Improving Global and Hurricane Predictions by Using Minimum-Cost Large Ensembles in GFS 4DEnVar Hybrid Data Assimilation System

Xuguang Wang, Bo Huang, Junkyung Kay

School of Meteorology
University of Oklahoma, Norman, OK, USA

Collaborators:

Rahul Mahajan (NOAA/EMC), Lili Lei (NOAA/ESRL), Daryl Kleist (UMD), Jeff Whitaker (NOAA/ESRL), John Derber (NOAA/EMC), Yuejian Zhu (NOAA/EMC)

NGGPS PI meeting, August 2-5 2016

Outline

- Background
- Methods
- Experiment design
- Results
- Summary and future work

Outline

- Background
- Methods
- Experiment design
- Results
- Summary and future work

Ensemble-Variational (EnVar) Hybrid Data Assimilation

•GSI-based 3DEnVar hybrid was operationally implemented for GFS at NCEP since 2012. GSI-based 4DEnVar hybrid is implemented operationally for GFS at NCEP May 2016. Significant improvement was found for global forecasts (e.g., Wang et al. 2013, Wang and Lei 2014; Kleist and Ide 2015; Mahajan et al. 2016) .

•Use of ensemble covariances allows estimating spatial, temporal and multivariate error covariances in a flow-dependent manner.

•Although flow-dependent, ensemble covariances may still have issues due to

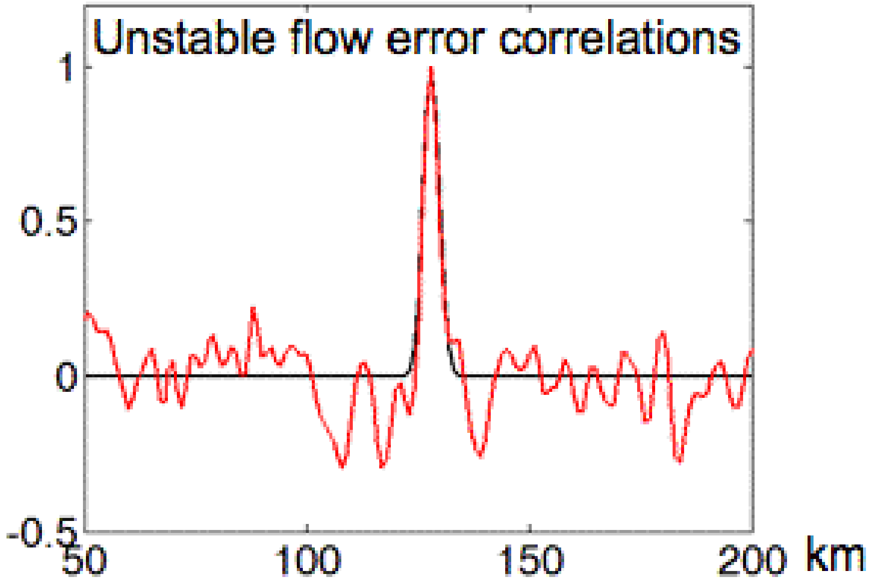
- ❑ Sampling error
- ❑ Model error

Sampling Error

- Caused by the use of a limited number of ensemble members

- An example: spurious correlation at distant locations

- Can lead to “filter divergence” if not treated



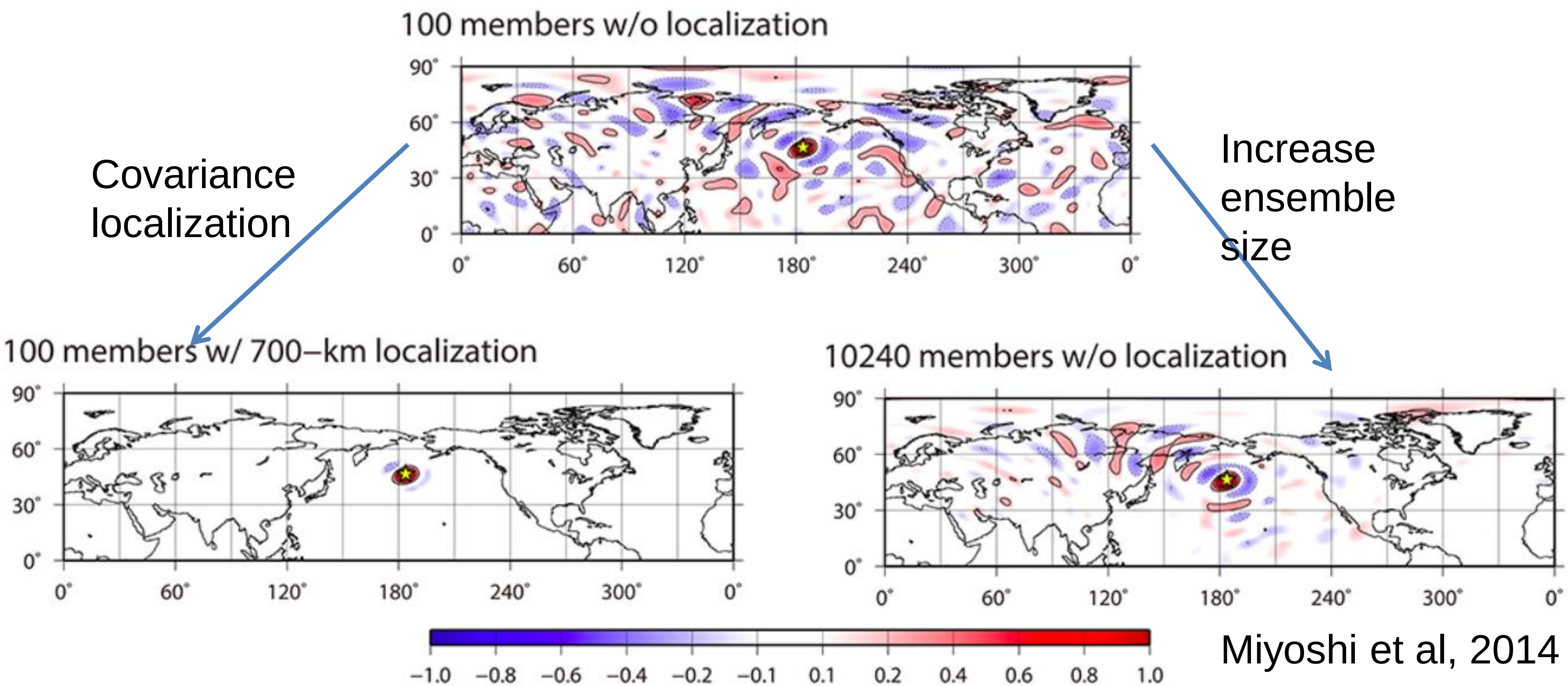
- Treatments:
  - ❑ Covariance localization
    - ○ Distance based localization (fixed or adaptive); scale-dependent localization; variable localization etc.

    Pros: reduce spurious correlation and increase the degree of freedom.
    Cons: may eliminate the remote realistic signal and/or incur additional imbalance

  - ❑ Increase ensemble size

    Pros: Capable of simulating the remote realistic signal, increasing the degree of freedom and alleviating the imbalance in the analysis;
    Cons: significant cost increase

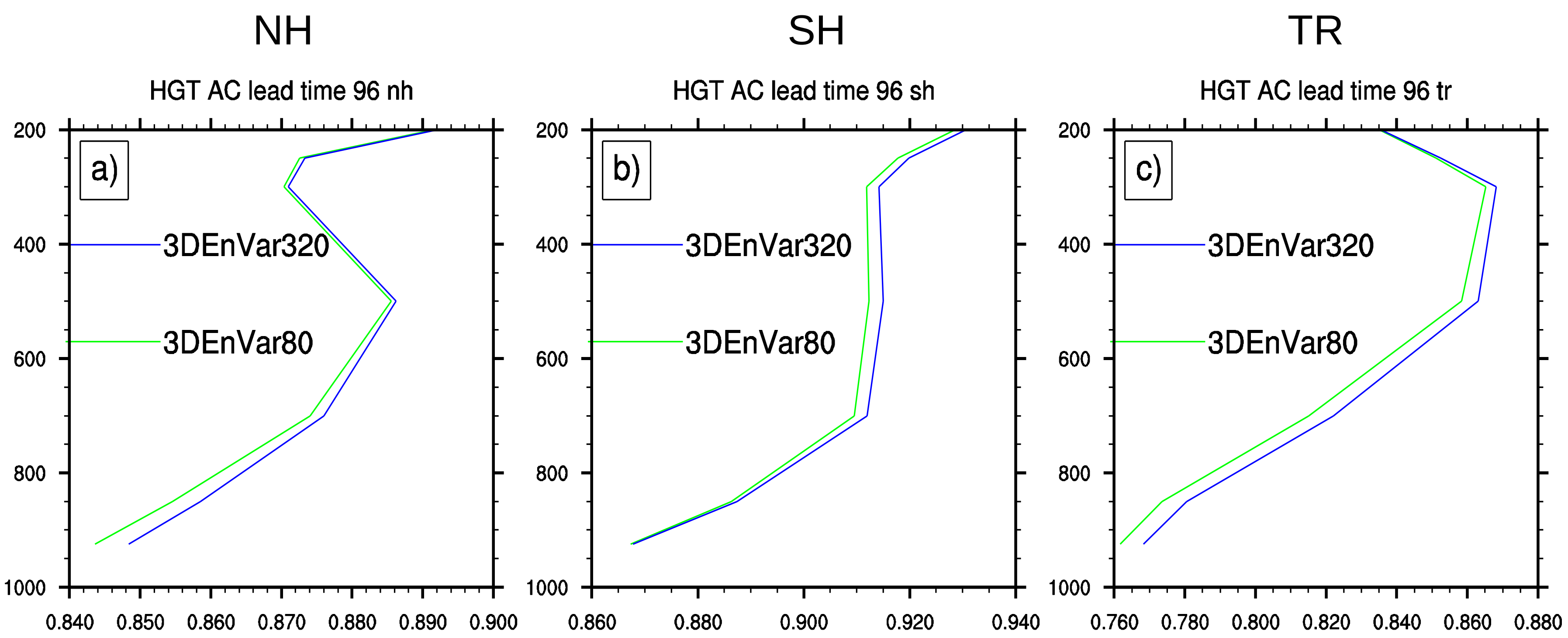
Sampling Error
Covariance localization vs increasing ensemble size



- Large-sized ensemble reveals long-range error covariance at the continental scales, while localization will remove this signal.
- Extremely expensive computations are required for large-sized ensemble.

Impact of Increasing Ensemble Size in GFS Hybrid 3DEnVar System

- Hybrid 3DEnVar in GFS was further improved by directly increasing the original 80 members to 320 members (Lei and Wang, 2016). Similar is found for 4DEnVar (Lei and Whitaker talk, EnKF workshop, 2016).



Objectives

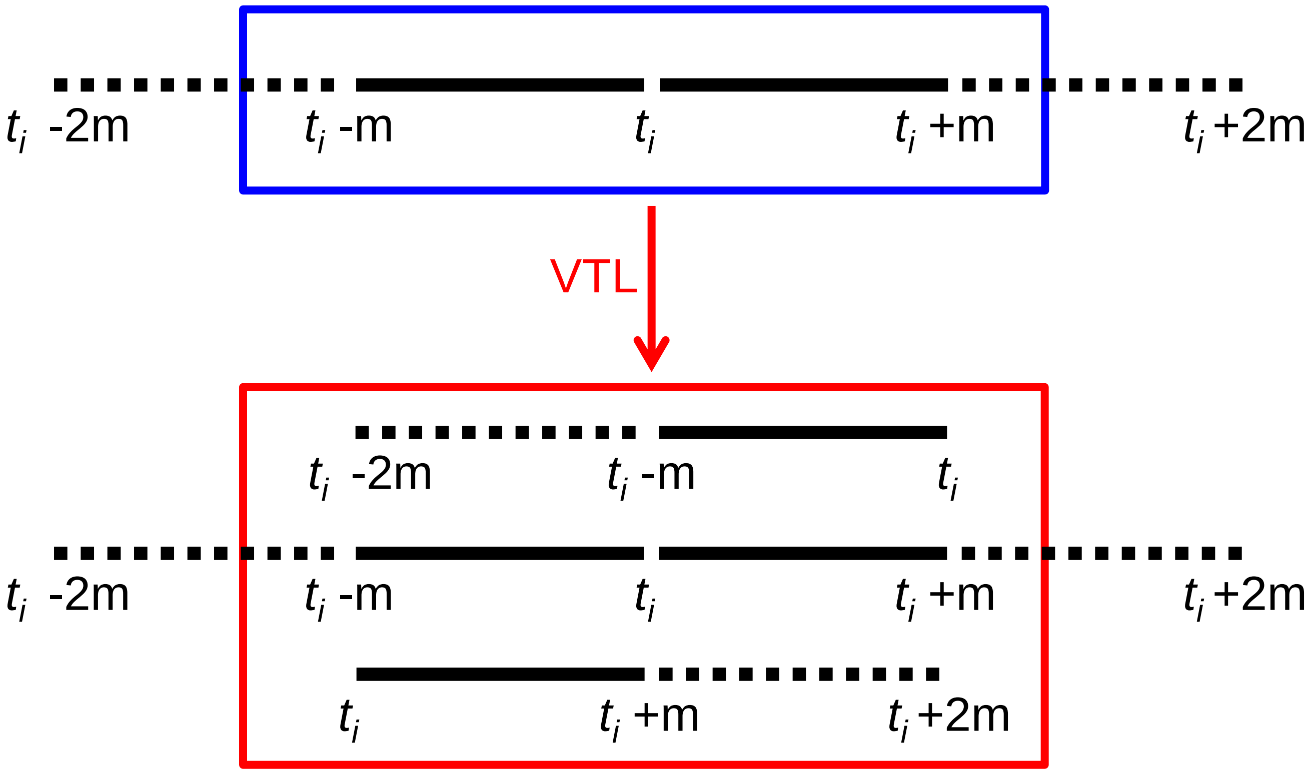
- Aside from increasing ensemble members directly, is there a cheaper way to increase the ensemble size while still improving the analysis and forecast?
- To what extent would such method help? Why is it helpful?
- What is the best way of using such cheaply generated ensemble?
- What is the best way of using ensemble resolving different range of scales?

Outline

- Background
- Methods
- Experiment design
- Results
- Summary and future work

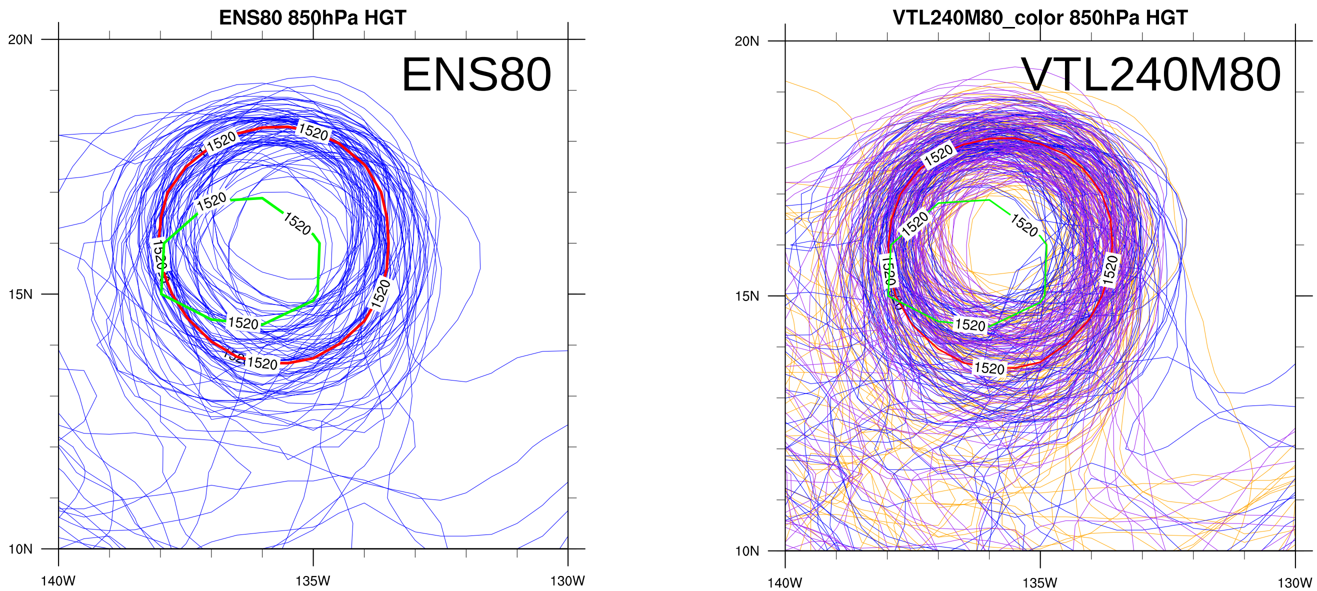
Valid Time Lagging method (VTL)

- Shift the the ensemble perturbations at *m* hours before and after t_i to t_i. E.g., m=3, then ensemble size will be tripled in current GFS 4DEnVar.



Why would VTL help?

- May more completely represent the forecasts errors related to the location, structure and orientation so that truth acts like one of the ensemble members.



80-mem 6h fcst

Ensemble mean from 80-mem 6h fcst

EC analysis

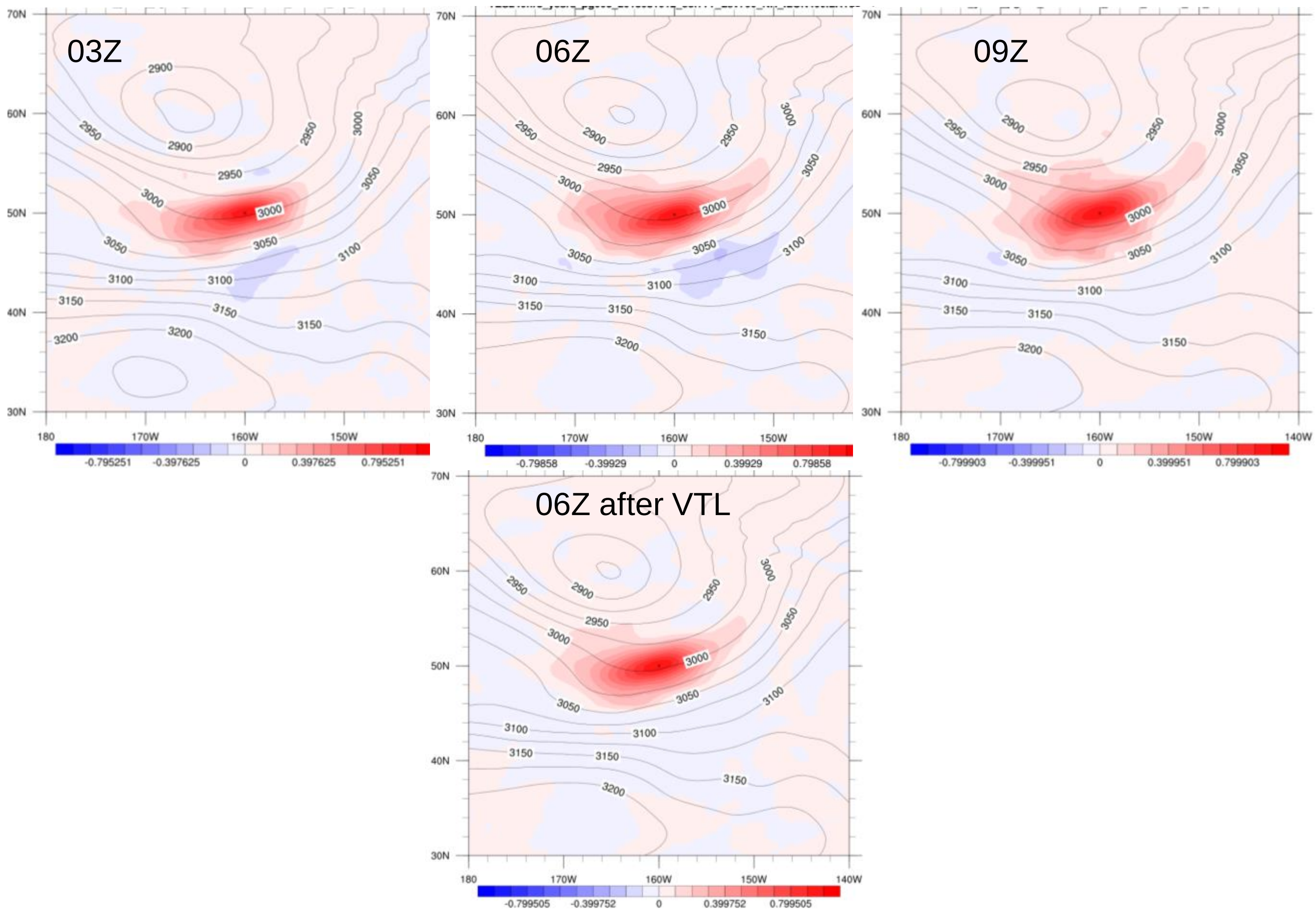
80-mem 3h perturbations plus 6h ensemble mean

80-mem 9h perturbations plus 6h ensemble mean

Hurricane Henriette (2013)

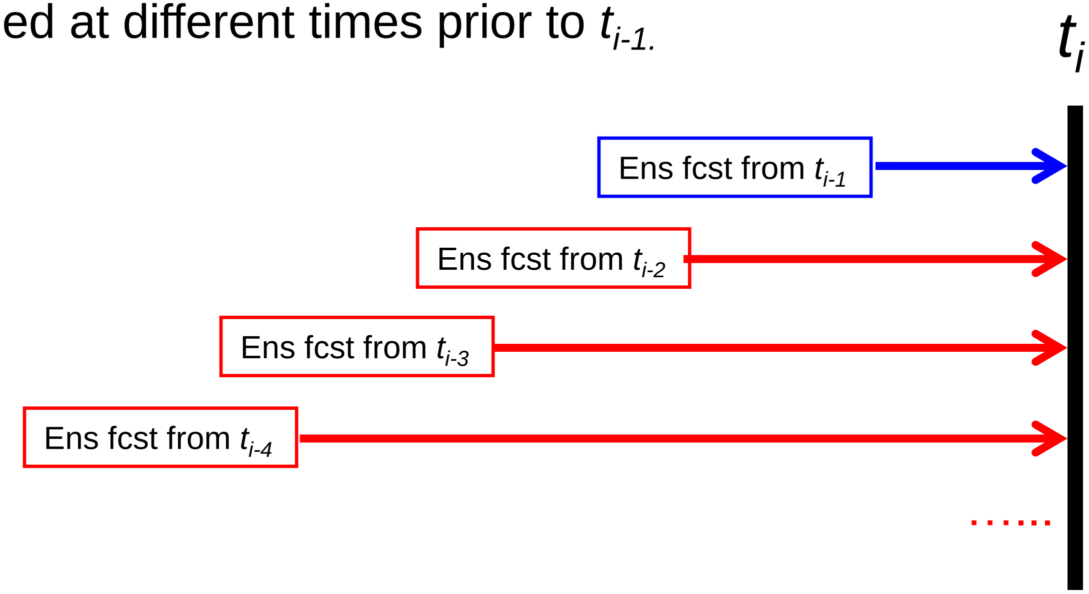
Why would VTL help ?

- VTL by design has smoothing effect.



Initial Time Lagging method (ITL)

- Increasing ensemble size by collecting ensemble perturbations valid at t_i but initialized at different times prior to t_{i-1}.



- May better represent the model error due to the use of forecasts with different leading times.

- Forecasts with different leading times may need to be assigned with different weights.

Outline

- Background
- Methods
- Experiment design
- Results
- Summary and future work

Experiment Design
GSI-4DEnVar for GFS

- **GSI-4DEnVar**: Naturally extended from and unified with GSI-based 3DEnVar hybrid formula. Conveniently avoid TLA.

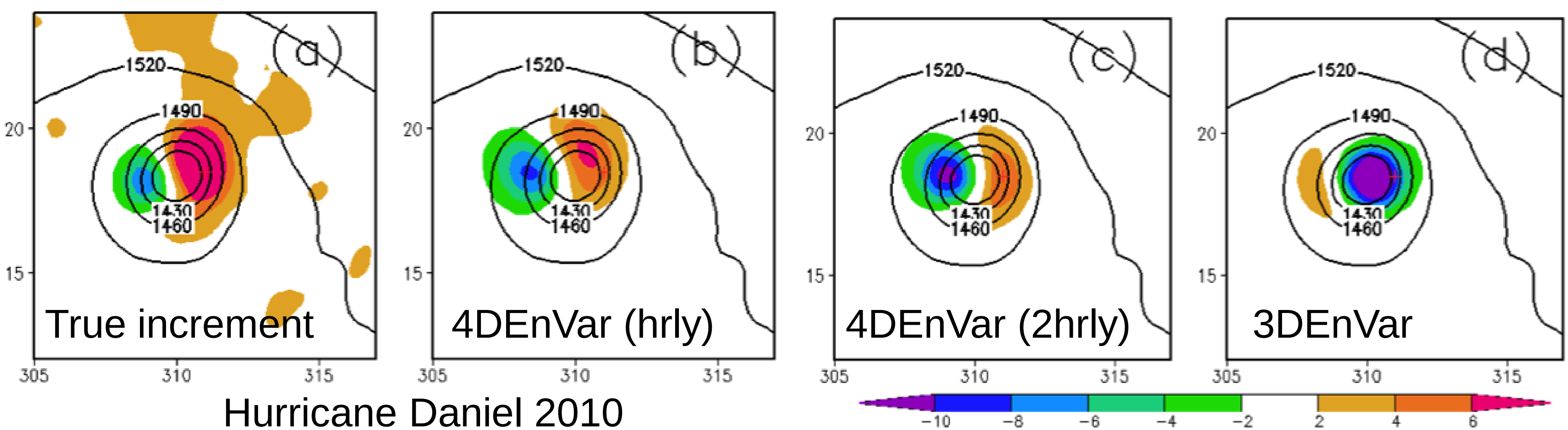
Add time dimension in 4DEnVar

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

$$=\beta_1\frac{1}{2}\mathbf{x}_1^{'T}\mathbf{B}_{static}^{-1}\mathbf{x}_1^{'}+\beta_2\frac{1}{2}\boldsymbol{\alpha}^T\mathbf{C}^{-1}\boldsymbol{\alpha}+\boxed{\frac{1}{2}\sum_{t=1}^{T}(y_t^{o}{'}\text{-}\mathrm{H}_t\mathbf{x}_t)^{\mathrm{T}}\mathrm{R}_t^{-1}(y_t^{o}{'}\text{-}\mathrm{H}_t\mathbf{x}_t)}$$

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

Wang and Lei 2014, MWR



Hurricane Daniel 2010

Experiment Design

Experiment Settings for Baseline Experiment (ENS80)

- 4DEnVar using 80 member ensemble.
- T670/T254 GFS model.
- 6-hourly assimilation of all operational conventional and satellite data.
- 12.5% weight on the static background error covariance and 87.5% weight on the ensemble covariance.
- Three-hourly ensemble perturbations.
- Level-dependent localization length scales.
- Multiplicative inflation with a relaxation coefficient 0.85 and stochastic parameterizations for the spread issue.
- Four-dimensional incremental analysis update (4DIAU) and tangent linear normal mode constraint (TLNMC) to alleviate the imbalance issue.
- Cycling experiment for Aug. 2013 and verify both global forecast and hurricane track forecast.

Experiment Design

Lagged Ensemble Experiment Design and Estimated Cost

Exps.	Total ensemble perts for 4DEnVar at t_i	Total mem for EnKF at t_i	Ensemble Fcsts at t_i	Cost estimate (rel. to ENS80)	Required storage (rel. to ENS80)
ENS80	80-mem perts from fcsts valid at t_i but initialized from t_{i-1}	80	80-mem 9-hour fcsts	1	1
ENS240	240-mem perts from fcsts valid at t_i but initialized from t_{i-1}	240	240-mem 9-hour fcsts	3	3
ENS320	320-mem perts from fcsts valid at t_i but initialized from t_{i-1}	320	320-mem 9-hour fcsts	4	4
VTL240M80	240-mem perts by shifting the 80-member perts. at the time 3-hour before and after t_i to t_i	80	80-mem 12-hour fcsts	1.1	1.3
ITL320M80	320-mem perts by adding additional 3 groups of 80-mem perts from fcsts valid at t_i but initialized from the analysis at t_{i-2}, t_{i-3} and t_{i-4}	80	80-mem 27-hour fcsts	2.05	2.5

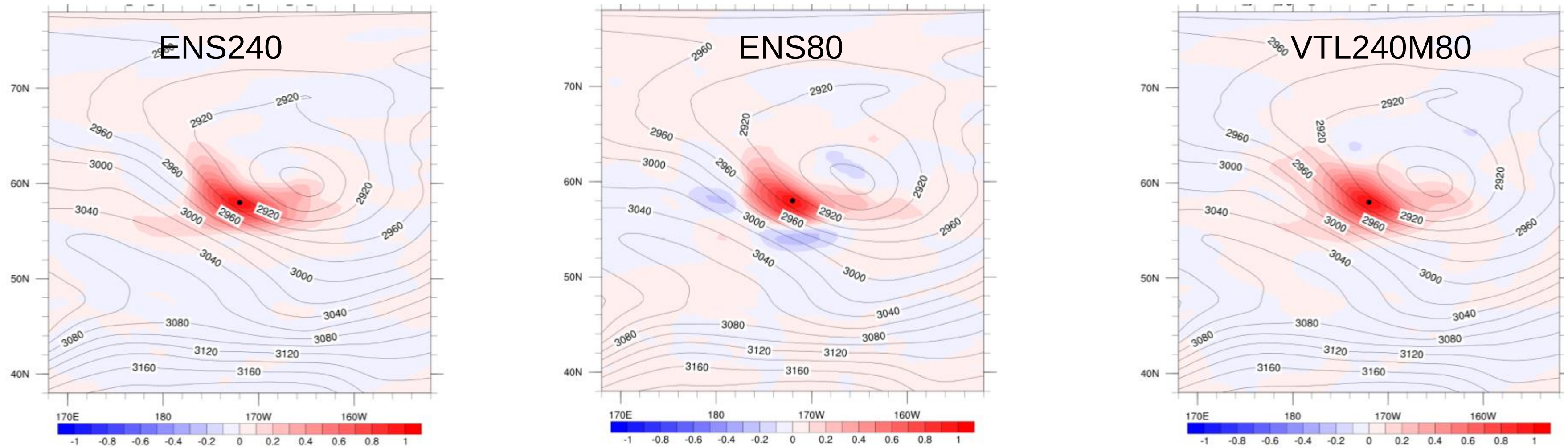
Outline

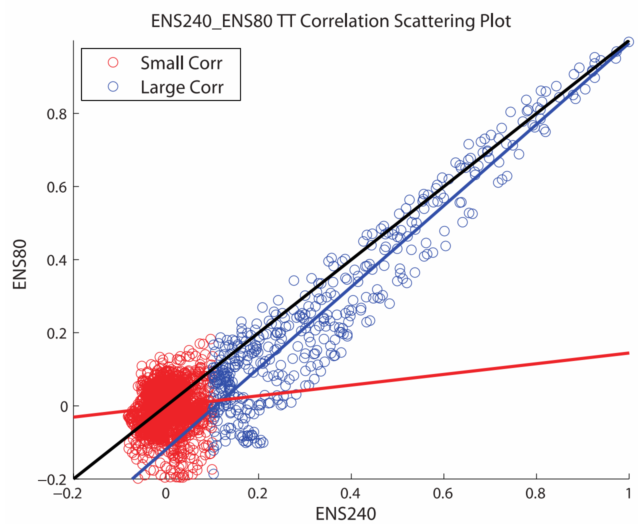
- Background
- Methods
- Experiment design
- Results
- Summary and future work

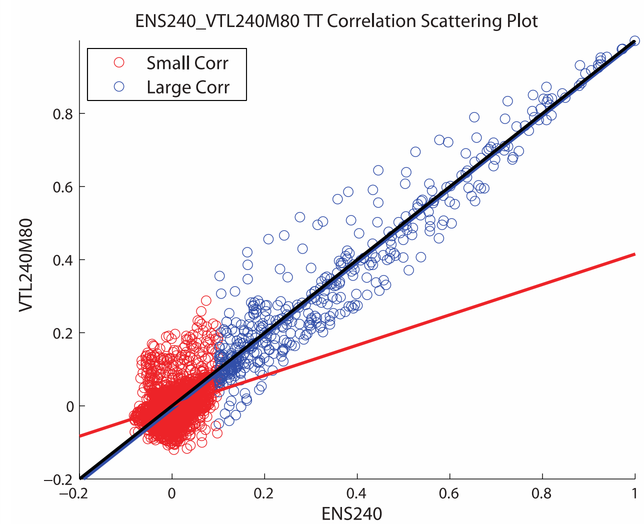
VTL Experiments

Exps.	Total ensemble perts for 4DEnVar at t_i	Total mem for EnKF at t_i	Ensemble Fcsts at t_i	Cost estimate (rel. to ENS80)	Required storage (rel. to ENS80)
ENS80	80-mem perts from fcsts valid at t_i but initialized from t_{i-1}	80	80-mem 9-hour fcsts	1	1
ENS240	240-mem perts from fcsts valid at t_i but initialized from t_{i-1}	240	240-mem 9-hour fcsts	3	3
VTL240M80	240-mem perts by shifting the 80-member perts. at the time 3-hour before and after t_i to t_i	80	80-mem 12-hour fcsts	1.1	1.3

VTL method
Self-correlation evaluation (TT at 700hpa)

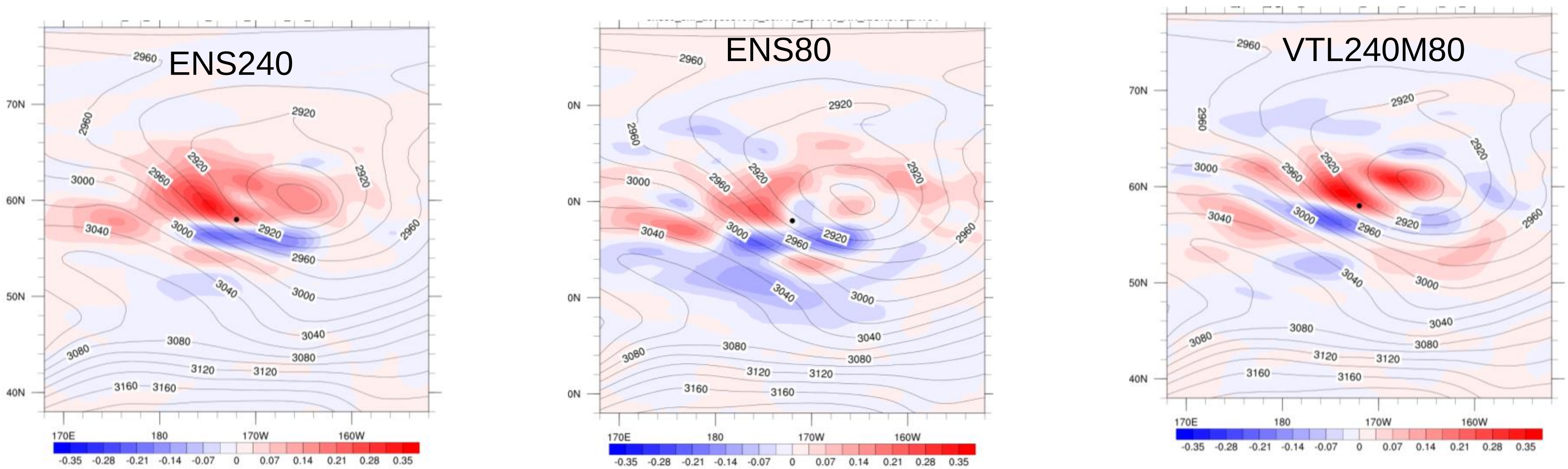




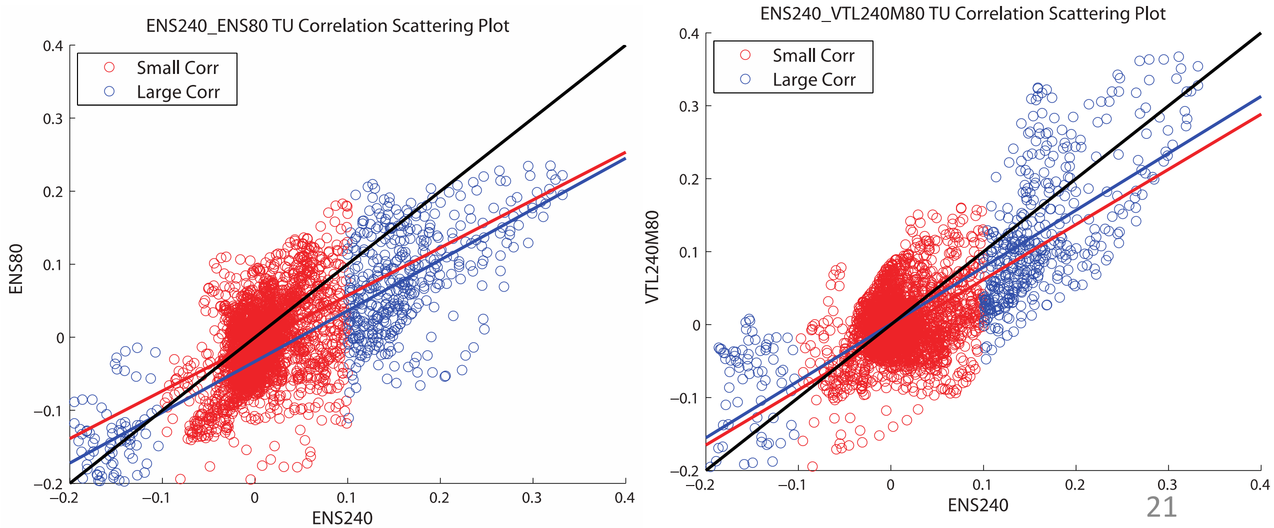


- VTL240M80 is qualitatively closer to ENS240 than ENS80.
- Scattering plots show that both the small ([0, 0.1], red) and large((0.1 1.0]) correlations of VTL240M80 are improved over ENS80.

VTL method
Cross-correlation evaluation (TU at 700hPa)



- VTL240M80 is qualitatively closer to ENS240 than ENS80.
- Scattering plots show that both the small ([0, 0.1], red) and large((0.1 1.0]) correlations of VTL240M80 are improved over ENS80.



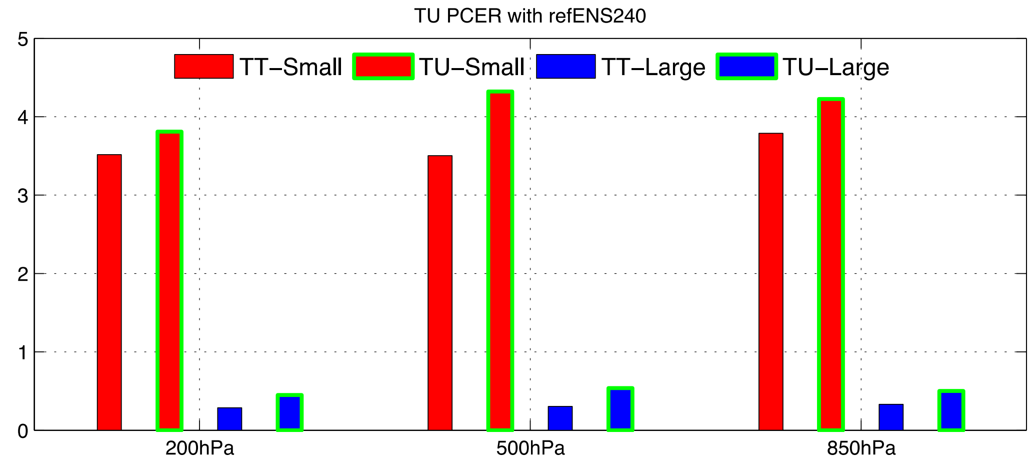
VTL method
Systematic evaluation of ENS80 correlation errors

- **Relative correlation errors (RCE)** of ENS80 was defined by comparing with ENS240.

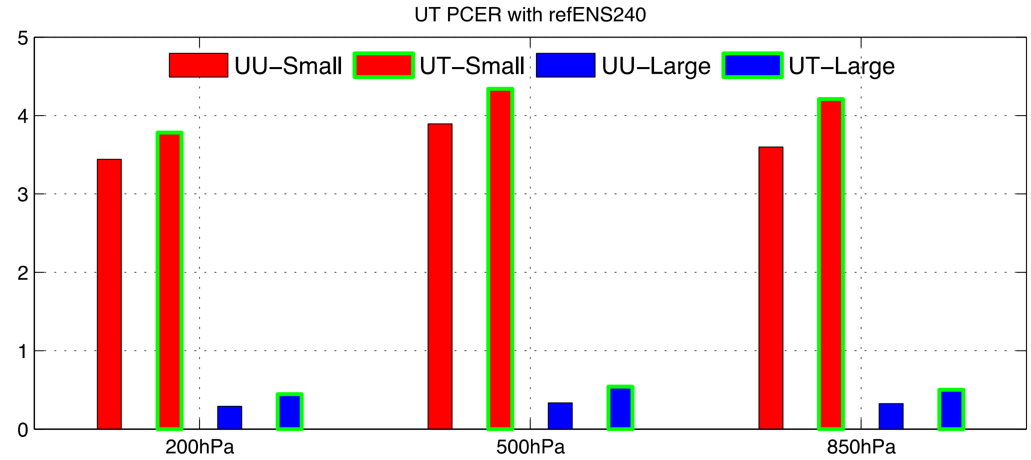
$$RCE(ENS80) = \frac{ABS[corr(ENS80) - corr(ENS240)]}{ABS[corr(ENS240)]}$$

- RCE was systematically evaluated for the small ([0 0.1), red) and large ([0.1 1.0],blue) absolute correlations and self-correlation (no green edge) and cross-correlations (with green edge)
- Errors for small correlations (red) are larger than large correlations (blue) for both the self- and cross-correlations.
- Errors for cross-correlations (with green edge) are larger than self-correlations (no green edge) for both the small and large correlations.

TT & TU



UU & UT



VTL method
Improvement of VTL correlation relative to ENS80

- **Percentage correlation errors reduction (PCER)** was defined by $$PCER = \frac{ABS[corr(ENS80) - corr(ENS240)] - ABS[corr(VTL240M80) - corr(ENS240)]}{ABS[corr(ENS80) - corr(ENS240)]} * 100\%$$

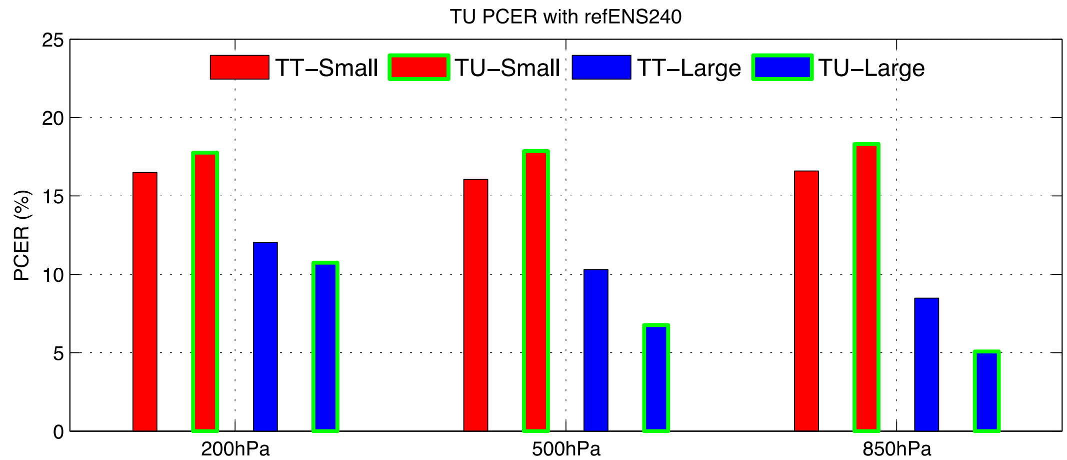
- PCER was systematically evaluated for the small ([0 0.1), red) and large ([0.1 1.0],blue) absolute correlations, and the self-correlation (no green edge) and cross-correlations (with green edge).

- VTL240M80 improves the correlation estimate compared to ENS80 for both the small/large and the self-/cross-correlation.

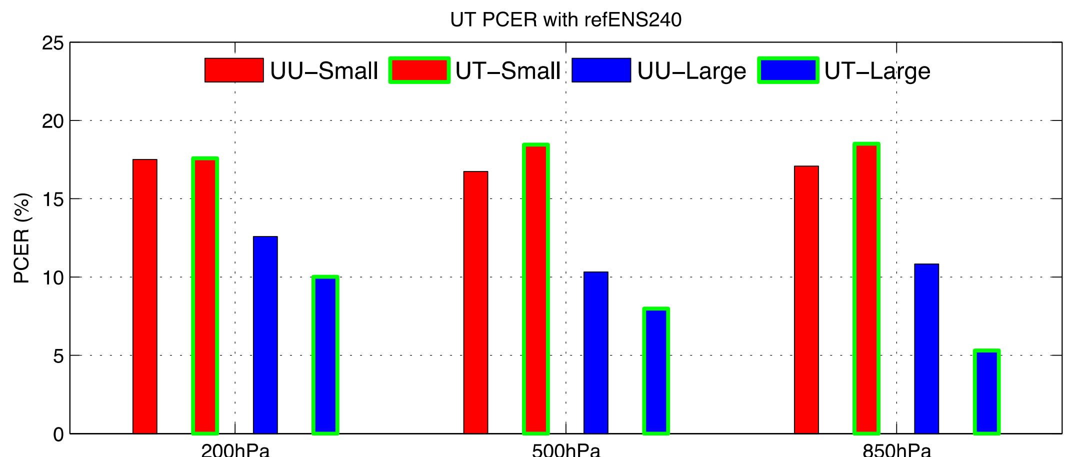
- VTL240M80 generally has a larger improvement for the small correlations than the large correlations.

- VTL240M80 has a larger improvement for the cross-correlation when correlation is small; it has a larger improvement for the self-correlation when the correlation is large.

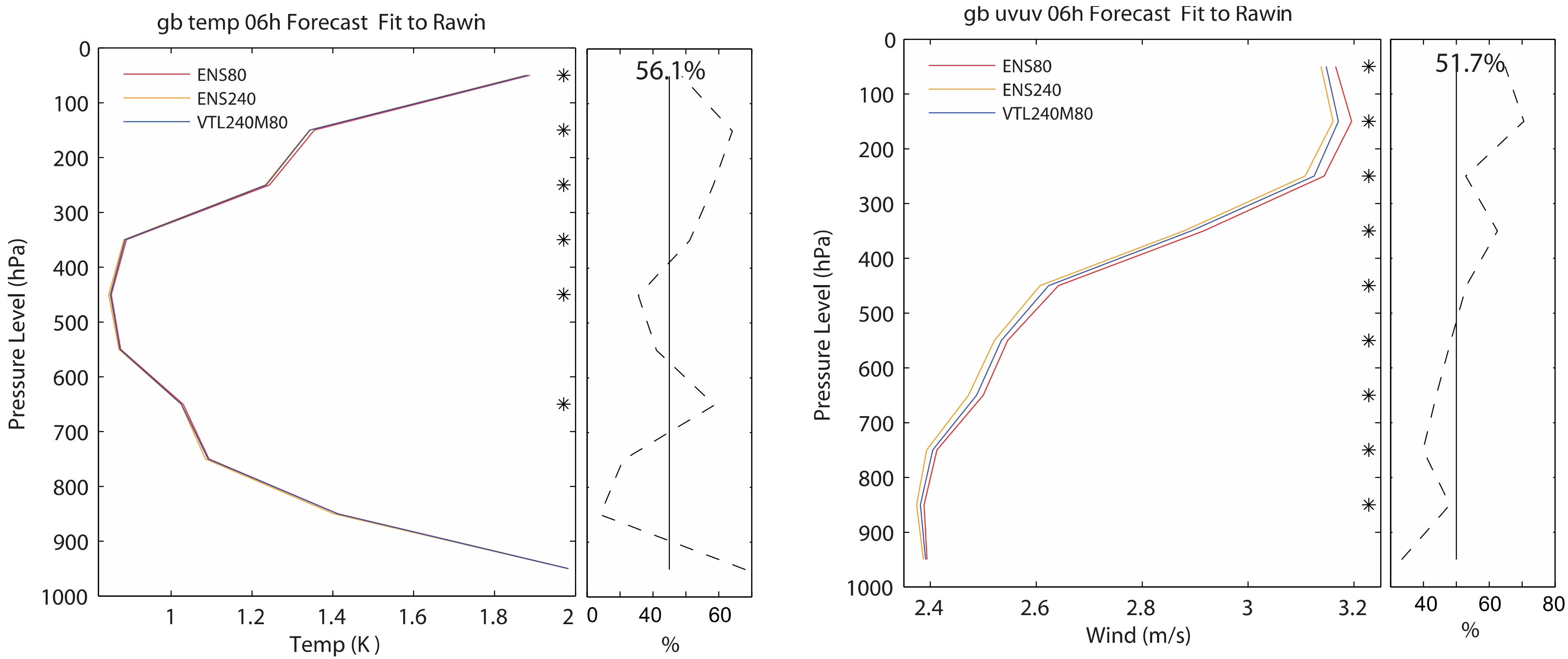
TT & TU



UU & UT



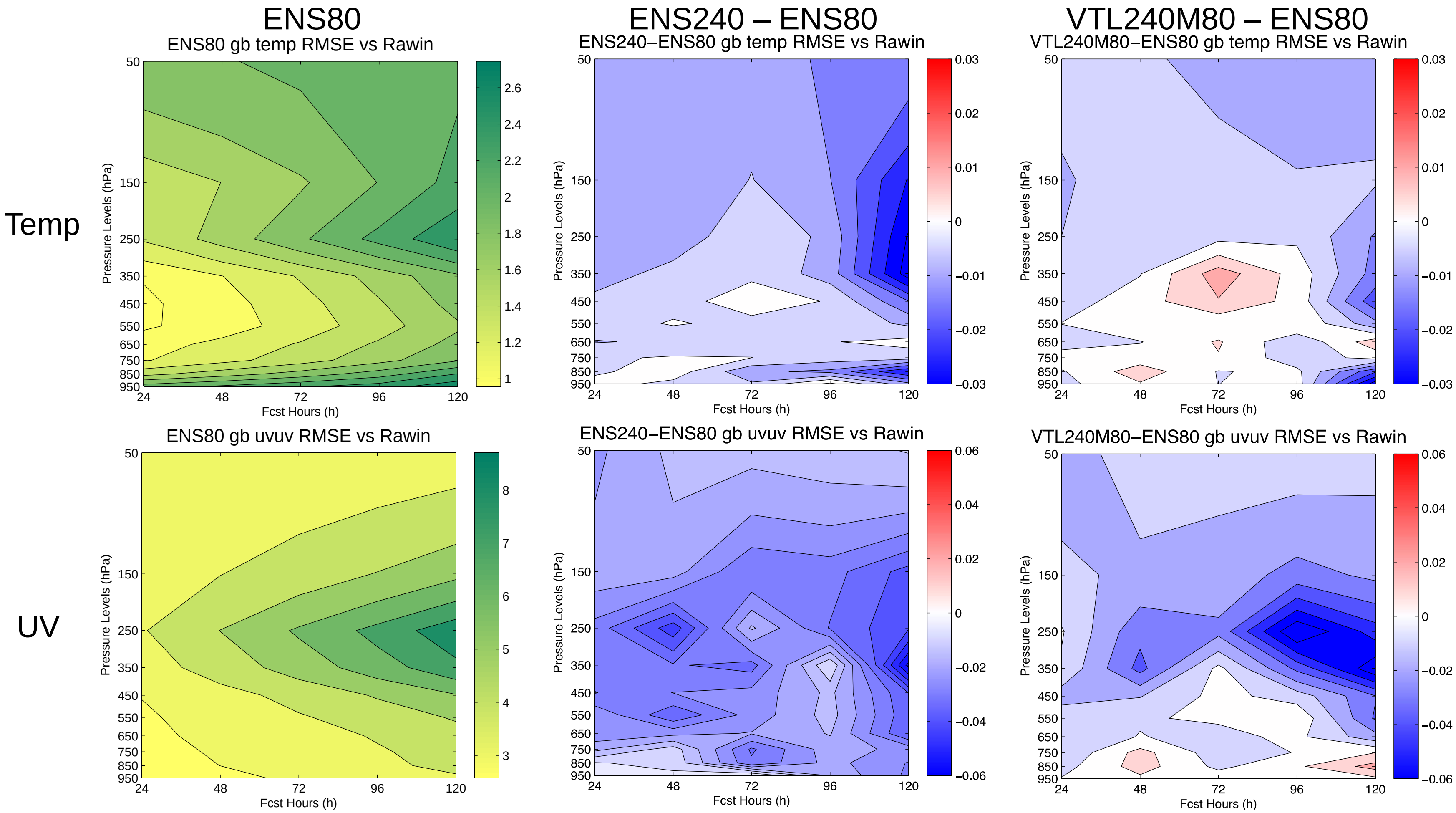
VTL method
6-hour forecast verification against rawinsondes



- Both ENS240 and VTL240M80 show a closer fit than ENS80, especially for the wind field. In most levels the difference VTL240M80 and ENS80 is significant at the 95% confident level.

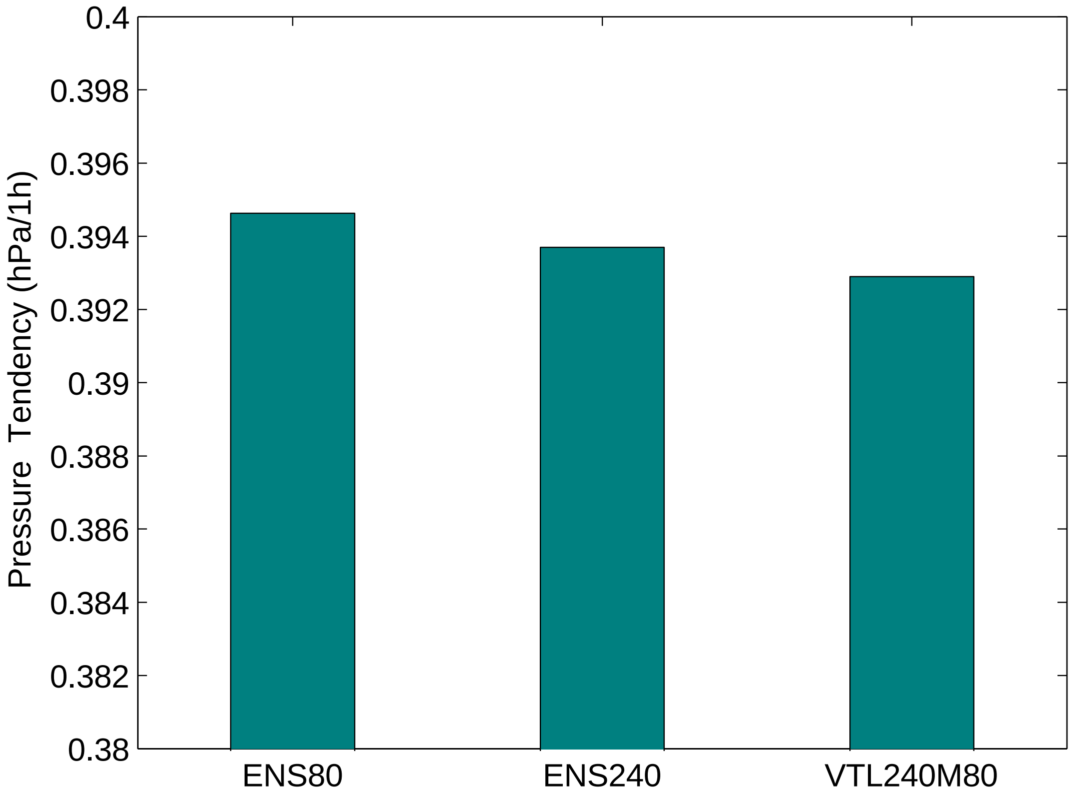
- Improvement of VTL240M80 is more than half of the improvement of ENS240.

VTL results
5-day forecast verification against rawinsonde



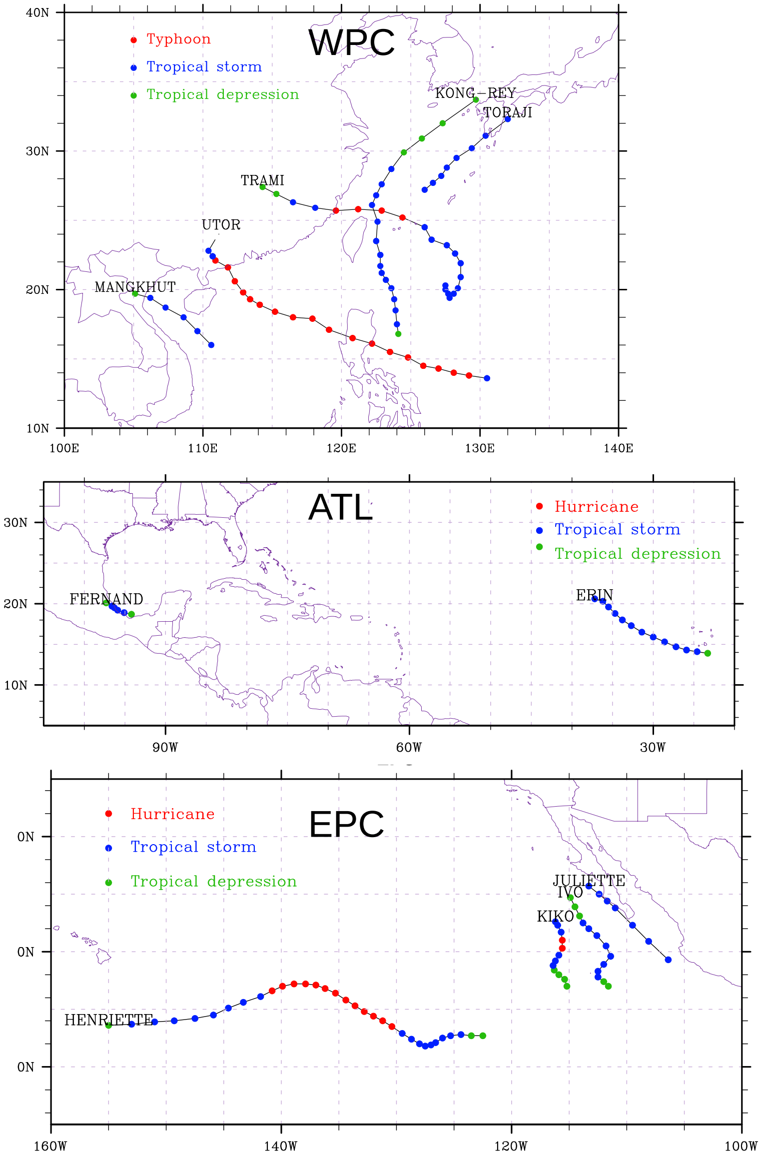
- VTL240M80 in general improves forecast during the 5-day period. Spatial pattern of improvement resembles that of ENS240.
- Degradation may be caused by the small samples (3-weeks). More DA cycling is ongoing.

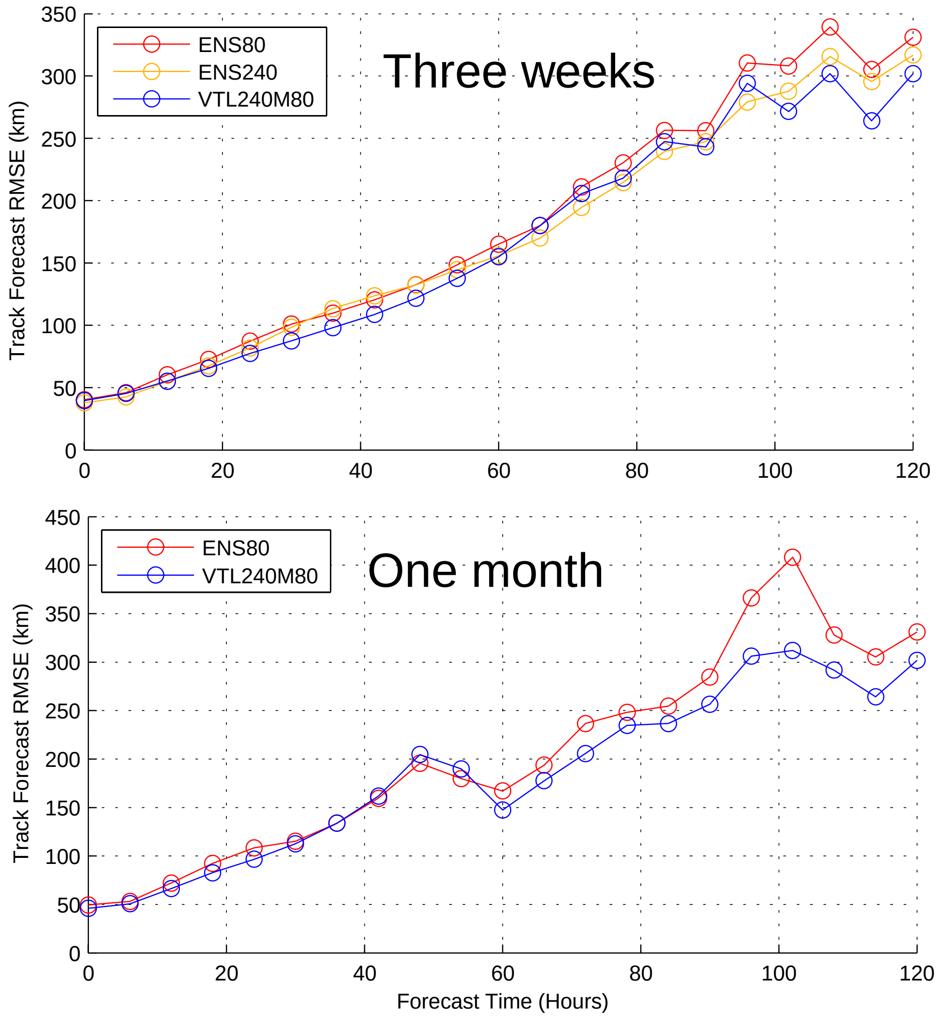
VTL method
Analysis balance evaluation



- VTL240M80 is most balanced.

VTL Method
Hurricane track verification





- VTL240M80 has a better track forecast compared with ENS80.
- VTL240M80 vs ENS240 is mixed (more sample is needed).

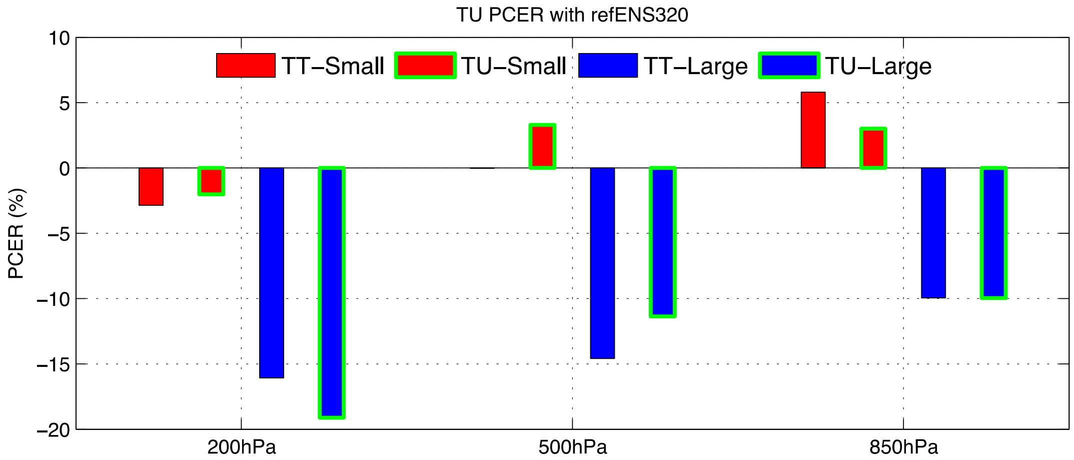
ITL experiments

Exps.	Total ensemble perts for 4DEnVar at t_i	Total mem for EnKF at t_i	Ensemble Fcsts at t_i	Cost estimate (rel. to ENS80)	Required storage (rel. to ENS80)
ENS80	80-mem perts from fcsts valid at t_i but initialized from t_{i-1}	80	80-mem 9-hour fcsts	1	1
ENS320	320-mem perts from fcsts valid at t_i but initialized from t_{i-1}	320	320-mem 9-hour fcsts	4	4
ITL320M80	320-mem perts by adding additional 3 groups of 80-mem perts from fcsts valid at t_i but initialized from the analysis at t_{i-2}, t_{i-3} and t_{i-4}	80	80-mem 27-hour fcsts	2.05	2.5

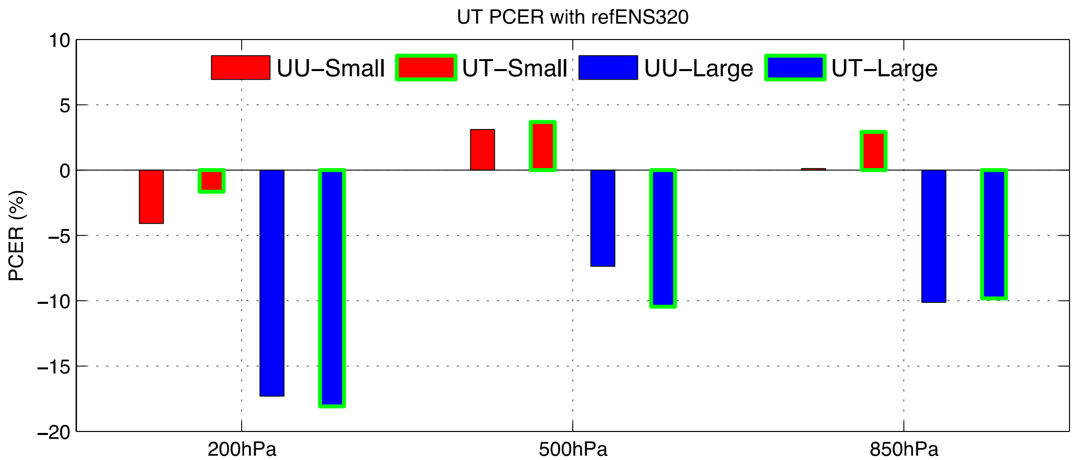
ITL method
Improvement of ITL correlation relative to ENS80

- ITL320M80 can only improve small correlation errors at certain levels, but degrade large correlation errors.

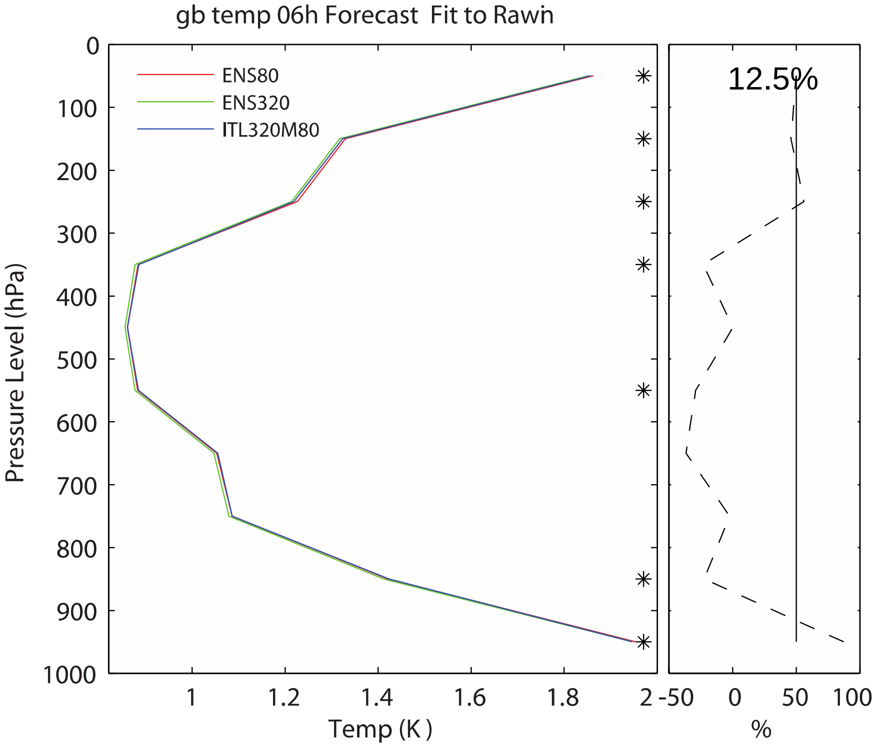
TT & TU

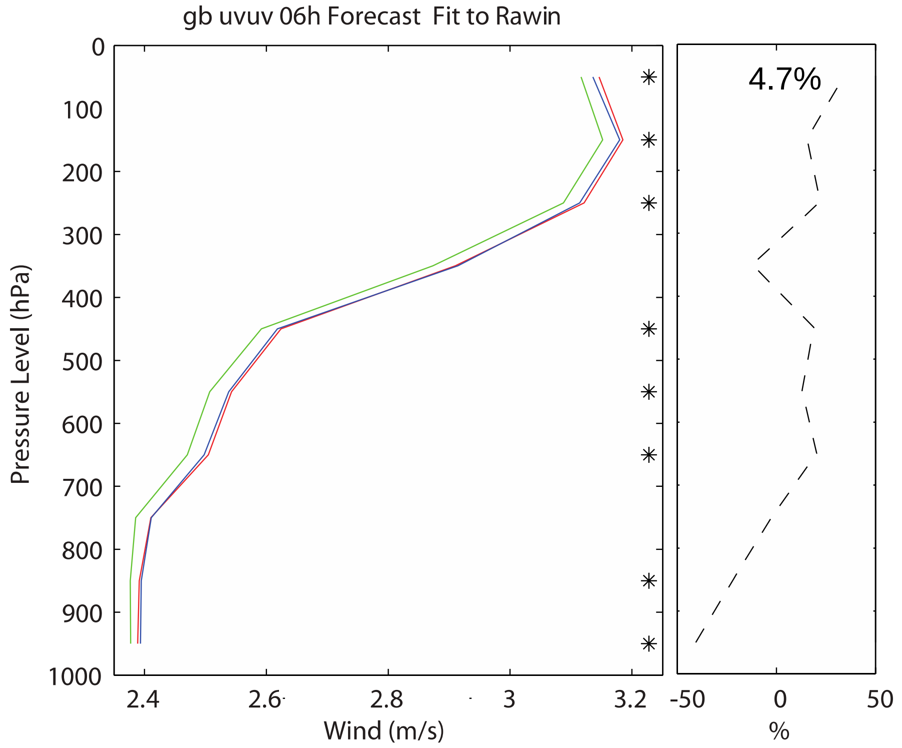


UU & UT



ITL method
6-hour forecast verification against rawinsondes

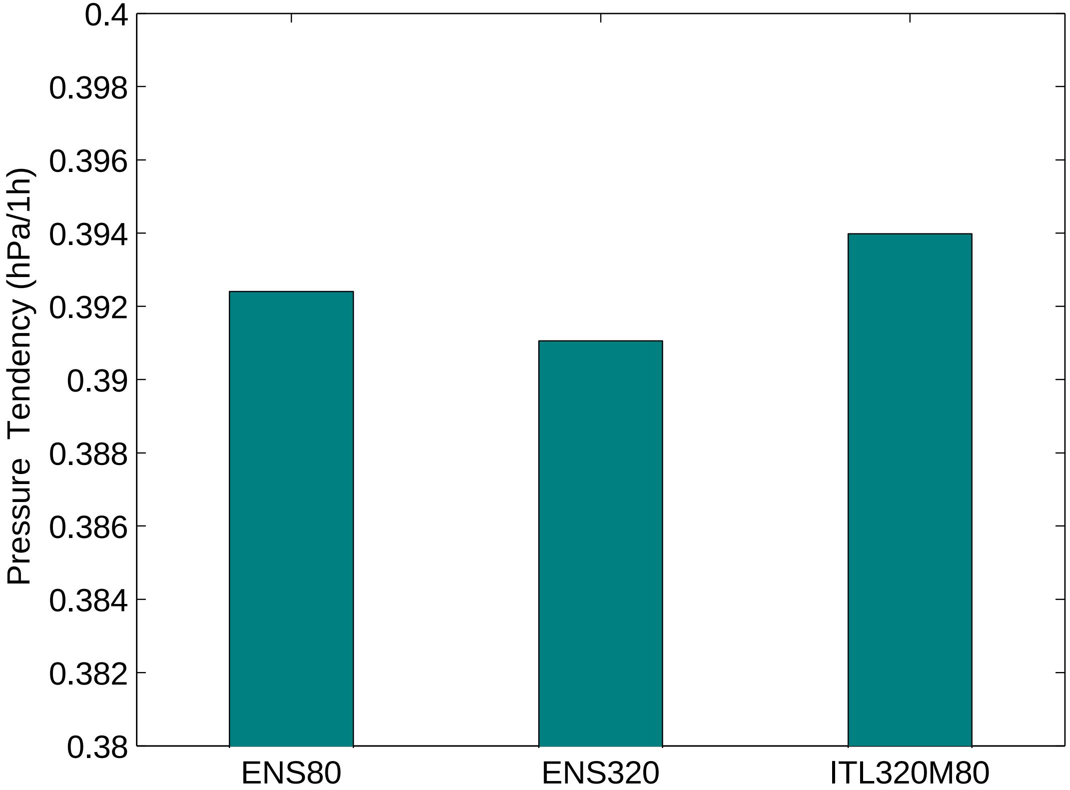




One-month experiments show that

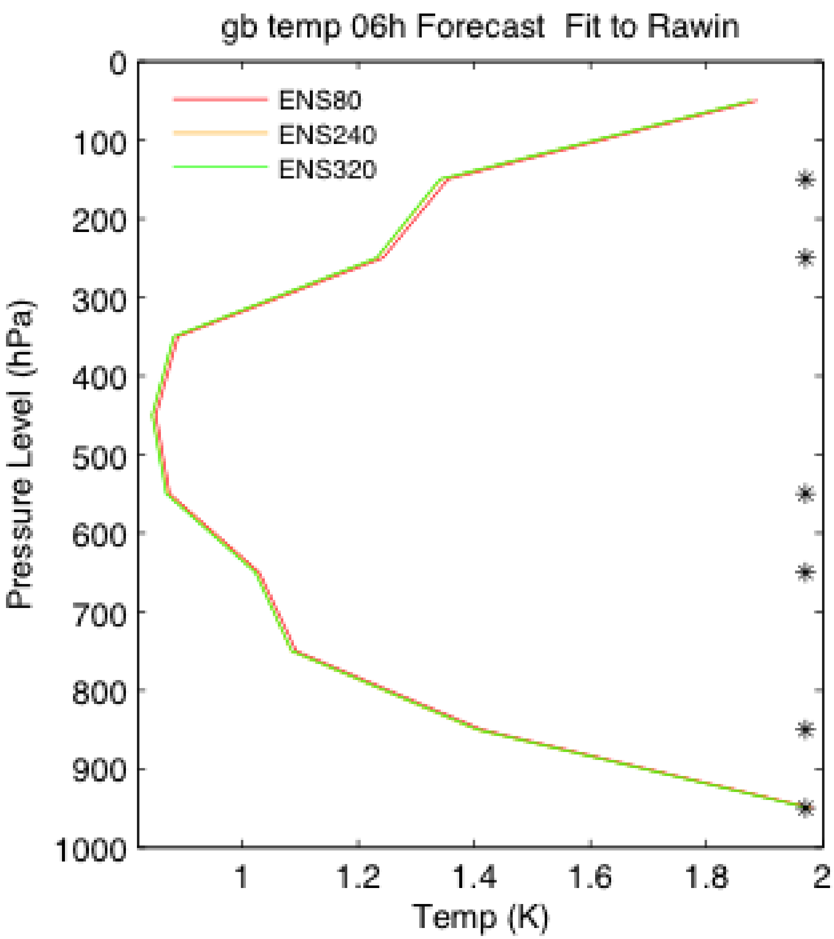
- Both ENS320 and ITL320M80 show a closer fit than ENS80, especially for the wind field;
- Improvement of ITL320M80 is much smaller than ENS320.

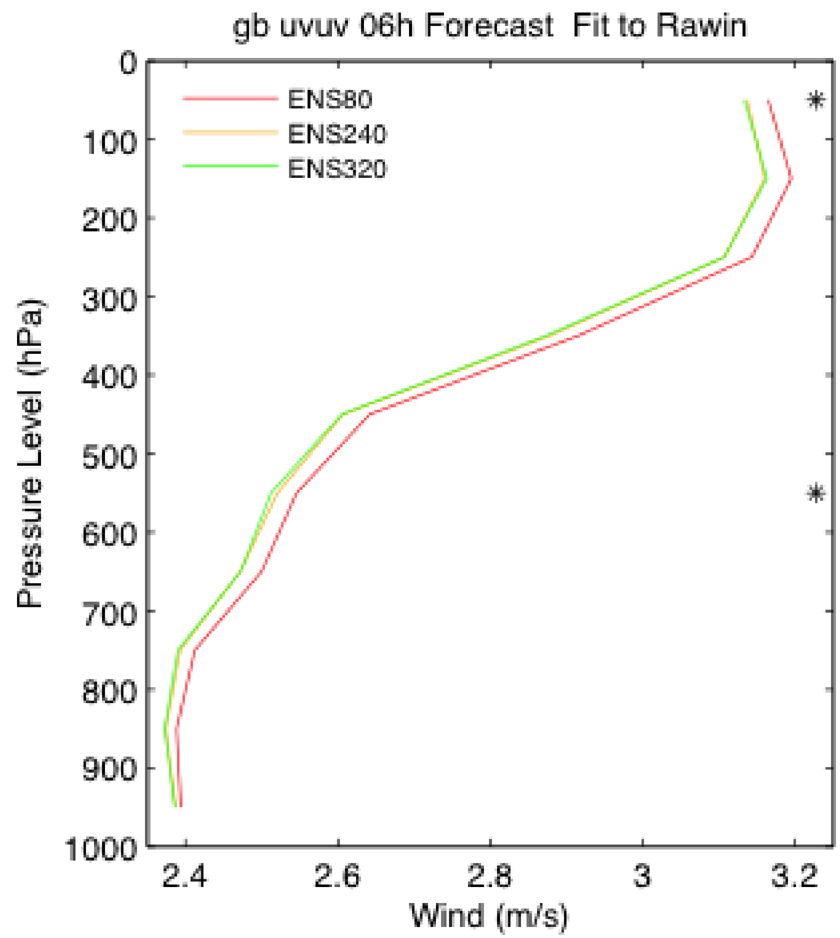
ITL method
Analysis balance evaluation



- ITL320M80 degrades the balance.

ENS80 vs ENS240 vs ENS320

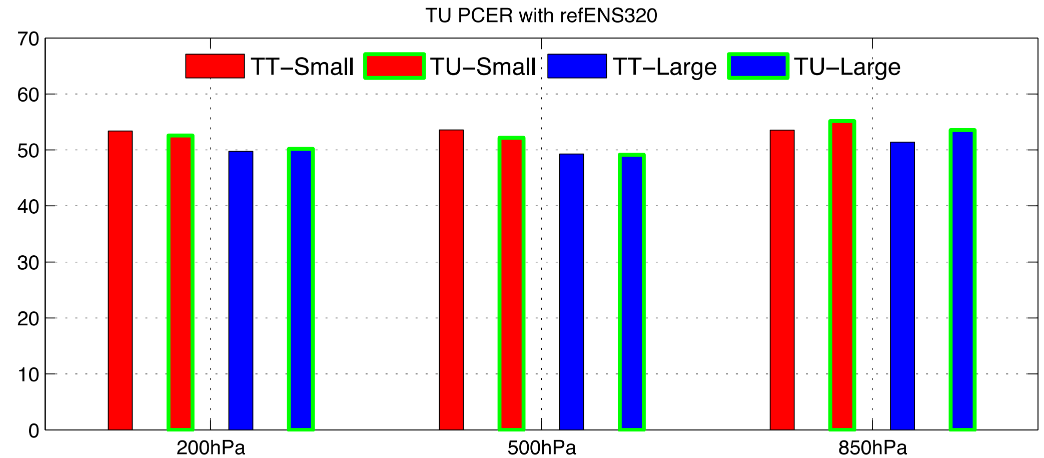




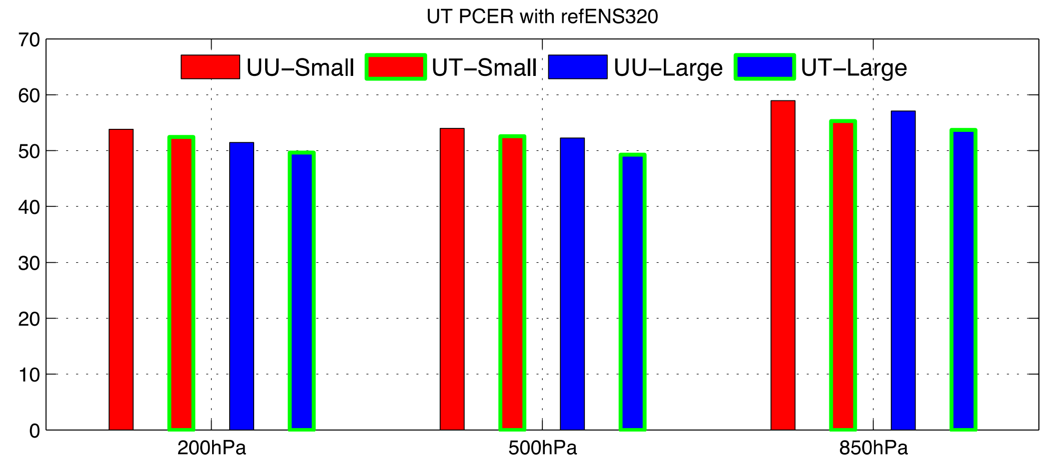
Improvement of ENS240 correlation relative to ENS80

- Compared to VTL240M80, the improvement of correlation seems more uniform.

TT & TU



UU & UT



Outline

- Background
- Method
- Experiment design
- Results
- Summary and future work

Summary

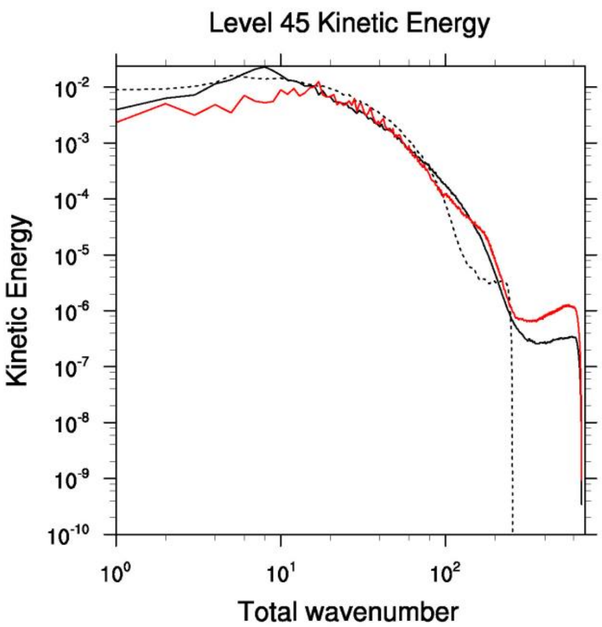
- Among all tested methods of increasing ensemble size, our results indicate VTL may be the most cost-effective method.
  - For correlations errors in ENS80, errors in small correlations are larger than the large correlations; errors for cross-correlations are larger than self-correlations.
  - VTL240M80 improves the correlation over ENS80 for both small and large correlation, with a larger improvement for the small correlation.
  - VTL240M80 show significant improvement in 6-hour forecast than ENS80, especially for the wind field.
  - The pattern of improvement by VTL240M80 out to 5-day forecasts is similar to ENS240. More samples are being collected.
  - VTL240M80 provides better hurricane track forecast compared to ENS80.

Ongoing and Future work

- Study the sensitivity of VTL240MA80 to the shifting period (m =1 or 2 hour(s) instead of 3 hours.
- Explore the benefit of VTL w/o IAU, TLMNC.
- Implement multi-resolution lagged ensemble in GSI 4DEnVar
- Implement and explore scale-dependent weighting on multi-resolution lagged ensemble in 4DEnVar

Implementation of multi-resolution lagged ensemble
Motivation

	Resolution	LAT/LON
High-resolution control forecast	T670	672/1344
Low-resolution ensemble forecast	T254	256/512



RMSE of high-resolution control forecast
Ensemble spread of high-resolution lagged ensemble forecast
Ensemble spread of low-resolution ensemble forecast

- High-resolution lagged ensemble perturbations can be generated by calculating the differences of the control forecasts initialized at different times.
- High-resolution lagged ensemble forecasts can sample forecast errors in the spectral regime that are not sampled in the low-resolution ensemble forecasts

Implementation of multi-resolution lagged ensemble Algorithm Development

Multi-resolution hybrid 4DEnVar

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

$$= \frac{1}{2}\beta_1 (\mathbf{x}_1')^T \mathbf{B}_1^{-1}(\mathbf{x}_1') + \boxed{\frac{1}{2}\beta_L (\boldsymbol{\alpha}^L)^T \mathbf{A}_L^{-1}(\boldsymbol{\alpha}^L)} + \boxed{\frac{1}{2}\beta_H (\boldsymbol{\alpha}^H)^T \mathbf{A}_H^{-1}(\boldsymbol{\alpha}^H)} + \frac{1}{2}(\mathbf{Hx}' - \mathbf{y}^{o\prime})^{\mathbf{T}} \mathbf{R}^{-1}(\mathbf{Hx}' - \mathbf{y}^{o\prime})$$

low-resolution ensemble high-resolution lagged ensemble

$$\mathbf{x}' = \mathbf{x}_1' + \boxed{\mathbf{u}\sum_{k=1}^{K_L}\left(\boldsymbol{\alpha}_k^L \circ \mathbf{x}_k^{eL}\right)} + \boxed{\sum_{k=1}^{K_H}\left(\boldsymbol{\alpha}_k^H \circ \mathbf{x}_k^{eH}\right)}$$

$\mathbf{B}_1$ 3DVAR static covariance; $\mathbf{R}$ observation error covariance; K ensemble size;

$\mathbf{A}$ correlation matrix for ensemble covariance localization; $\mathbf{x}_k^e$ kth ensemble perturbation;

$\mathbf{x}'$ total (hybrid) increment; $\mathbf{x}_1'$ 3DVAR increment; $\mathbf{y}^{o\prime}$ innovation vector;

$\mathbf{H}$ linearized observation operator; $\boldsymbol{\alpha}$ extended control variable;

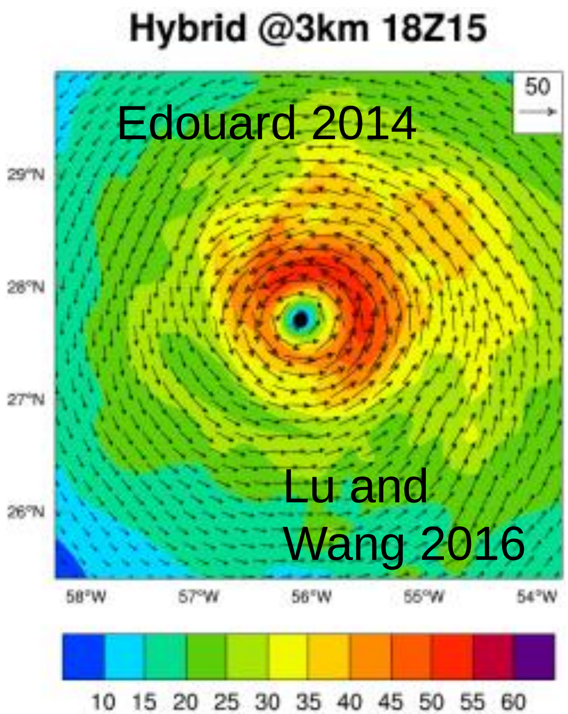
$\mathbf{u}$ linear transform matrix from low resolution to high resolution;

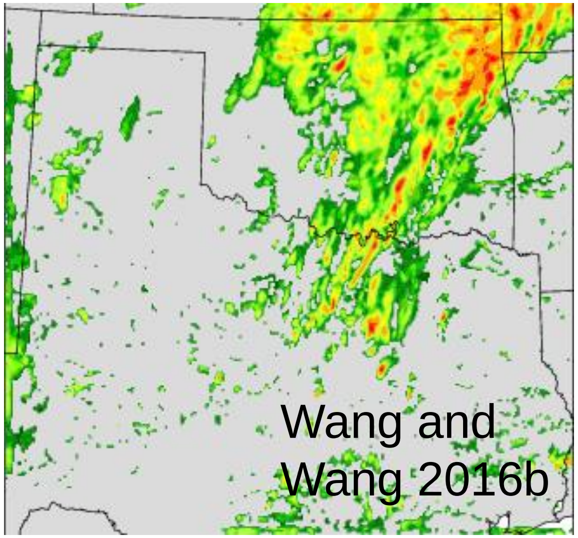
β_1 weighting coefficeint for static covariance;

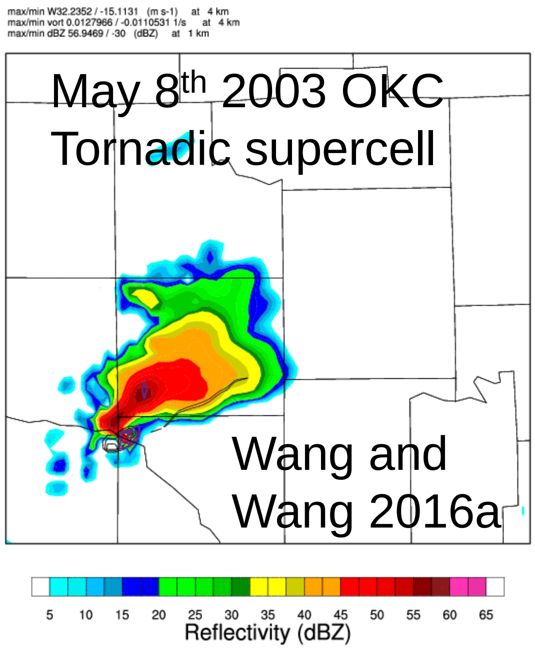
β_L weighting coefficeint for low-resolution ensemble covariance;

β_H weighting coefficeint for High-resolution ensemble covariance;

GSI based hybrid DA has been further developed for convective scale NWP supported by other sources







- Given GFS becoming non-hydrostatic and convection resolving, global DA now is a multiscale problem. DA experiences gained from global NWP as well as from regional, convection resolving NWP should be leveraged to form foundation for further R&D.