Hybrid 4D EnVar for the NCEP GFS: Implementation plans and ~~potential for outer loops~~ comments on initialization

Daryl Kleist[1]

Rahul Mahajan[2], Catherine Thomas[1,2], Deng-Shun Chen[2,3]

[1]Univ. of Maryland-College Park, Dept. of Atmos. & Oceanic Science

[2]NOAA/NCEP/Environmental Modeling Center

[3]Central Weather Bureau and National Central University

Workshop on Sensitivity Analysis and Data Assimilation
(10th Adjoint Workshop)
Roanoke, WV, 1-5 June 2015

Hybrid 4D-Ensemble-Var [H-4DEnVar]

The Hybrid EnVar cost function can be easily extended to 4D and include a static contribution

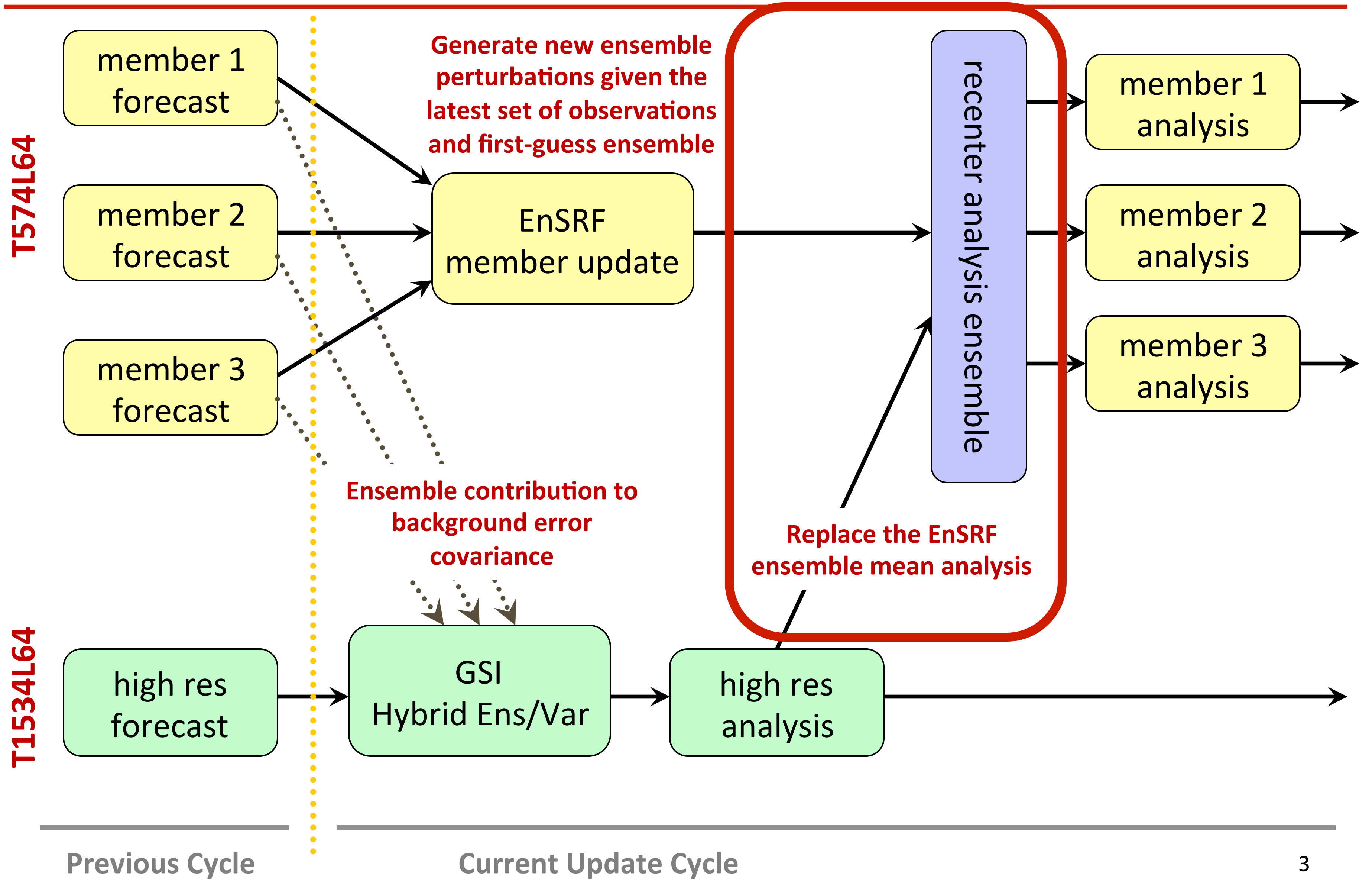
$$J\left(\mathbf{x}_\mathrm{f}',\boldsymbol{\alpha}\right)=\beta_\mathrm{f}\frac{1}{2}\left(\mathbf{x}_\mathrm{f}'\right)^\mathrm{T}\mathbf{B}_\mathrm{f}^{-1}\left(\mathbf{x}_\mathrm{f}'\right)+\beta_\mathrm{e}\frac{1}{2}\sum_{n=1}^{N}\left(\boldsymbol{\alpha}^n\right)^\mathrm{T}\mathbf{L}^{-1}\left(\boldsymbol{\alpha}^n\right)+\frac{1}{2}\sum_{k=1}^{K}\left(\mathbf{H}_k\mathbf{x}_k'-\mathbf{y}_k'\right)^\mathrm{T}\mathbf{R}_k^{-1}\left(\mathbf{H}_k\mathbf{x}_k'-\mathbf{y}_k'\right)$$

Where the 4D increment is prescribed through linear combinations of the 4D ensemble perturbations plus static contribution

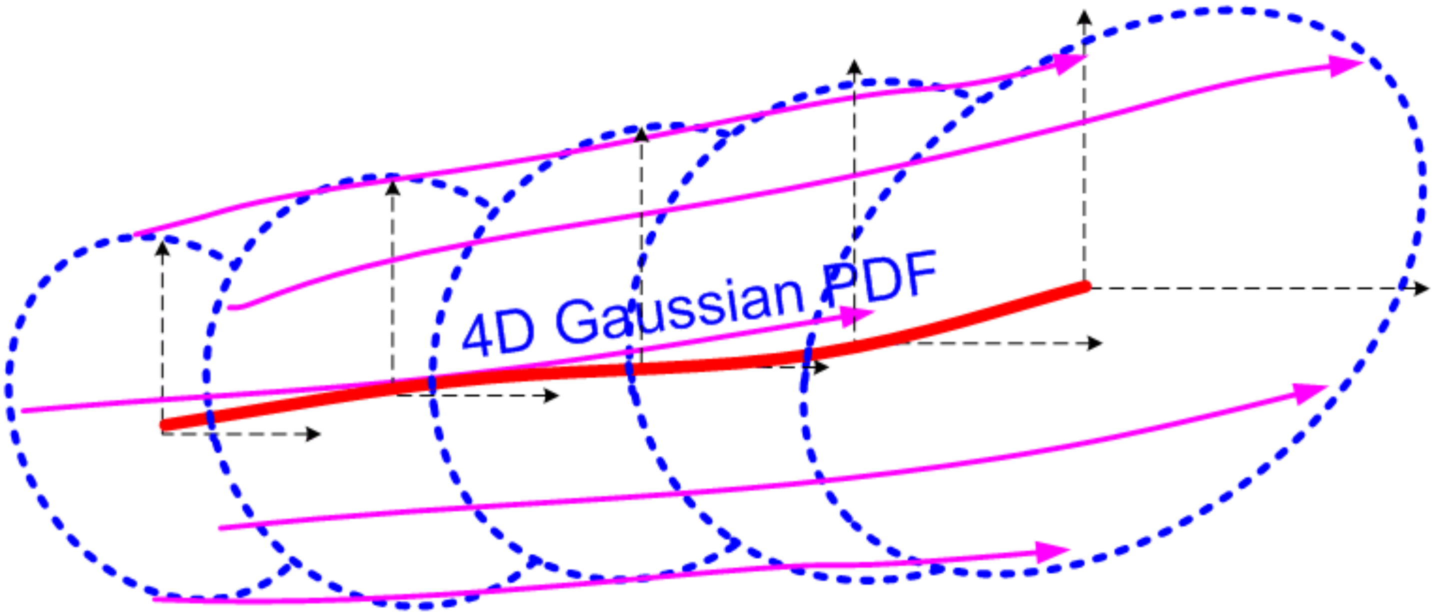
$$\mathbf{x}_k'=\mathbf{x}_\mathrm{f}'+\sum_{n=1}^{N}\left(\boldsymbol{\alpha}^n\circ\left(\mathbf{x}_\mathrm{e}\right)_k^n\right)$$

Here, the static contribution is considered time-invariant (i.e. from 3DVAR-FGAT). Weighting parameters exist just as in the other hybrid variants. No need for tangent linear/adjoint models.

Dual-Res Hybrid EnVar/EnSRF Cycling at NOAA



4D EnVar



Trajectories of perturbations from ensemble mean

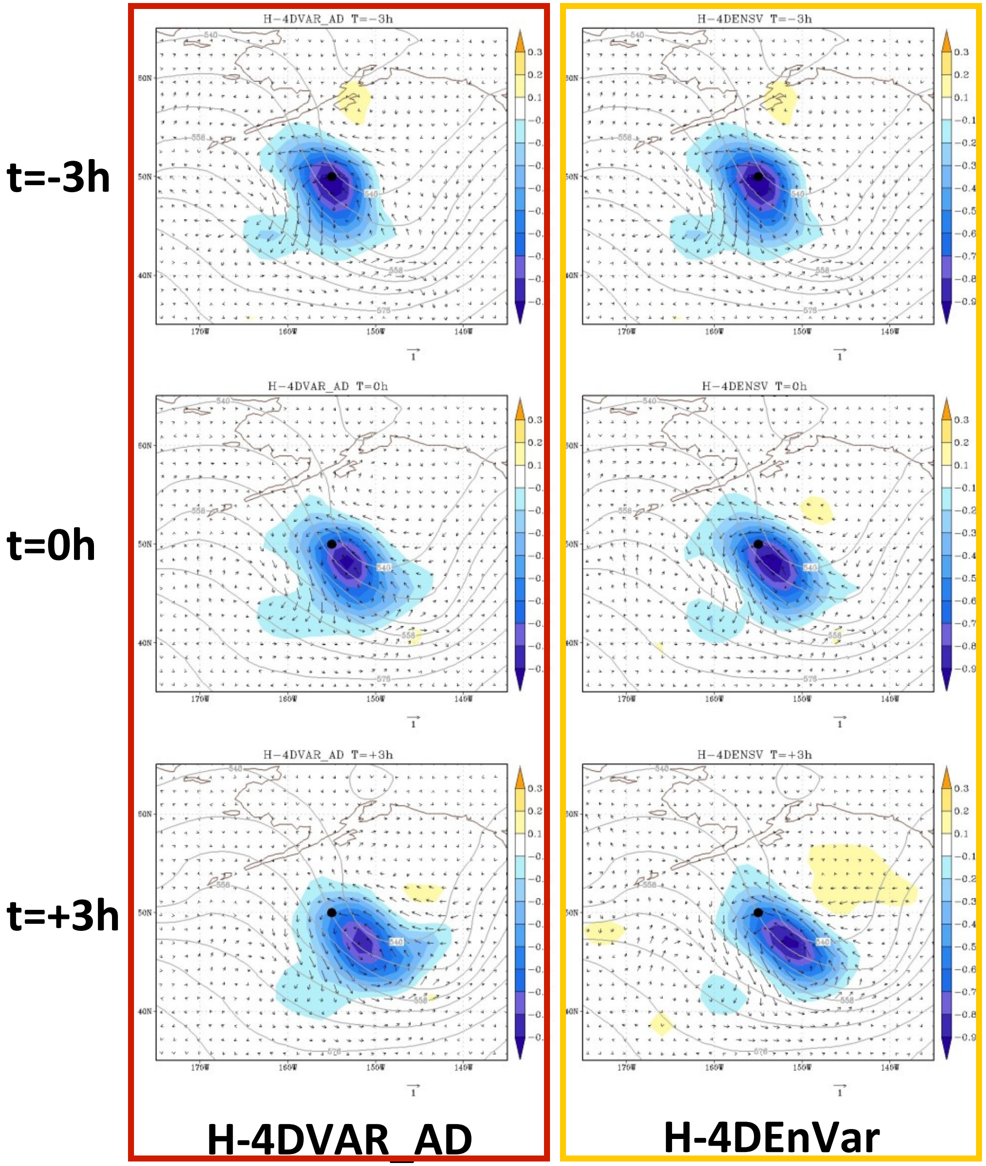
Full model evolves mean of PDF

Localised trajectories define 4D PDF of possible increments

4D analysis is a (localised) linear combination of nonlinear trajectories. It is not itself a trajectory.

Courtesy: Andrew Lorenc

Single Observation (-3h) Example From Kleist and Ide (2015)



GFS/GDAS Cycling
Experiments with real observations

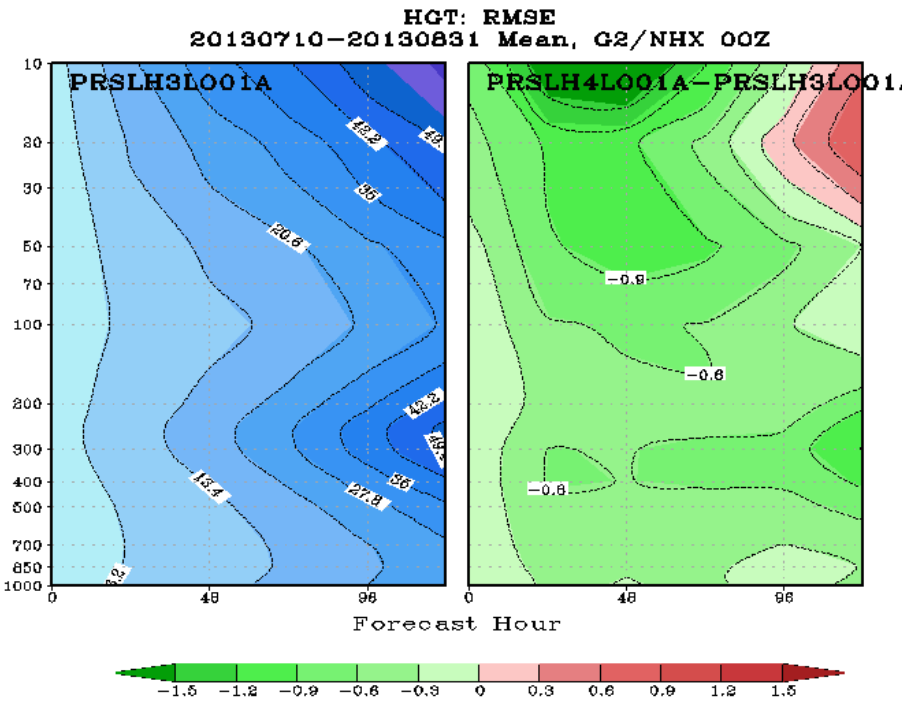
- Basic configuration
  - T670L64 Semi-Lagrangian GFS, operational observations, GFS/GDAS cycles

- **Hybrid 3D EnVar**
  - 80 member T254L46 ensemble with fully coupled (two-way) EnSRF update
  - Incremental normal mode initialization (TLNMC) on total increment
  - 87.5% ensemble & 12.5% static
  - Multiplicative inflation and stochastic physics for EnSRF perturbations

- **Hybrid 4D EnVar**
  - As in 3D Experiment, but extended to 4D
  - TLNMC on all time levels
  - Hourly TC relocation, O-G, binning of observations (not 3-hourly)

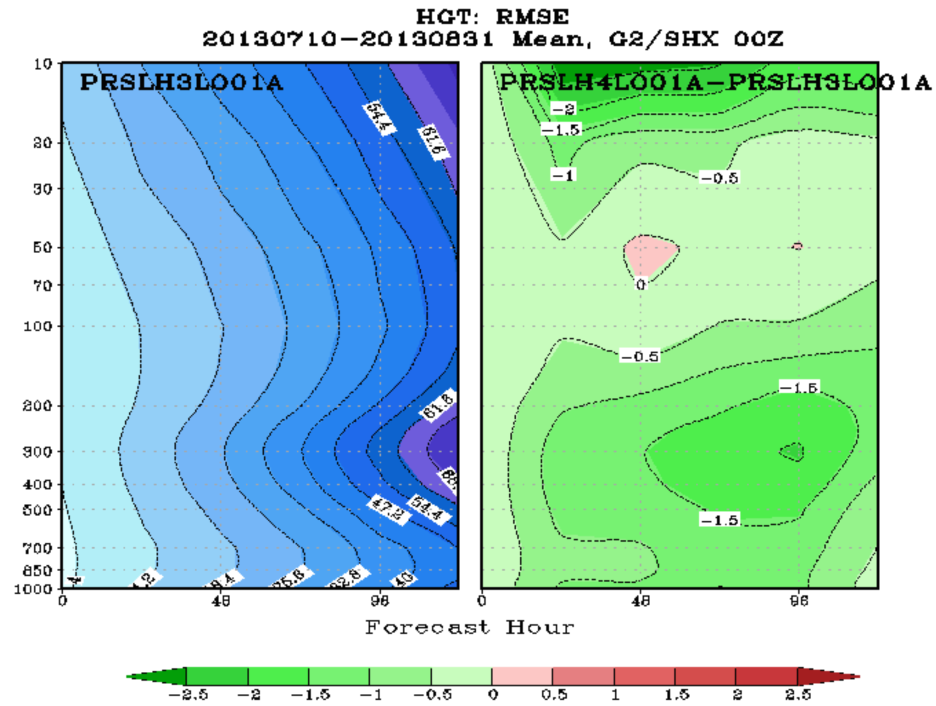
3D v 4D hybrid in SL GFS 20130710-20130831

Hyb 4DEnVar-3DEnVar



Northern Hemisphere Height RMSE and Difference

Hyb 4DEnVar-3DEnVar



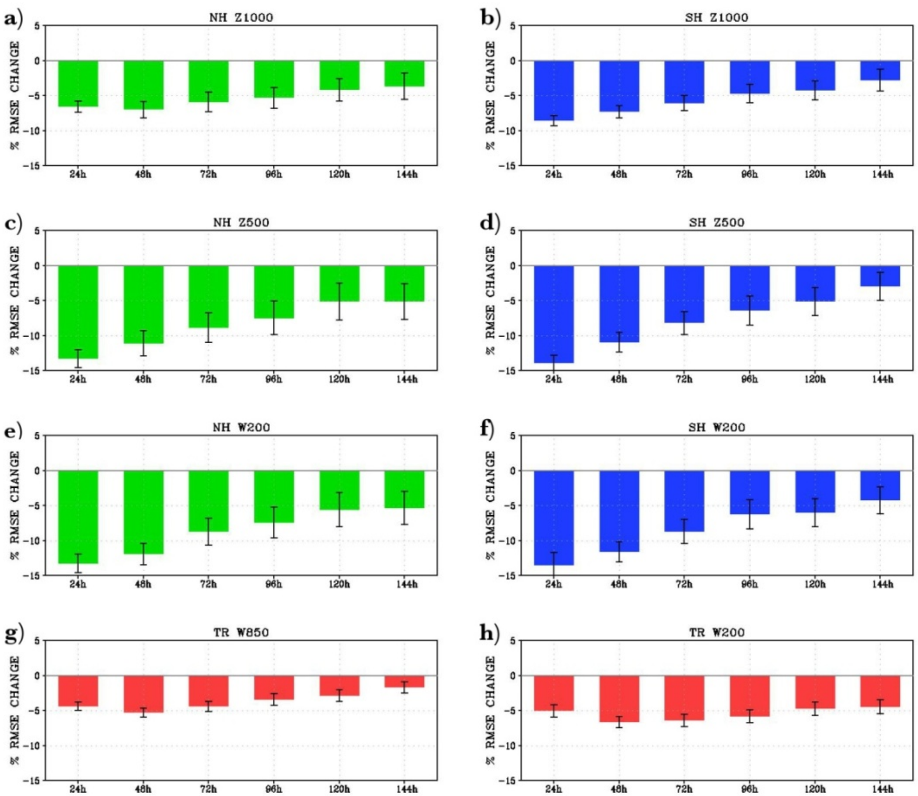
Southern Hemisphere Height RMSE and Difference

Hybrid Assimilation Trials

representative of many metrics/lead including tropical cyclone track

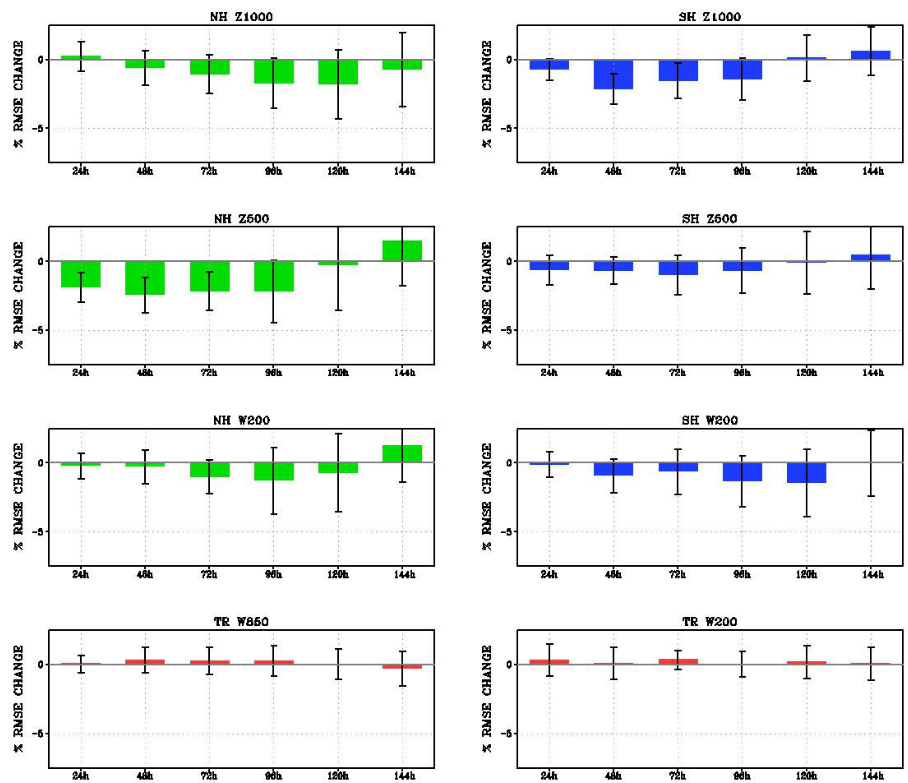
* Not an apples to apples comparison

H 3DEnVar – 3DVAR



Trial run 3D Hybrid minus 3DVAR for various metrics using the T574/T254 Eulerian GFS configuration, Feb 1 through May 15 2012 [with additive inflation, old tuning]

H 4DEnVar – H 3DEnVar



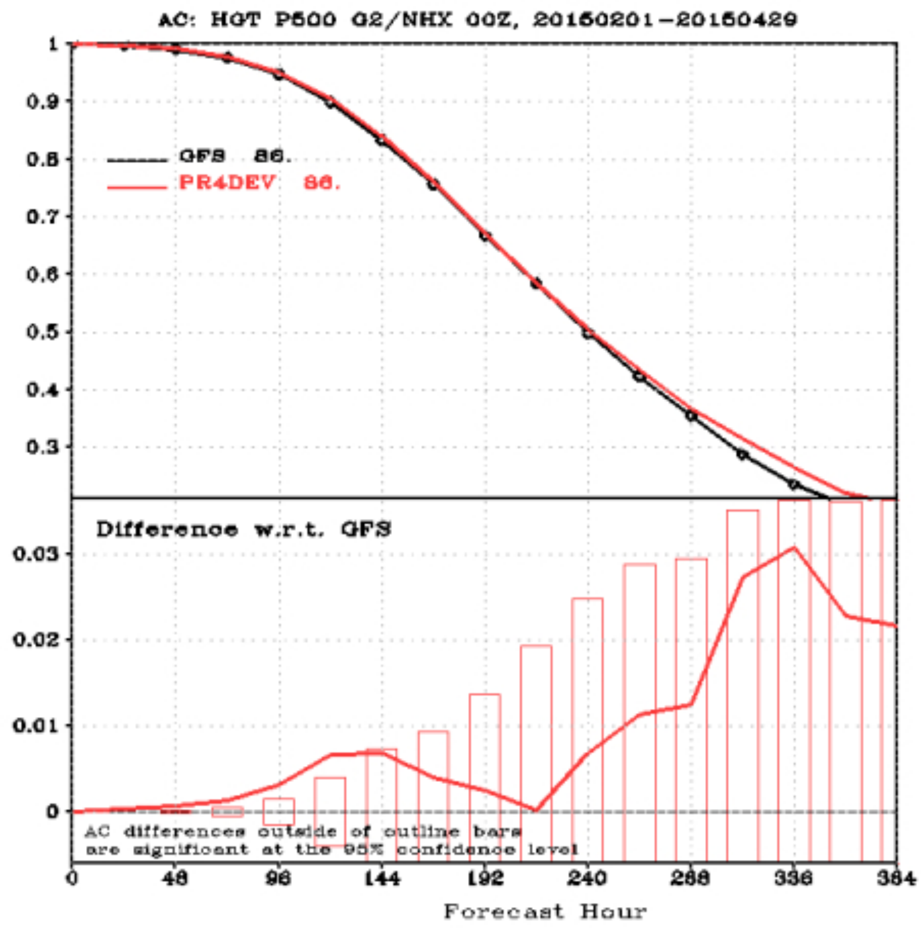
Trial run 4D Hybrid minus 3D Hybrid [EnVar] for various metrics using the T670/T254 Semi-Lagrangian GFS configuration, December 5, 2013 through January 5, 2014 [stochastic physics, new tuning]

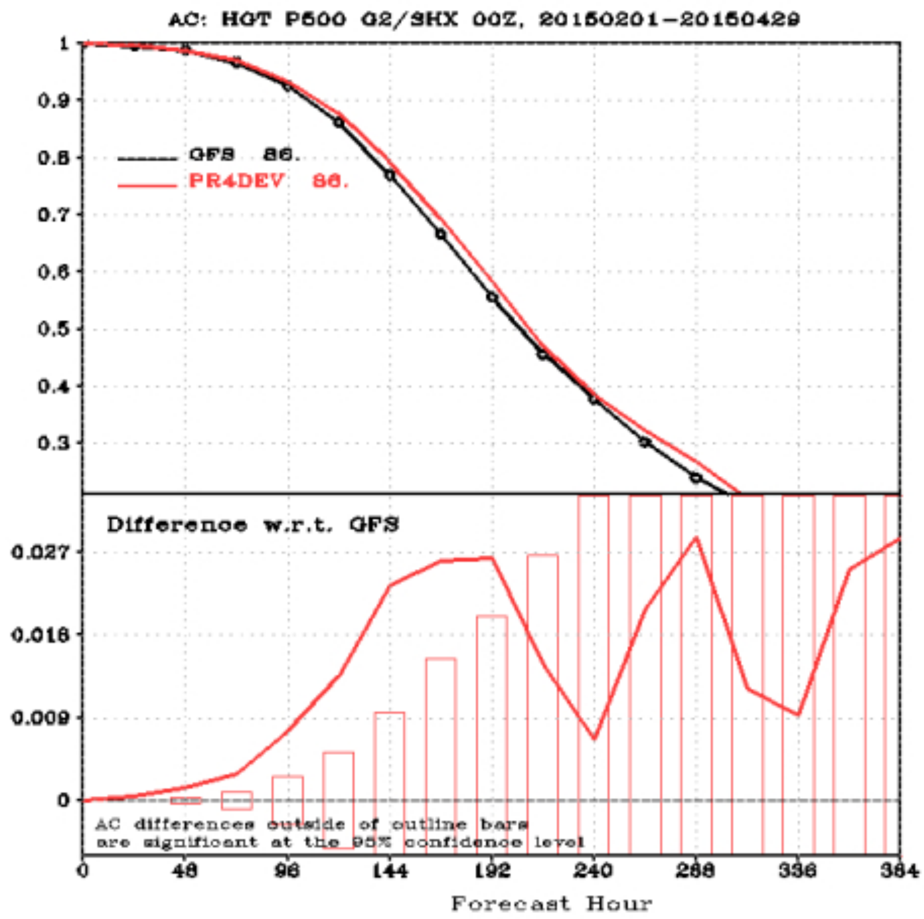
(Planned) Implementation

- Hybrid 4D EnVar to become operational for GFS/GDAS by January 2016 (tentative)
  - Basic package already frozen
  - Real time and retrospectives already underway

- **Package Configuration**
  - T1534 deterministic GFS with 80 member T574L46 ensemble with fully coupled (two-way) EnKF update
  - Incremental normal mode initialization (TLNMC) on total increment
  - 87.5% ensemble & 12.5% static
  - Multiplicative inflation and stochastic physics for EnKF perturbations
  - Full field digital filter
  - All-sky radiance assimilation
  - Minor model changes

Full Resolution (T1534/T574) Trials: 500 hPa AC





500 hPa AC for the Operational GFS (Black, 3D Hybrid) and Test 4D configuration (Red) for the period covering 02-01-2015 through 04-29-2015.

Full Resolution (T1534/T574) Trials: Summary Scorecard (02-01 through 04-29 2015)

			N. American						N. Hemisphere						S. Hemisphere						Tropics					
			Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10	Day 1	Day 3	Day 5	Day 6	Day 8	Day 10
Anomaly Correlation	Heights	250hPa	▲	▲					▲	▲	▴				▲	▲	▲	▲								
		500hPa	▲	▴					▲	▲	▴				▲	▲	▲	▲	▴							
		700hPa	▲	▲					▲	▲	▲				▲	▲	▲	▲								
		1000hPa	▲	▲	▴				▲	▲	▲				▲	▲	▲	▲								
	Vector Wind	250hPa	▲	▲	▴				▲	▲	▲				▲	▲	▲	▲								
		500hPa	▲	▲	▴				▲	▲	▲	▴			▲	▲	▲	▲	▴							
		850hPa	▲	▲	▲				▲	▲	▲	▴			▲	▲	▲	▲								
	Temp	250hPa	▲	▲					▲	▲	▲	▴			▲	▲	▴	▲	▴							
		500hPa	▲	▲					▲	▲	▴				▲	▲	▲	▲								
		850hPa	▲	▲	▴				▲	▲	▲				▲	▲	▲	▲	▴							
	MSLP	MSL	▲	▴	▴				▲	▲	▲				▲	▲	▲	▲								
RMSE	Heights	10hPa		▾					▼	▼	▼				▼	▼	▼	▼	▼		▼	▼	▼	▼	▼	▼
		20hPa							▼	▼					▼						▼					
		50hPa							▼										▴			▲	▲	▲	▲	▲
		100hPa									▴	▴			▴		▴	▴								
		200hPa	▲	▴					▲	▲	▲				▲	▲	▲	▲			▴		▲	▴		▴
		500hPa	▲	▲	▴				▲	▲	▲				▲	▲	▲	▲	▴		▾	▼				
		700hPa	▲	▲	▴				▲	▲	▲				▲	▲	▲	▲			▾	▼				
		850hPa	▲	▲	▴				▲	▲	▲				▲	▲	▲	▲								
		1000hPa	▲	▲	▴				▲	▲	▲	▴			▲	▲	▲	▲								
	Vector Wind	10hPa	▲	▲	▲	▲	▲		▲	▲	▴	▴	▴		▲	▲	▲	▴			▲	▲	▲	▲	▲	▲
		20hPa	▲	▲					▲	▲					▲	▲									▼	▼
		50hPa	▲	▲					▲	▲	▲	▴			▲	▲	▲	▴				▲	▲	▴		
		100hPa	▲	▲					▲	▲	▲	▲			▲	▲	▴	▴								
		200hPa	▲	▲	▴				▲	▲	▲				▲	▲	▲	▴			▲	▲				
		500hPa	▲	▲	▲				▲	▲	▲	▴			▲	▲	▲	▲	▴		▲	▲	▲	▲	▴	
		700hPa	▲	▲	▲				▲	▲	▲	▴			▲	▲	▲	▲	▴		▲	▲	▴	▲		
		850hPa	▲	▲	▲				▲	▲	▲	▲			▲	▲	▲	▲	▴		▲	▲				
		1000hPa	▲	▲	▲				▲	▲	▲	▲			▲	▲	▲	▲	▴		▲	▲				▲
	Temp	10hPa		▼	▼	▼	▾			▼	▼	▼				▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼
		20hPa														▼	▼	▼	▼	▾	▼	▼	▼	▼	▼	▼
		50hPa																▴			▼					
		100hPa	▲	▴					▲	▲					▴	▲	▲	▲			▼					▾
		200hPa	▲	▲	▲				▲	▲	▲	▴			▲	▲	▲	▲	▴		▴	▲				
		500hPa	▲	▲					▲	▲	▲	▴			▲	▲	▲	▲			▲	▲	▲	▲	▴	
		700hPa	▲	▲	▴				▲	▲	▲	▴			▲	▲	▲	▲	▴		▲	▲	▲	▴	▴	▲
		850hPa	▲	▲	▲				▲	▲	▲	▴			▲	▲	▲	▲	▴			▴				
		1000hPa	▲	▲	▴				▲	▲	▲	▴			▲	▲	▲	▲								

Hybrid 4D EnVar: Summary

- Natural extension to operational 3D EnVar
  - Thanks to Yannick Tremolet and efforts toward 4D capability in GSI
- Scalable
  - And can be improved even further
  - Aligns with technological/computing advances
- Stochastic Physics
  - Suitable replacement for additive inflation ("NMC perturbations")
- Computationally inexpensive relative to 4DVAR (with TL/AD)
  - Estimates of improved efficiency by 10x or more
  - Caveat: Still not better than hybrid 4DVAR

Potential Corrections for Noise and/or Imbalance

- Noise in the background (first guess/model forecast)
  - Full field digital filters ** (currently used in GFS)
  - Initialization (Nonlinear Normal Mode Initialization)
    - Analysis draws to data, Initialization pushes away from observations

- Noise in the analysis increment
  - Improved multivariate variable definition
  - Penalty terms
  - *Incremental normal mode initialization*

- Discrepancy in passing increment to model
  - Incremental analysis update

Constraint Options Explored in NCEP System

- Tangent Linear Normal Mode Constraint
  - Based on past experience and tests with 3D hybrid, default configuration includes TLNMC over all time levels (quite expensive)

$$\mathbf{x}_k' = \mathbf{C}_k \left[\mathbf{x}_\mathrm{f}' + \sum_{n=1}^{N} \left(\boldsymbol{\alpha}^n \circ \left(\mathbf{x}_\mathrm{e} \right)_k^n \right) \right]$$

- Weak Constraint "Digital Filter"
  - Construct filtered/initialized state as weighted sum of 4D states

$$J_{dfi} = \chi \left\langle \mathbf{x}_m - \mathbf{x}_m^\mathrm{i}, \mathbf{x}_m - \mathbf{x}_m^\mathrm{i} \right\rangle$$

$$\mathbf{x}_m^\mathrm{i} = \sum_{k=1}^{K} \mathrm{h}_{k-m} \mathbf{x}_k^\mathrm{u}$$

- Combination of the two
  - Apply TLNMC to center of assimilation window only in combination with JcDFI (Cost effective alternative?)

Tangent Linear Normal Mode Constraint
Kleist et al. (2009)

$$J(\mathbf{x}_c') = \frac{1}{2}(\mathbf{x}_c')^T \mathbf{C}^{-T}\mathbf{B}^{-1}\mathbf{C}^{-1}(\mathbf{x}_c') + \frac{1}{2}(\mathbf{y}_o' - \mathbf{H}\mathbf{x}_c')^T \mathbf{R}^{-1}(\mathbf{y}_o' - \mathbf{H}\mathbf{x}_c') + J_c$$

$$\mathbf{x}_c' = \mathbf{C}\mathbf{x}'$$

- analysis state vector after incremental NMI
  - **C** = Correction from incremental normal mode initialization (NMI)
    - represents correction to analysis increment that filters out the unwanted projection onto fast modes

- No change necessary for **B** in this formulation
- Based on:
  - Temperton, C., 1989: "Implicit Normal Mode Initialization for Spectral Models", MWR, vol 117, 436-451.

TLNMC Procedure

$C=[I-DFT]x'$

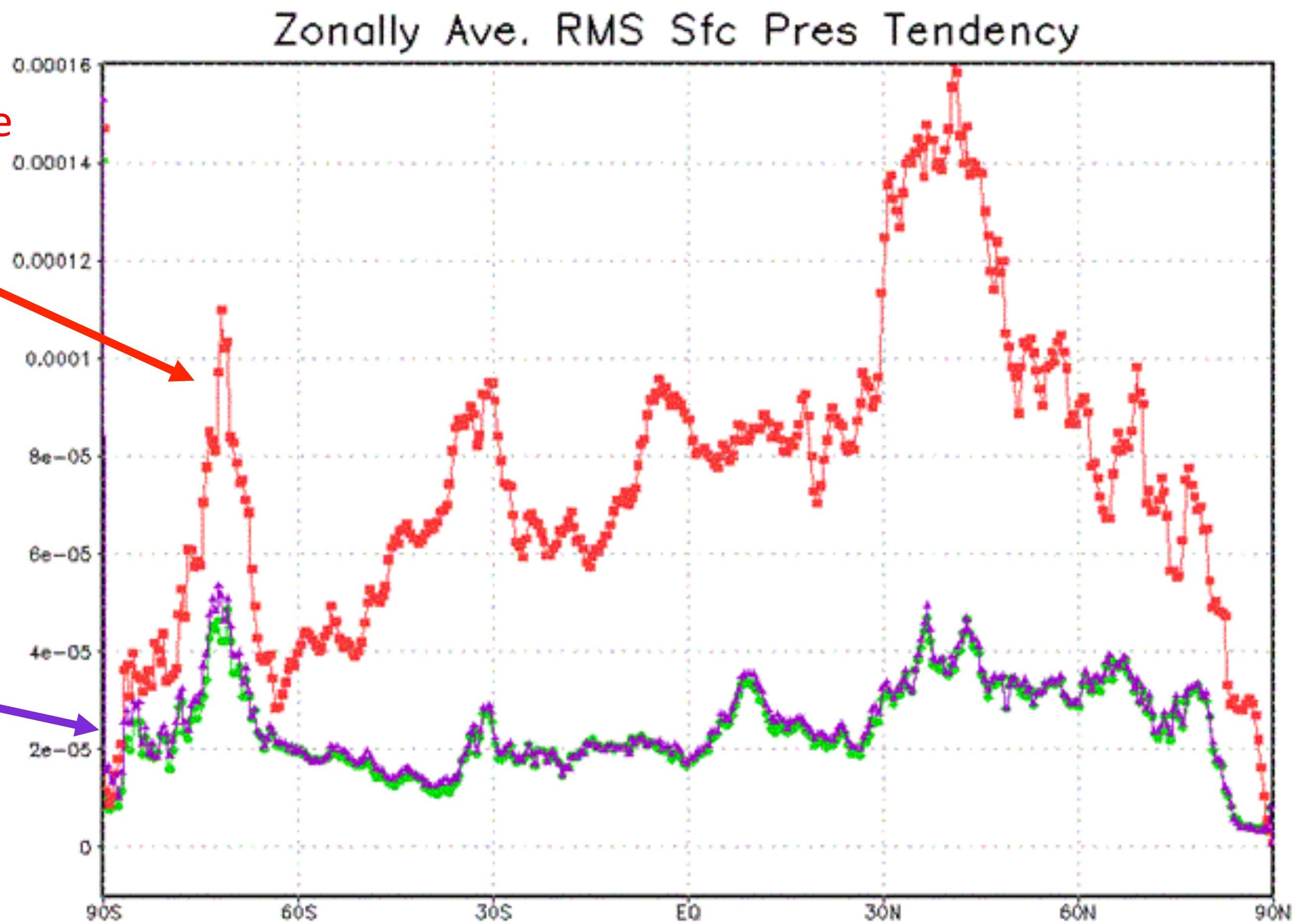
T	F	D
n x n	m x n	n x m
$\frac{d\mathbf{x}'}{dt}$	$\frac{d\mathbf{x}_g'}{dt}$	$-\Delta\mathbf{x}'$
Dry, adiabatic tendency model	**Projection onto m gravity modes** **m-2d shallow water problems**	**Correction matrix to reduce gravity mode Tendencies** **Spherical harmonics used for period cutoff**

- Practical Considerations:
  - **C** is operating on **x**' only, and is the tangent linear of NNMI operator
  - Only need one iteration in practice for good results
  - ***Adjoint of each procedure needed as part of minimization/variational procedure***

Impact of TLNMC analysis

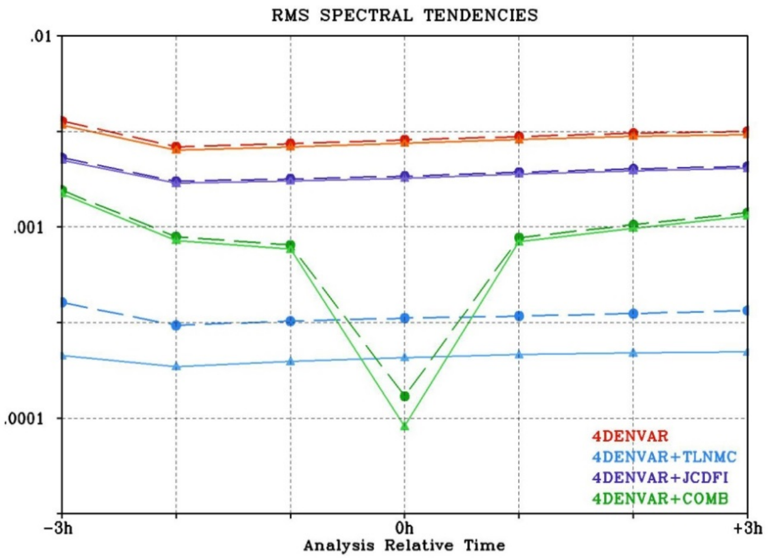
Substantial increase without constraint

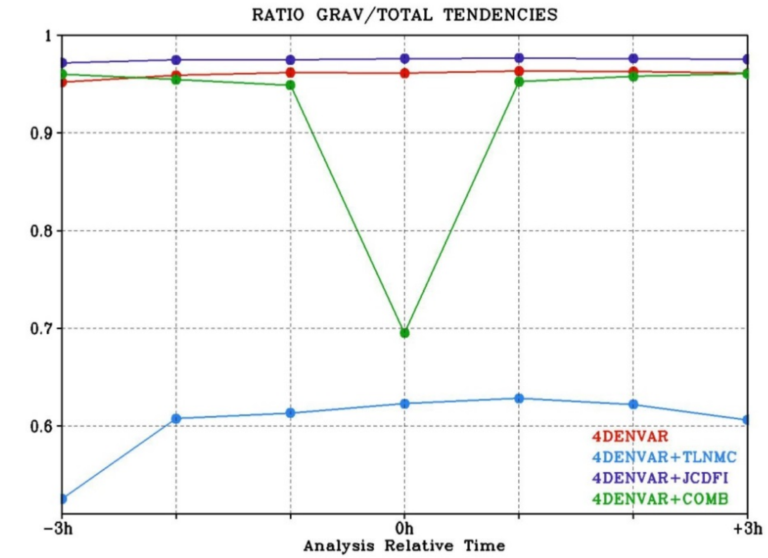
Minimal increase with TLNMC



Zonal-average surface pressure tendency for background (green), unconstrained 3DVAR analysis (red), and 3DVAR analysis with TLNMC (purple).

Constraint impact (single case)

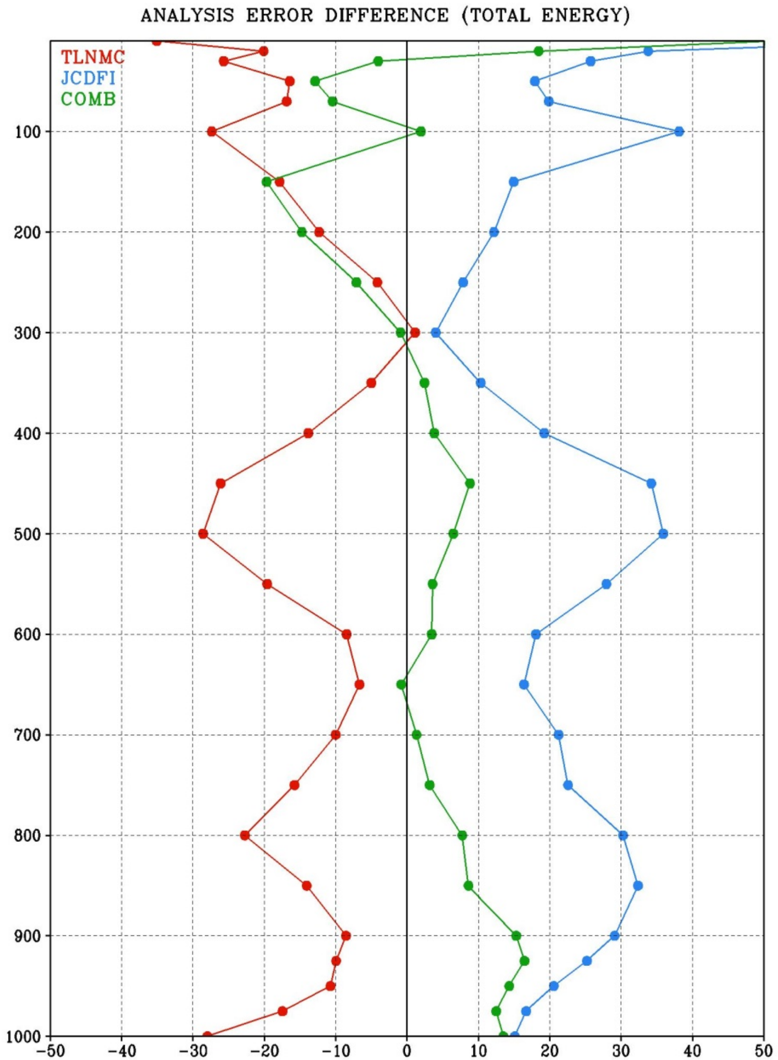




- **Impact on tendencies**
  - Dashed: Total tendencies
  - Solid: Gravity mode tendencies
  - All constraints reduce incremental tendencies

- **Impact on ratio of gravity mode/ total tendencies**
  - JcDFI increases ratio of gravity mode to total tendencies
  - TLNMC most effective (but most expensive)
  - Combined constraint potential (cost effective alternative)

Analysis Error Relative to Control for various options(cycled OSSE)

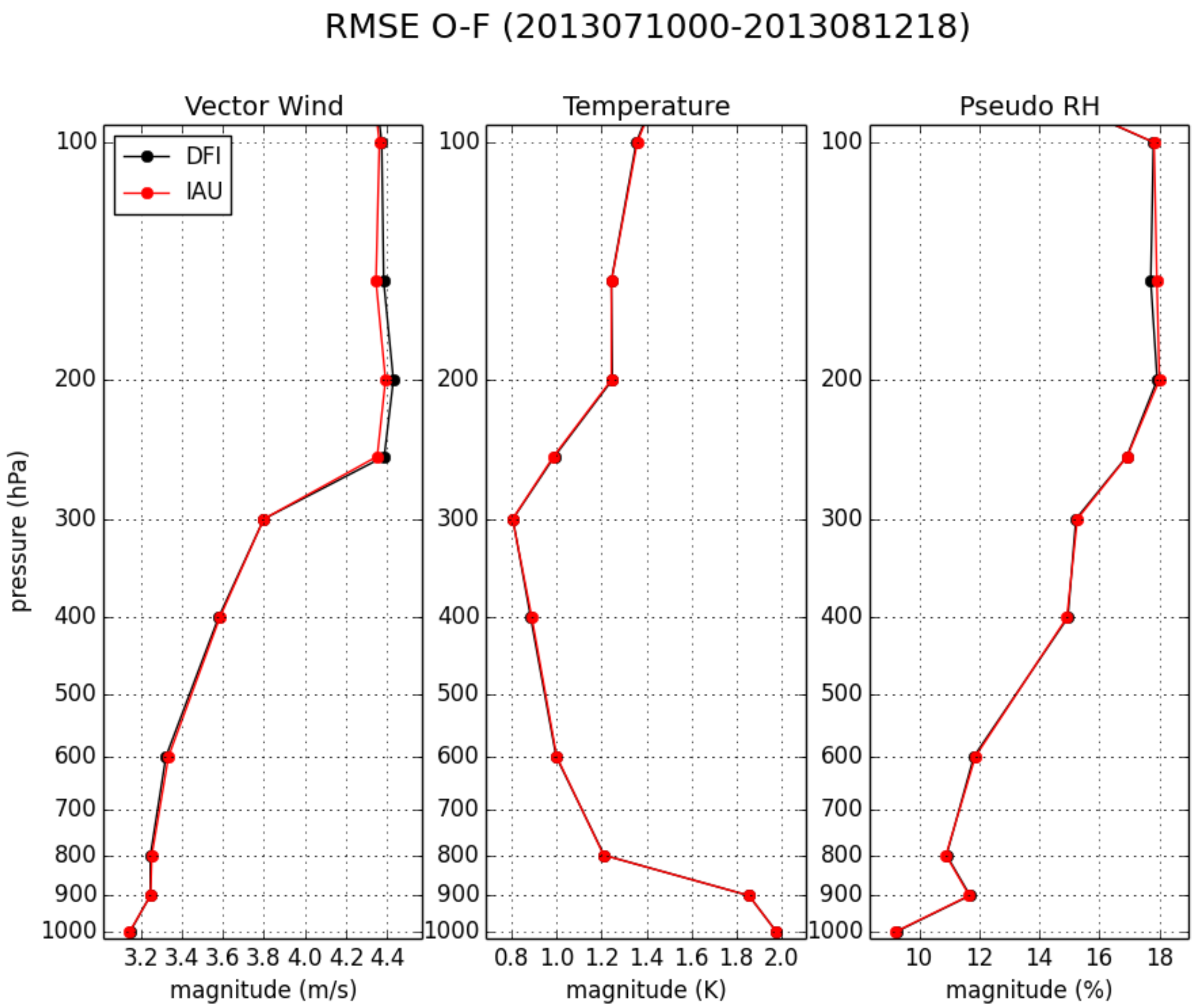


- From Joint OSSE (ECMWF T511 NR, Simulated Obs from Nikki)
- Control (zero) is 4D hybrid without constraint/initialization
- 6 week trial
- Time mean (August) change in analysis error (total energy) ***relative*** to 4D hybrid EnVar experiment that utilized no constraints at all
    - TLNMC universally better
    - Combined constraint mixed
    - JcDFI increases analysis error

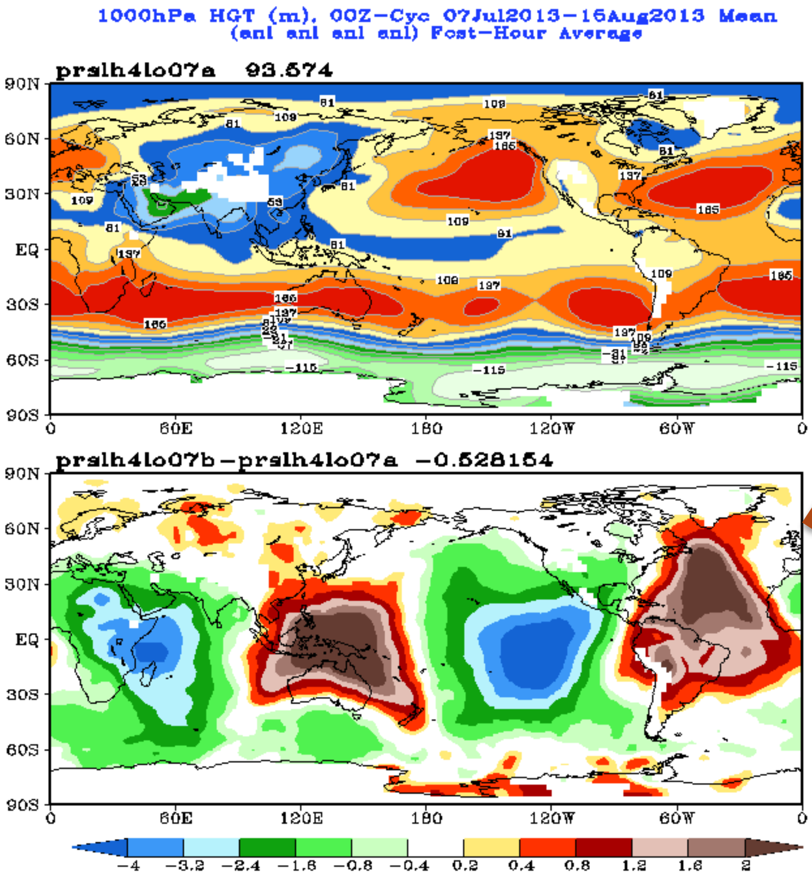
4D Incremental Analysis Update

- Incremental Analysis Update (Bloom, 1996) helps by using model to distribute a (single) increment over a time window with constant weights (we call this 3DIAU).
  - Propagation of increment neglected, might be significant for fast-moving weather systems.
  - May help spin up unobserved/non-updated state variables
- 4D version of IAU has been proposed by UK Met Office
  - Positive Impact in UKMO and Canadian 4D EnVar
- Approximation of "mollified" time-continuous formulation EnKF proposed by Bergemann & Reich (2010).

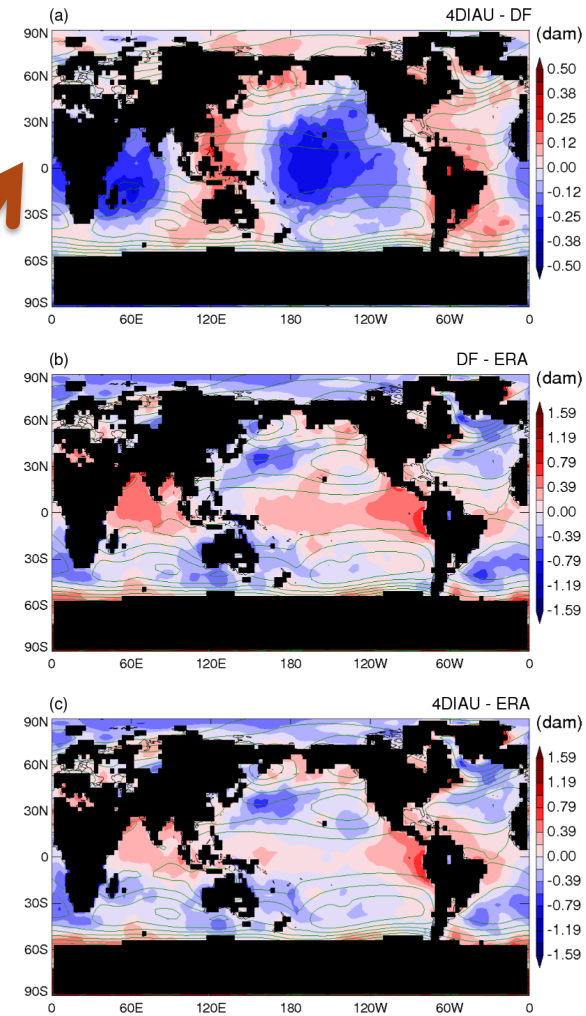
Preliminary Low Res. Results Courtesy: Rahul Mahajan



IAU Impact on 1000 hPa heights relative to DFI



GFS 4D Hybrid IAU-DFI at 00 UTC



Buehner et al. (2015)

What Next?

- 4D IAU likely replacement for DFI
- TLNMC needs work (or replacement)
  - Tropics degraded
- **Can we get 4D EnVar to beat hybrid 4DVAR?**
- Future work on 4D EnVar DA at NCEP/UMD:
  - Scale-dependent weighting (visitor Deng-shun Chen)
  - Localization: Wave-band/scale-dependent
  - Incorporate ensemble update into GSI (EVIL, d-EVIL, mean-pert)
  - Nonlinearity, ***outer loops***, variable choices
  - What to do about static (time-invariant) static error covariance
    - Low order modeling
    - FOTO