right)_{k}^{n}\right)$$