Data assimilation using a hybrid variational-ensemble Kalman filter approach

- ❑Meso/regional-scale TC track forecast
- ❑ Convective-scale radar DA applied for TC
- ❑Global-scale forecast

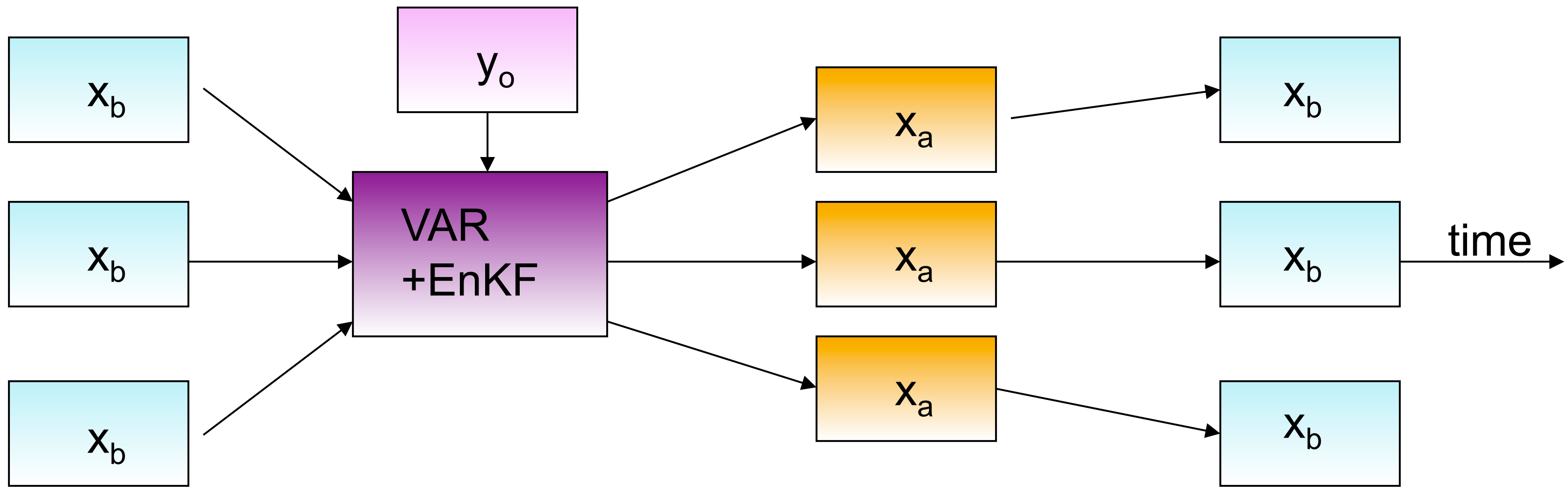
Xuguang Wang

School of Meteorology
and Center for Analysis and Prediction of Storms,
University of Oklahoma, Norman, OK, USA

Yongzuo Li (OU), Ming Xue (OU), Jeff Whitaker (NOAA/ESRL), Dave Parrish (NOAA/NCEP), Daryl Kleist (NOAA/NCEP)

The 4th EnKF workshop, Albany, NY, USA, April 6-9, 2010

What's Hybrid?



X_b: forecast
X_a: analysis
y_o: observation
VAR: variational (3DVAR, 4DVAR) method
EnKF: ensemble Kalman filter

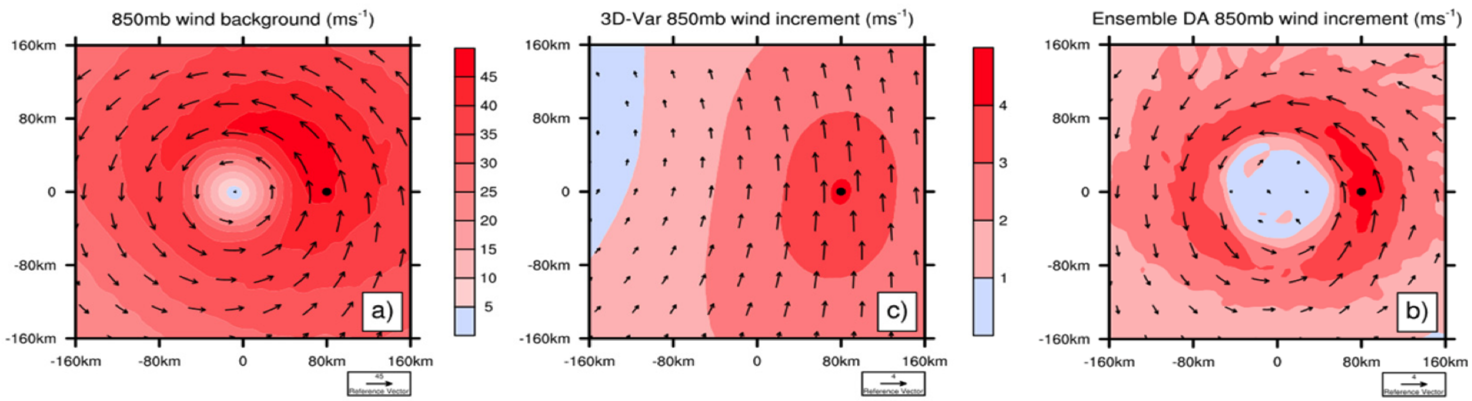
HYBRID

- Hybrid of two frameworks: unlike EnKF, data assimilation part adopts variational framework.

- Hybrid of background error covariances: unlike VAR, ensemble forecasts involve in the estimate of background error covariance.

Why Hybrid?
“Best of both worlds”

- Studies (e.g., Wang et al. 2007b,2008ab, 2009b, Buehner et al. 2009ab) have demonstrated that a hybrid VAR-EnKF can significantly improve upon a standalone VAR system due to the inclusion of flow dependent ensemble covariance in the estimate of the background error covariance.



Wang 2010b

- Studies (Wang et al. 2007b, 2009b; Buehner et al. 2009ab) also showed hybrid can improve upon a standalone EnKF due to the limited ensemble members used by EnKF.

- Compared to EnKF, hybrid adopts model space rather than observation space covariance localization, more appropriate for nonlocal observation operator (e.g., satellite radiance, GPS refractivity, radar obs. with attenuation) (Campbell et al. 2009).

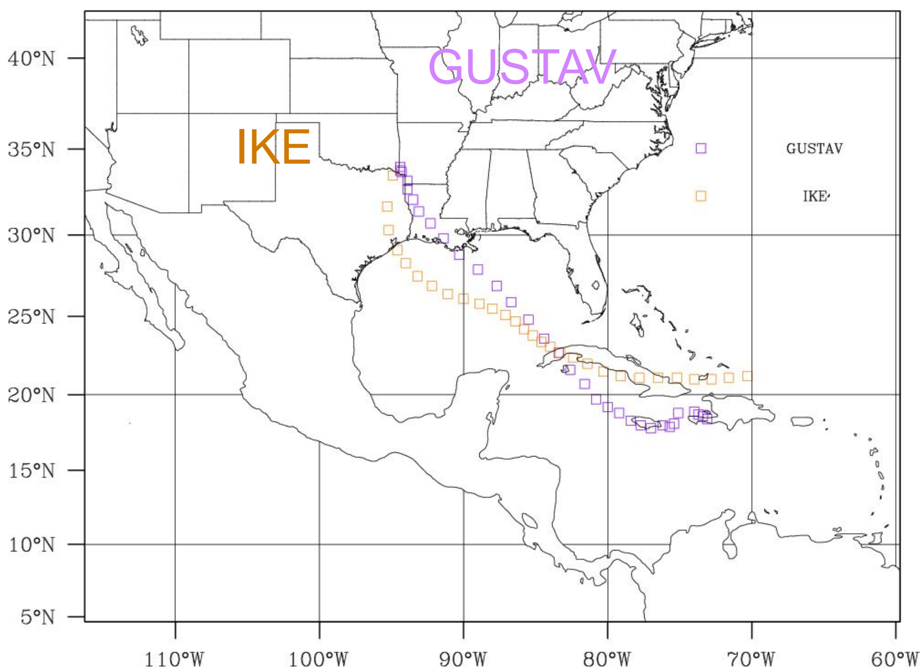
Why Hybrid?
"Best of both worlds"

▪Good choice for already having an established VAR and ensemble forecast system (Wang et al. 2007ab, 2009b)
✓Minor changes to the existing operational variational framework
✓Take advantages of existing capability of VAR, e.g., variational data quality control
✓ Static covariance model in advanced VAR can provide sophisticated method to reduce sampling errors in ensemble covariance (e.g., anisotropic recursive filter for adaptive covariance localization).
✓VAR framework can also easily add in dynamic constraint.

▪VAR framework provides maximum likelihood solution and thus allows non-Gaussian errors (Zupanski 2005). Outer-loop in VAR can take care of nonlinearity (Kalnay et al. 2009)

▪ Study (Caya et al. 2005) shows that for radar DA, 4DVAR spins up faster than EnKF, but EnKF is better in later stage of the DA cycles. Hybrid can take advantage of both. Kalnay (et al. 2009) suggested ways to adapt this advantage of 4DVAR in EnKF.

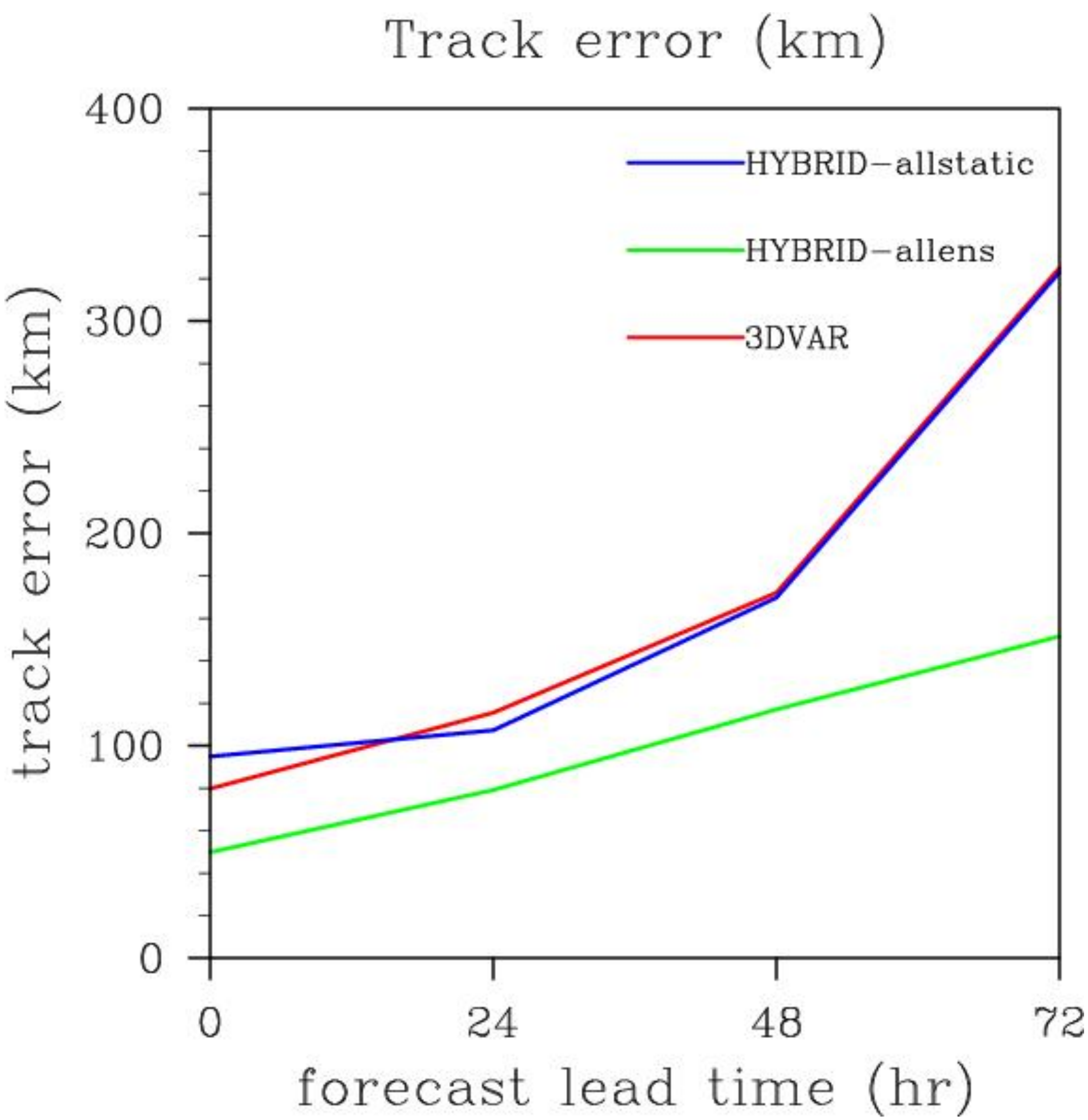
Hurricane track forecast: IKE, GUSTAV 2008



Wang 2010b

- DA system: hybrid ETKF-3DVAR (Wang et al. 2008ab)
- WRF Model: Δx=30km; 35 levels
- Observations: from GTS; all conventional in-situ data plus cloud wind, QuikScat wind, satellite derived temperature profile. No vortex relocation or bogus or position assimilation.
- Ensemble size: 32 members
- DA, forecast and verification:
  - 3hrly DA cycling;
  - forecasts after 2-day/3-day spin up for IKE/ GUSTAV every 12h.
  - Compare forecasts initialized by WRF hybrid with WRF 3DVAR

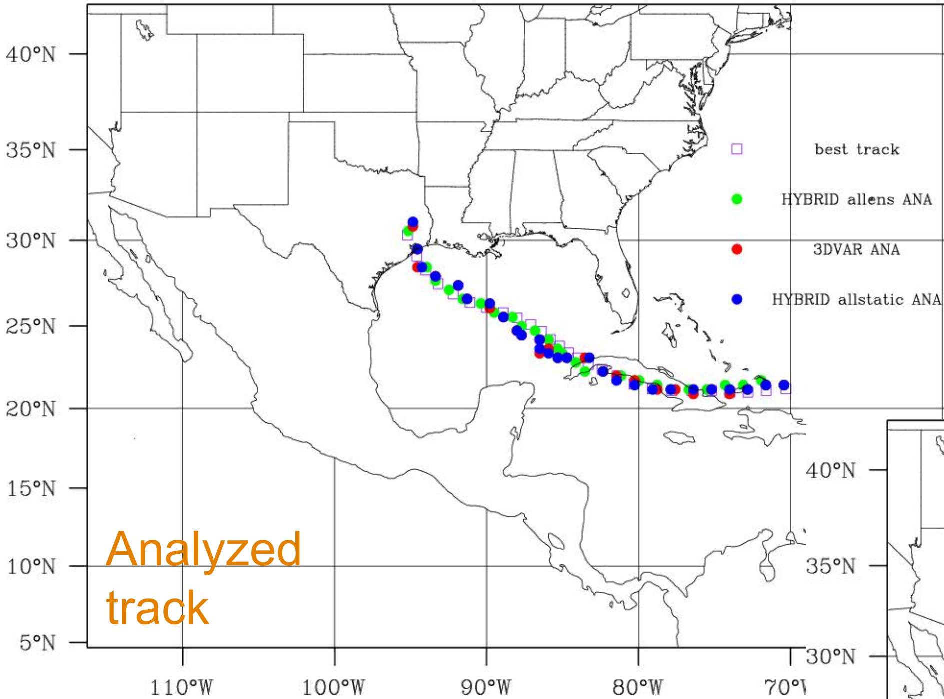
Avg. track forecast error



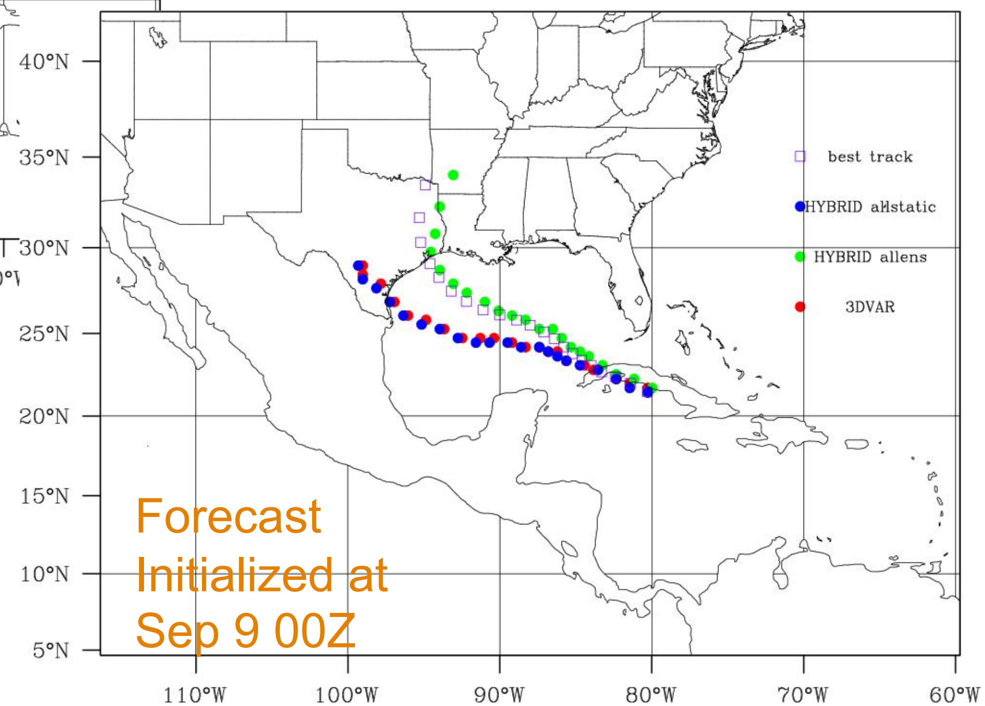
- Averaged track forecast initialized by analyses generated by the hybrid with ETKF ensemble covariance is more accurate than 3DVAR

- The advantage of hybrid relative to 3DVAR is mainly from the flow-dependent ensemble covariance, not ensemble averaging.

IKE track: analysis and forecast



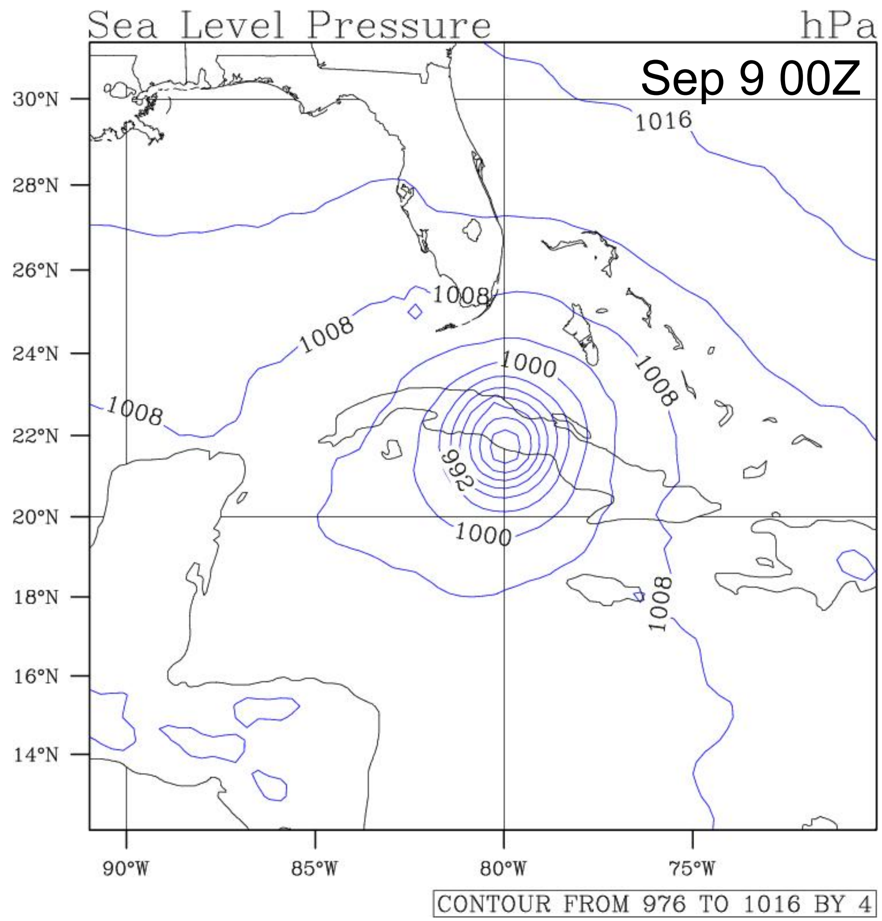
• Though analyzed tracks close, track forecast by hybrid with ETKF ensemble covariance is much better than 3DVAR.



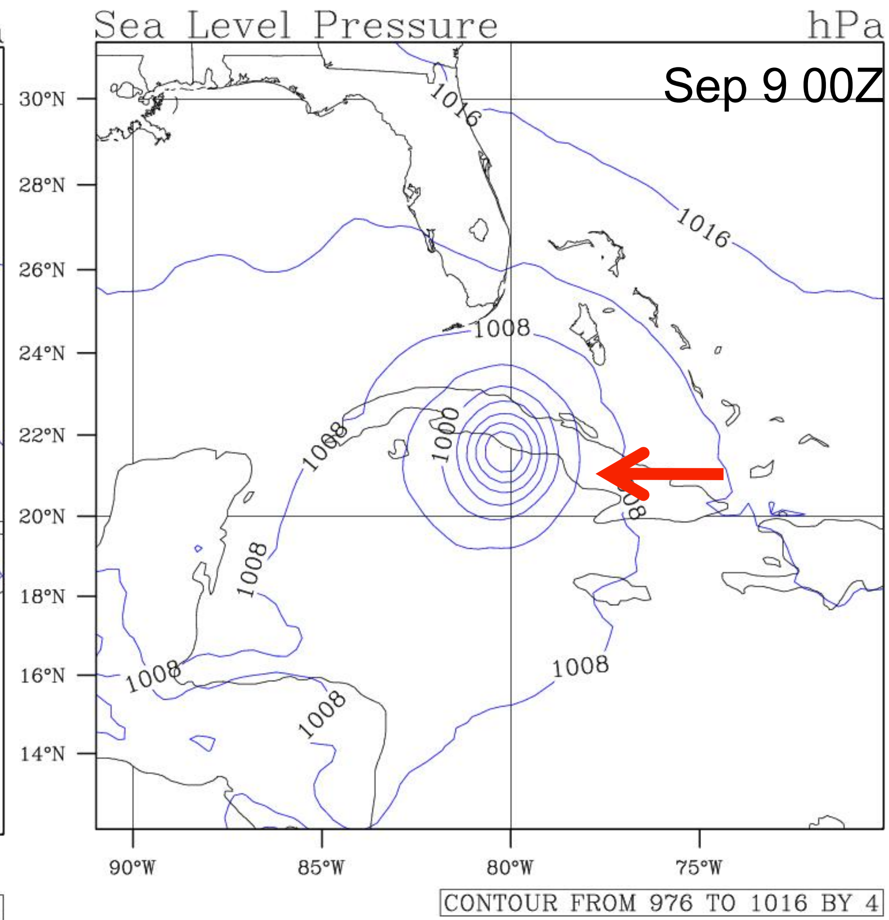
• Most of the time when 3dvar forecast performed poorly, it is biased toward the west.

What analyzed differently?

HYBRID (all ensemble)



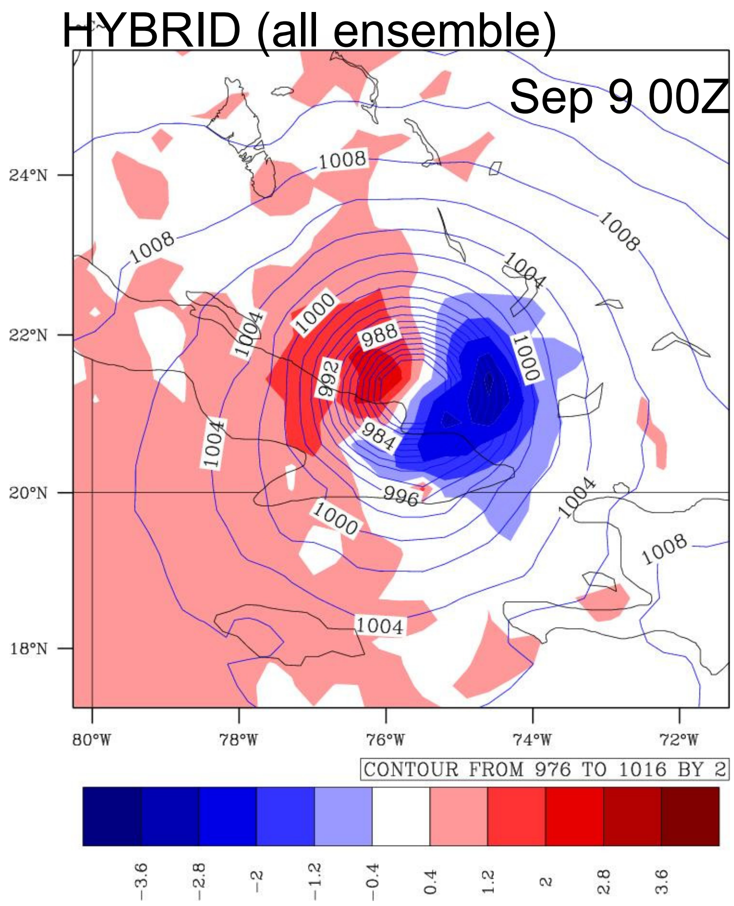
3DVAR

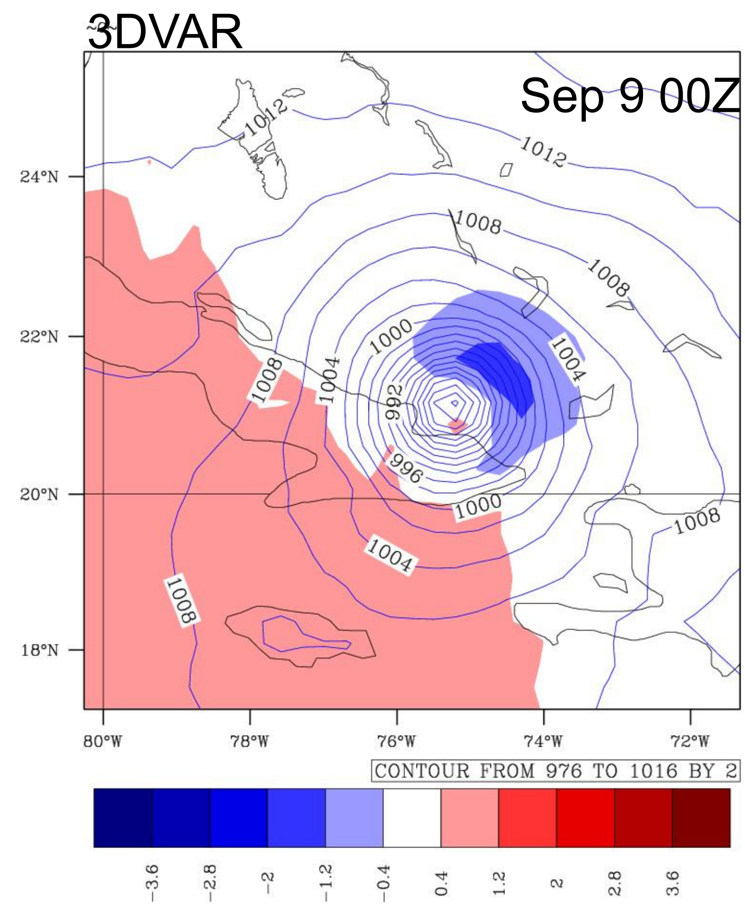


• Hybrid (all ensemble) analyzed TC is bigger and stronger.

• Stronger easterly in TC environment (e.g. 500mb) by 3DVAR.

SLP Increment differences

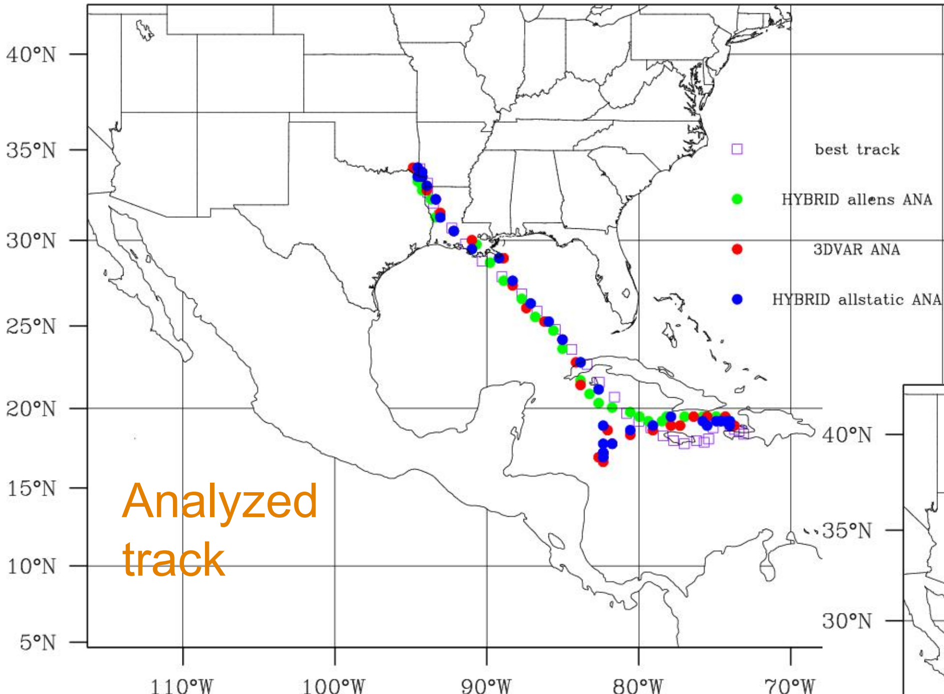




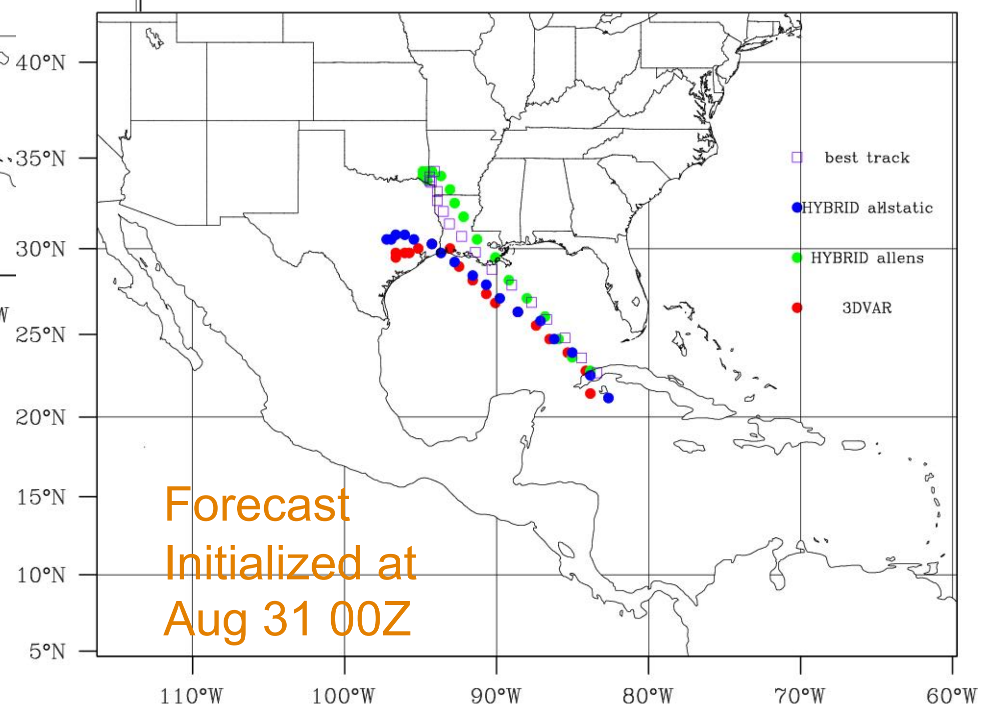
- Coherent position and intensity increment by hybrid

- Asymmetric/localized increment by 3DVAR

GUSTAV track: analysis and forecast

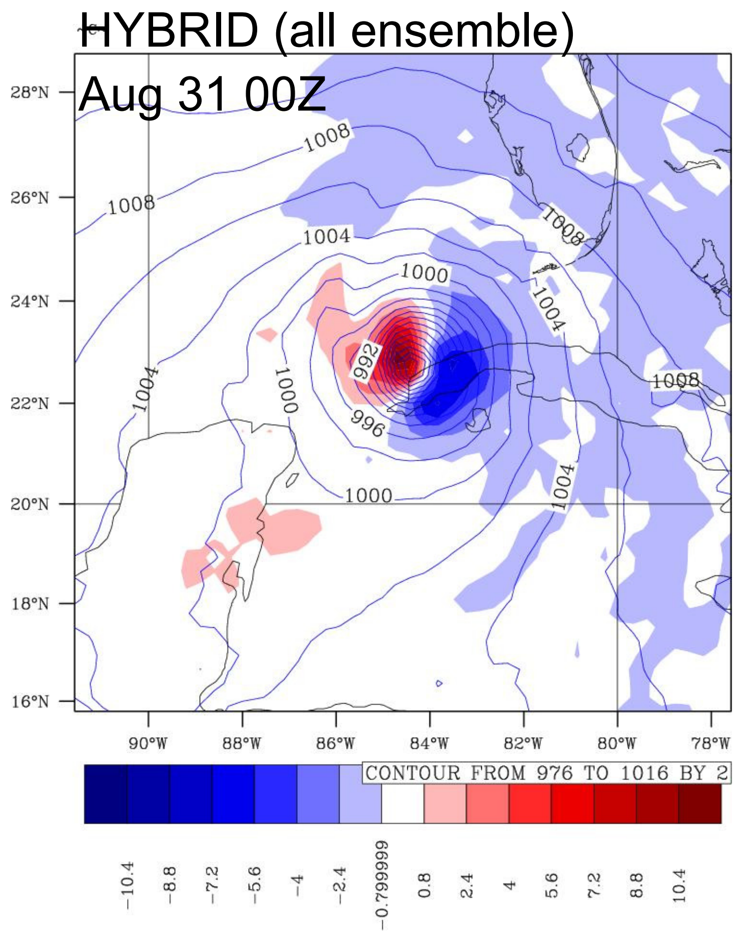


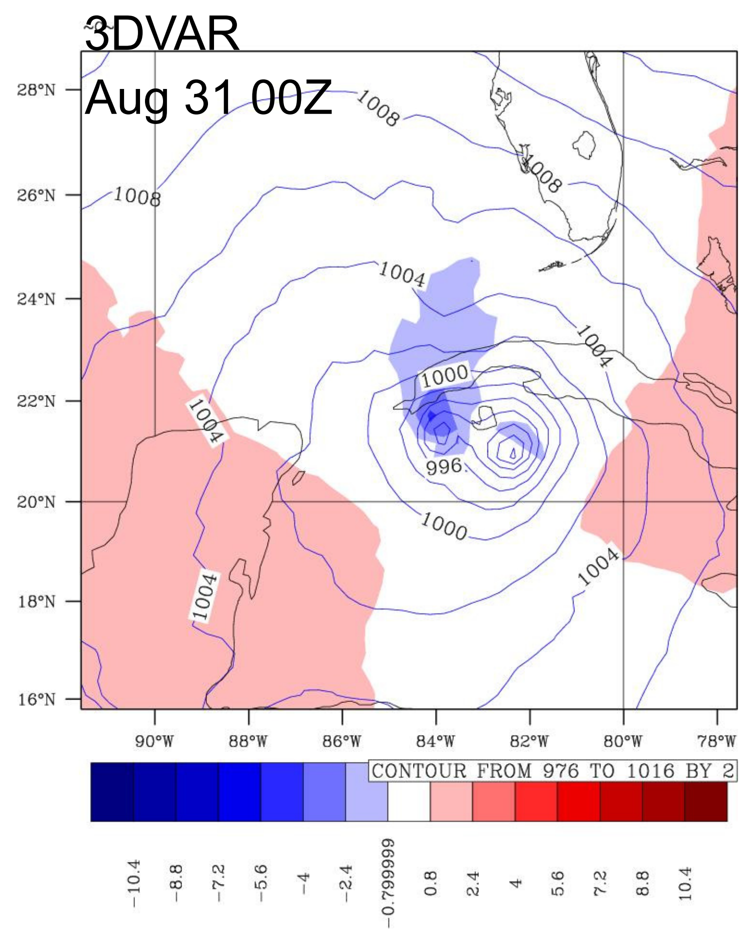
• Analyzed track by 3DVAR made a wrong loop before reaching Cat 4. 3DVAR track veered to the south before corrected by TC-specific obs.



• Track forecasts by 3DVAR were still worse than hybrid (e.g., westward bias) even the analyzed position started to catch the right place.

SLP Increment differences

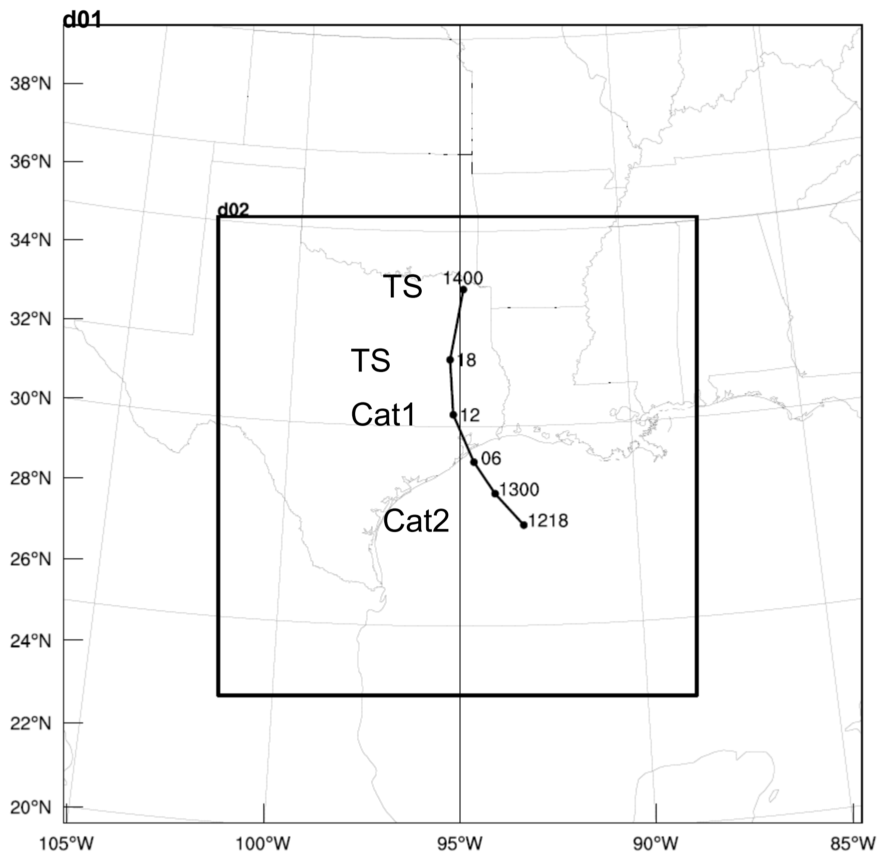




- Coherent position and intensity increment by hybrid

- Double-Vortex in 3DVAR (occurred during the spurious loop)
- Asymmetric/localized increment

Hurricane IKE 2008 radar data assimilation experiment setup



•Nested domains: Δx=15km (d01)/ 5km(d02)

•DA system: WRF VAR hybrid with ensemble generated by perturbed obs.

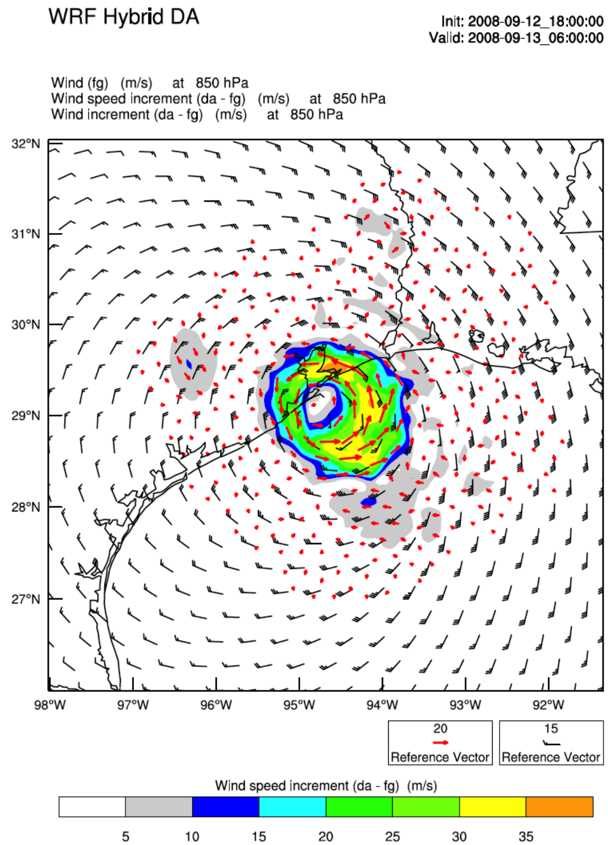
•Ensemble size: 40 members

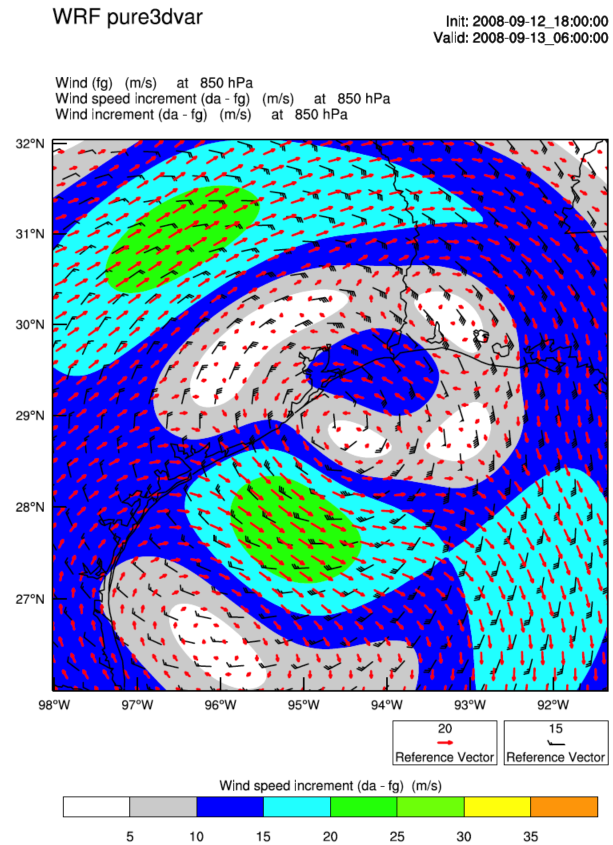
•DA and forecast:

- ➢12h ensemble initialized at 18ZSep12
- ➢Radar DA at 06ZSep13 and no cycling
- ➢18h deterministic forecast from 06ZSep13
- • 6h ensemble initialized at 18ZSep 12
- • DA cycling started 00ZSep13 every 30 min for 3 hrs
- • 21 hr deterministic forecast from 03ZSep13

•Observations: radial velocity from two WSR88D radars (KHGX, KLCH)

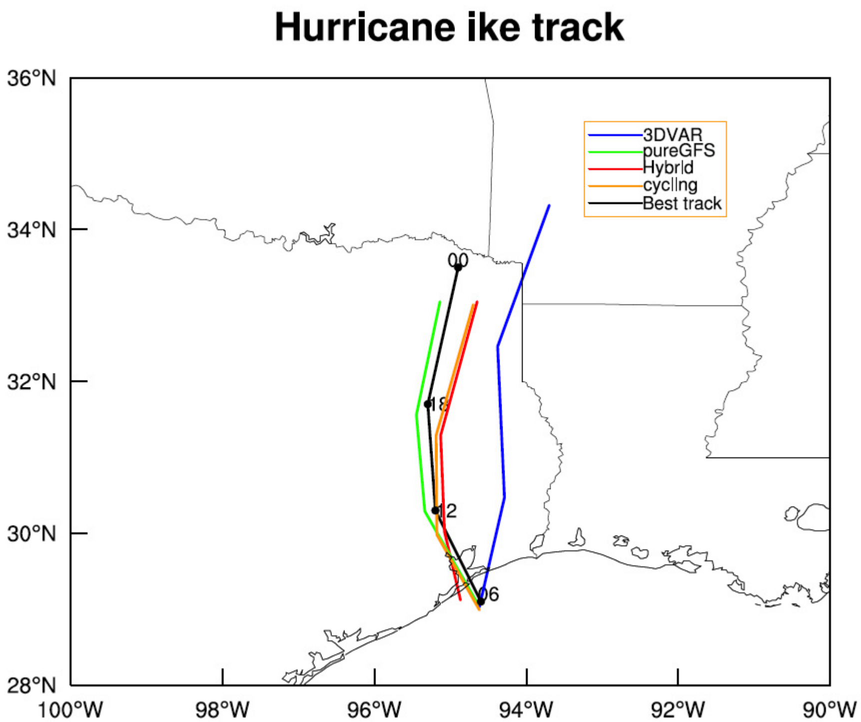
Analysis Increment: 850mb wind

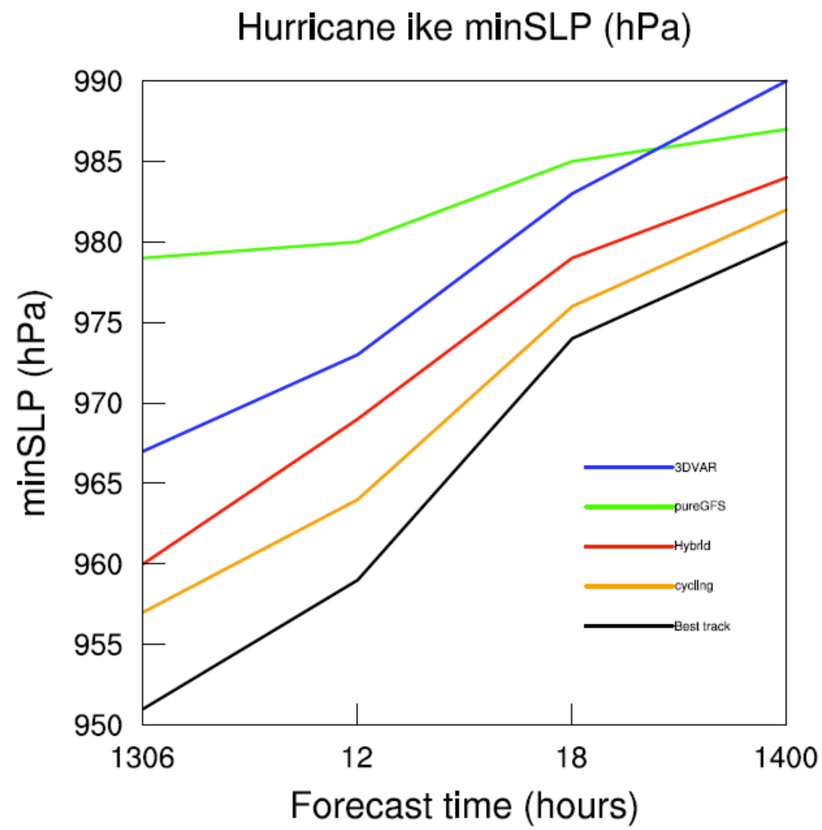




- Hybrid increment is around the eye of IKE, which suggested a stronger IKE than the first guess.
- 3DVAR increment is not recognizing IKE.

Track and Intensity forecast





- Track forecast by hybrid was better than WRF 3DVAR and similar to GFS.
- Hybrid analyzed and predicted a stronger IKE (closer to the best track) than both WRF 3DVAR and GFS.
- Further improvement by cycling in the hybrid (need to run 3DVAR cycling experiment).

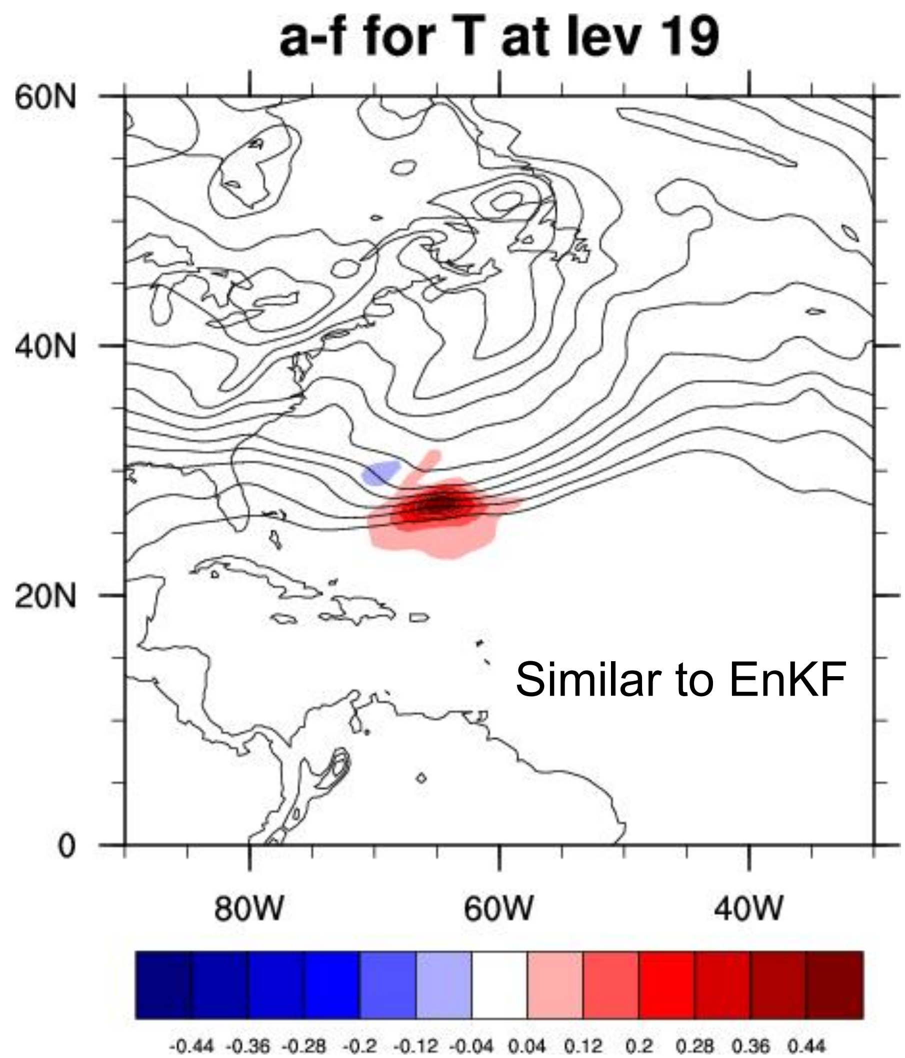
GSI based hybrid DA for GFS

• GSI based hybrid DA is being developed and tested using the extended control variable method (Wang 2010a)

• 1 ob tests (right) show ensemble covariance is correctly ingested

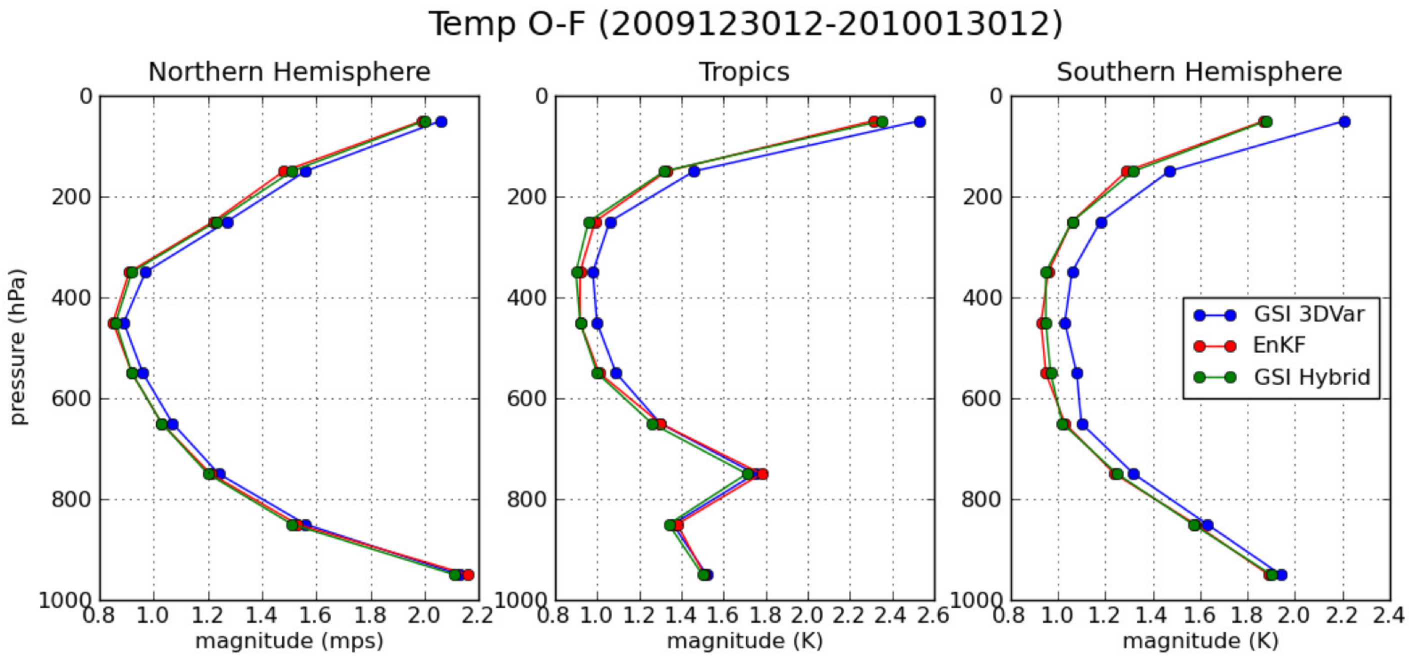
• Preliminary cycling experiment where GSI is **two-way** coupled with EnSRF were conducted.

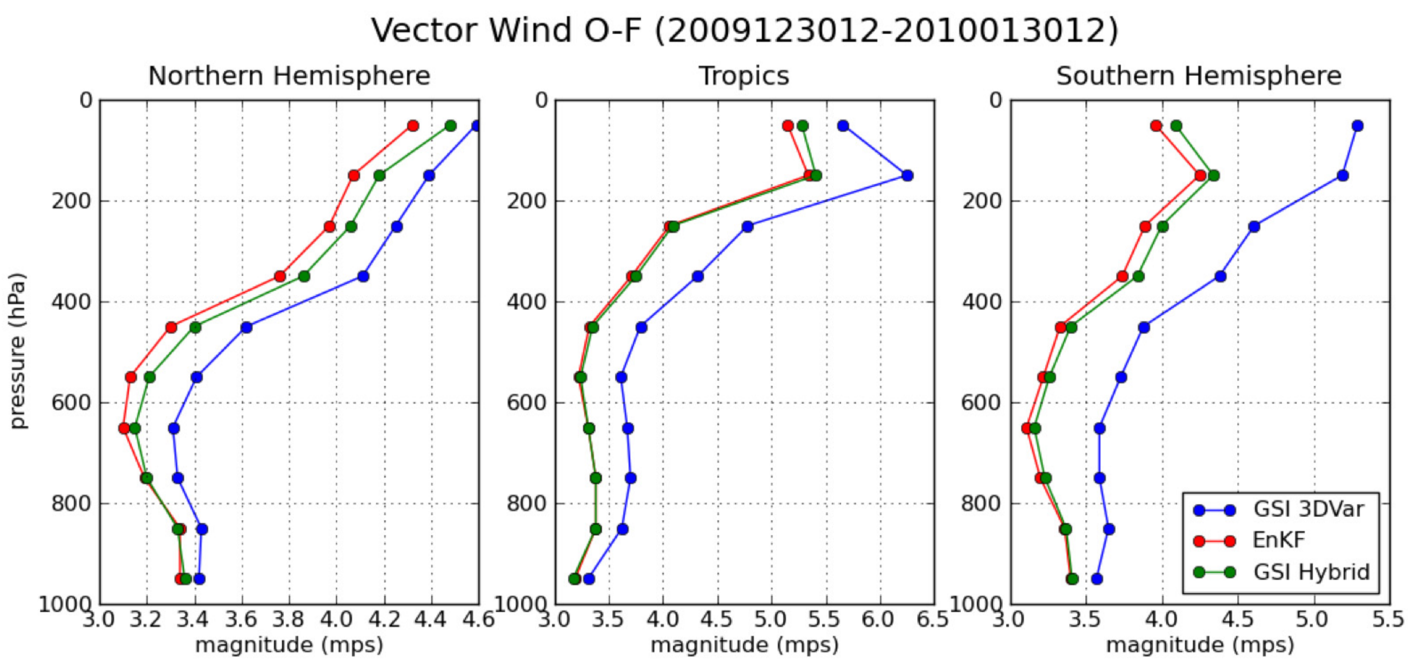
➢T190/L64 resolution, operational obs
➢Two-Way coupling:
•Mean of EnKF used as the background and ensemble covariance provided by EnKF
•EnKF recentered on the analysis by the hybrid
➢ Preliminary set up:
•Half static, half ensemble covariance
•localization chosen for hybrid was less tight than EnKF
•vertical localization in hybrid done in model grids, different from EnKF in scale height



• Hybrid is better than GSI.

• Experiments underway to understand the difference of hybrid and EnKF.





Conclusion and discussion

- Hybrid VAR-EnKF has been developed and applied for various scales assimilating various data and was demonstrated to improve the analyses and subsequent forecasts.

- Experiments have shown that flow-dependent covariance provided by the ensemble contributed to the better performance by the hybrid.

- Keep using WRF VAR and GSI based hybrid DA system (including both 3DVAR and 4DVAR) to test the hypothesis proposed early in the talk to understand the differences between VAR, EnKF and their hybrid for various scales.

References cited

Campbell, W. F., C. H. Bishop, D. Hodyss, 2009: Vertical Covariance Localization for Satellite Radiances in Ensemble Kalman Filters. *Mon. Wea. Rev*., in press.

James C. N., and R. A. Houze, 2001: A real—time four-dimensional Doppler dealiasing scheme. J. Atmos. Oceanic Tech., 18, 1647-1683.

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.

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013-1043.

Wang, X., and C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.*, **60**, 1140-1158.

Wang, X., C. H. Bishop, and Simon J. Julier, 2004: Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble? *Mon. Wea. Rev.*, **132**, 1590-1605.

Wang, X., C. Snyder, and T. M. Hamill, 2007a: On the theoretical equivalence of differently proposed ensemble/3D-Var hybrid analysis schemes. *Mon. Wea. Rev.*, **135**, 222-227.

Wang, X., T. M. Hamill, J. S. Whitaker and C. H. Bishop, 2007b: A comparison of hybrid ensemble transform Kalman filter-OI and ensemble square-root filter analysis schemes. *Mon. Wea. Rev.*, **135**, 1055-1076.

Wang, X., D. Barker, C. Snyder, T. M. Hamill, 2008a: A hybrid ETKF-3DVar data assimilation scheme for the WRF model. Part I: observing system simulation experiment. *Mon. Wea. Rev.*, **136**, 5116-5131.

Wang, X., D. Barker, C. Snyder, T. M. Hamill, 2008b: A hybrid ETKF-3DVar data assimilation scheme for the WRF model. Part II: real observation experiments. *Mon. Wea. Rev.*, **136**, 5132-5147.

Wang, X., T. M. Hamill, J. S. Whitaker, C. H. Bishop, 2009: A comparison of the hybrid and EnSRF analysis schemes in the presence of model error due to unresolved scales. *Mon. Wea. Rev.*, **137**, 3219-3232.

Wang, X., 2010a: Incorporating ensemble covariance in the Gridpoint Statistical Interpolation (GSI) variational minimization: a mathematical framework. *Mon. Wea. Rev.*, in press.

Wang, X., 2010b: Impact of flow-dependent ensemble covariance for hurricane track forecast: a study using the hybrid ETKF-3DVAR data assimilation system for WRF. *Mon. Wea. Rev*.

Hybrid Data Assimilation Theory

- Ensemble covariance is included in the VAR cost function through augmentation of control variables (Lorenc 2003, Buehner 2005, Wang et al. 2007a, 2008a, Wang 2010a).

Extra term associated with extended control variable

$$J(\mathbf{x}_1^{'}, \boldsymbol{\alpha}) = \beta_1 J_1 + \beta_2 J_e + J_o$$

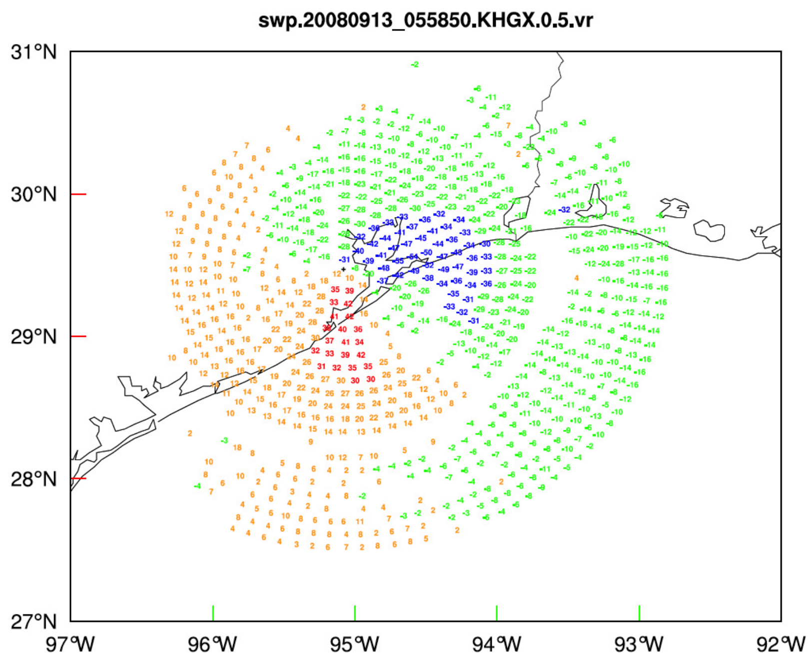
$$= \beta_1 \frac{1}{2} \mathbf{x}_1^{'T} \mathbf{B}^{-1} \mathbf{x}_1^{'} + \underline{\beta_2 \frac{1}{2} \boldsymbol{\alpha}^T \mathbf{C}^{-1} \boldsymbol{\alpha}} + \frac{1}{2} \left(\mathbf{y}^{o'} - \mathbf{H}\mathbf{x}^{'}\right)^T \mathbf{R}^{-1} \left(\mathbf{y}^{o'} - \mathbf{H}\mathbf{x}^{'}\right)$$

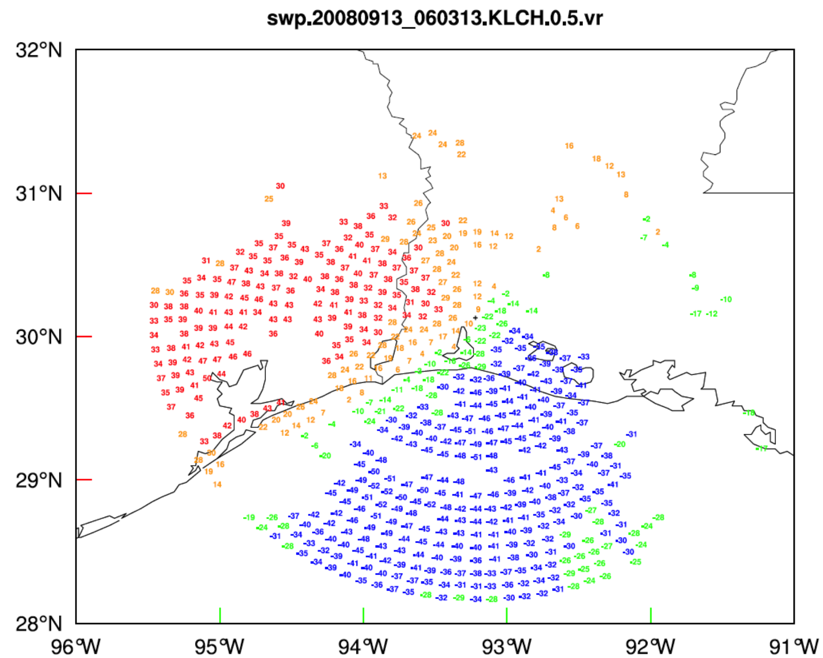
$$\mathbf{x}^{'} = \mathbf{x}_1^{'} + \underline{\sum_{k=1}^{K} \left(\boldsymbol{\alpha}_k \circ \mathbf{x}_k^e\right)}$$

Extra increment associated with ensemble

$\mathbf{B}$ 3DVAR static covariance; $\mathbf{R}$ observation error covariance; K ensemble size;
$\mathbf{C}$ correlation matrix for ensemble covariance localization; $\mathbf{x}_k^e$ kth ensemble perturbation;
$\mathbf{x}_1^{'}$ 3DVAR increment; $\mathbf{x}^{'}$ total (hybrid) increment; $\mathbf{y}^{o'}$ innovation vector;
$\mathbf{H}$ linearized observation operator; β_1 weighting coefficient for static covariance;
β_2 weighting coefficient for ensemble covariance; $\boldsymbol{\alpha}$ extended control variable.

Radar data: Vr

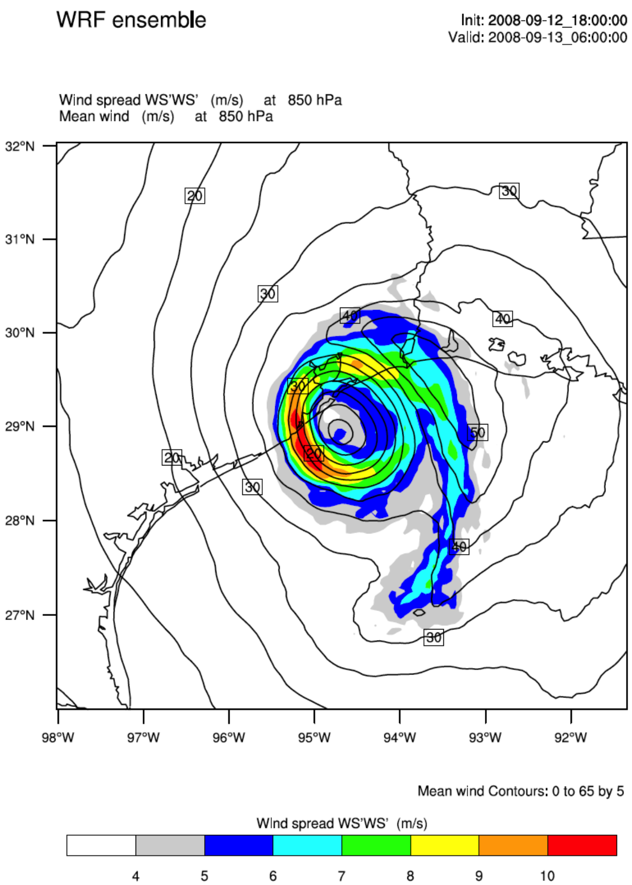


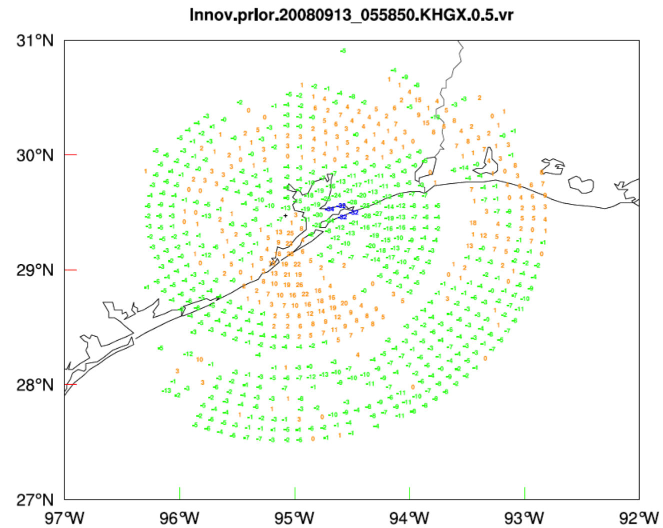


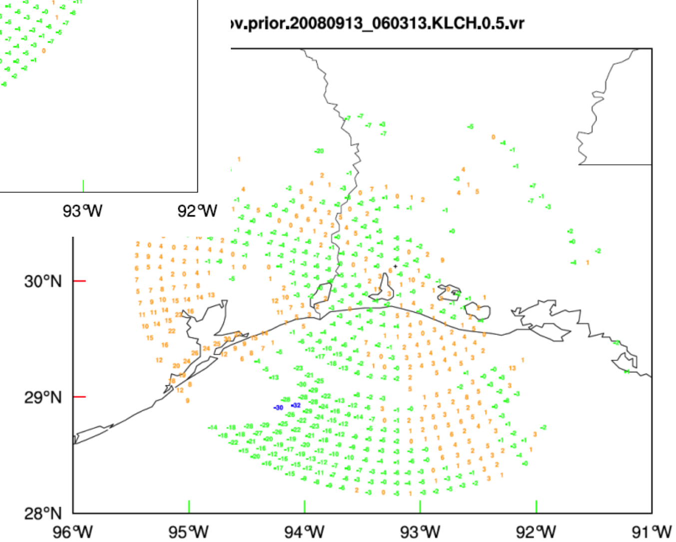
Preprocessing:

- Raw data: WSR88D Level II
- Use wind profile based on RAOB or GFS grid data to create the background
- De-aliasing using a modified version of Four-Dimensional Doppler Dealiasing Scheme (4DD) (James and Houze, 2001).
- Thinning: 500m in vertical and 5-10 km in horizontal

Ensemble Spread: 850mb wind



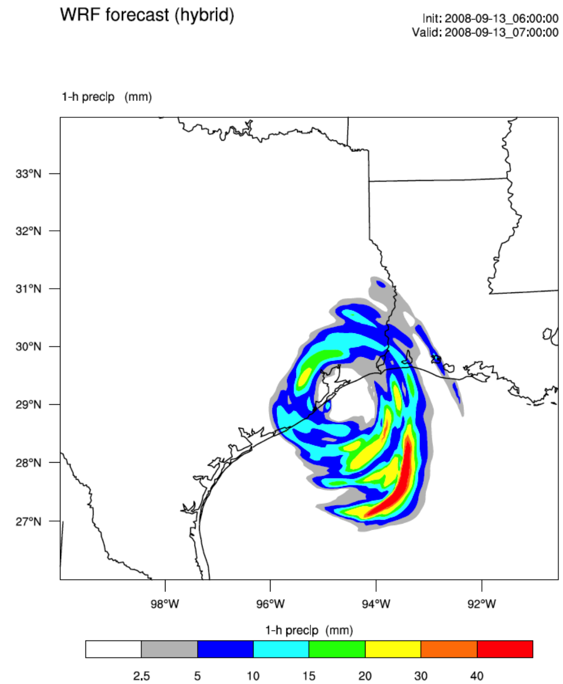


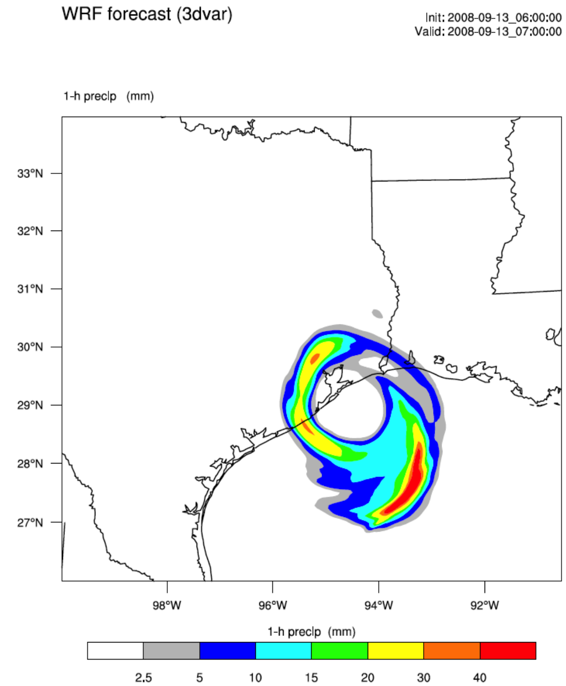


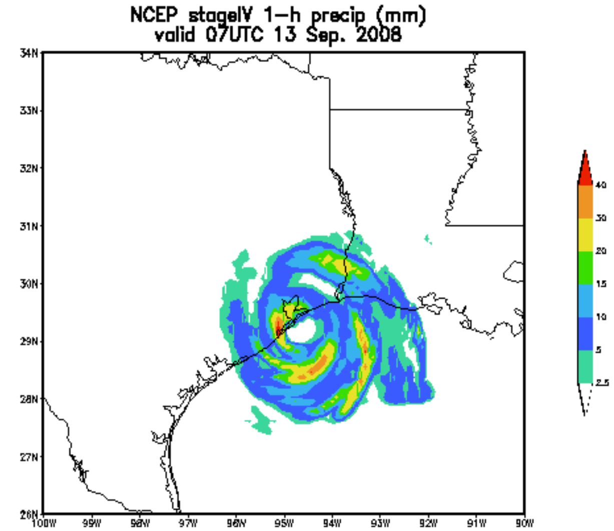
- Large ensemble spread at maximum wind speed gradient in the first guess around the eye

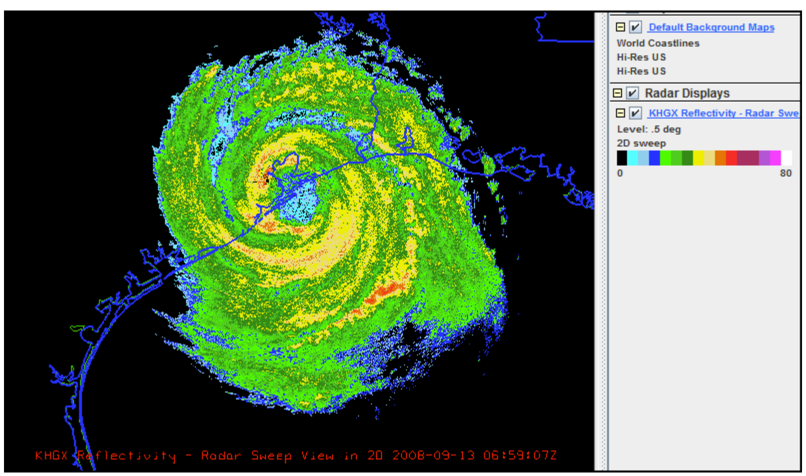
- The ring of the ensemble spread is associated with relatively large innovation, which suggests the spread can distinguish the first guess errors around the eye from other places.

1h Precipitation forecast









- Precipitation forecast by the hybrid revealed more detailed structure of the rain-band than 3DVAR.