Adjusting the ensemble for EnKF assimilation: applications to regional EnKF and severe weather prediction

Shu-Chih Yang[1]

Eugenia Kalnay[2], Takeshi Enomoto[3]
Jhe-Wei Chou[1], Ping-Yin Wu[1] and Chi-Chien Chang[1]

[1] Depart. Of Atmospheric Sciences, National Central University, Taiwan
[2]University of Maryland, USA
[3] Kyoto University, Japan

4th international symposium on data assimilation, Kobe, 2015

Motivation

The performance of EnKF are related to both the structures of the ensemble perturbations and the accuracy of the mean state.

- Enhancing the EnKF ensemble perturbations with the ensemble singular vectors (ESV)
  - Results from a QG-LETKF system (YangKalnayEnomoto, 2015)
  - Application to typhoon prediction
- Improving the accuracy of the mean state with the mean-recentering scheme.
  - Application to typhoon prediction
  - Application convective-scale nowcasting

Background

- Ensemble singular vectors (ESVs) indicates the directions of the fastest growing forecast errors
  - Given the choice of the perturbation norm and forecast interval, the leading ESV maximizes the growth of the perturbations.
  - No need for the TLM/ADJ models (Bishop and Toth, 1999, Enomoto et al. 2006).
  - ESVs can be derived and used in the ensemble assimilation/prediction system
- Additive covariance inflation aims to perturb the subspace spanned by the ensemble vectors and better capture the sub-growing directions that may be missed in the original ensemble.
  - Random perturbations may introduce non-growing or irrelevant error structures.
  - Houtekamer et al., (2005) generated additive errors according to the 3D-Var errors covariance structures.
  - Whitaker et al. (2008) generated additive noise by selecting random differences between adjacent 6-houly analyses
  - ESVs are applied as the “flow-dependent” covariance inflation

Ensemble singular vector (ESV)

A set of initial (I) and final (F) perturbations:

$$\mathbf{X}_{t-\Delta t}^{I} = \left[\delta\mathbf{x}_{1,t-\Delta t}, \cdots, \delta\mathbf{x}_{i,t-\Delta t}, \cdots, \delta\mathbf{x}_{K,t-\Delta t}\right]; \quad \mathbf{X}_{t}^{F} = \left[\delta\mathbf{x}_{1,t}^{F}, \cdots, \delta\mathbf{x}_{i,t}^{F}, \cdots, \delta\mathbf{x}_{K,t}^{F}\right]$$

Find the linear combination of initial perturbations that will grow fastest given a optimization time period (Δt)

Initial ESV: $\delta\mathbf{x}_{t-\Delta t}^{I} = \mathbf{X}_{t-\Delta t}^{I}\mathbf{p}$

Final ESV: $\delta\mathbf{x}_{t}^{F} = \mathbf{X}_{t}^{F}\mathbf{p}$

By defining the initial and final perturbation norms ($\mathbf{C}_I$ and $\mathbf{C}_F$), we can solve $\mathbf{p}$ (Enomoto et al. 2006).

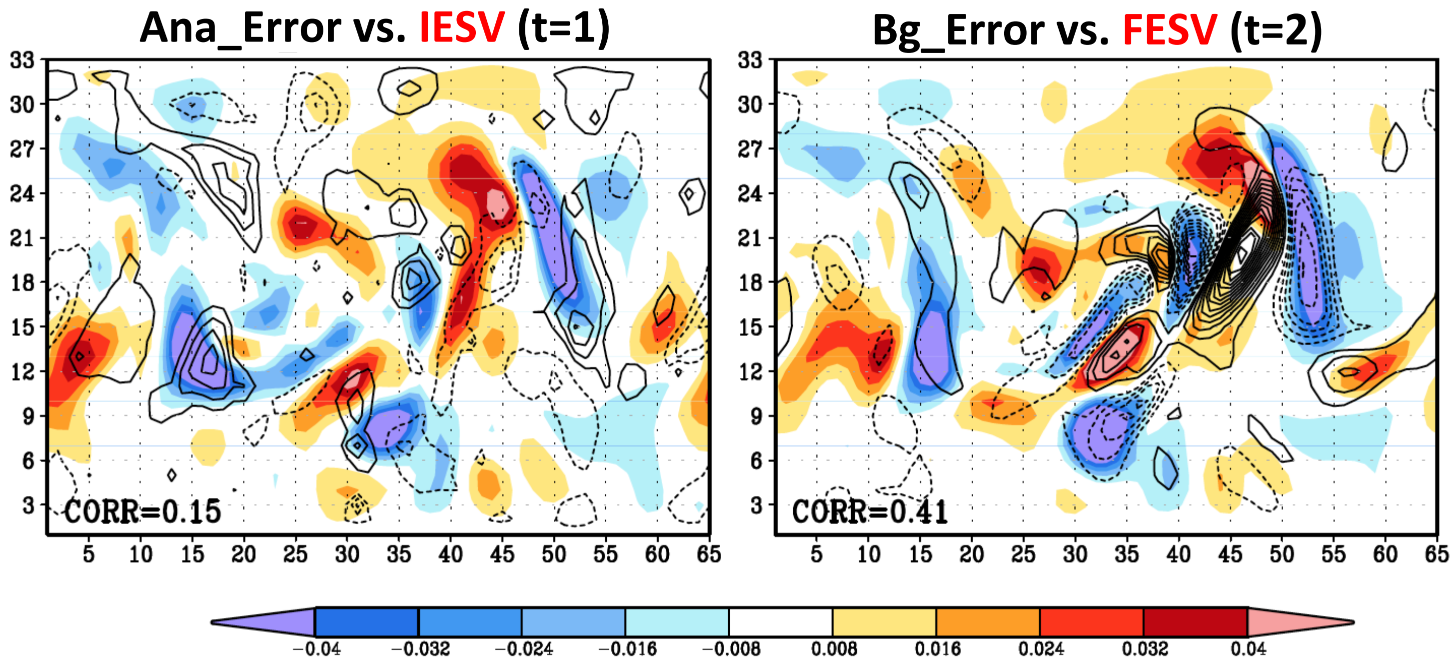
$$\boxed{\left(\mathbf{X}_{t-\Delta t}^{I}{}^{T}\mathbf{C}_{I}\,\mathbf{X}_{t-\Delta t}^{I}\right)^{-1}\left(\mathbf{X}_{t}^{F}{}^{T}\mathbf{C}_{F}\,\mathbf{X}_{t}^{F}\right)\mathbf{p} = \lambda\,\mathbf{p}}$$

We can find K pairs of IESV and FESV with $\left(\lambda^{i}, \mathbf{p}^{i} \quad i = 1, \cdots, K\right)$

ESV1 in a Quasi-geostrophic model

(12-hr analysis cycle, QG-LETKF with 20 mem)

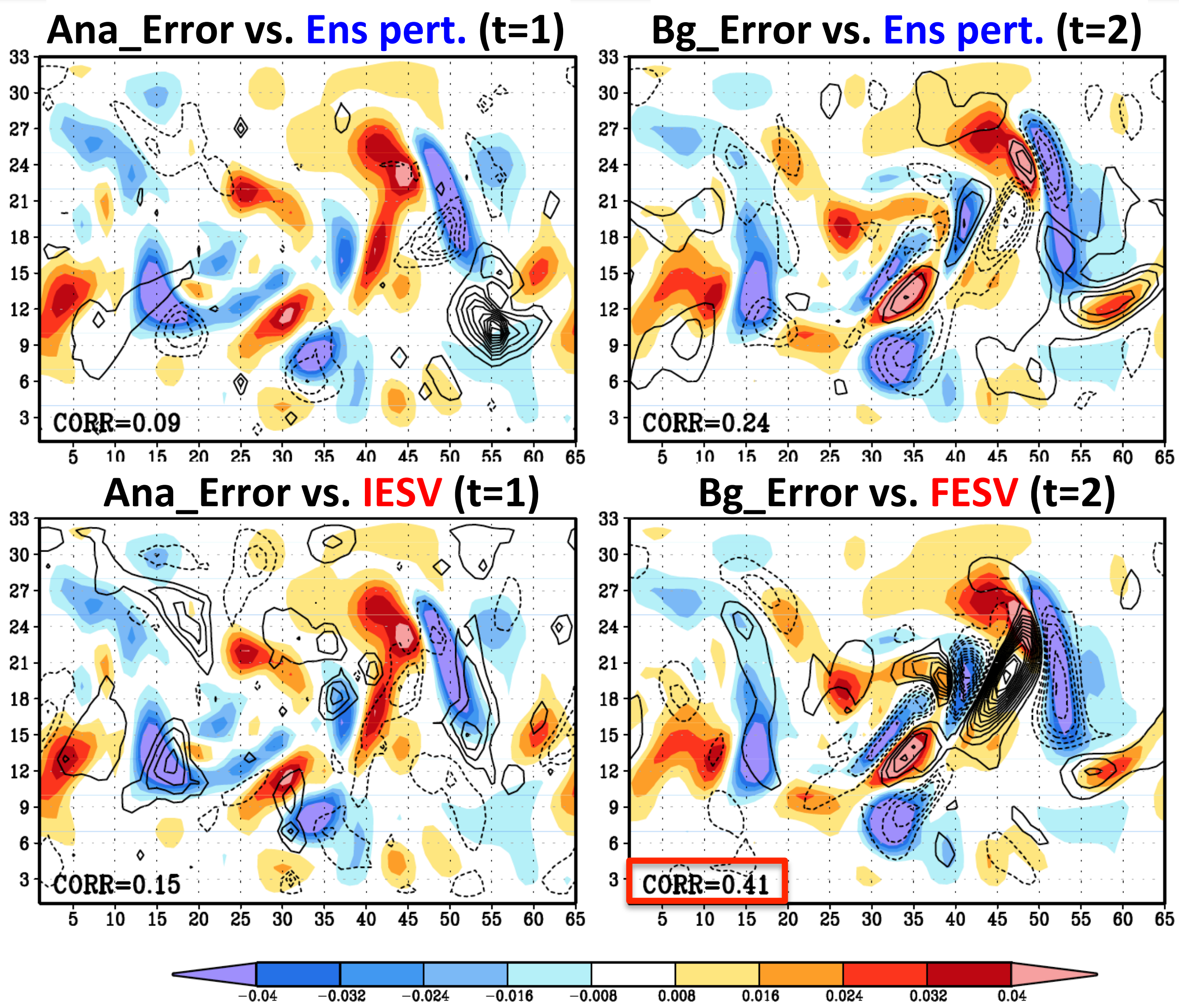
$\mathbf{X}^{I}_{t-\Delta t} = \mathbf{X}^{a}_{t-\Delta t}$ (LETKF Ana. Ens); $\mathbf{X}^{F}_{t} = \mathbf{X}^{b}_{t}$ (LETKF Back. Ens)



The fast growing perturbation (contours) is very closely related to the background errors (color). The IESV (an initial Singular Vector) is NOT related to the initial errors.

Yang et al. (2015)

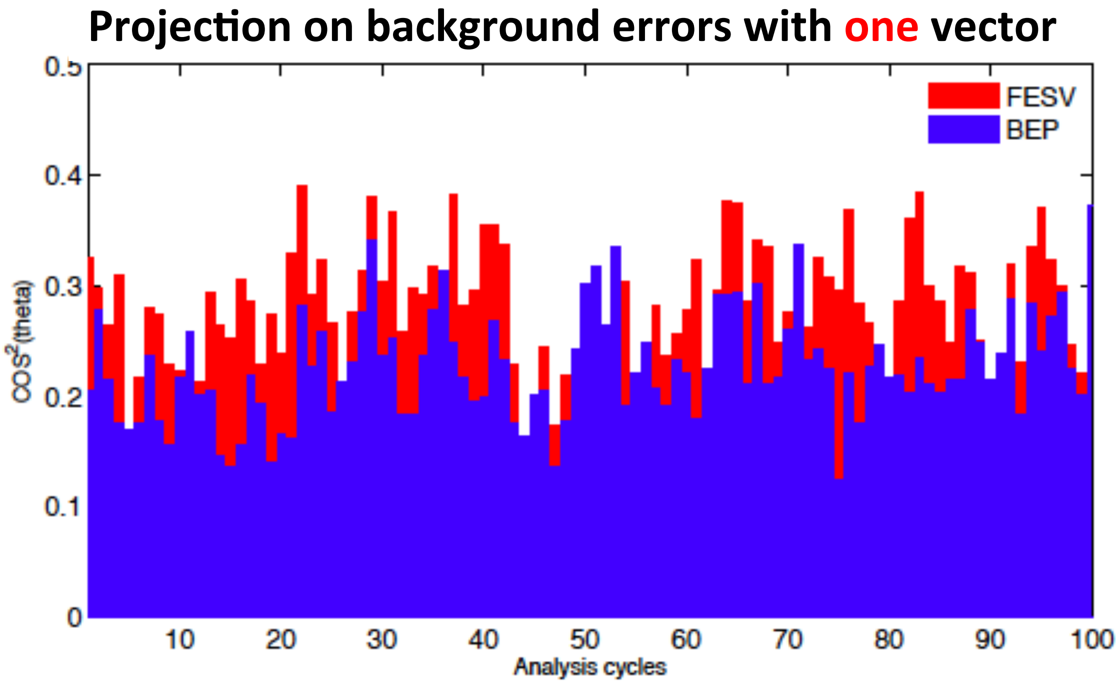
Comparisons between the ESV, ensemble perturbations and errors during the spin-up



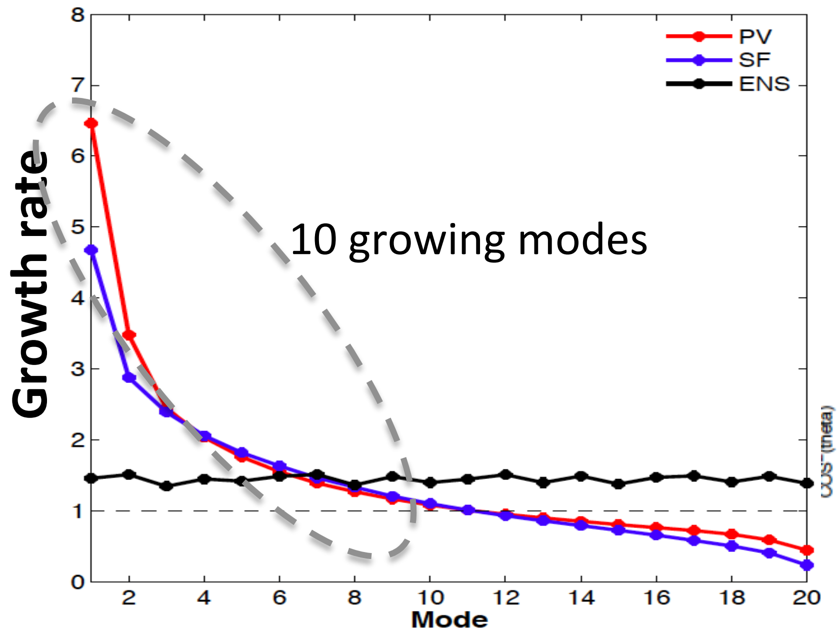
Background ensemble perturbations are still under development during LETKF's spin-up

The final ESV1 effectively captures the fast growing errors.

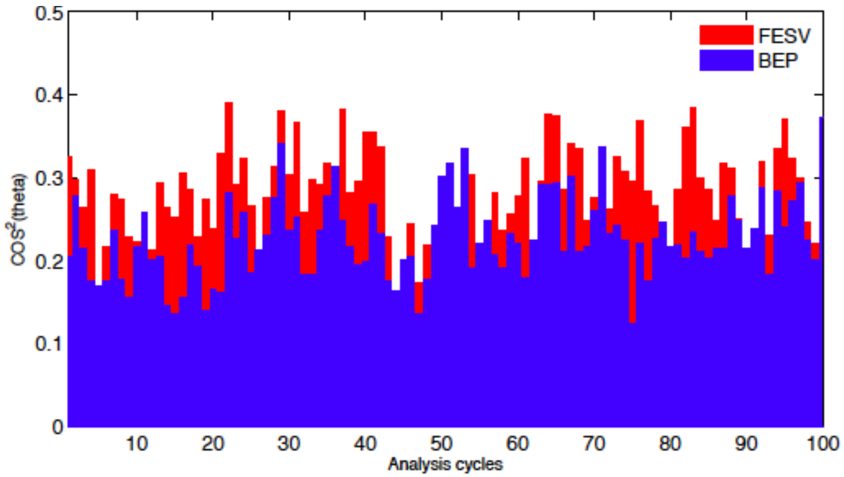
ESV, ensemble perturbations and errors



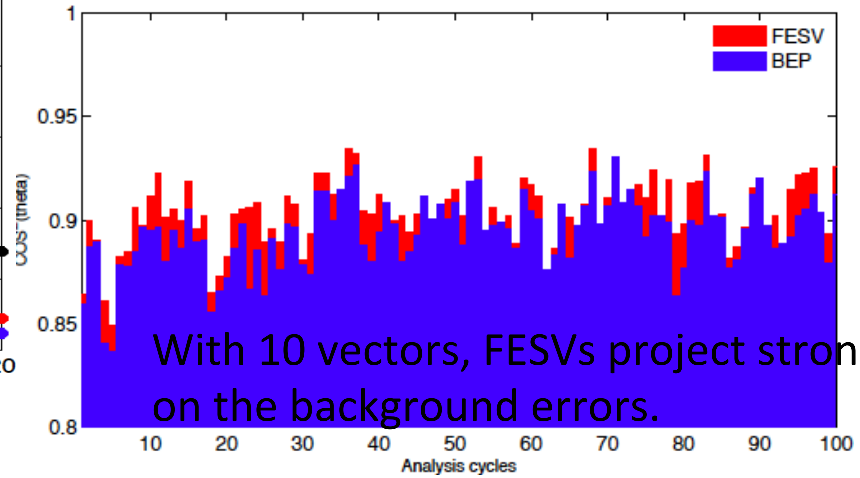
ESV, ensemble perturbations and errors



Projection on background errors with one vector



Projection on background errors with 10 vectors



With 10 vectors, FESVs project stronger on the background errors.

Use ESV as additive inflation in EnKF

$$\mathbf{P}_b' = (1+\alpha)\mathbf{P}_b$$

CNTL: standard LETKF with multiplicative inflation
($\alpha(z)$~ 8%,with vertical dependence)

$$\mathbf{X}_a' = \mathbf{X}_a' + c\mathbf{X}_{IESV}'$$

RDM: LETKF with random perturbations as additive inflation

IESVall: use all IESVs as additive inflation

IESV10: use growing ESVs as additive inflation
(+/- IESVs, randomize the order)

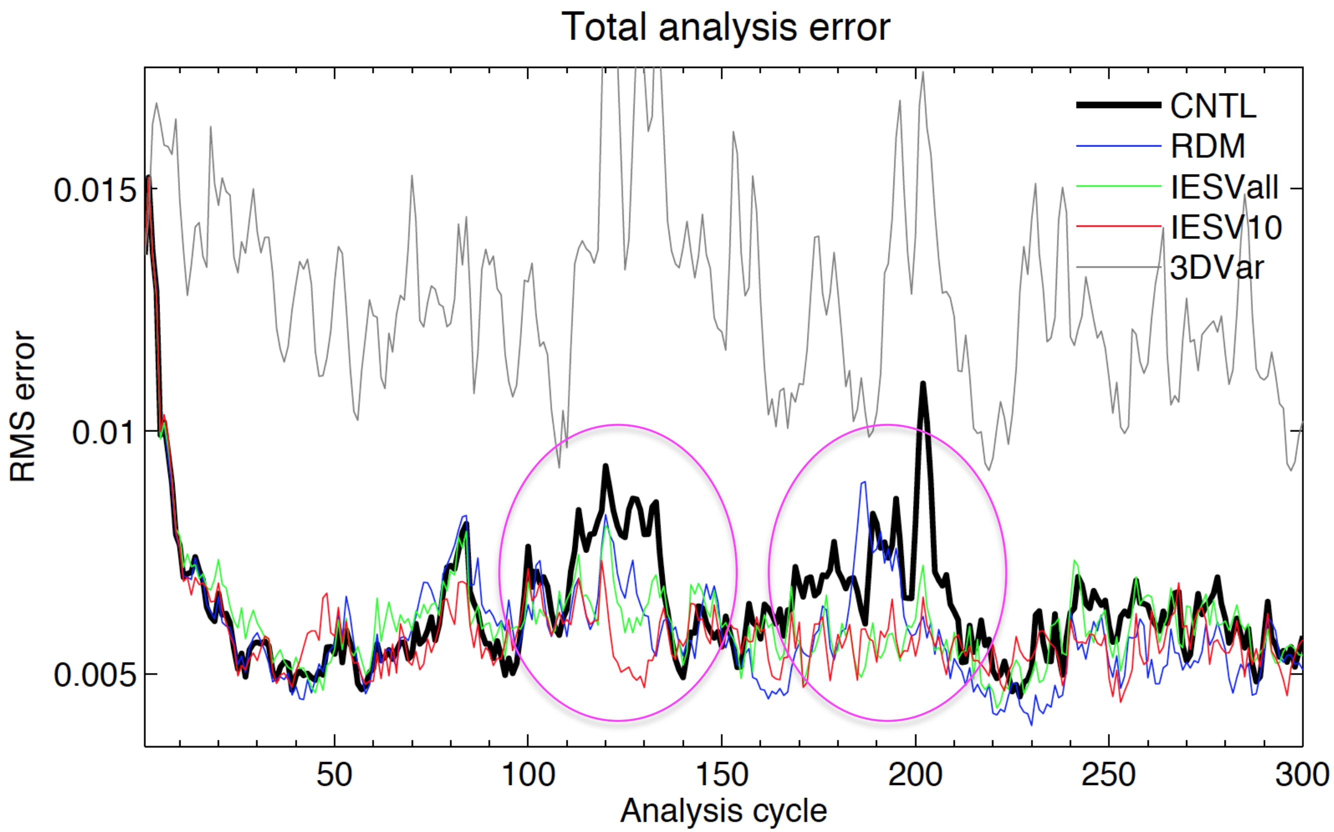
Analysis error with additive inflation

CNTL: standard LETKF with multiplicative inflation

RDM: LETKF with random perturbations as additive inflation

IESVall: LETKF with all IESVs as additive inflation

IESV10: LETKF with 10 IESVs as additive inflation



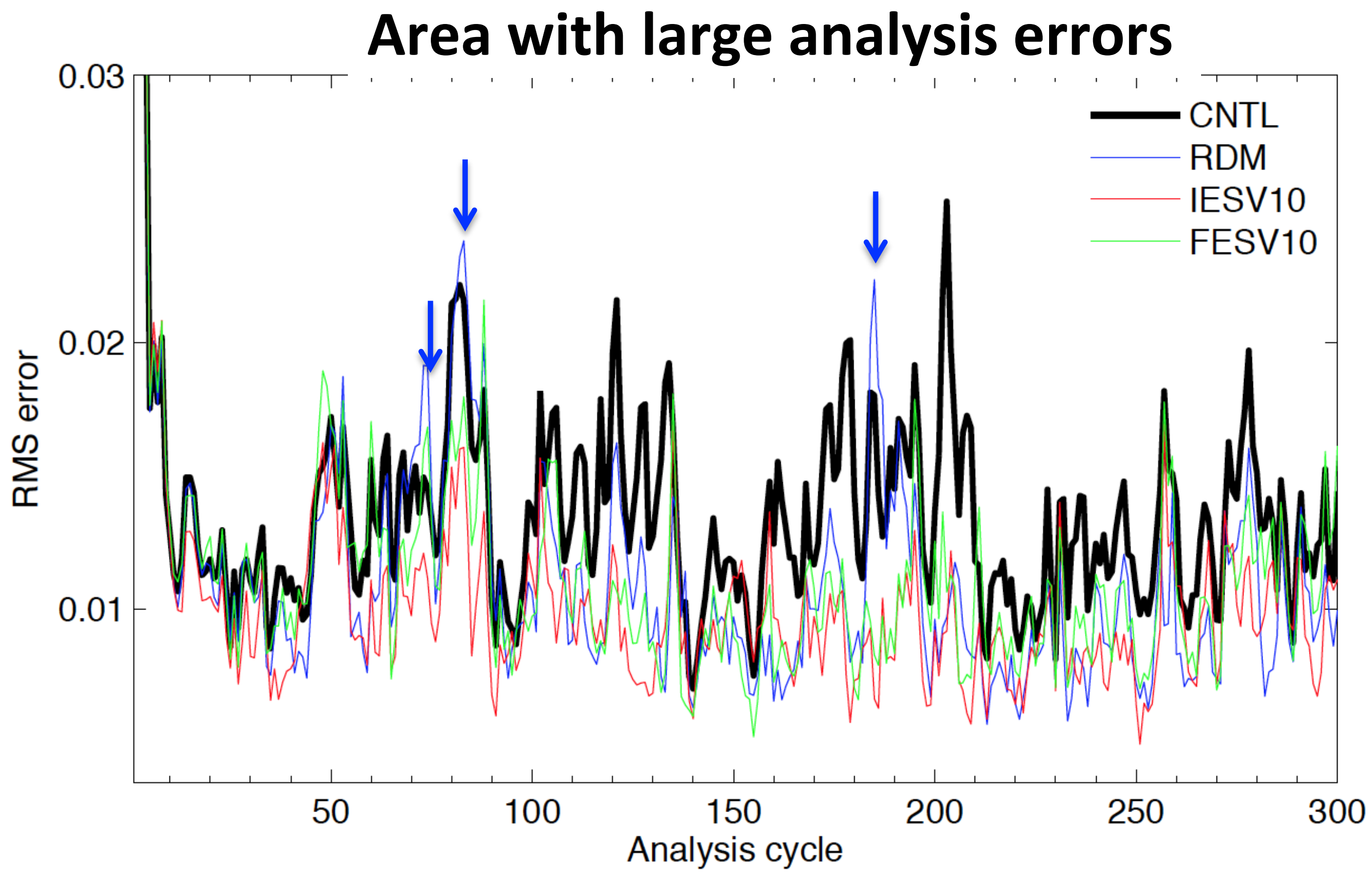
ESVs are particularly useful for large errors!

Analysis error with additive inflation

CNTL: standard LETKF with multiplicative inflation

RDM: LETKF with random perturbations as additive inflation

IESV10: LETKF with 10 IESVs as additive inflation

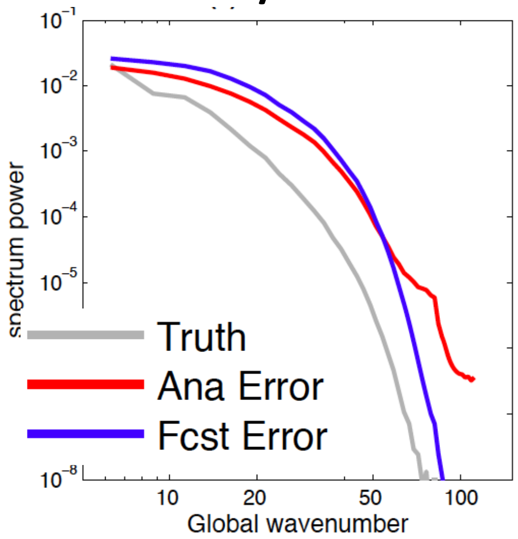


ESVs are particularly useful for large errors!

RDM occasionally has large error spikes.

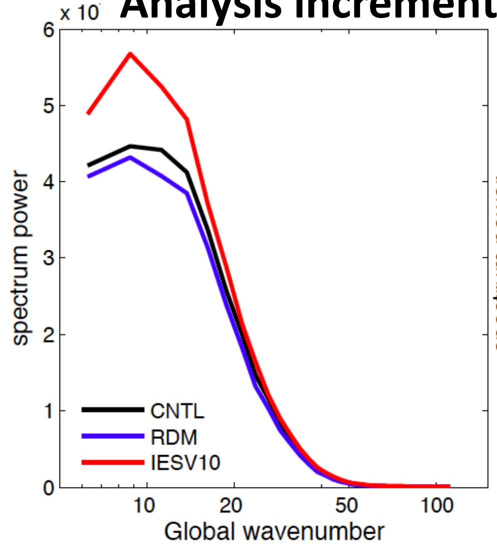
Spectrum analysis of errors/increments

Truth/Errors

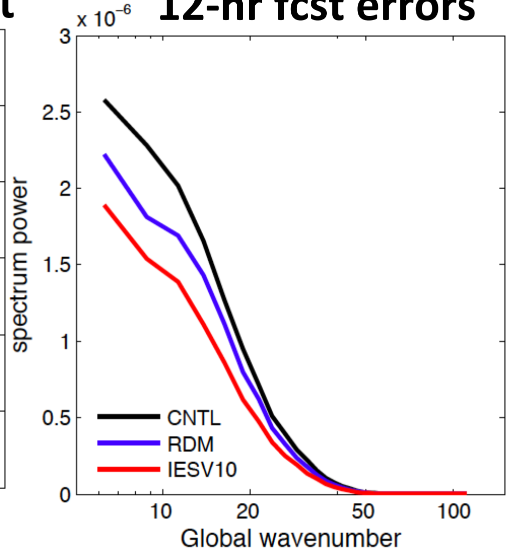


- Growing errors are characterized by structures with wv# <20
- Additional corrections with ESVs are related to the growing errors

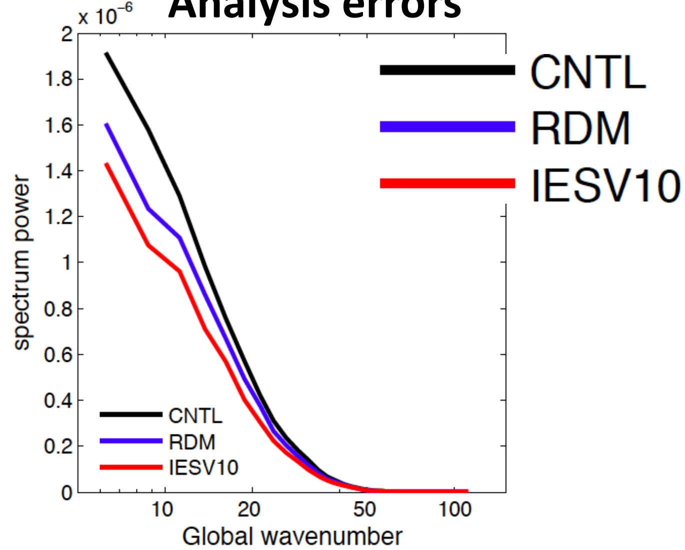
Analysis increment



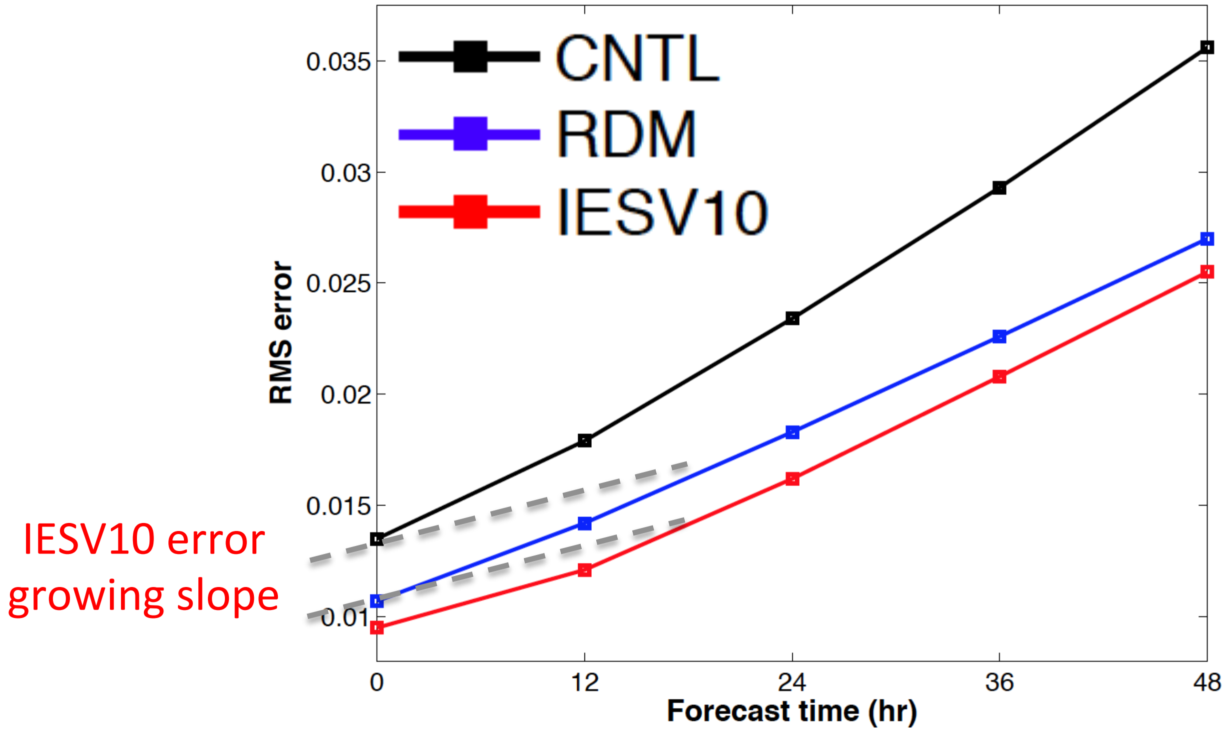
12-hr fcst errors



Analysis errors



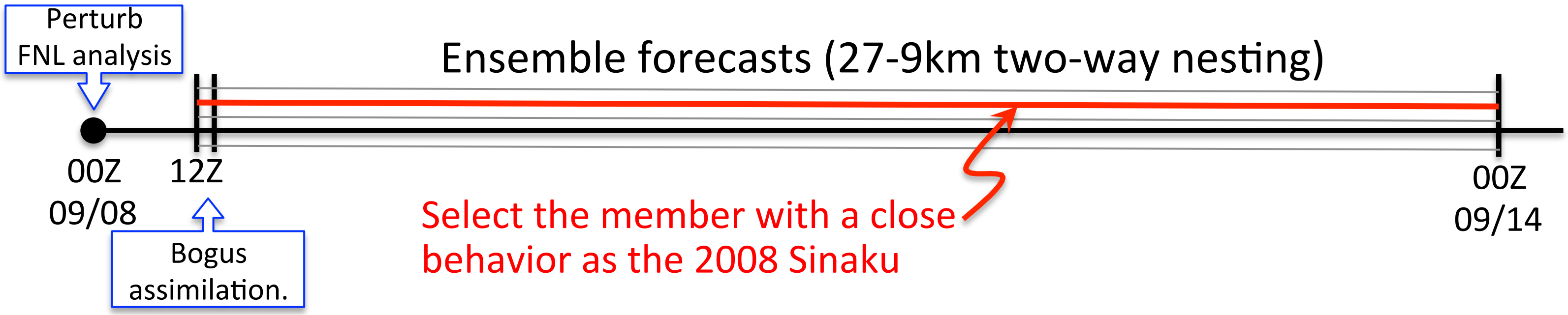
Forecast errors in time

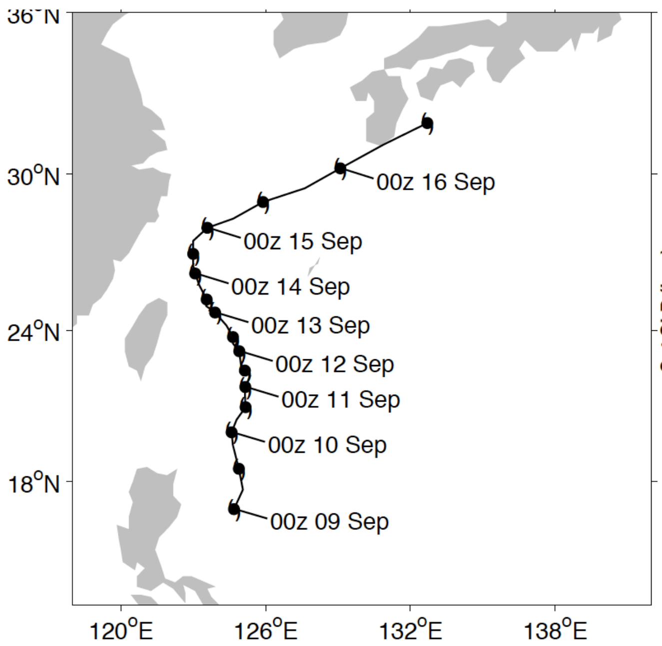


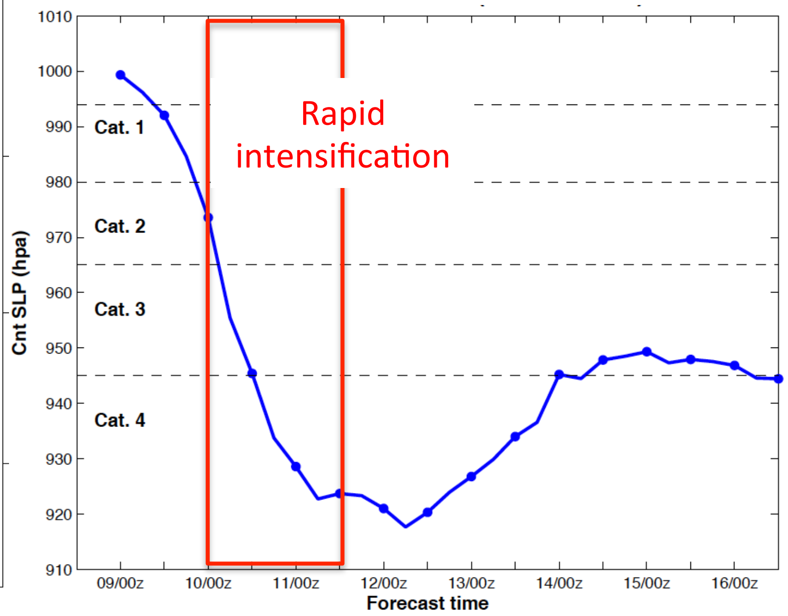
- In CNTL, the incompletely removed growing errors **amplify** at later forecast time.
- With ESVs, growing errors are more effectively removed, especially during the first 12 hours.

ESVs for typhoon assimilation and prediction: results from OSSE (Chou et al. 2015)

Simulated typhoon for OSSE

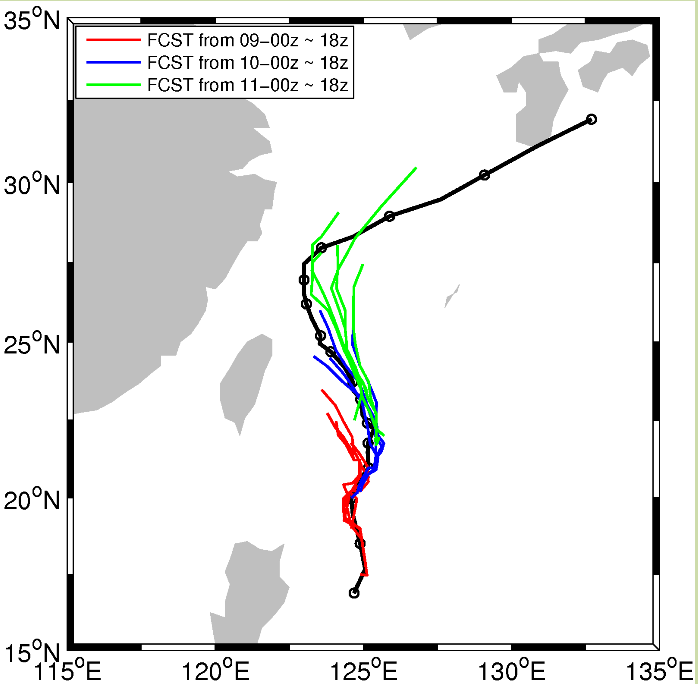






Experiment setup

Track prediction initialized from LETKF analysis

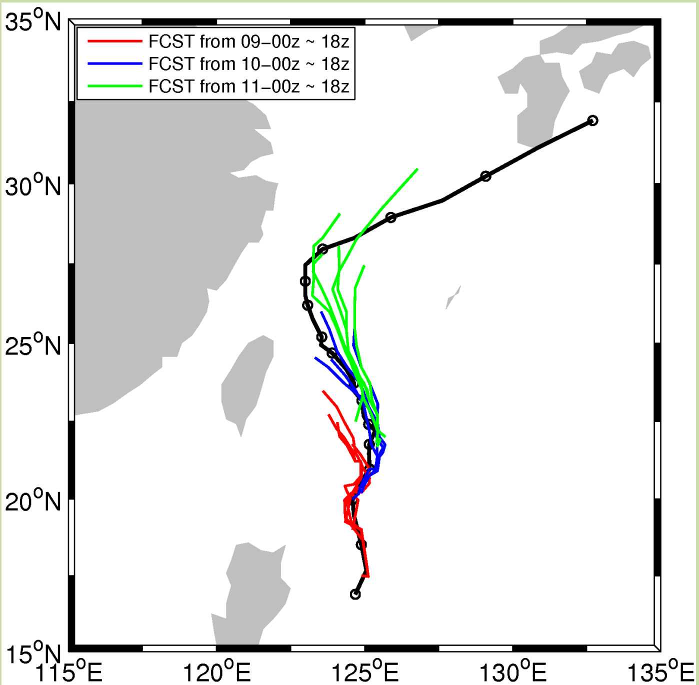


Large track errors when initialized from 18Z 09/10 18Z to 0911 06Z.

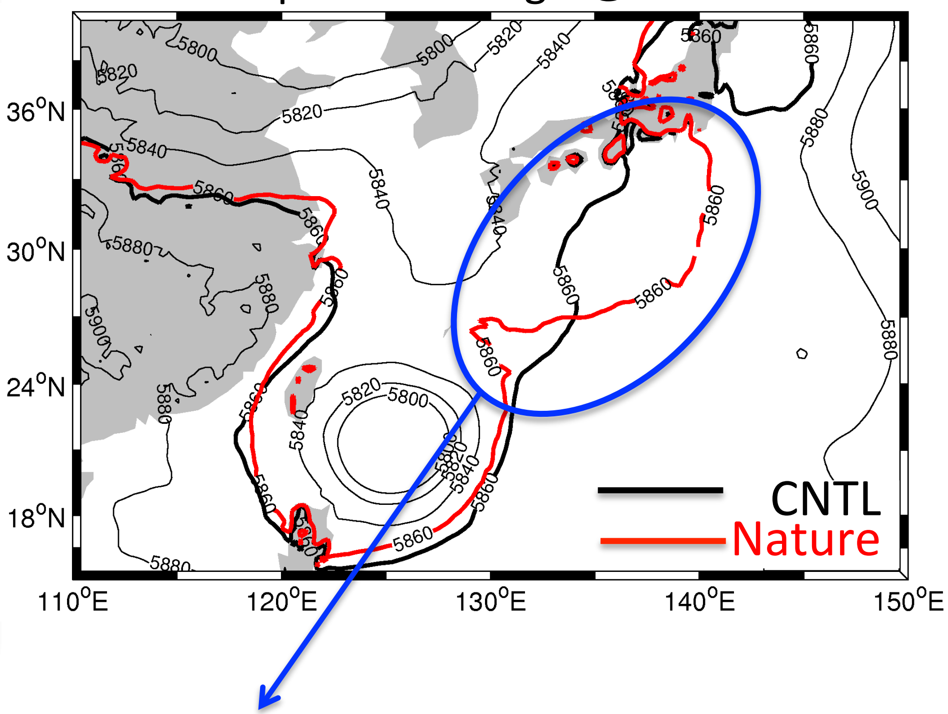
- WRF-model resolution: 27km (200x180x27)
- WRF-LETKF with 32 members
  - Conventional sondes over land and dropsondes from penetrating flight (no information from the ocean)
- ESVs are derived with the LETKF analysis ensemble and 6-hr forecast ensemble.
  - perturbation norm: kinetic energy
  - Target domain: a 1500kmx1500km box centered at the TC center

CNTL Track prediction

Track prediction initialized from LETKF analysis



Geopotential height@500hPa

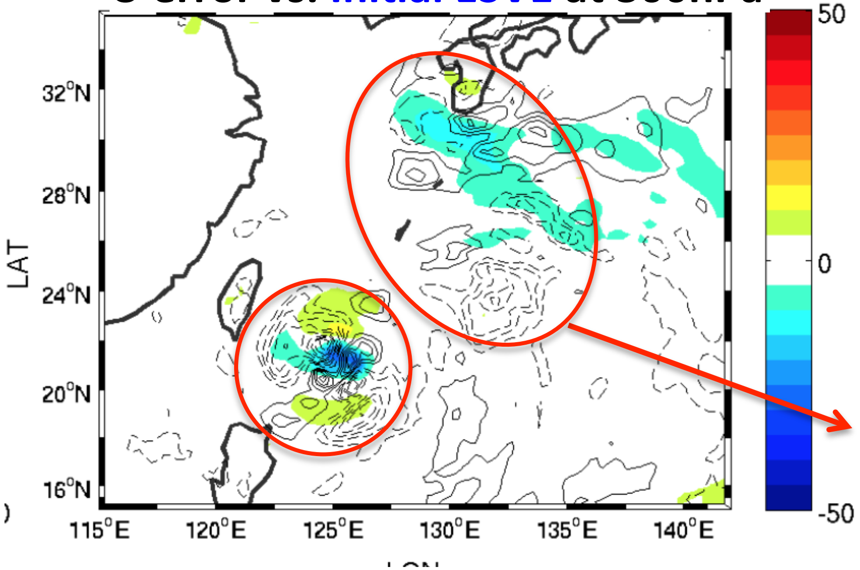


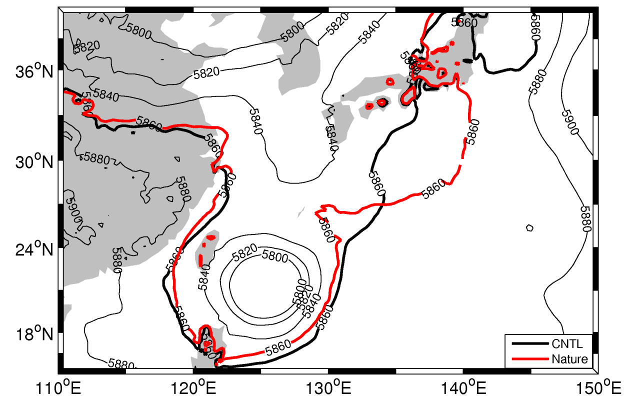
Large track errors when initialized from 18Z 09/10 18Z to 0911 06Z.

Errors associated with the wrong northeastward TC movement.

6-hr ESV1 (contour) vs. U errors (color)

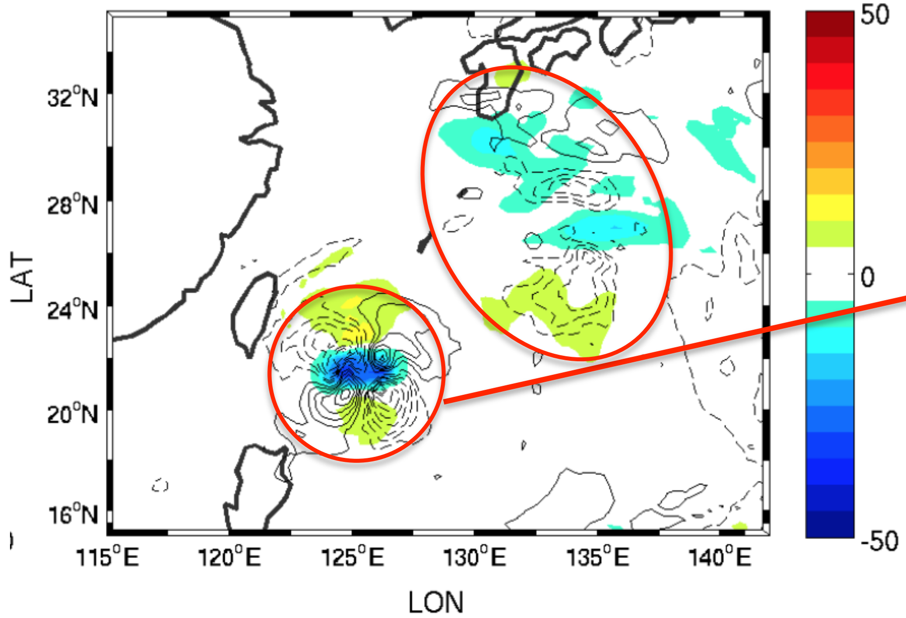
U error vs. Initial ESV1 at 300hPa





Dynamical sensitivity
associated with TC movement
(environment, subtropical high)

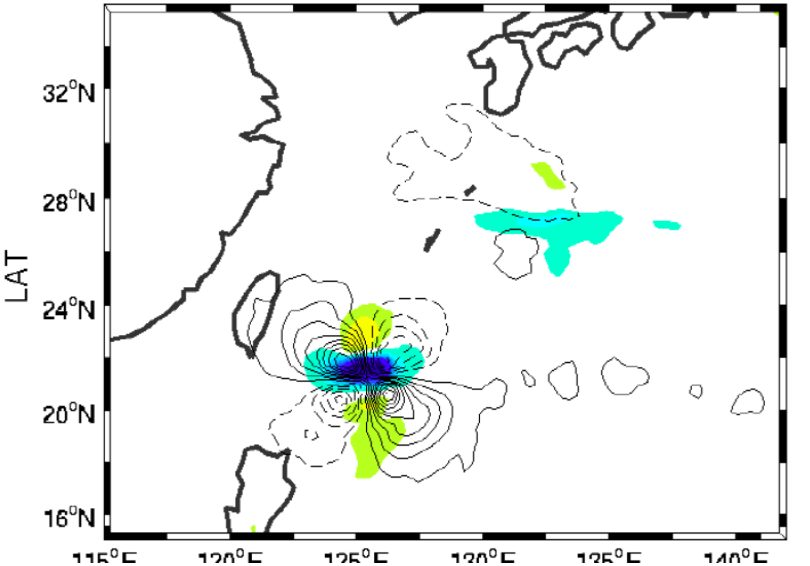
U error vs. Final ESV1 at 300hPa



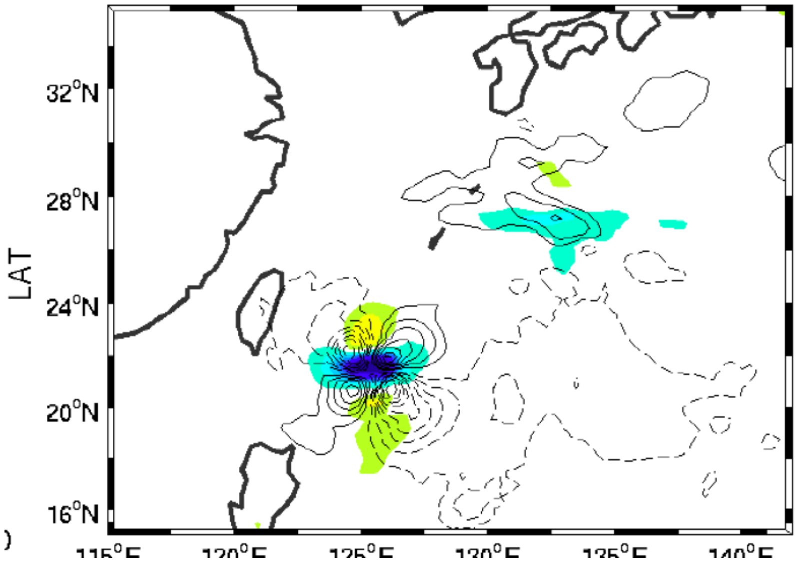
Dynamical sensitivity
associated with TC-circulation

FESV vs. ensemble perturbation

U error vs. background pert#1 at 500hPa



U error vs. final ESV1 at 500hPa

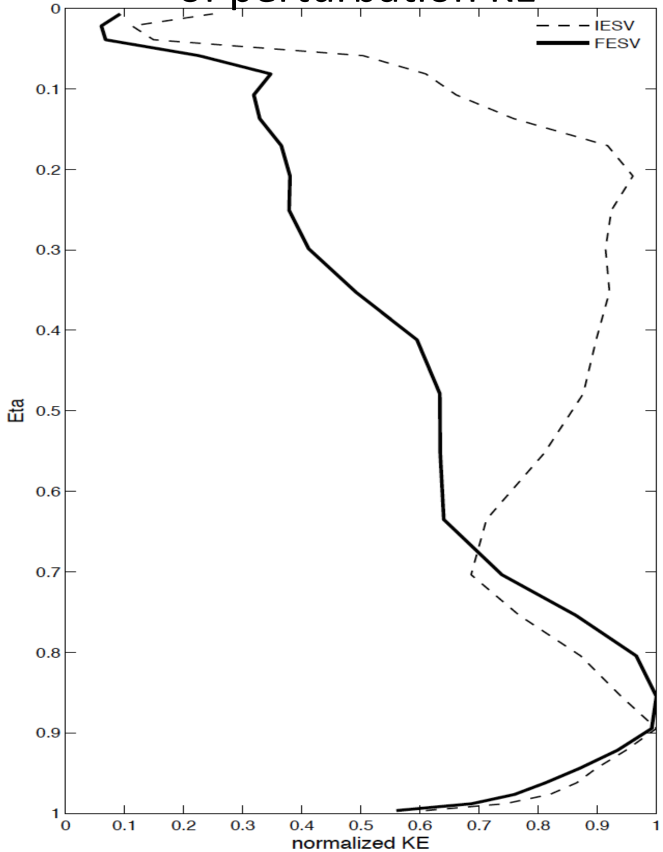


Ensemble perturbations are dominated by TC circulation related structures.

ESV1 shows dynamic sensitivities in the area of TC and its environment.

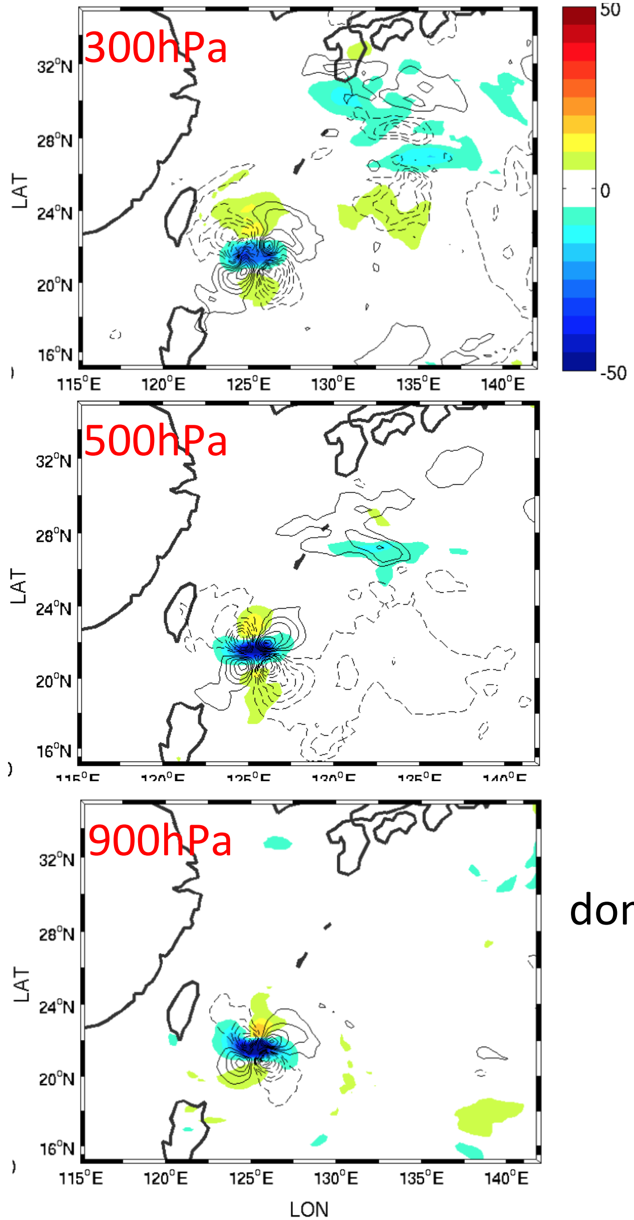
Vertical structure of TC-associated ESV1

Vertical distribution of perturbation KE



Initial ESV1

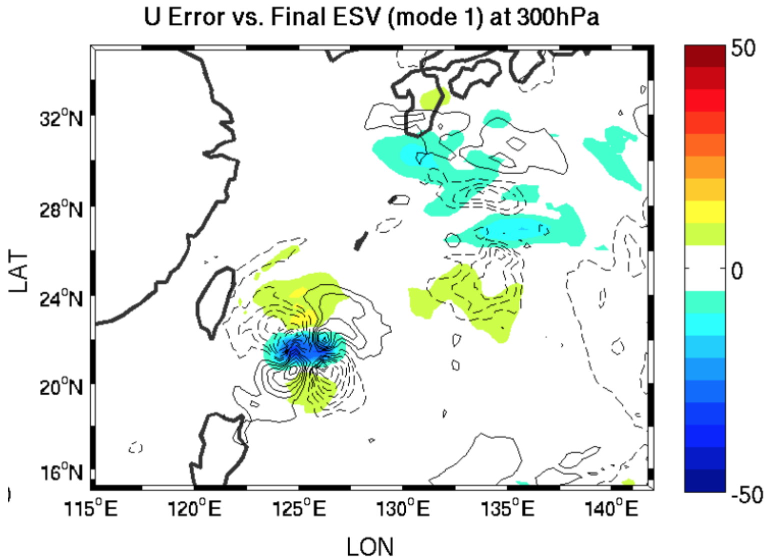
Final ESV1

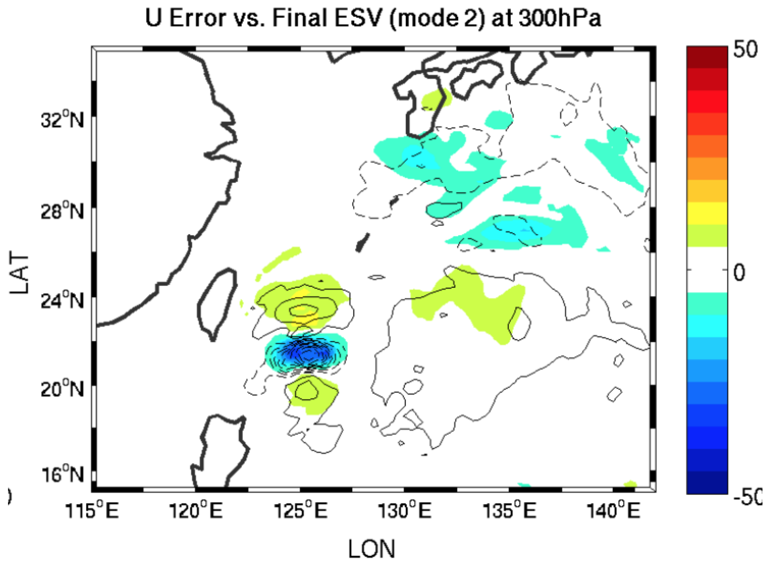


Environmental sensitivity appears in the mid-upper levels

dominated by TC-circulation

FESV1 vs. FESV2

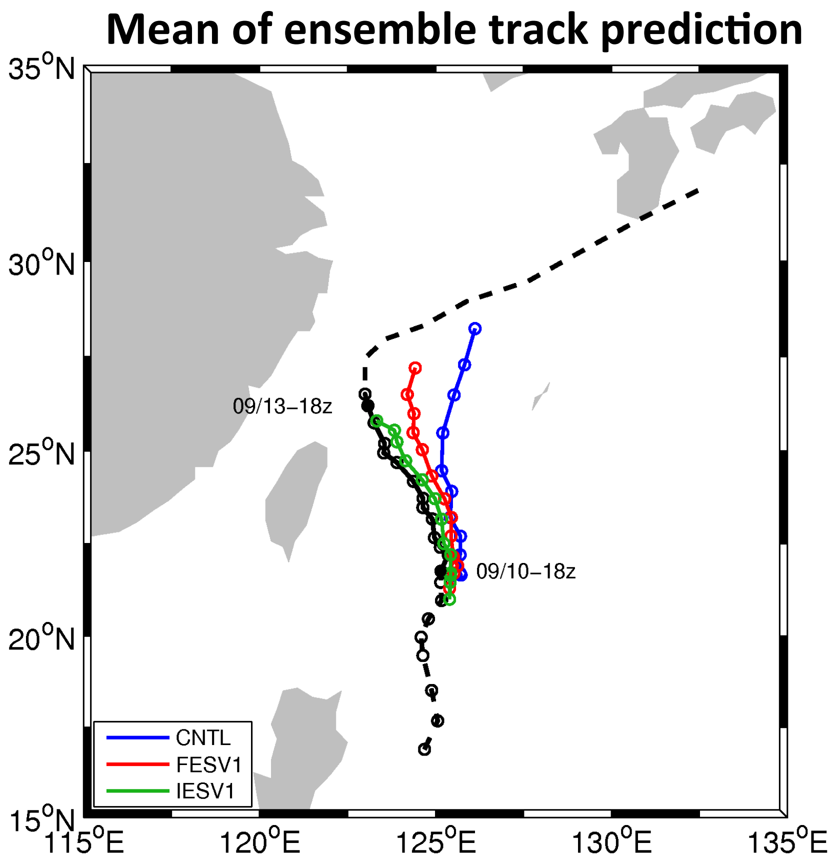




- In both FESV1 and FESV2, the sensitivity region are related to TC circulation and movement.

Apply ESV to TC track prediction

CNTL: ensemble forecast initialized from the WRF-LETKF analysis ensemble
FESV1: add/minus the final ESV1 to the 6-hr forecast ensemble
IESV1: add/minus the initial ESV1 to the analysis ensemble



With either initial ESV or final ESV1, the track prediction is significantly improved!

Summary (I)

- ESVs representing the fast growing errors can be derived without the use of tangent linear/adjoint model.
  - Constructing ESVs with a EnKF framework is almost cost-free.
- ESVs better correlated with the background errors than the ensemble perturbations.
  - When ensemble perturbations are still developing the flow-dependent structure, ESVs are able to capture the fast growing modes already.
- ESVs can be used as the additive covariance inflation for EnKF.
  - Positive impact is particularly identified for area with the large errors.
- ESVs can be used to represent dynamical sensitivities related to TC circulation and movement.
  - Applying the TC-associated ESV to the WRF-EPS can improve the TC track prediction.
  - TC-associated ESV will be used in the WRF-LETKF system as the additive covariance inflation.

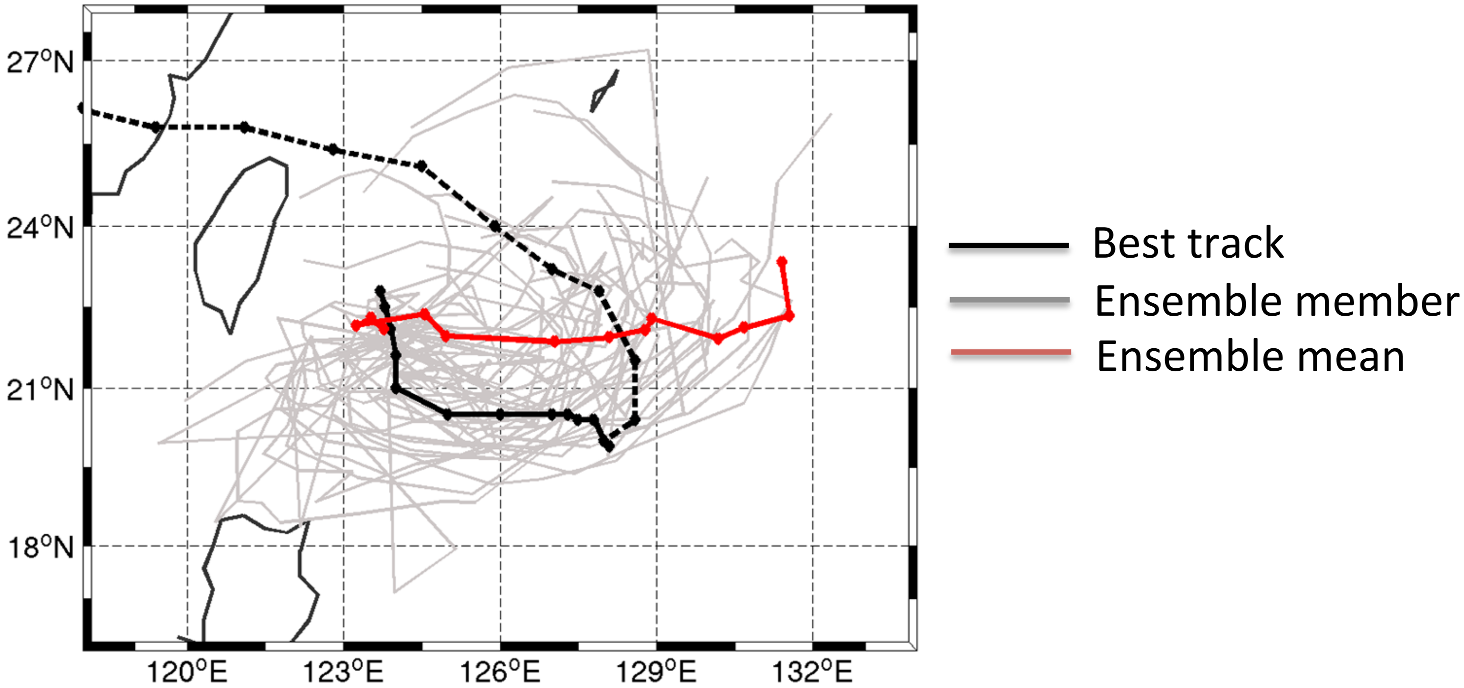
Turning a poor ensemble into a useful ensemble

Mean re-centering method

- If the ensemble is normally distributed, the mean state is the most likely state and can be used as the optimal estimation of the atmospheric state.
- However, if the ensemble violates the Gaussian distribution, the mean state is not representative for the best estimate.
- With a poor ensemble, the ensemble mean cannot well represent the behavior of the realistic state. BUT, it is possible that some ensemble members can!
- The purpose of mean recentering scheme (MRC) is to use the information from the good members to 'recenter' the ensemble and aims to improve the nonlinear evolution of ensemble.
- Coupled the MRC method with EnKF can bring positive feedback for improving the EnKF performance (Chang et al. 2014).
- The MRC method shares a similar idea proposed by the ensemble recentering Kalman filter (ERKF, Keppenne 2013).

Ensemble forecast for Trami (2013)

Ensemble forecast initialized at 00z 16 Aug. 2013



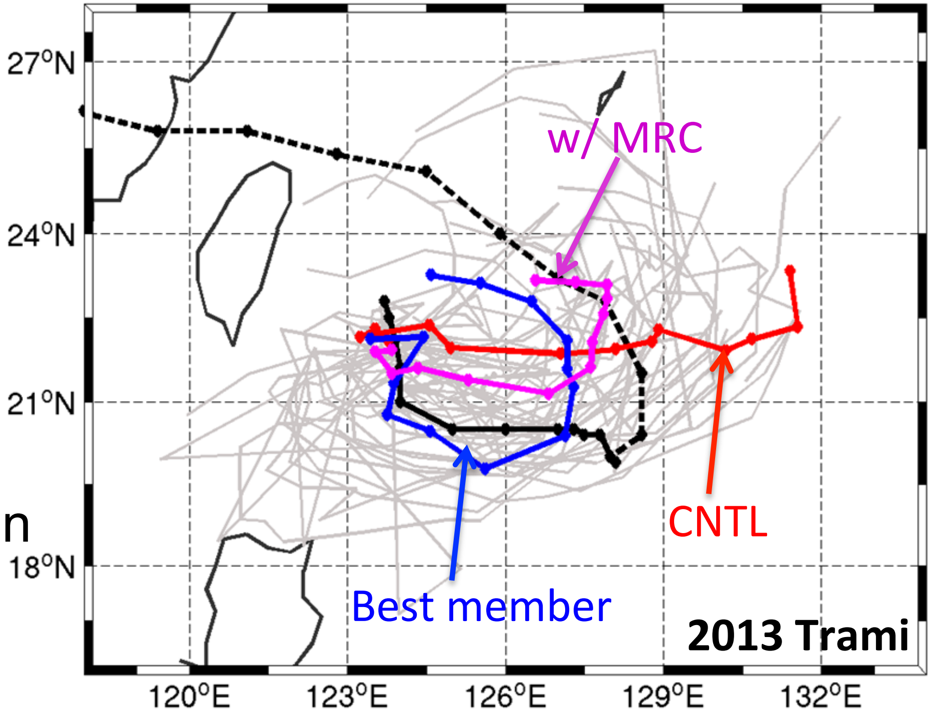
- Cold-started WRF-based ensemble prediction system(EPS) with perturbations sampled according to B_{3dvar}. Useless ensemble prediction?
- High uncertainties in the TC environment.
- Can we turn the poor ensemble into the useful ensemble?

Mean Recentering Scheme (MRC)

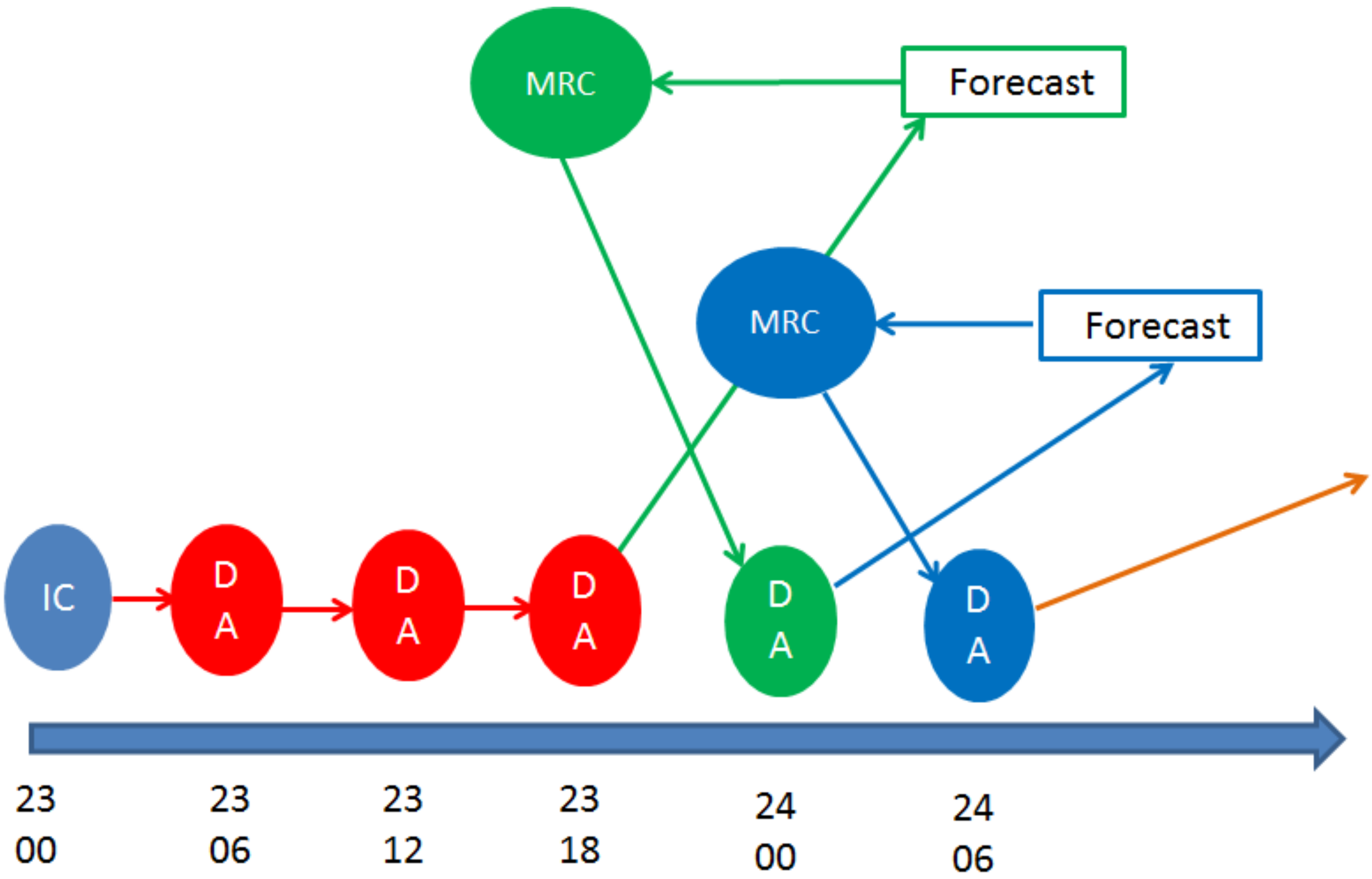
Ensemble forecast initialized at 00z 16 Aug.

Adjust the ensemble evolution

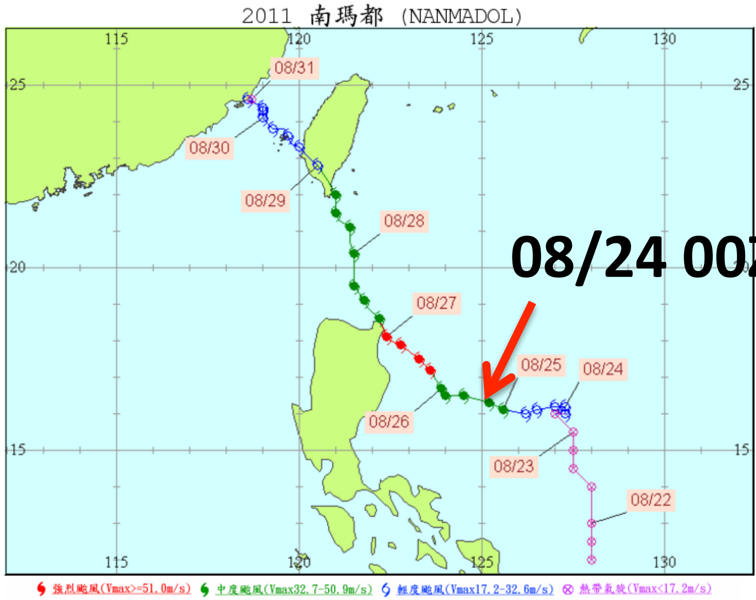
1. Select the best member based on the accumulated track error.
2. Re-center the ensemble at the initial state of the best member
3. Re-evolve the ensemble upon the best member



The MRC cycle

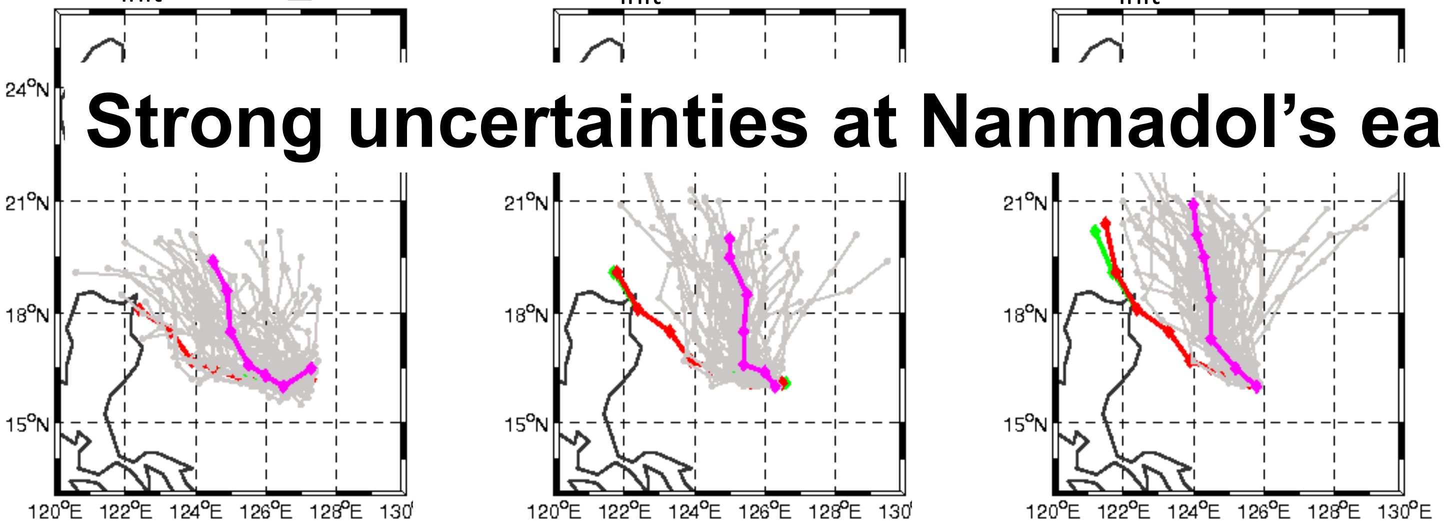


Typhoon Nanmadol (2011)



— TIGGE EC ens. fcst (51)
— Mean forecast
— Analysis
— CWB best track

t_{init}=00Z 8/24 t_{init}=12Z 8/24 t_{init}=00Z 8/25

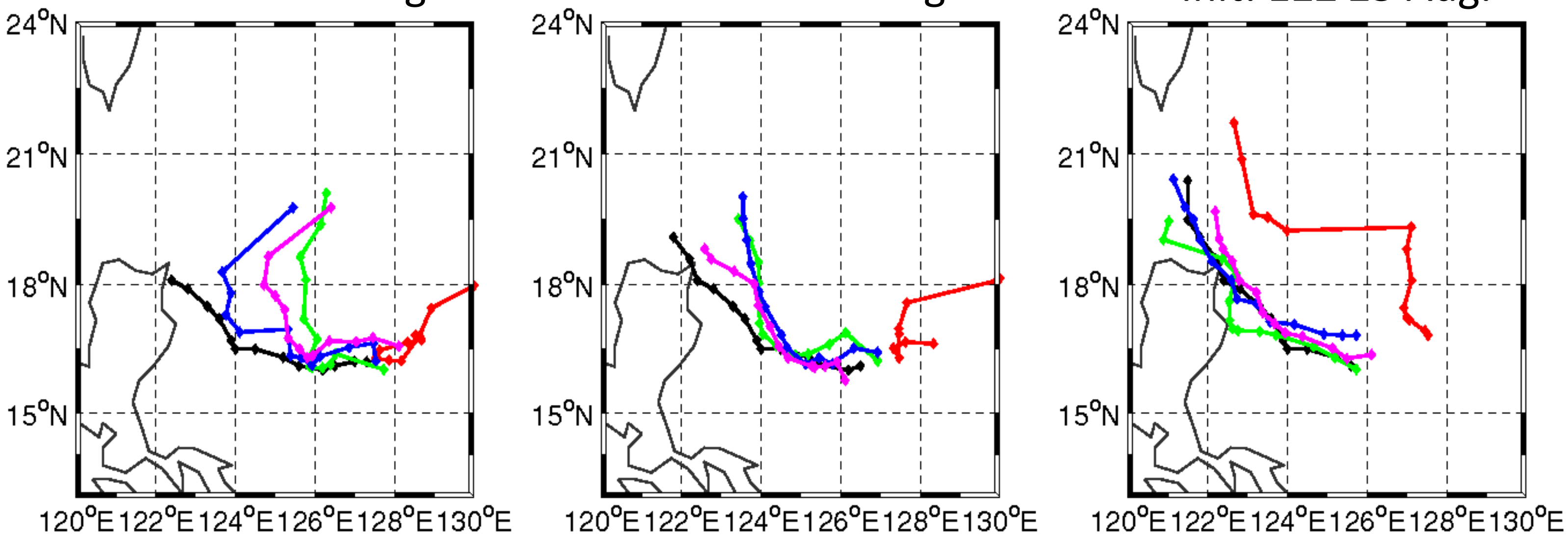


Strong uncertainties at Nanmadol's early stage

WRF-LETKF (warm-started) EPS for 2011 Namado

- CWB best track
- DA_CNT: regular Data assimilation cycle
- MRC_DA24a:One best member with smallest track error.
- MRC_DA24b: Use average of the best five members as best member.
- MRC_DA24c: Use average of the best group derived by cluster analysis.

Init: 00Z 24 Aug. Init: 12Z 24 Aug. Init: 12Z 25 Aug.

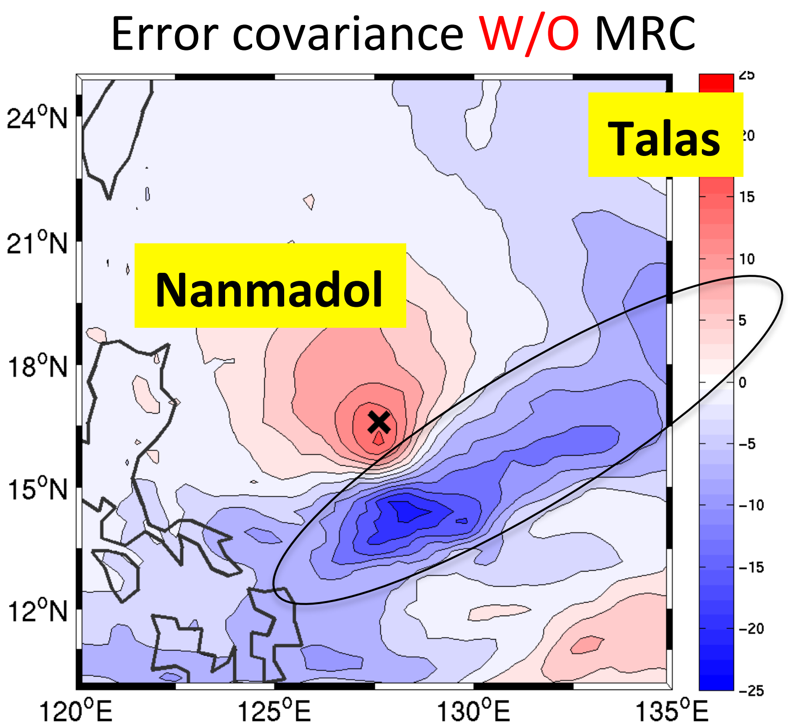


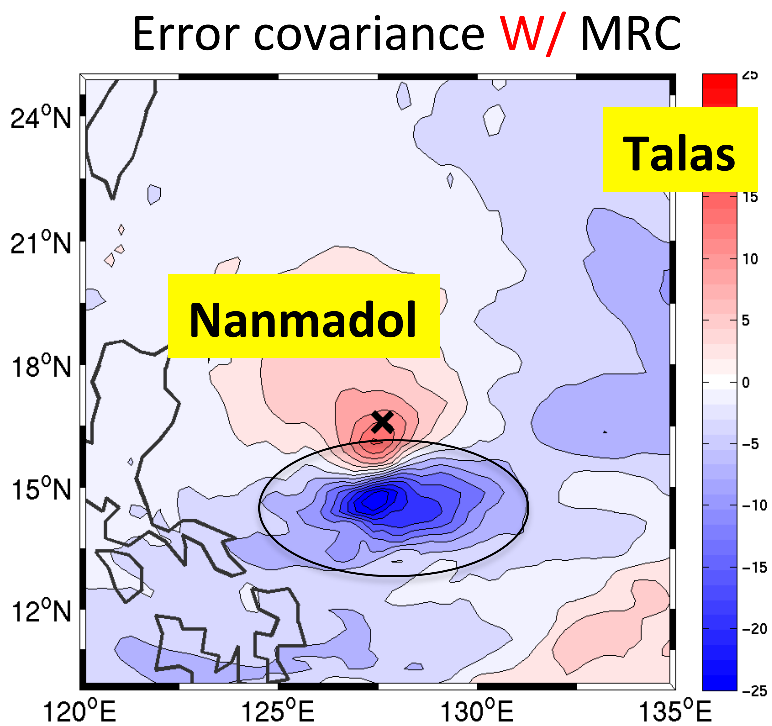
2nd DA/MRC cycle 4th DA/MRC cycle 6th DA/MRC cycle

Positive feedback further improves the analysis and forecast!

Chang et al. (2015)

Error covariance structure



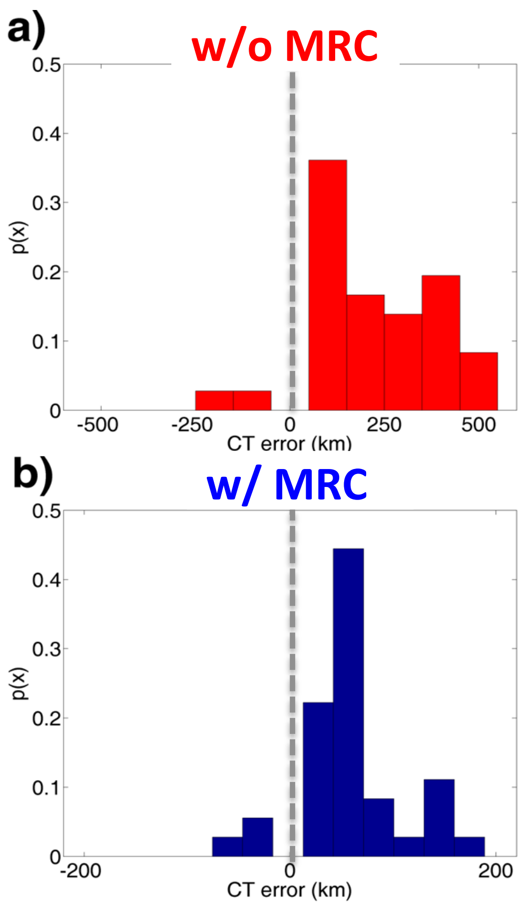


- Without MRC, the negative covariance structure is very board and link to another TC, Talas.
- With MRC, the covariance structure is more symmetric.

Gaussianity

The Probability distribution of TC cross track error

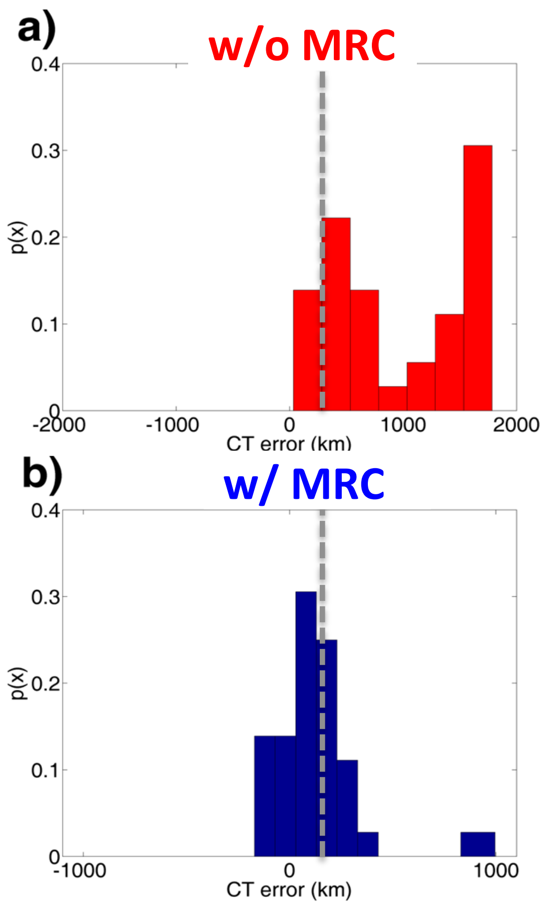
Analysis ensemble at 00Z 8/24



Skewed ensemble with a large uncertainty

Gaussianity is significantly improved.

2-day **forecast** ensemble initialized at 00Z 8/25

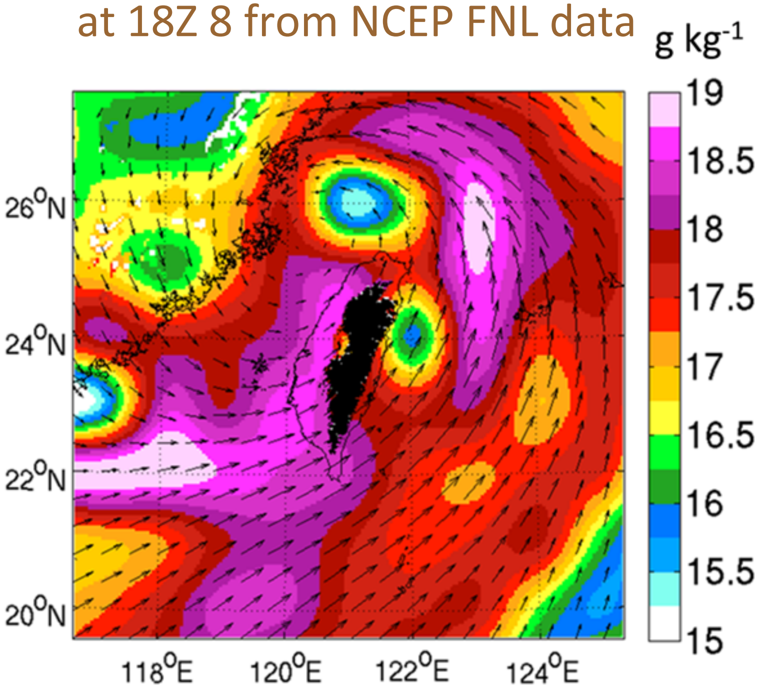


Application of the MRC method to the convective-scale nowcasting

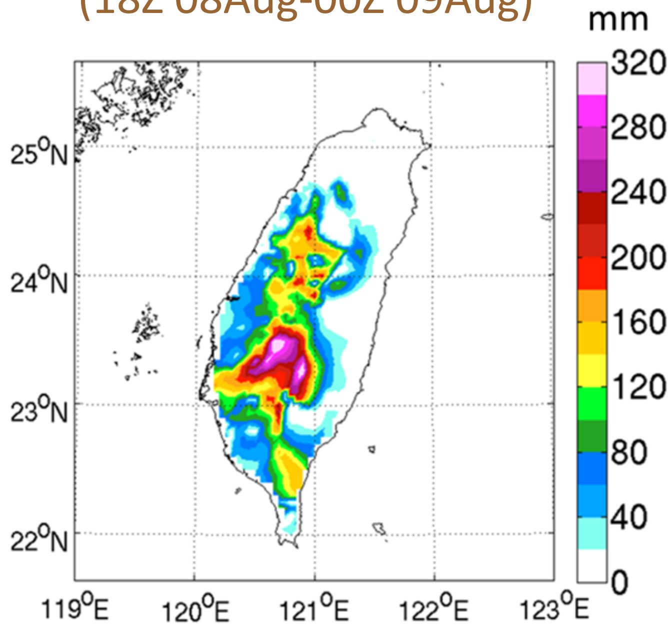
(Wu et al. 2015)

Heavy rainfall event (2008 Morakot)

1-km horizontal wind and q_v
at 18Z 8 from NCEP FNL data



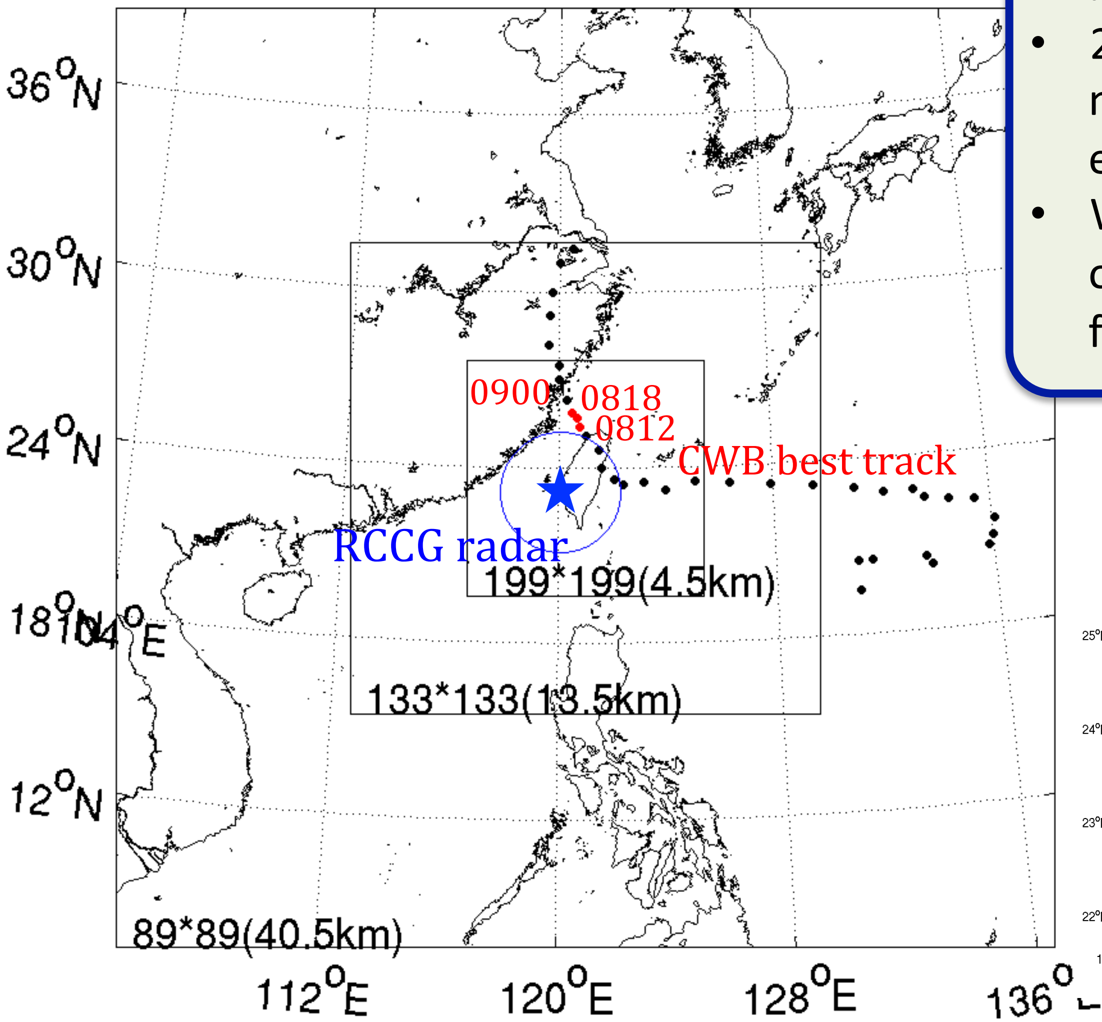
6-hour accum. rainfall
(18Z 08Aug-00Z 09Aug)

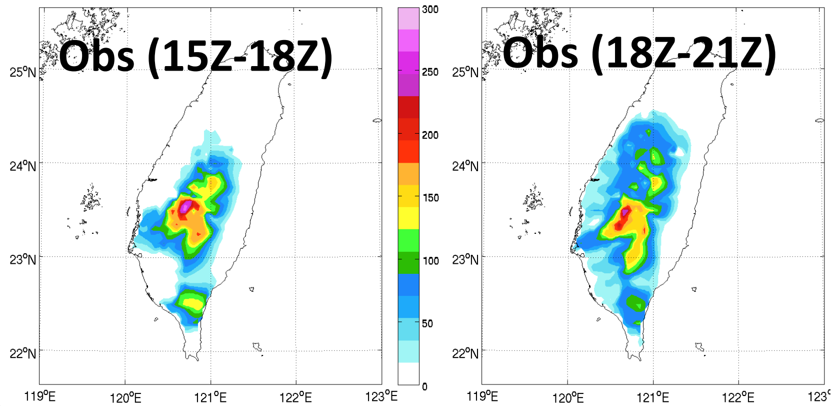


- Reach a minimum moving speed of 2.6 m/s
- Convergence between moist southwest monsoon and typhoon circulation
- Strong terrain effect that enhances rainfall

WRF-LETKF Radar assimilation system for improving QPN

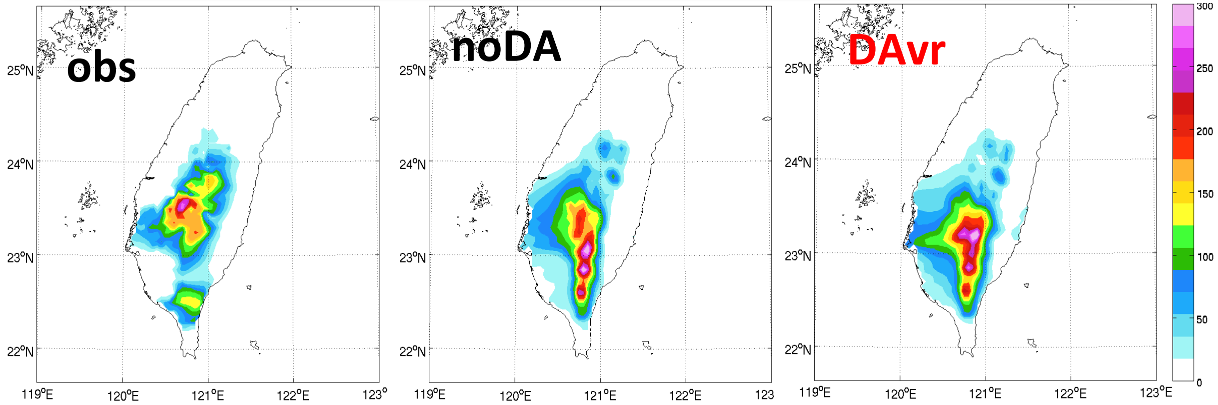
- Assimilate the $\boldsymbol{V_r}$ **and** $\boldsymbol{Z_h}$ from the RCCG radar (Tsai et al. 2014)
- 2-hour assimilation period with a 15-min cycle interval ; initialize 6-hr ensemble forecasts
- When initialized at 15Z 08Aug, observation impact beyond 3-hr forecast is limited.



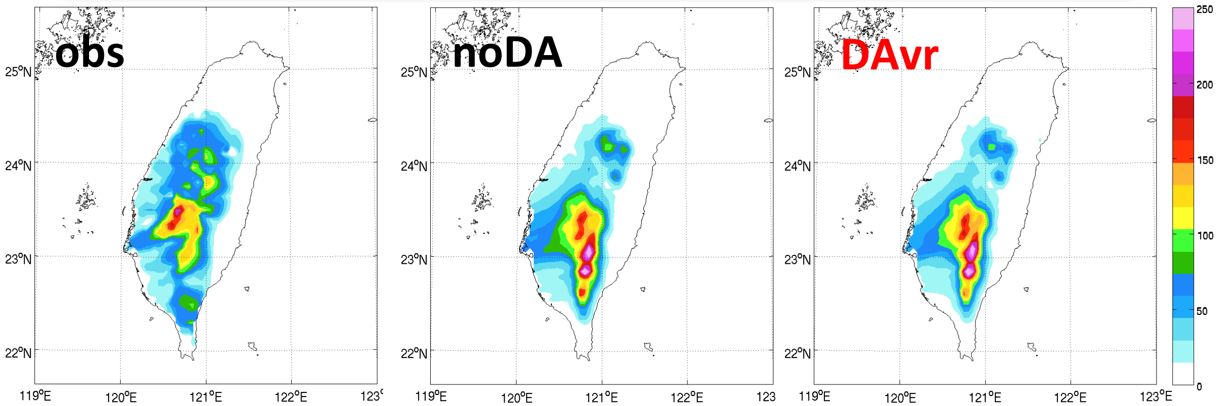


Accumulated rainfall

Accumulated rainfall (first 3 hrs, 15Z-18Z)

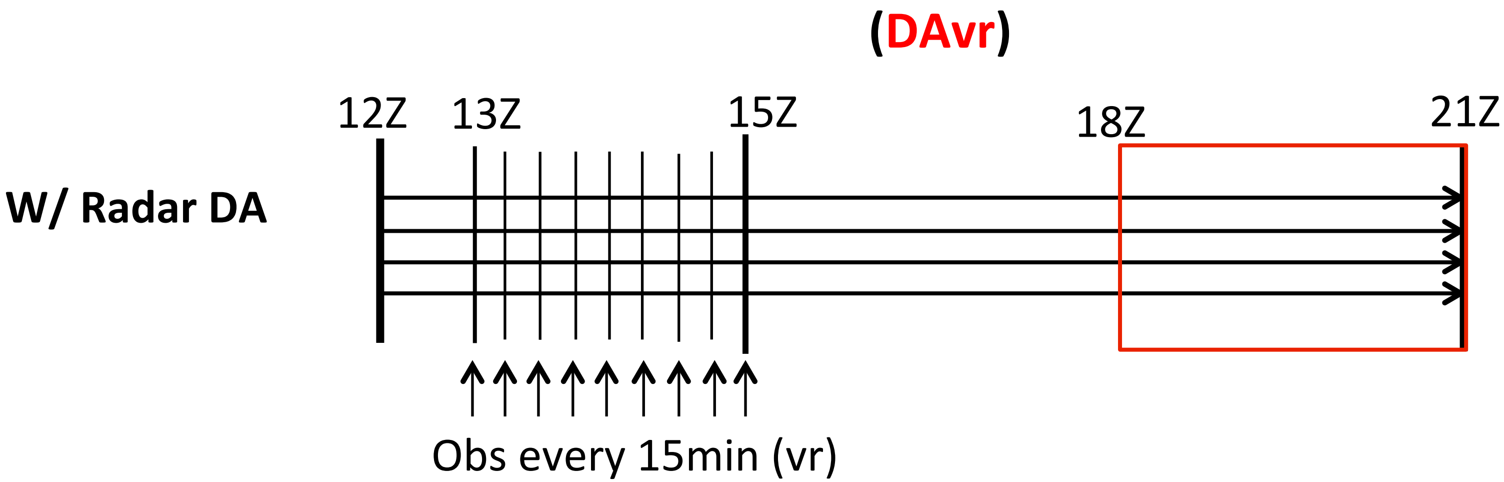


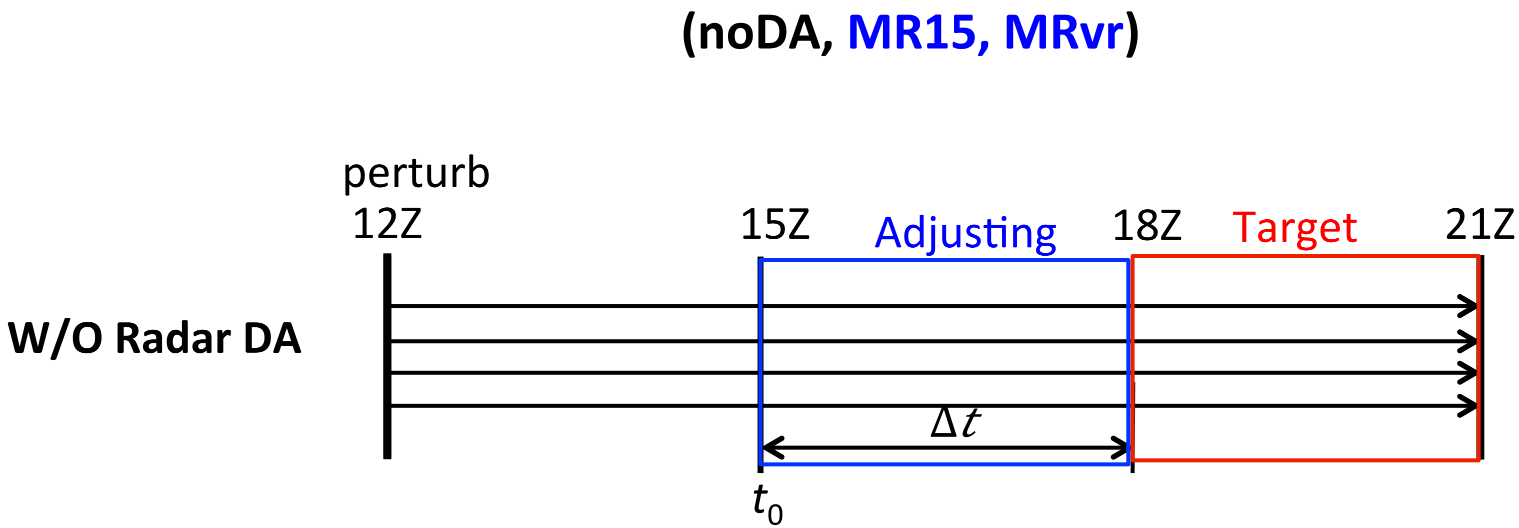
Accumulated rainfall (later 3 hrs, 18Z-21Z)



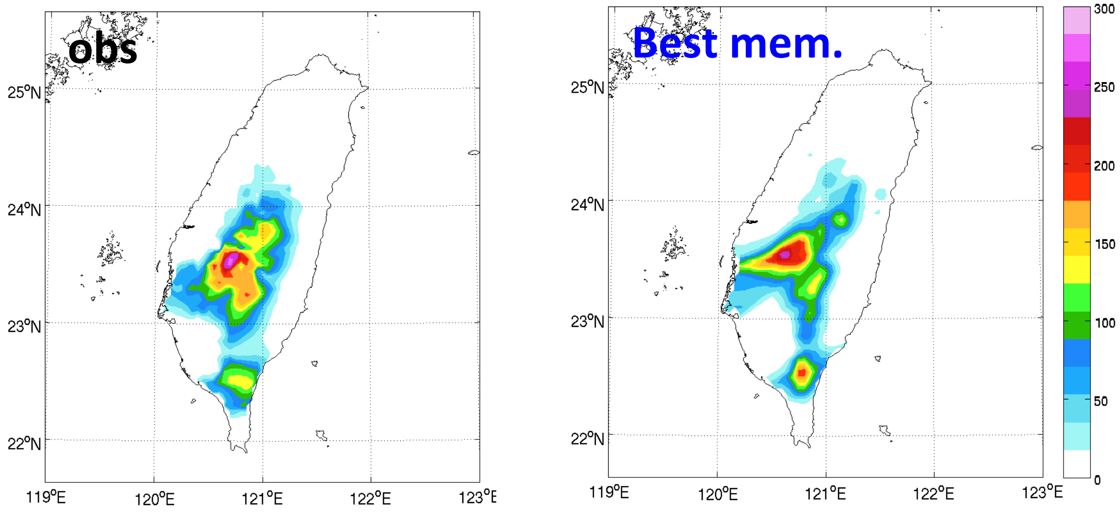
- Ensemble mean derived from the PM method
- Excessive rainfall on the southern mountainous area.
- Impact from radar assimilation on PF is limited after 3-hr forecast.
- Apply the MRC method to improve the PF at 18-21Z.

Experiment setup





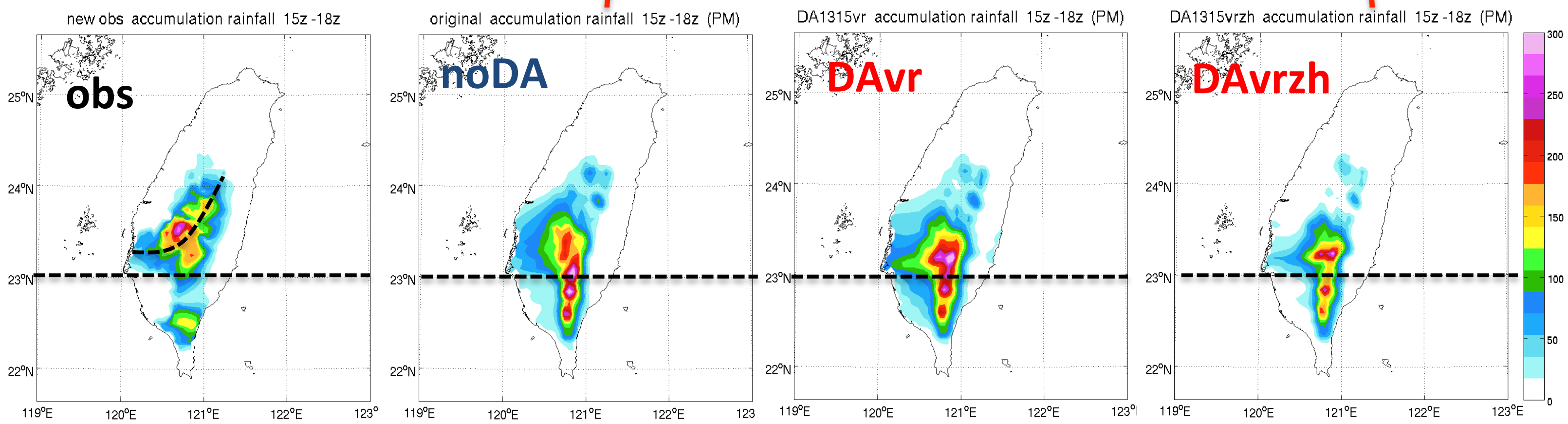
Selection of the best member



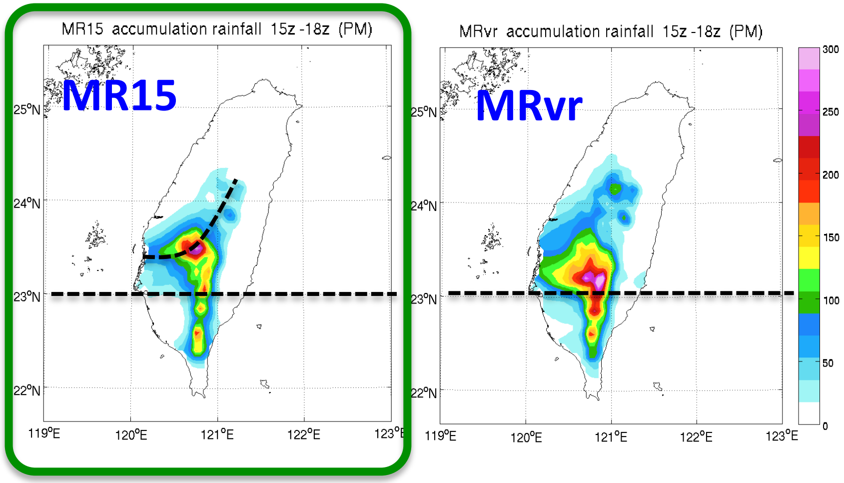
- The best member is selected based on the 15-18Z accumulated rainfall.
- Even with model bias, it's possible to select a member with a similar precipitation pattern as the observation.

Accumulated rainfall during the adjusting period

Excessive rainfall in the southern mountain area

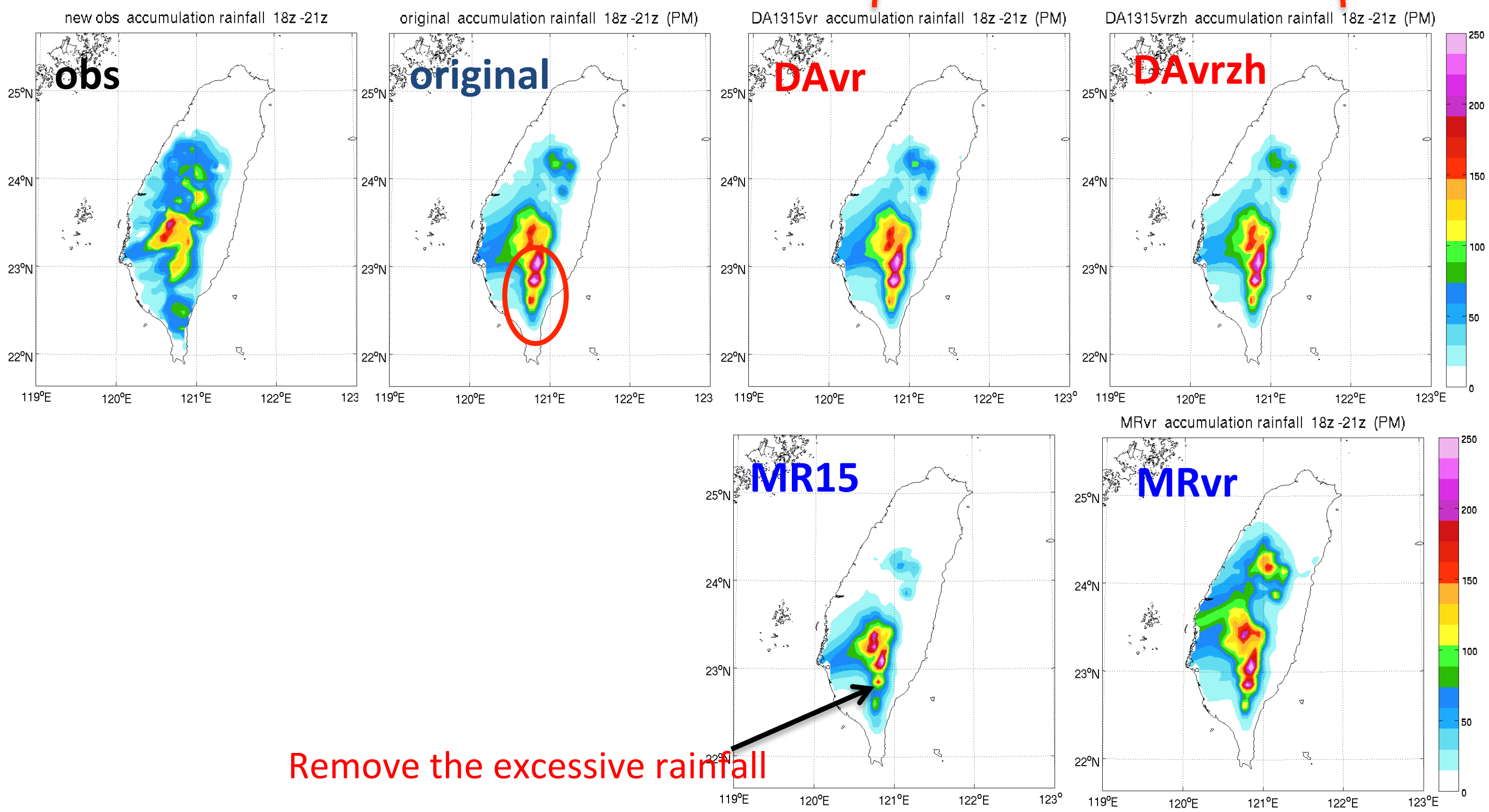


Adjust the ensemble with the best member can improve the ensemble evolution



Accumulated rainfall during the target period (18Z-21Z)

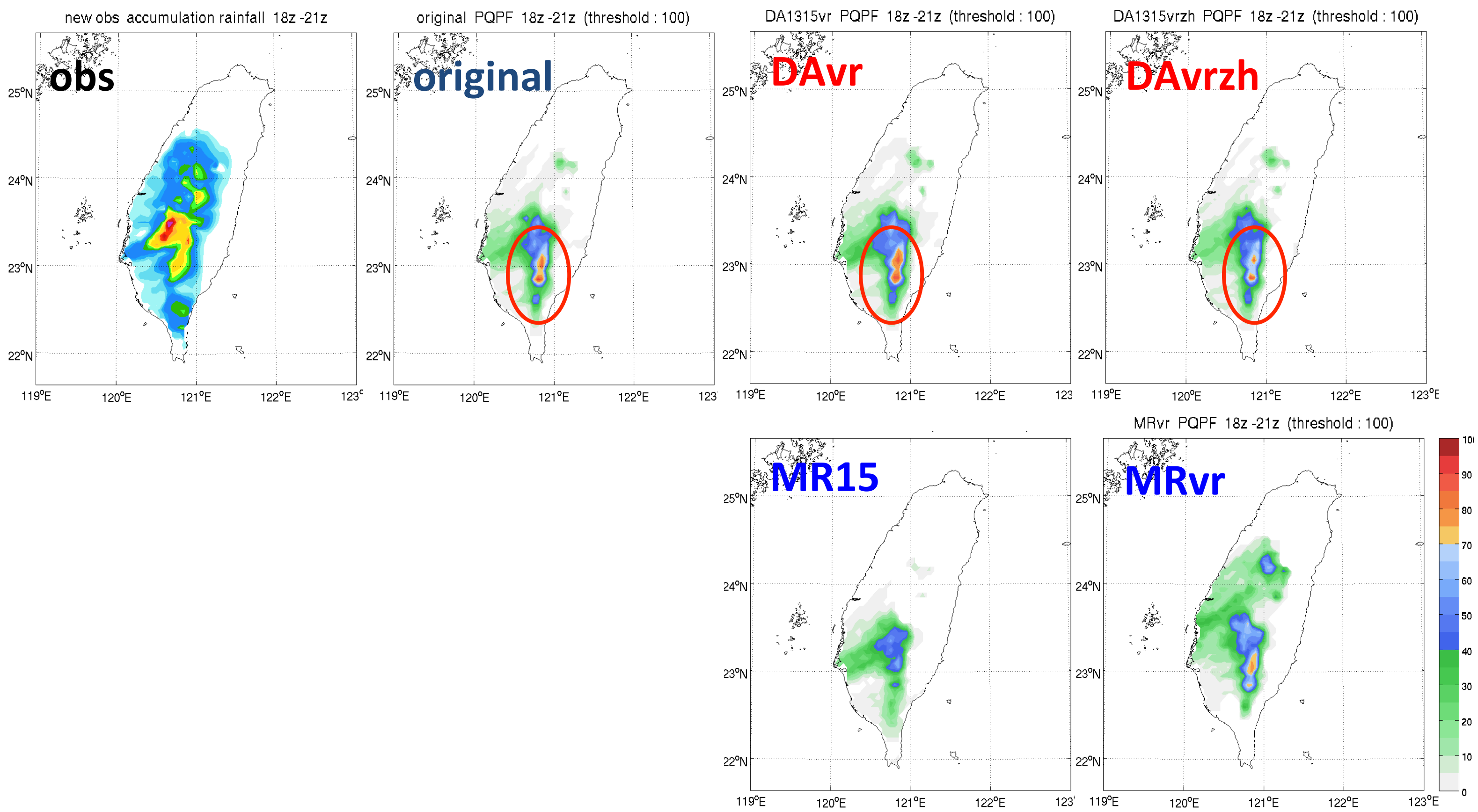
Limited impact from assimilating radar data



Remove the excessive rainfall

The rainfall on western coast is enhanced

Probability Quantitative Precipitation nowcasting during target period (18Z-21Z)

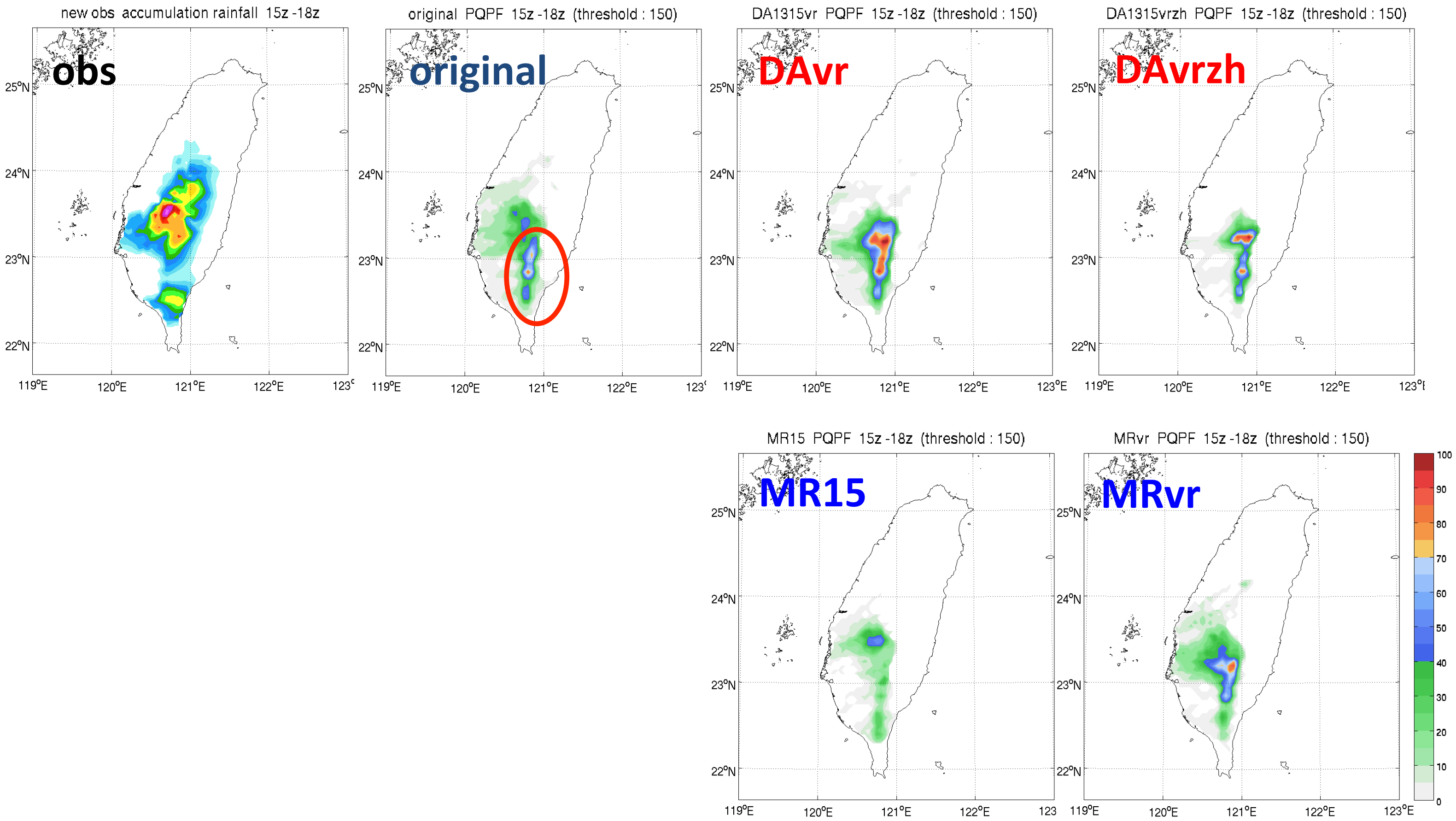


- With MRC, the ensemble evolution can be better adjusted toward the observed state.
- With a proper metric for selecting the best member, the issue of model error can be alleviated.

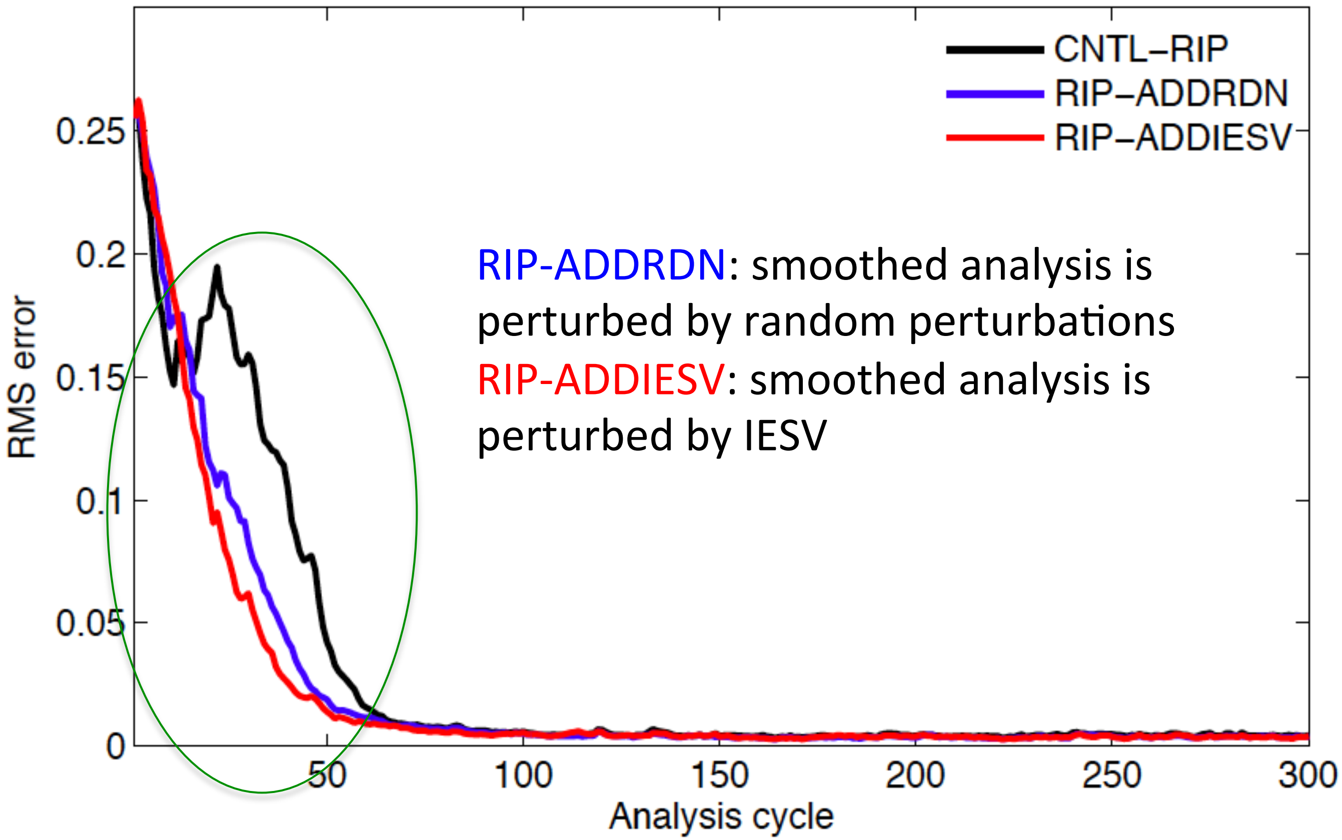
Summary (II)

- The MRC method can improve a poor ensemble, which suffers from strong uncertainties and non-Gaussian distribution.
- The positive impact from MRC can feedback to the DA system, further improving the DA spin-up, background error covariance and analysis accuracy.
- Although MRC requires future information to determine the best member, for operational purpose, this could be an valuable trade off to adjust the ensemble at the early TC developing stage and alleviate the impact from model errors.
- With the MRC method, the usefulness of the regional EPS on the convective-scale QPN can be extended.

Probability Quantitative Precipitation Forecast during adjusting period (15Z-18Z)

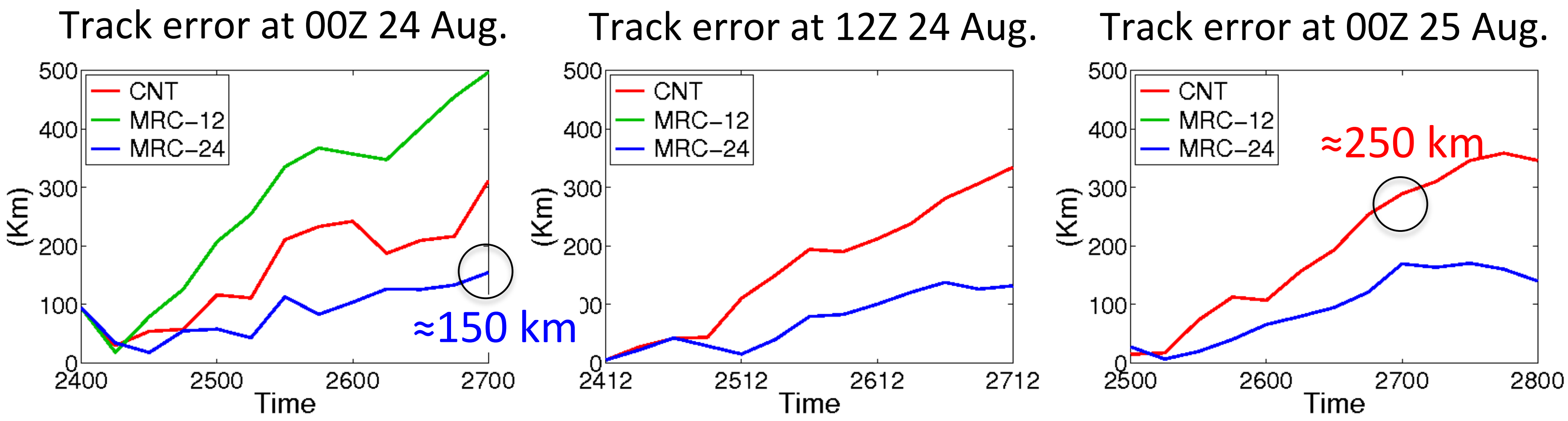


Apply IESV in LETKF-RIP



IESVs point out the fast growing directions of the errors and can further accelerate the spin-up of LETKF!

Track error of cold-started EPS



Proper adjustment on ensemble can significantly improve the forecast skill

- It is excepted that MRC_24 outperforms CNT since it contains next 24-h future information.
- However, comparing MRC_24 initialized at 00Z 24 and CNT initialized at 00Z 25, MRC shows better forecast skill than CNT when the observation information is comparable.