# Hybrid 4DEnVar at NCEP


Rahul Mahajan[1], Daryl Kleist[3], Catherine Thomas[1], Yanqiu Zhu[1], and John Derber[2]

[1]IMSG at [2]NOAA/NWS/NCEP Environmental Modeling Center, [3]University of Maryland, College Park


## Hybrid EnVar

The use of hybrid covariance models, which combine a fixed climatological estimate with an ensemble-based representation, has become quite popular for numerical weather prediction (NWP). One such method for incorporating localized covariances from an ensemble within the variational framework utilizes an augmented control variable (ensemble-var, or EnVar), and has been implemented in the operational NCEP data assimilation system (Gridpoint Statistical Interpolation, GSI). As first proposed in Lorenc (2003) and ignoring preconditioning for simplicity, the cost function (with an extension for 4D) for such an algorithm can be written as:

$$J(\mathbf{x}'_f,\boldsymbol{\alpha}) = \beta_f \frac{1}{2}(\mathbf{x}'_f)^T \mathbf{B}_f^{-1}(\mathbf{x}'_f) + \beta_e \frac{1}{2}\sum_{n=1}^{N}(\boldsymbol{\alpha}^n)^T \mathbf{L}^{-1}(\boldsymbol{\alpha}^n) + \frac{1}{2}\sum_{k=1}^{K}(\mathbf{H}_k\mathbf{x}'_k - \mathbf{y}'_k)^T \mathbf{R}_k^{-1}(\mathbf{H}_k\mathbf{x}'_k - \mathbf{y}'_k).$$

Where $\mathbf{B}_f$ is the static background error covariance, **R** the observation error covariance, **H** the observation operators, $\mathbf{y}'$ the observation innovation, $\alpha^n$ the ensemble control variable for the *n*th member of ensemble size *N*, **L** the "error covariance" for the control variable used to enforce localization. The total analysis increment is then constructed to be a linear combination of static ($\mathbf{x}_f$) and ensemble contributions:

$$\mathbf{x}'_t = \mathbf{M}_k\mathbf{x}'_f + \mathbf{T}\sum_{n=1}^{N}\left(\boldsymbol{\alpha}^n \circ (\mathbf{x}_e)_k^n\right),$$

where $\mathbf{M}_k$ is chosen to be the identity model and the single static contribution is valid through the entire assimilation window as in hybrid 3DEnVar (e.g. it is time invariant). The ensemble contribution is prescribed as a linear combination of ensemble perturbations weighted by the control variable. The operator **T** represents a transform (interpolation) between the ensemble perturbations and deterministic resolution if using a dual-resolution configuration. The contribution of static and ensemble respectively is controlled by the weighting parameters $\beta_f$ and $\beta_e$. NCEP employs an 80 member ensemble updated with an Ensemble Kalman Filter in the global data assimilation system (GDAS) and global forecast system (GFS). Pre-implementation tests showed that going from 3DVAR to hybrid 3DEnVar to hybrid 4DEnVar reduced forecast errors significantly for many metrics, regions, and lead times (Fig. 1).

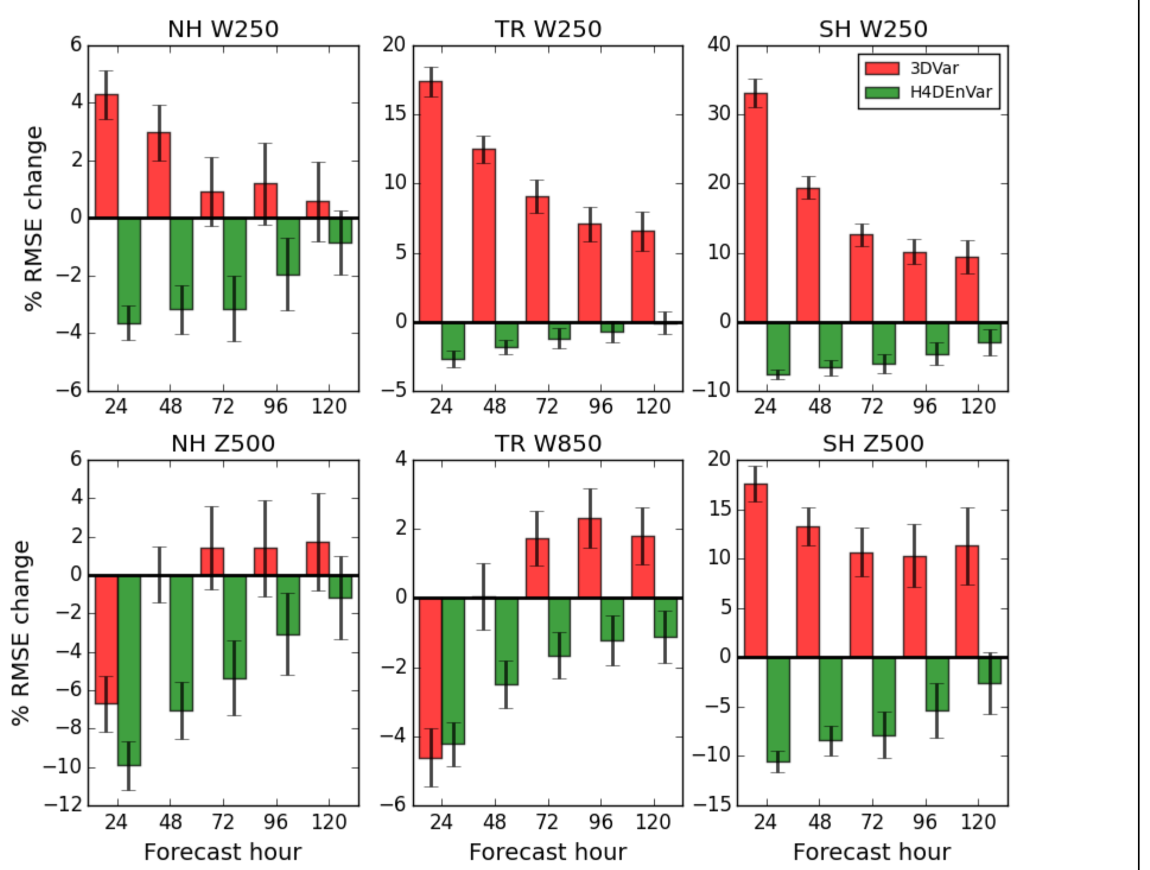
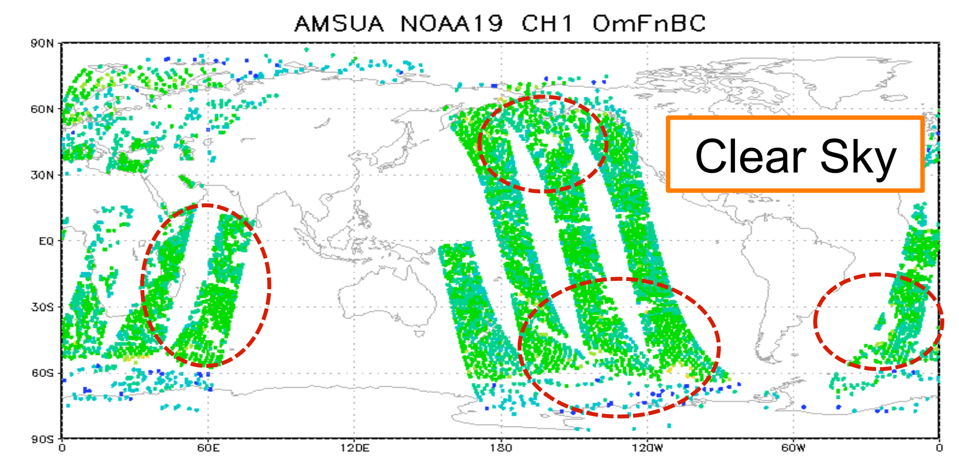


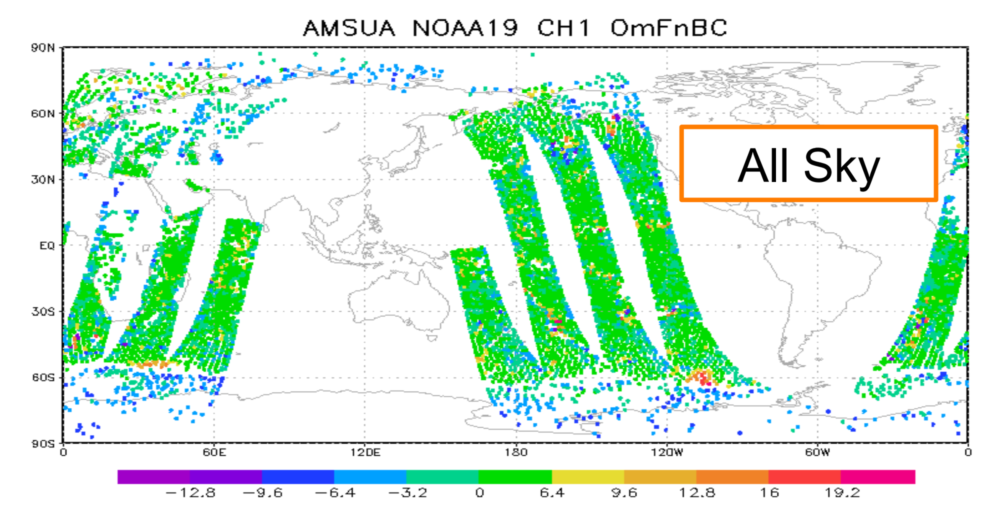

*Figure 1: Percent change in root mean square error from hybrid 3DEnVar and 3DVar (red) and hybrid 4DEnVar (green) for the period covering 15 July 2013 through 15 September 2013 for selected variables as a function of forecast lead time. The forecast variables include 250hPa vector winds in the NH, Tropics and SH (top row), 500 hPa geopotential heights and 850 hPa vector winds. All verification is performed using self-analysis. The error bars represent the 95% confidence threshold for a significance test.*

## All-Sky MW Radiance Assimilation

In the previous operational hybrid 3DEnVar, the clear-sky approach of radiance data assimilation was employed. With the development effort made for all-sky conditions on modifying and assessing quality control, observation error assignment, bias correction, and background error covariance in the EnVar framework, the capability of all-sky microwave radiance assimilation in the Gridpoint Statistical Interpolation (GSI) analysis system has been developed at the NCEP. Because the GFS output does not provide snow and precipitation profiles, the AMSU-A cloudy radiances affected by non-precipitating cloud over ocean are assimilated in this study. In this section, we will present the configuration of the all-sky microwave radiance assimilation and discuss in more detail about observation error, bias correction, cloud water background error variance and ensemble spread. Outstanding issues and ongoing/future work are also presented.

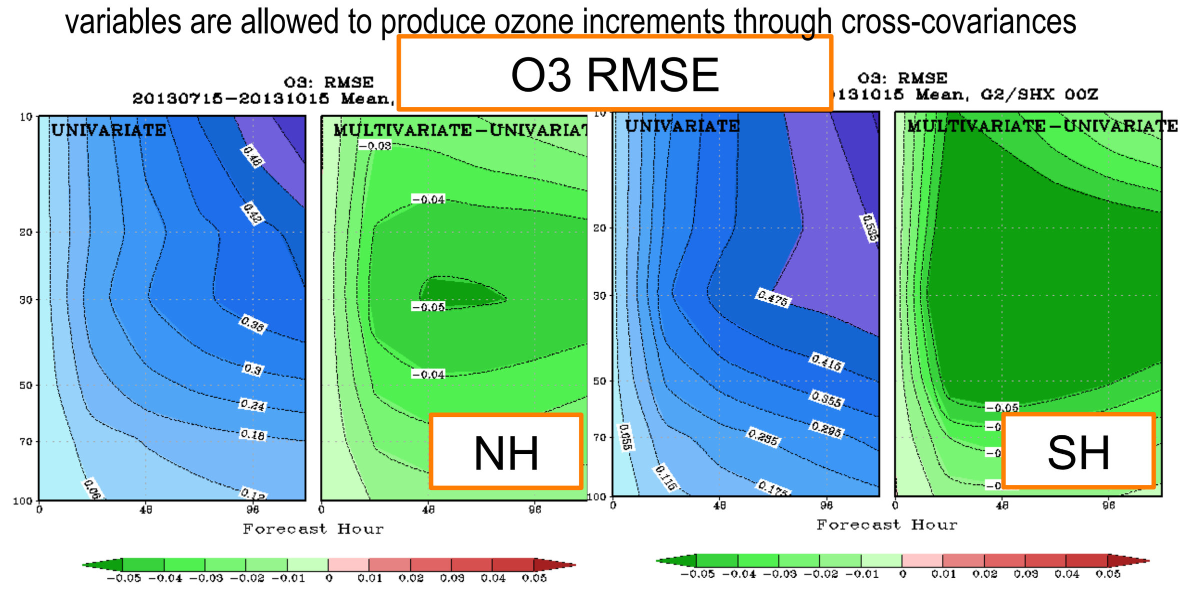
- Clear sky data + radiances affected by **thick** and **thin** clouds
- **AMSU-A channels 1-5, 15** currently assimilated
- New all-sky radiance bias correction strategy (Zhu et. al 2014)
- Situation-dependent observation error inflation; AMSU-A observation error re-tuned
- Additional quality control: cloud effect (Geer et al. 2013) and emissivity sensitivity screening
- Precipitation regions are excluded
- Only initial implementation of this capability
- Approximately **10% increase** in observations




## Multivariate Ozone Update

Ozone observations are allowed to update other state variables, and other state variables are allowed to produce ozone increments through cross-covariances




## FY16 Implementation

Putting together all the above mentioned components,

For most metrics and lead times, 4D hybrid is significantly better than the 3D counterpart (Fig. 4 and Fig. 5). The forecast improvements are not nearly as large as what was found in moving from 3DVAR to hybrid 3DEnVar, but still better.

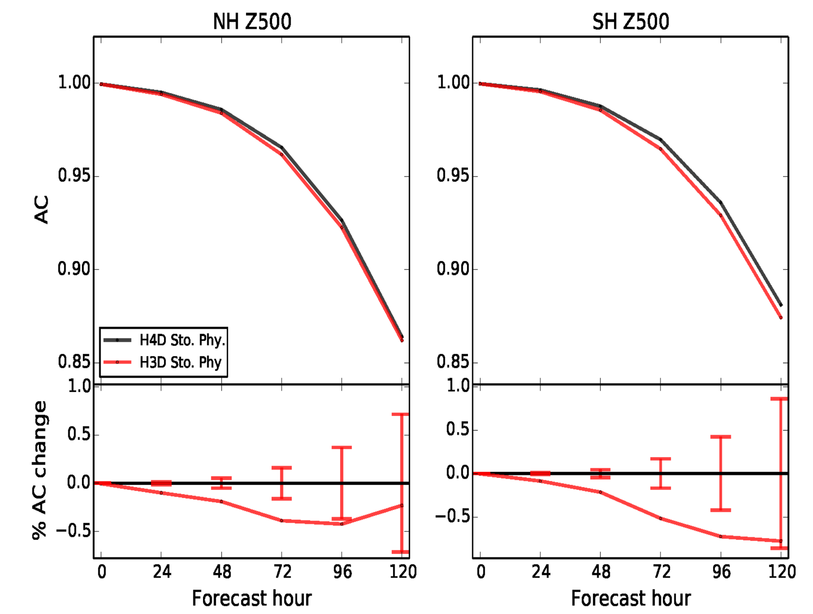



*Figure 4: Time-averaged 500 hPa anomaly correlation (upper panels) for the Northern Hemisphere (left) and Southern Hemisphere (right) for the Hybrid 3DEnVar (red) and Hybrid 4DEnVar (black) experiments for forecasts from the 00 UTC analyses as a function of lead time as well as the difference (lower panels) 3D minus 4D for the entire three month trial run. The 95% confidence threshold for a significance test (derived from a standard t-test) is also plotted in the lower panels.*

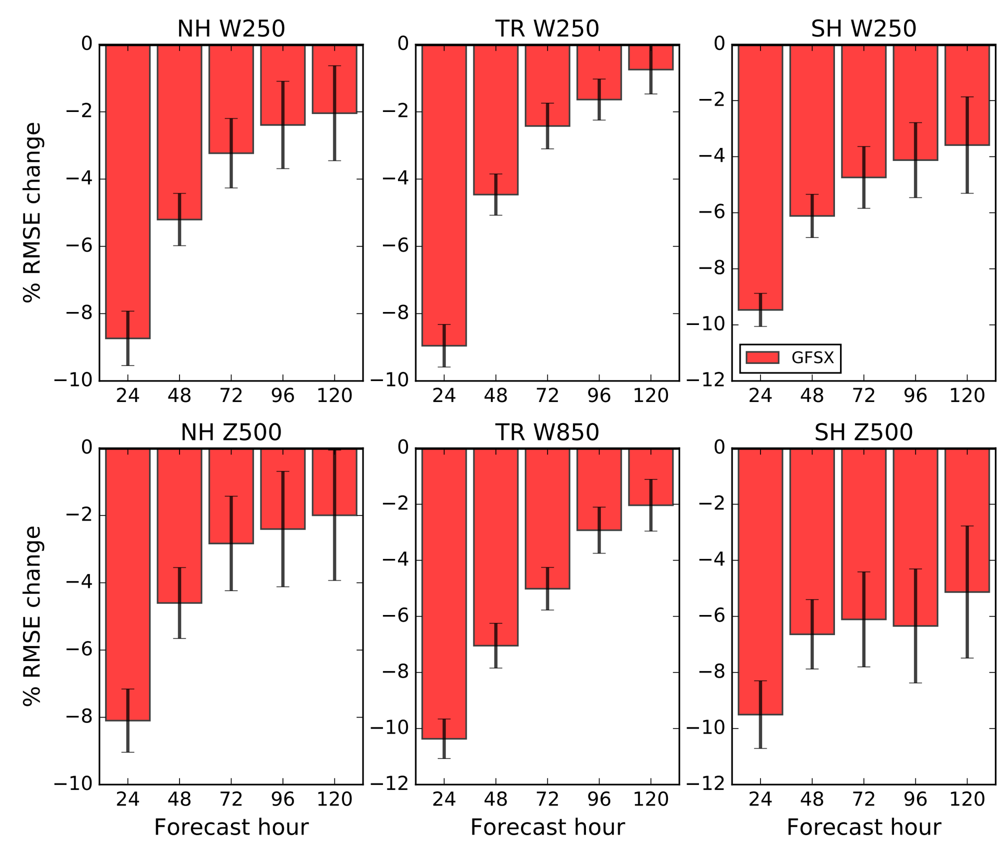



*Figure 5: As in Fig. 1, but for the percent change in the Hybrid 3DEnVar experiment relative to the Hybrid 4DEnVar experiment for Northern Hemisphere 250 hPa vector wind (upper left) and 500 hPa geopotential height (lower left), tropical 250 hPa vector wind (upper middle) and 850 hPa vector wind (lower middle), and Southern Hemisphere 250 hPa vector wind (upper right) and 500 hPa geopotential height (lower right). All verification is performed using self-analysis. The error bars represent the 95% confidence threshold for a significance test.*

## Looking Forward

There is still significant room for improvement relative to the 4D configurations carried out thus far. Work is under way to improve the initialization (using 4D incremental analysis update), explore the use of an legitimate outer loop as is done on 4DVAR, and reducing the negative impacts from the static, time-invariant contribution to the hybrid solution. Additional testing with increasing the ensemble (analysis) resolution is also underway.

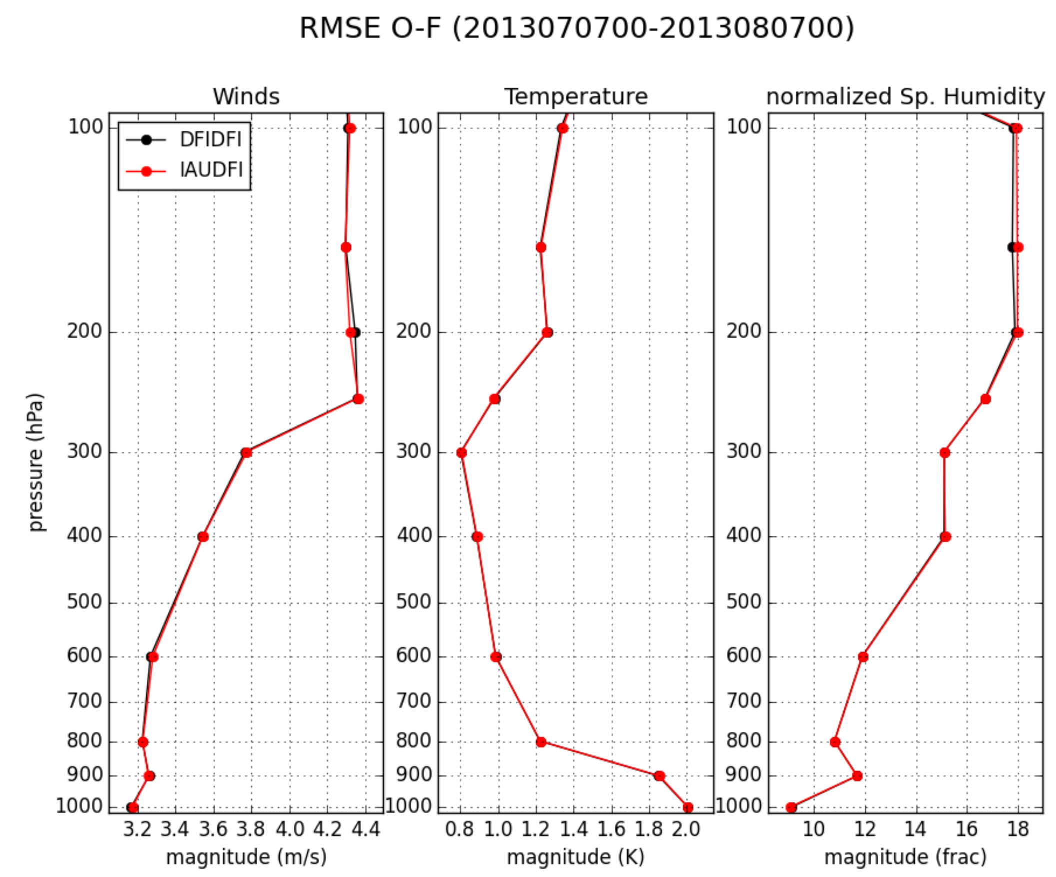



*Figure 6: Time-averaged OmF RMSE profiles with respect to Radiosondes comparing the use of full-field digital filter (black) with 4DIAU (red). The IAU is only used in the high-resolution forecasts. 4DIAU is also being tested in the ensemble.*

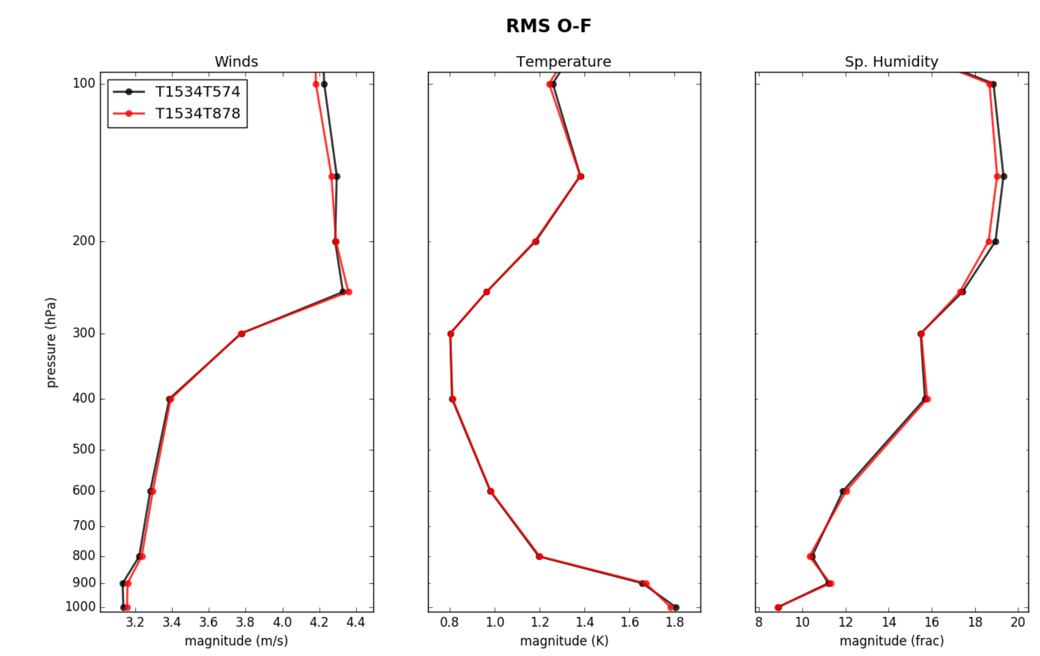



*Figure 6: Time-averaged OmF RMSE profiles with respect to Radiosondes . The control (black) is the current operational configuration with ensemble resolution at T574. The experiment (red) has an increased ensemble resolution of T878. Further increase in resolution of the ensemble is being tested.*

## References


Kleist, D. T., and K. Ide, 2015: An OSSE-based Evaluation of Hybrid Variational-Ensemble Data Assimilation for the NCEP GFS, Part II: 4DEnVar and Hybrid Variants. Mon. Wea. Rev., in press (doi: 10.1175/MWR-D-13-00350.1).

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP – a comparison with 4D-VAR. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183-3203 (doi: 10.1256/qj.02.132).

Zhu et. al., 2016: All-Sky Microwave Radiance Assimilation in the NCEP's GSI. Mon. Wea. Rev., in review.