January 18, 2013, DA Workshop, Tachikawa, Japan

Advances and Challenges in Ensemble-based Data Assimilation in Meteorology

Takemasa Miyoshi

RIKEN Advanced Institute for Computational Science

Takemasa.Miyoshi@riken.jp

With many thanks to

E. Kalnay, K. Ide, B. Hunt, S. Greybush, S. Penny, G.-Y. Lien, UMD Weather-Chaos group, Y. Ota (NCEP/EMC, JMA)

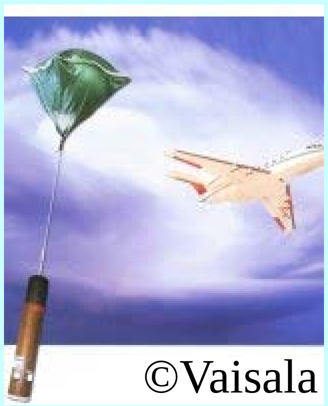
J. Ruiz (Argentina), S.-C. Yang (Taiwan),

M. Kunii, K. Kondo, T. Enomoto, and N. Komori (Japan)

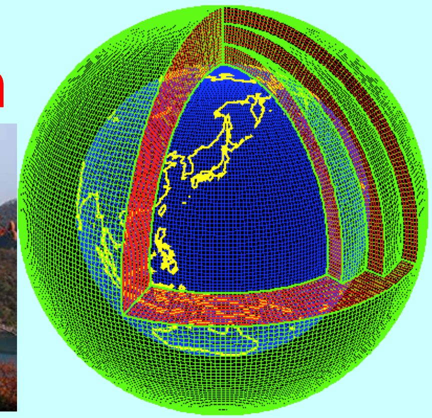
Data Assimilation (DA)

Observations

Numerical models



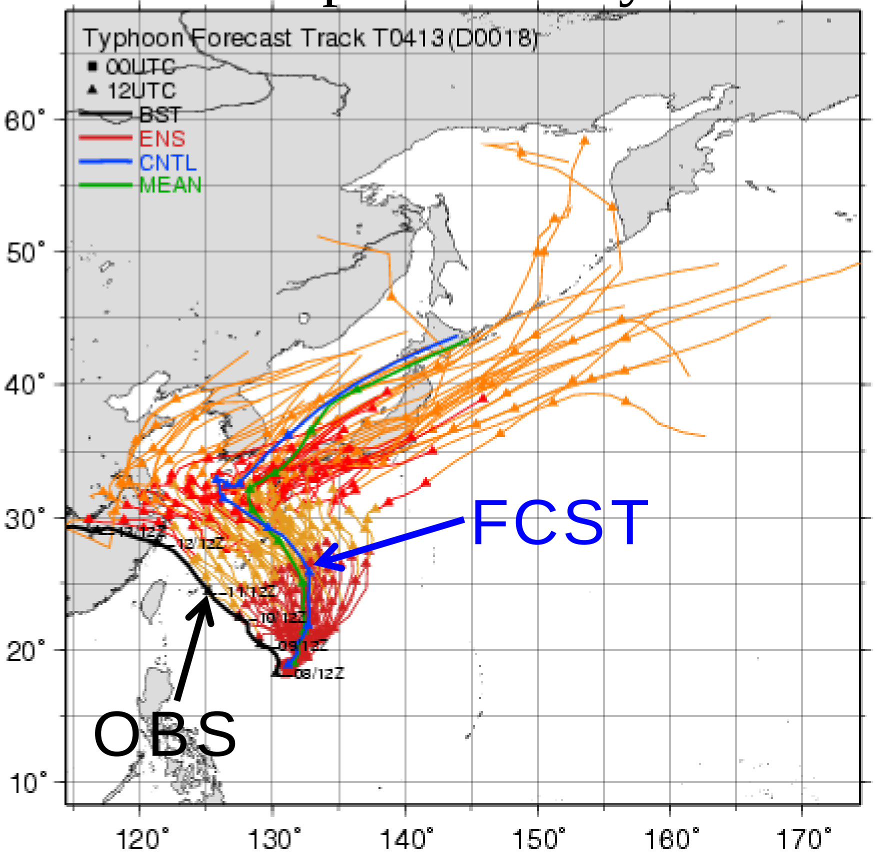
Data Assimilation



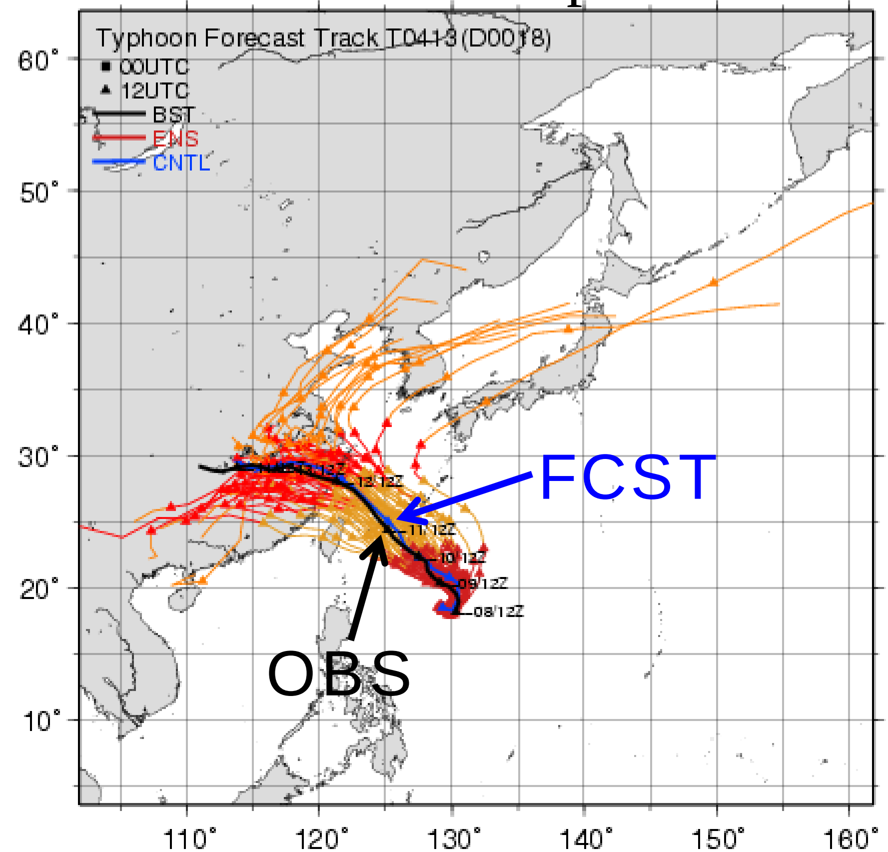
Data assimilation best combines observations and a model, and brings synergy.

DA has an impact.

SV w/ 4D-Var
JMA operational system



LETKF
under development



Miyoshi and Sato (2007)

Using the same NWP model and observations.
DA matters!

WRF-LETKF studies at UMD

National Oceanographic Partnership Program

Goal: Improve TC forecasts by improving the initial conditions

- WRF-LETKF system (Miyoshi and Kunii 2012)
  - LETKF: Local Ensemble Transform Kalman Filter (Hunt et al. 2007)
  - Adaptive inflation method (Miyoshi 2011)

- Several techniques to improve TC forecasts have been explored.
  - Running-In-Place method (Yang, Miyoshi and Kalnay 2012)
  - Including SST uncertainties (Kunii and Miyoshi 2012)
  - Assimilating AIRS retrievals (Miyoshi and Kunii 2012)
  - Estimating observation impact using the ensemble-based method (Kunii, Miyoshi and Kalnay 2012)
    - Based on the ensemble sensitivity method of Liu and Kalnay (2008)
  - Two-way nested WRF-LETKF for higher-resolution experiments

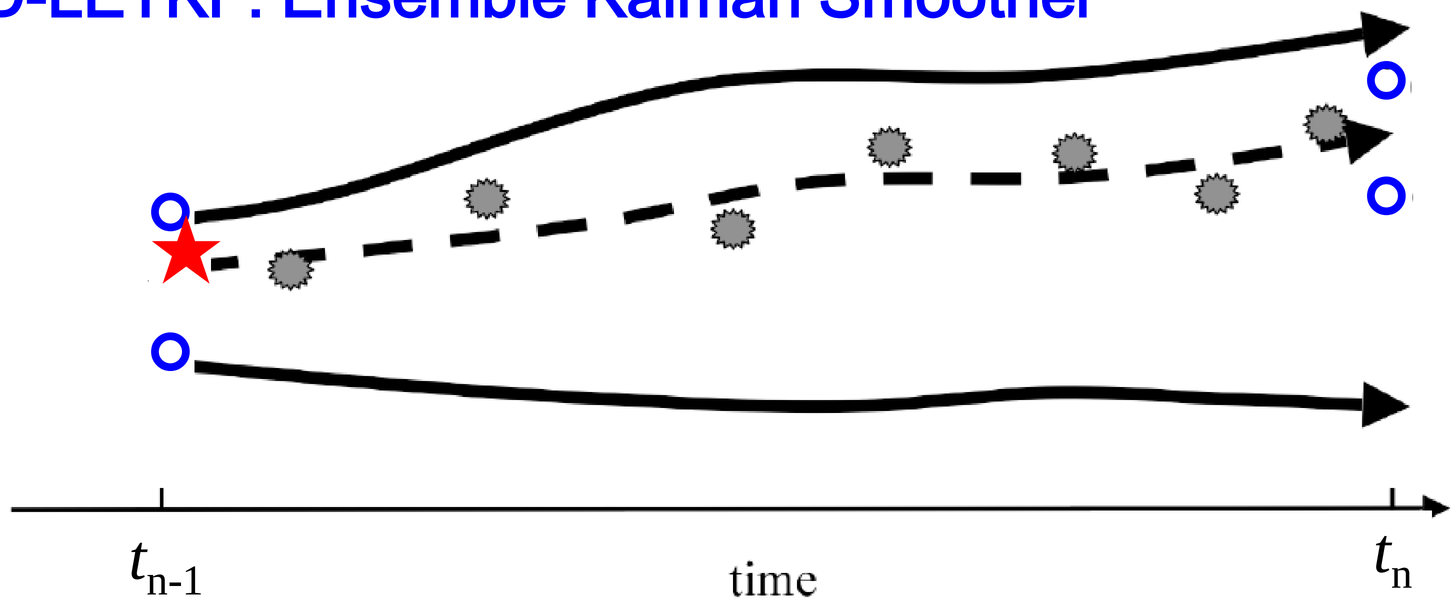
Studies on methods: towards optimal use of available observations

RUNNING-IN-PLACE (RIP)

Yang, Kalnay, and Hunt (2012, in press)
Yang, Miyoshi and Kalnay (2012, in press)
Yang, Lin, Miyoshi, and Kalnay (in progress)

Running-In-Place (RIP, Kalnay and Yang 2008)

4D-LETKF: Ensemble Kalman Smoother



$$\tilde{\bar{\mathbf{x}}}_a(t_{n-1}) = \bar{\mathbf{x}}_a(t_{n-1}) + \mathbf{X}_a(t_{n-1})\bar{\mathbf{w}}_a(t_n)$$

$$\tilde{\mathbf{X}}_a(t_{n-1}) = \mathbf{X}_a(t_{n-1})\mathbf{W}_a(t_n)$$

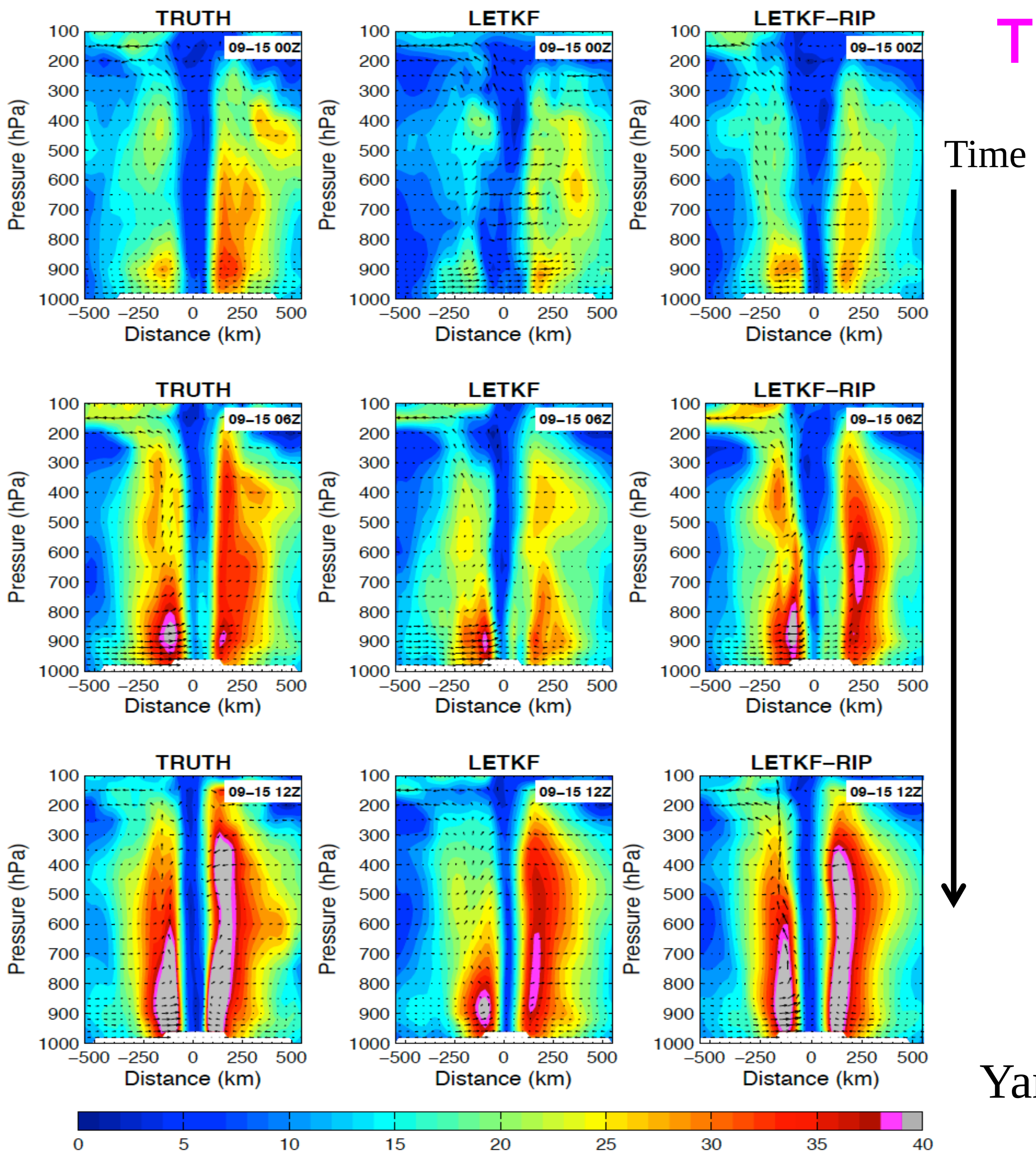
$$\bar{\mathbf{w}}_a = \tilde{\mathbf{P}}_a \mathbf{Y}_b^T \mathbf{R}^{-1}(\mathbf{y} - H(\bar{\mathbf{x}}));$$

$$\mathbf{W}_a = [(K-1)\tilde{\mathbf{P}}_a]^{\frac{1}{2}}$$

Running-In-Place (RIP) method:

1. Update the state (★) at t_{n-1} using observations up to t_n (smoother)
2. Assimilate the same observations again (dealing with nonlinearity)
3. Repeat as long as we can extract information from the same obs.

In OSSE, RIP is very promising



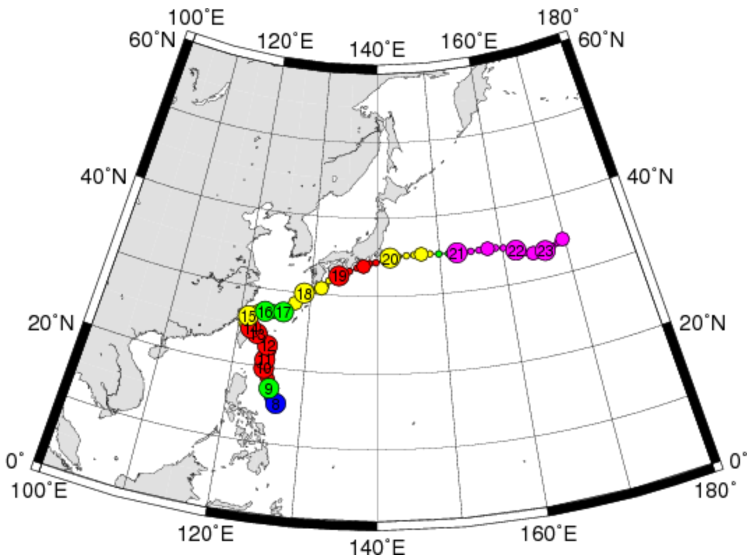
This is a simulation study.

Realistic observing systems are assumed, including dropsondes near the TC.

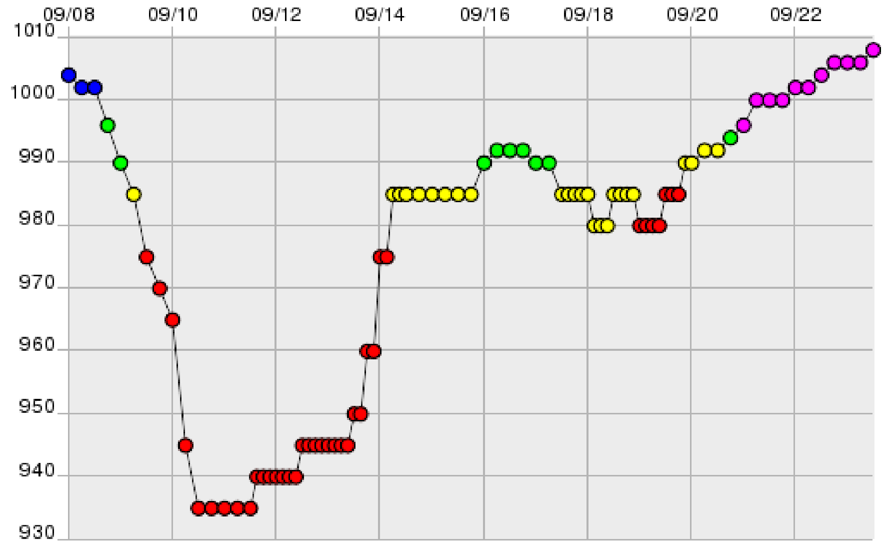
Vortex strength and structure are clearly improved.

Yang, Miyoshi and Kalnay (2012)

Typhoon Sinlaku (2008)



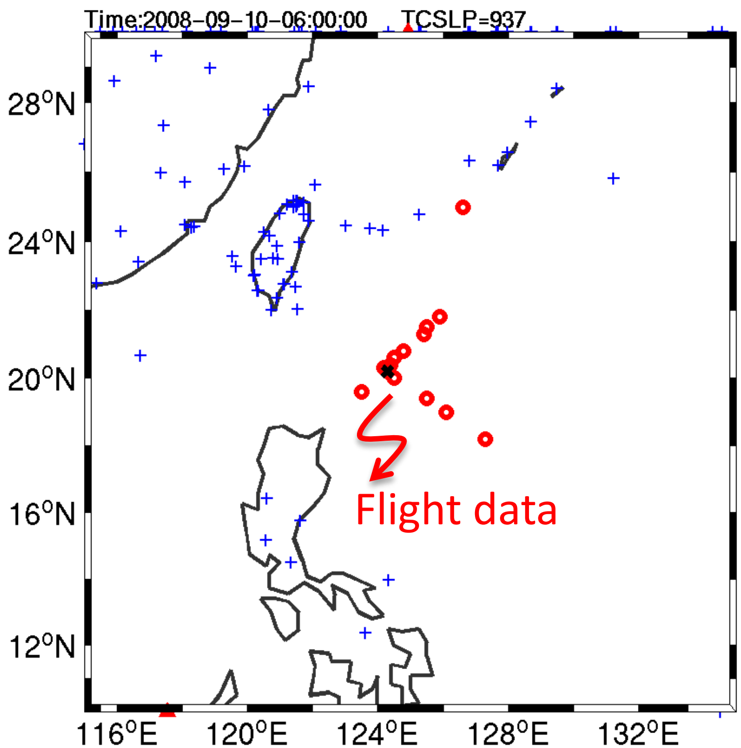
Track



MSLP

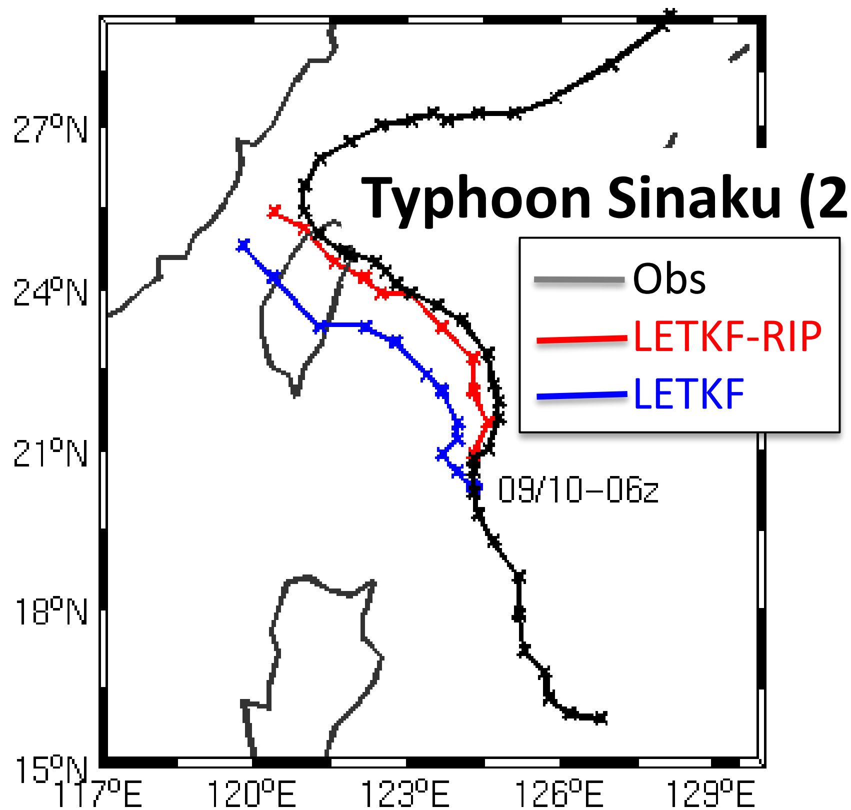
RIP impact on Sinlaku track forecast

This is the real case.



SYNOP(+),SOUND(△),
DROPSONDE(○),
Typhoon center (X)

3-day forecast



RIP better use the "limited observations"!

S.-C. Yang (2012)

WRF-LETKF: including additional sources of uncertainties

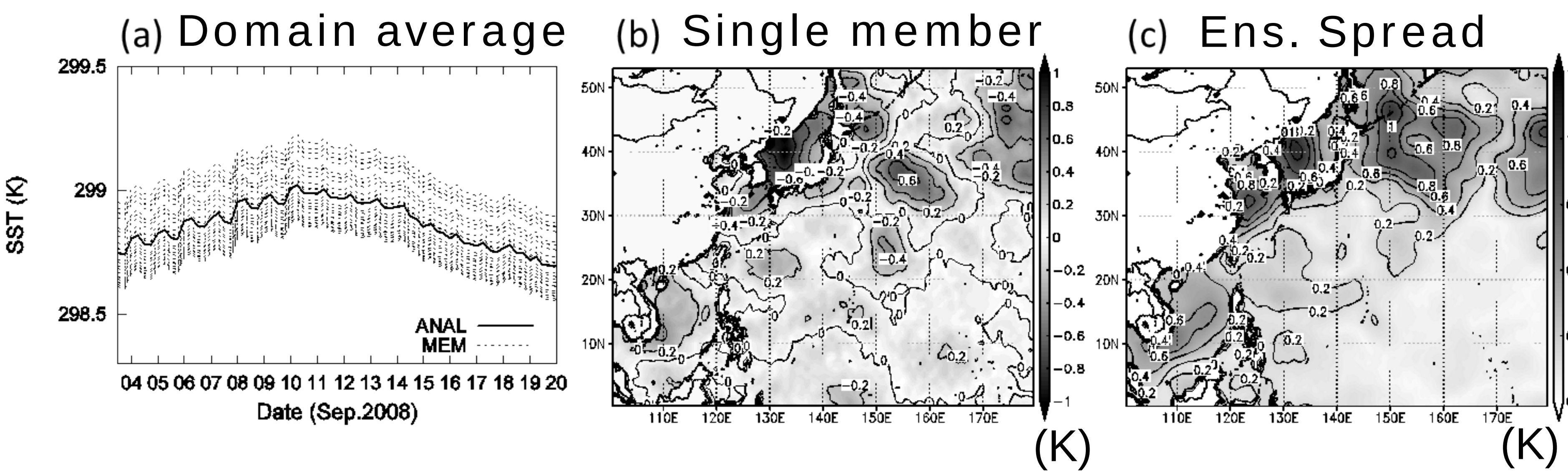
SST UNCERTAINTIES

Kunii and Miyoshi (2012, *Weather and Forecasting*)

SST ensemble perturbations

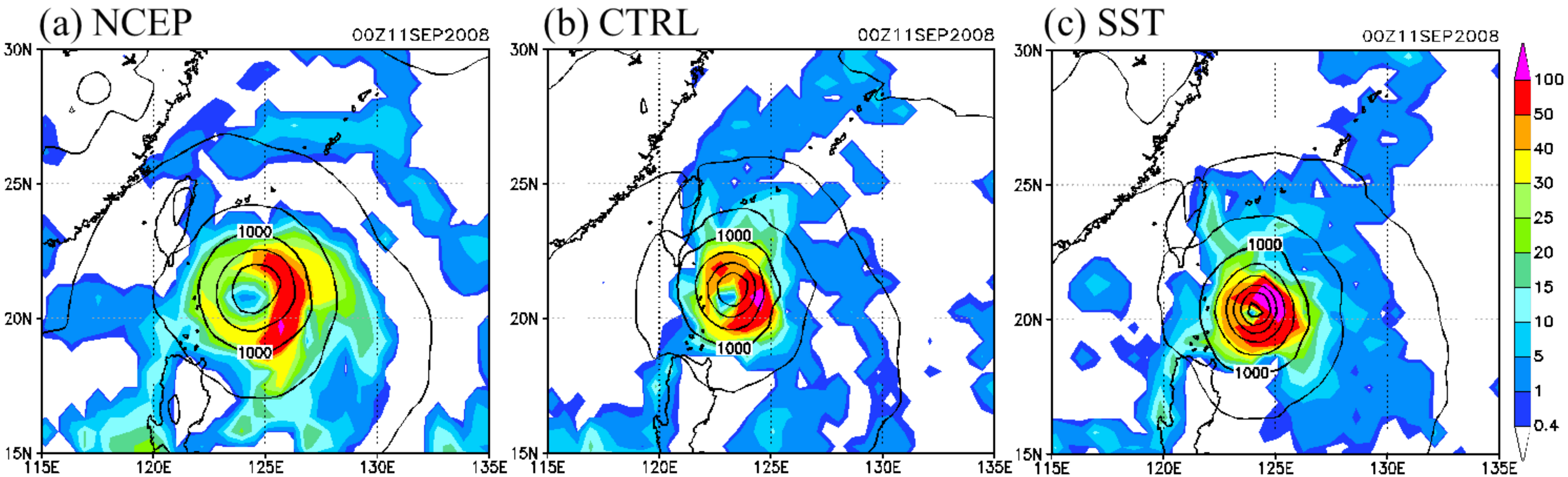
SST is randomly perturbed around the SST analysis in the WRF-LETKF cycle.

The SST perturbations are the differences between SST analyses on randomly chosen dates. The perturbation fields are fixed in time.



Kunii and Miyoshi (2012)

6-h forecast fields



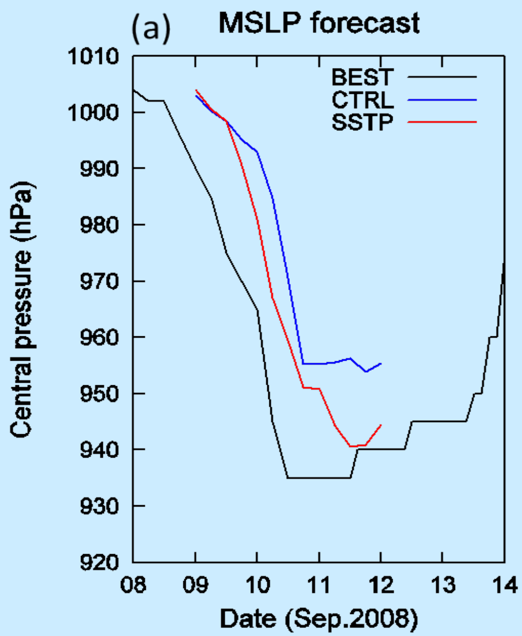
The location and intensity are the best with the SST perturbation.

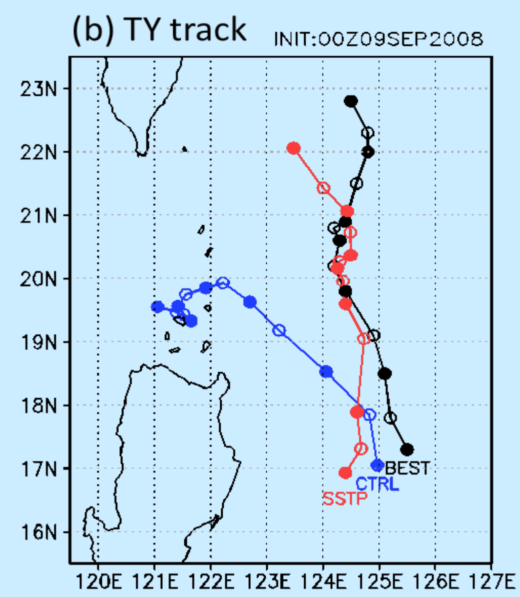
NOTE: There is no SST perturbation in the forecast, but only in the DA cycle.

Improvement in TC forecasts

1. TC intensity and track forecasts are greatly improved.

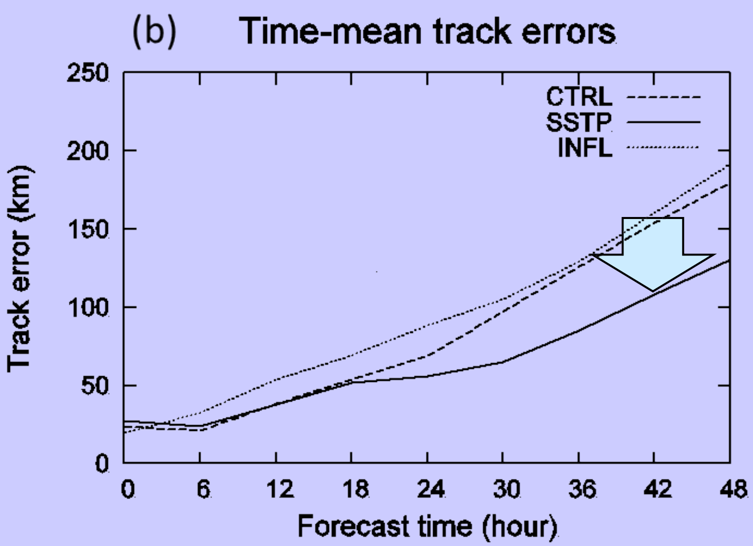
(NO SST perturbations in the forecast)





2. Improvement is not only in the single case.

(NO SST perturbations in the forecast)



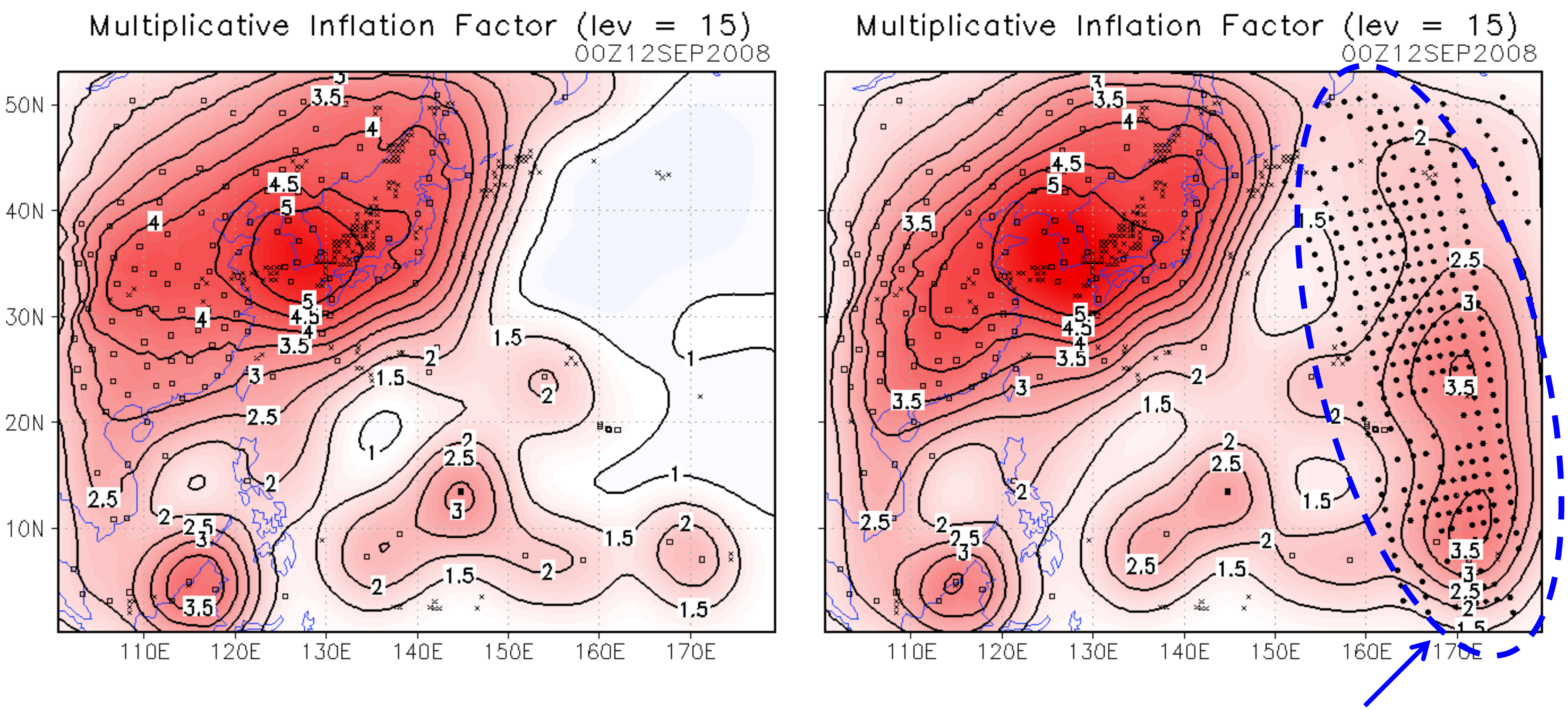
WRF-LETKF: using satellite data

ASSIMILATION OF AIRS DATA

Miyoshi and Kunii (2012, *Tellus*)

Assimilation of AIRS retrievals

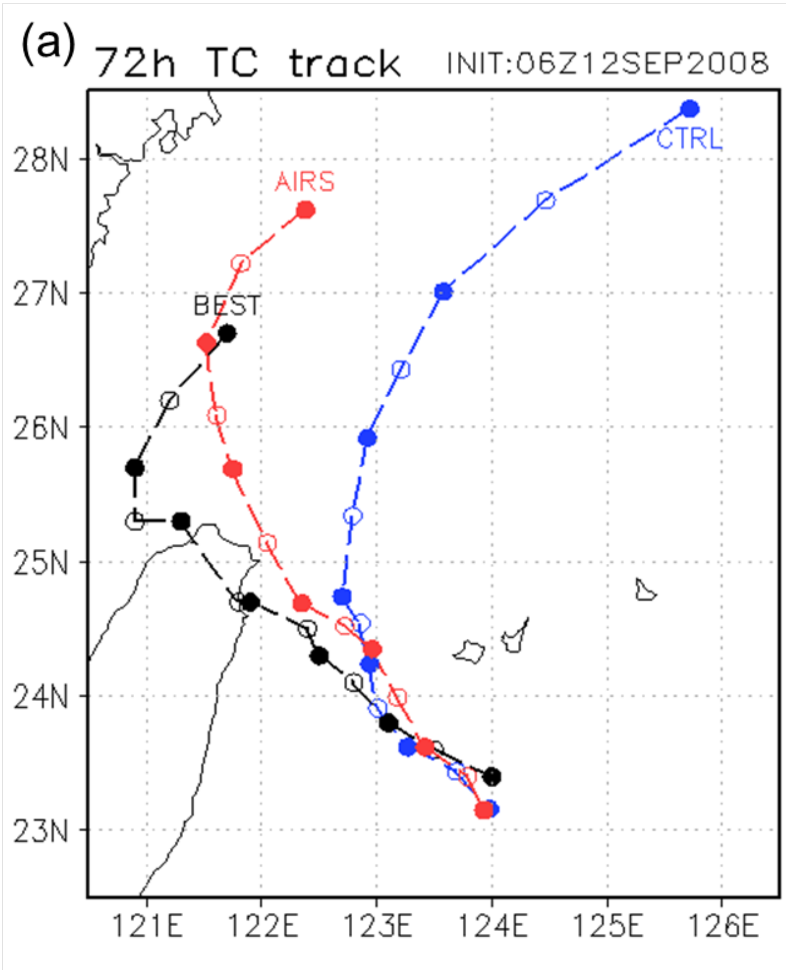
CTRL	AIRS
Conventional (NCEP PREPBUFR)	Conv. + AIRS retrievals (AIRX2RET - T, q)

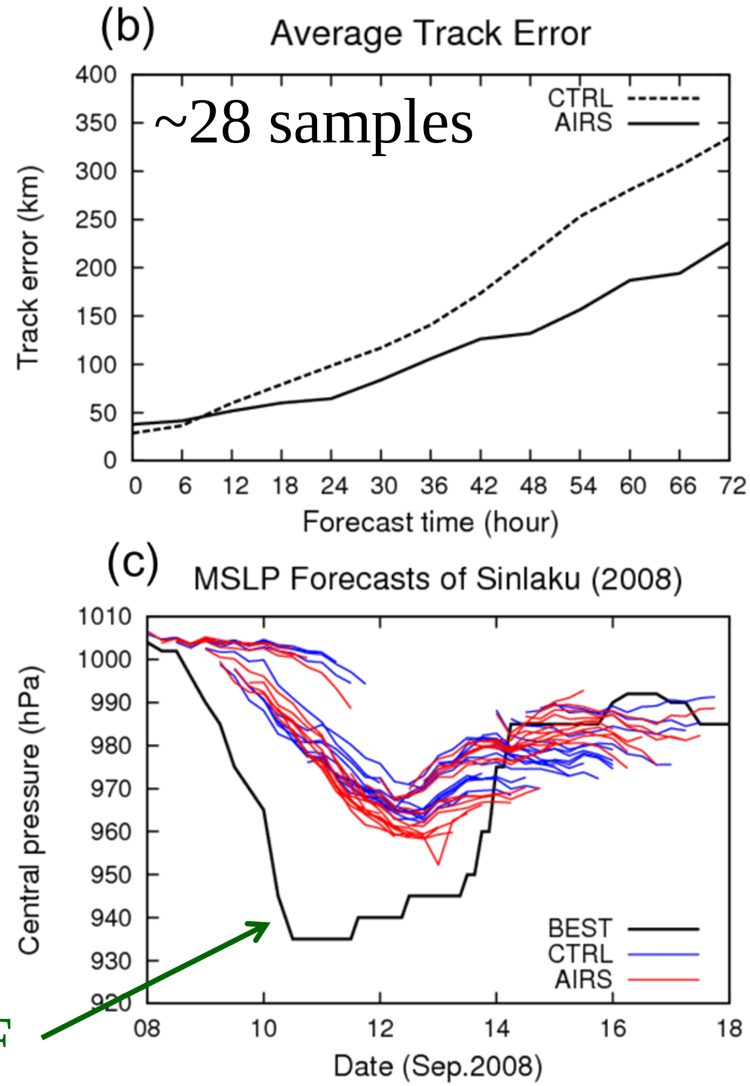


Larger inflation is estimated due to the AIRS data.

- August-September 2008, focusing on Typhoon Sinlaku

AIRS impact on TC forecasts





Too deep to resolve by 60-km WRF

- TC track forecasts for Typhoon Sinlaku (2008) were significantly better, particularly in longer leads.

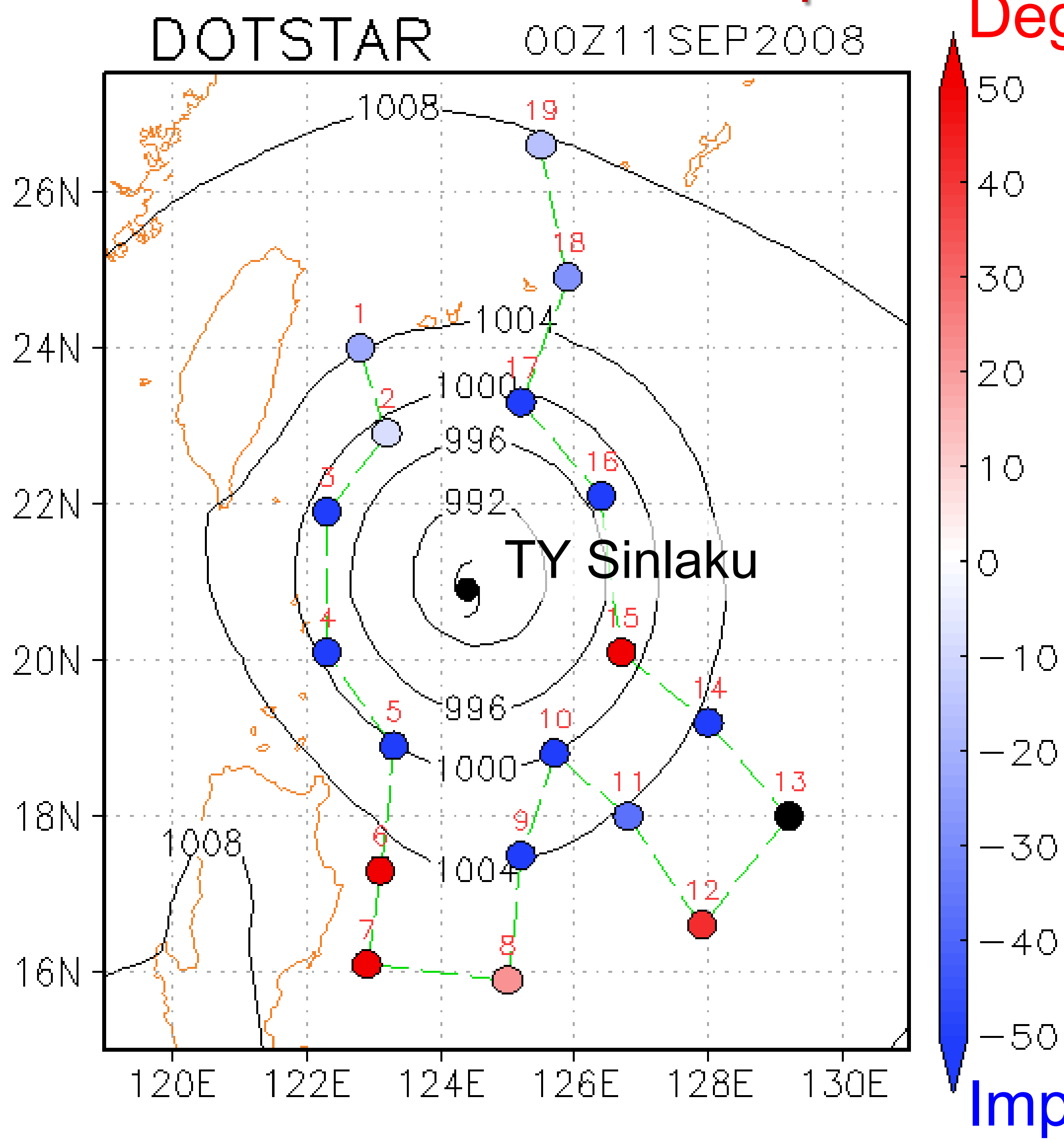
ENSEMBLE-BASED OBS IMPACT

Kunii, Miyoshi and Kalnay (2012, *Mon. Wea. Rev.*)

Ota, Kalnay, Miyoshi and Derber (under review)

Forecast Sensitivity to Observations (FSO)

Estimated observation impact

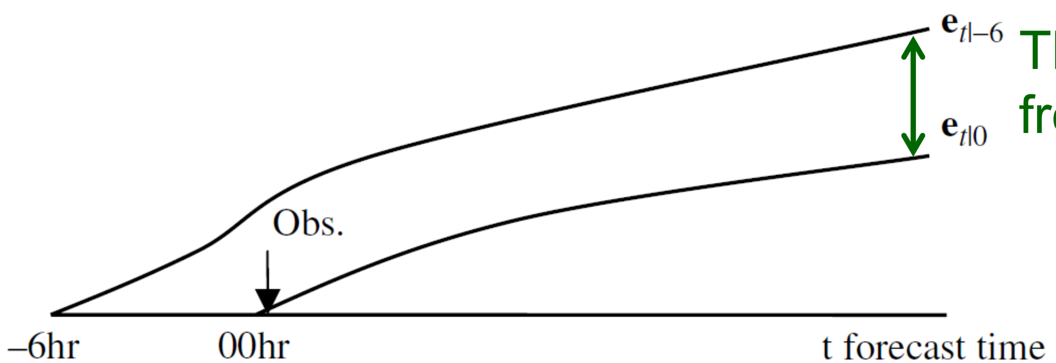


With FSO approaches, observation impacts can be estimated without performing expensive data denial experiments (or OSEs).

Kunii, Miyoshi, Kalnay (2012)

Forecast sensitivity to observations

Observation impact can be calculated using an adjoint model (*Langland and Baker 2004*)



This difference comes from obs at 00hr

$$\mathbf{e}_{t|0} = \boldsymbol{x}^f_{t|0} - \boldsymbol{x}^a_t$$

The error reduction (or increase) due to obs at 00 (*i.e.*, *obs impact*):

$$J = \left(\mathbf{e}^T_{t|0} C \mathbf{e}_{t|0} - \mathbf{e}^T_{t|-6} C \mathbf{e}_{t|-6}\right) = \left(\underbrace{\mathbf{e}_{t|0} - \mathbf{e}_{t|-6}}\right)^T C \left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right)$$

$$\boldsymbol{x}^f_{t|0} - \boldsymbol{x}^f_{t|-6} \approx \mathbf{M}(\underbrace{\boldsymbol{x}^a_0 - \boldsymbol{x}^f_{0|-6}}_{\text{analysis increment!}})$$

$$\mathbf{K}(\boldsymbol{y}_0 - H\boldsymbol{x}^f_{0|-6})$$

$$\boxed{J \approx \delta \mathbf{y}^T \mathbf{K}^T \mathbf{M}^T C \left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right)}$$

Forecast sensitivity to observations

Observation impact can be calculated without an adjoint model

(*Liu and Kalnay 2008; Li et al. 2009; Kalnay et al. 2012*)

$$J \approx \delta\mathbf{y}^T\mathbf{K}^T\mathbf{M}^T C\left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right) \qquad \text{(\textit{Langland and Baker 2004})}$$

In the ensemble Kalman filter (EnKF),

$$\mathbf{K} = \frac{\mathbf{X}_0^a \mathbf{X}_0^{a\,T}}{N_{ens} - 1}\mathbf{H}^T\mathbf{R}^{-1} = \frac{\mathbf{X}_0^a \mathbf{Y}_0^{a\,T}\mathbf{R}^{-1}}{N_{ens} - 1}$$

$$J \approx \delta\mathbf{y}^T\mathbf{K}^T\mathbf{M}^T C\left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right) = \frac{\delta\mathbf{y}^T\mathbf{R}^{-1}\mathbf{Y}_0^a \mathbf{X}_0^{a\,T}\mathbf{M}^T}{N_{ens} - 1} C\left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right)$$

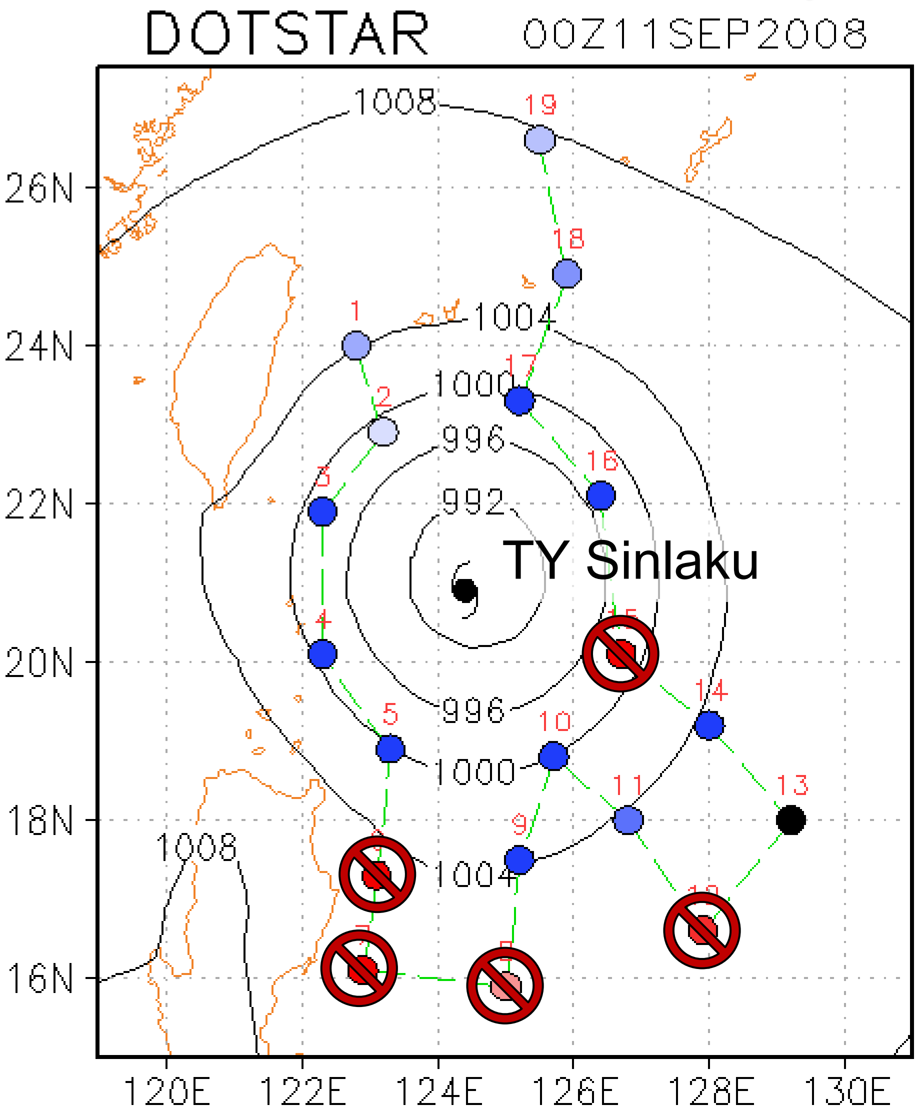
$$\boxed{J \approx \frac{1}{N_{ens} - 1}\delta\mathbf{y}^T\mathbf{R}^{-1}\mathbf{Y}_0^a \mathbf{X}_{t|0}^{f\;T} C\left(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}\right)}$$

We just need an ensemble of forecasts.

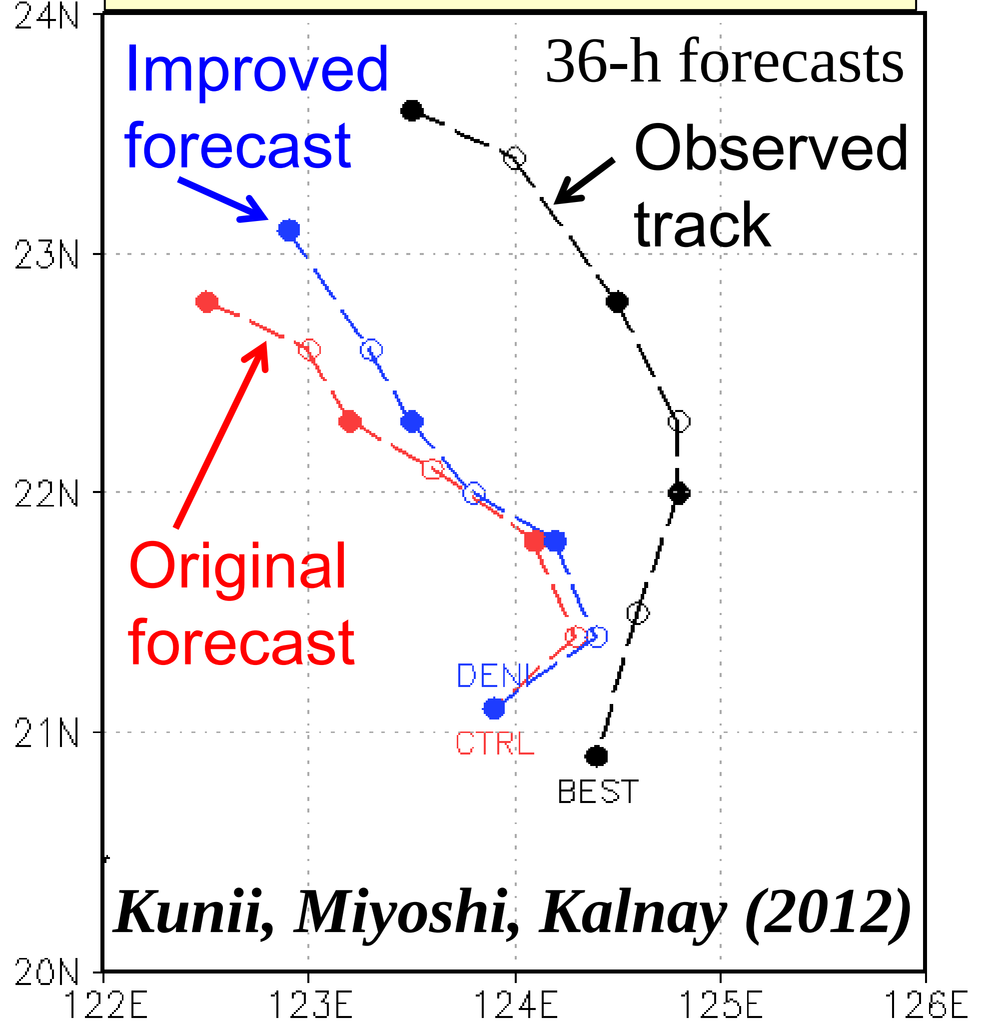
Kalnay et al. (2012)

Denying negative impact data improves forecast!

Estimated observation impact

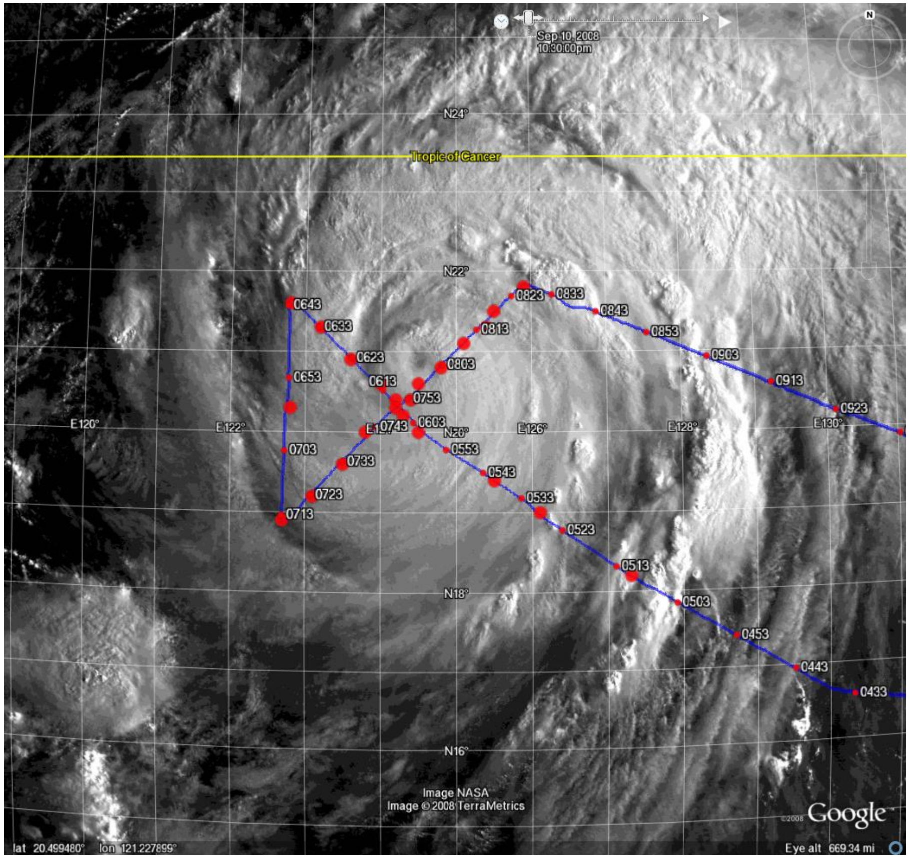


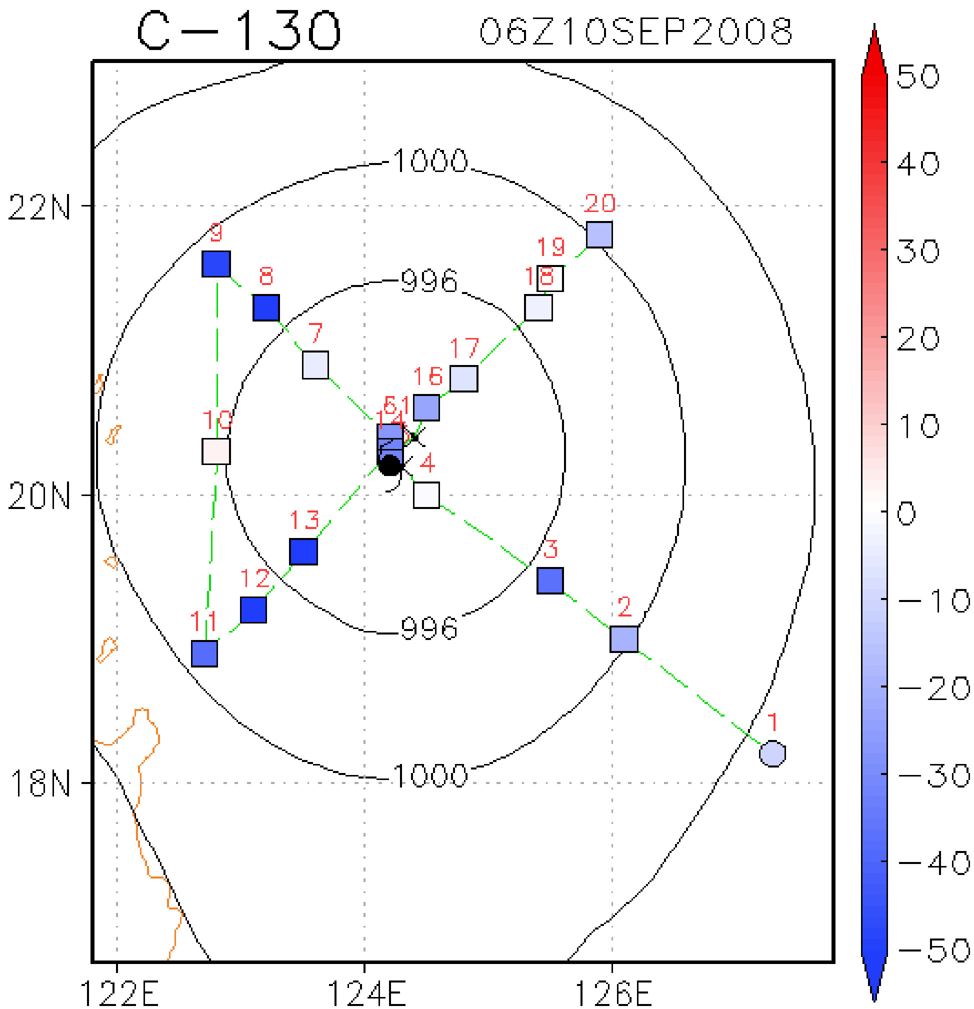
Typhoon track forecast is actually improved!!



Kunii, Miyoshi, Kalnay (2012)

Impact of WC-130J dropsondes



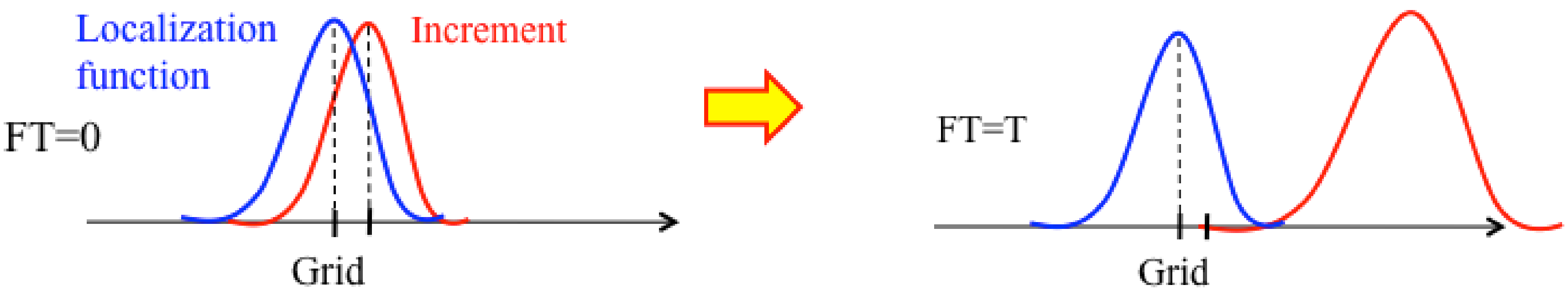


Kunii, Miyoshi, Kalnay (2012)

An issue on localization

$$J \approx \frac{1}{N_{ens}-1}\delta\mathbf{y}^T\mathbf{R}^{-1}\underline{\mathbf{Y}_0^a\mathbf{X}_{t|0}^f}^T C\left(\mathbf{e}_{t|0}+\mathbf{e}_{t|-6}\right)$$

The ensemble-based covariance needs localization.

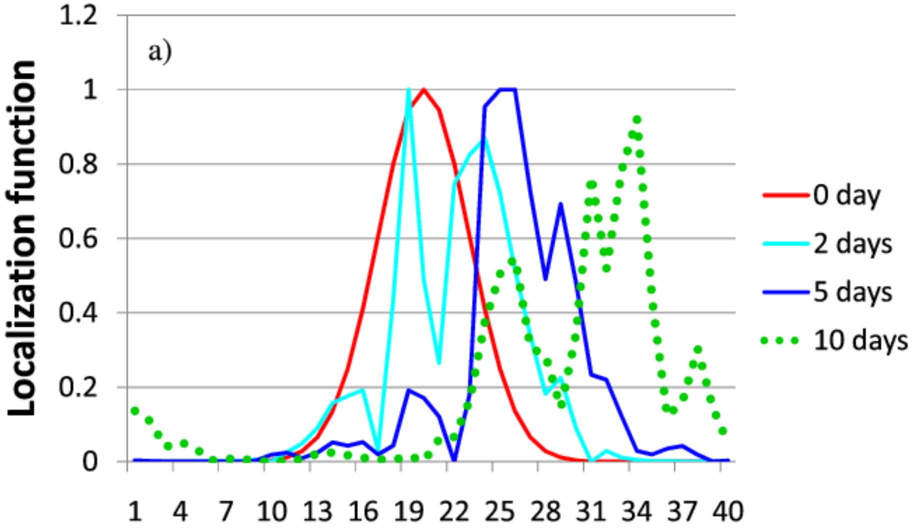


We need to consider a “mobile” localization function.

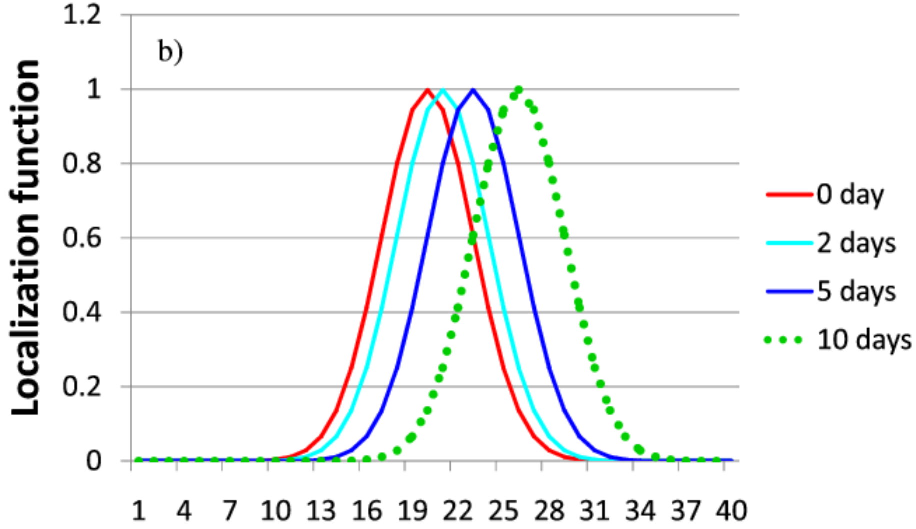
cf. *Bishop and Hodyss (2009)*

Ideas for "mobile" localizaiton

a) Nonlinear incremental evolution



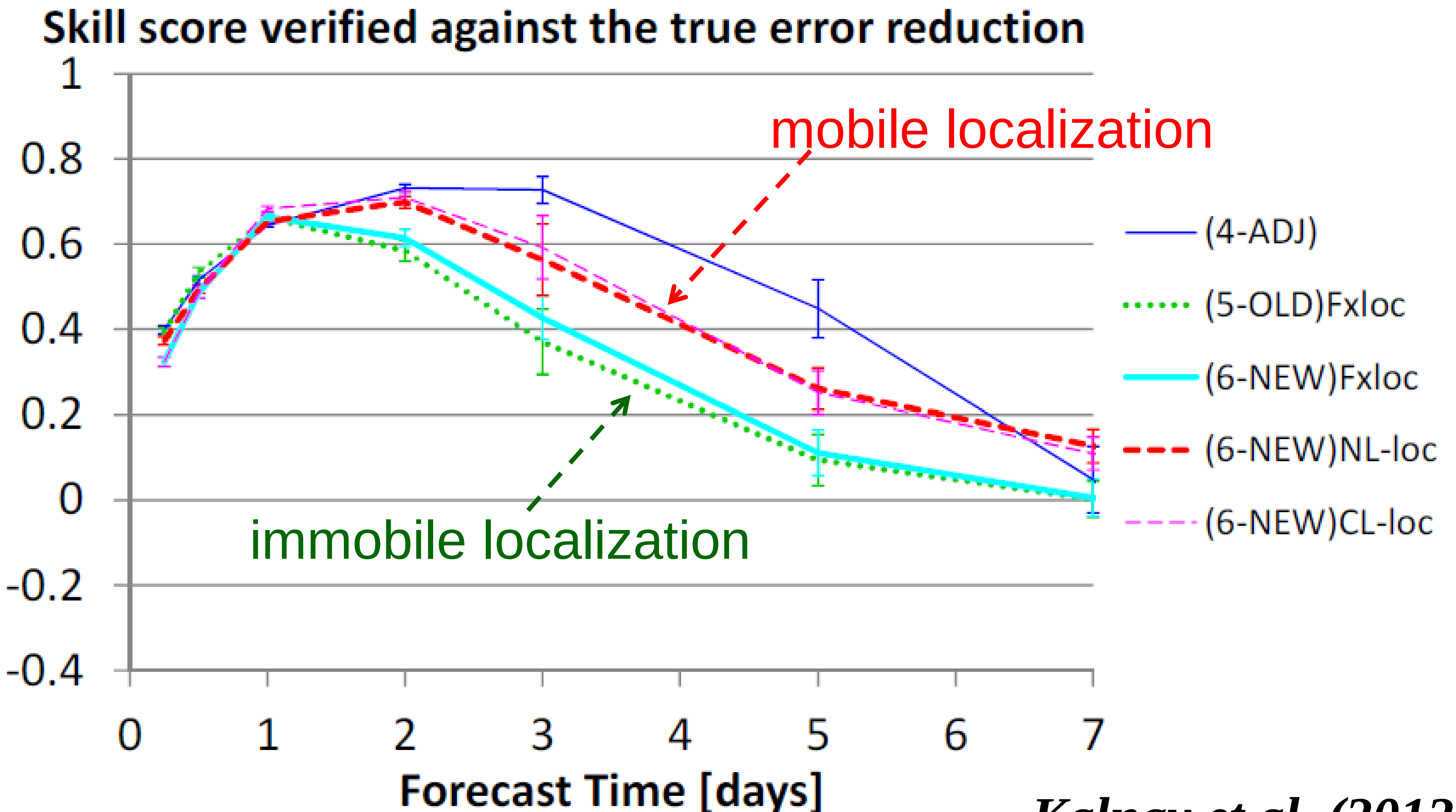
b) Constant advection



Kalnay et al. (2012)

Impact of mobile localization

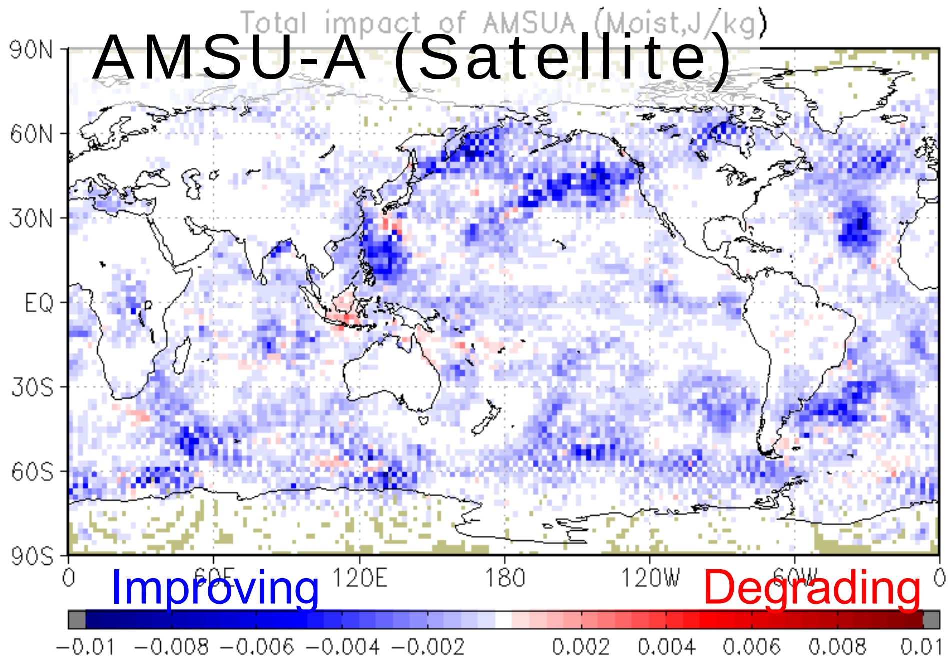
Results from idealized experiments with the Lorenz-96 model.

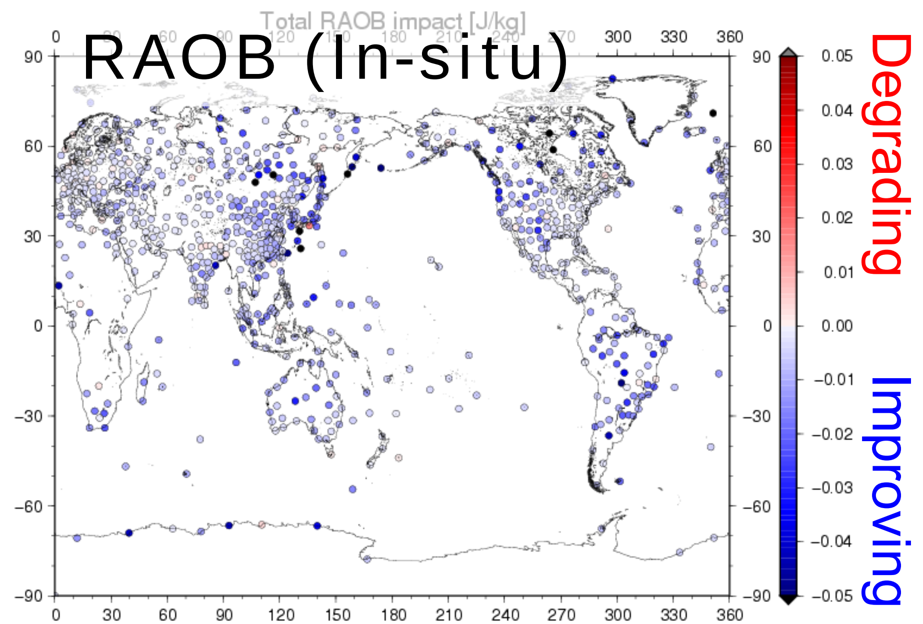


Kalnay et al. (2012)

Impact estimates with NCEP GFS

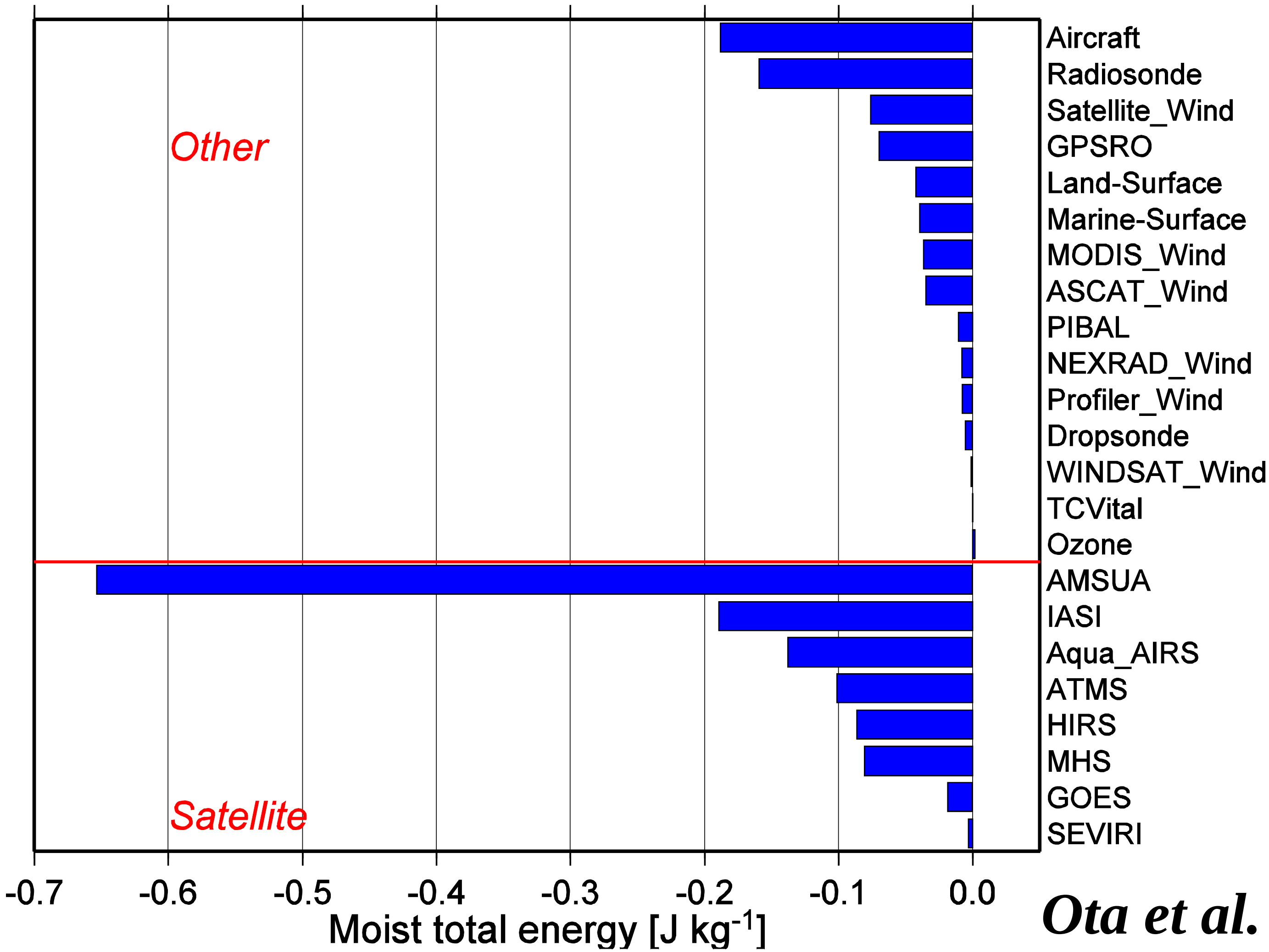
Ota et al. (2012)





Impact estimates with NCEP GFS

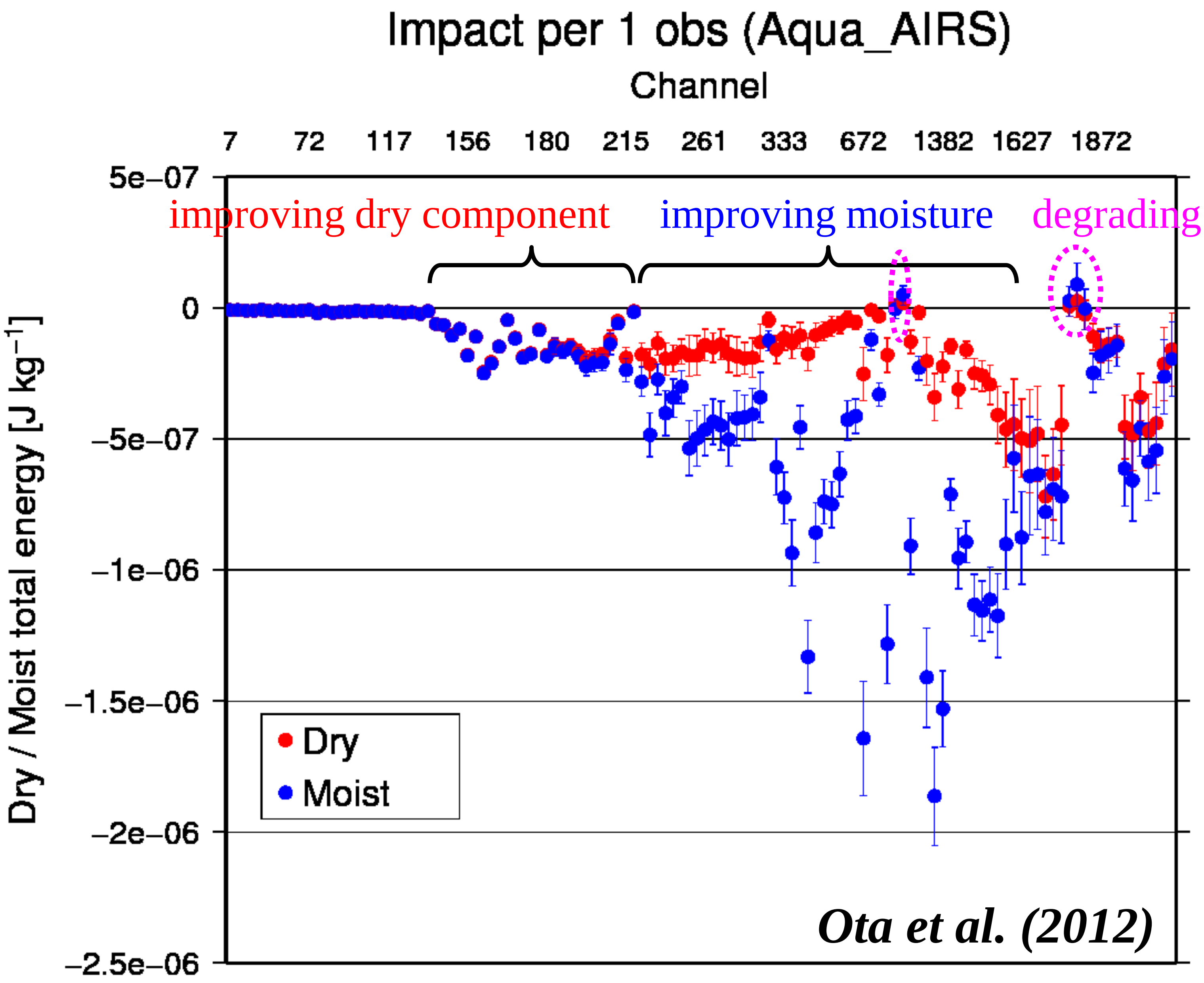
a) Total observation impacts on 1 analysis



Ota et al. (2012)

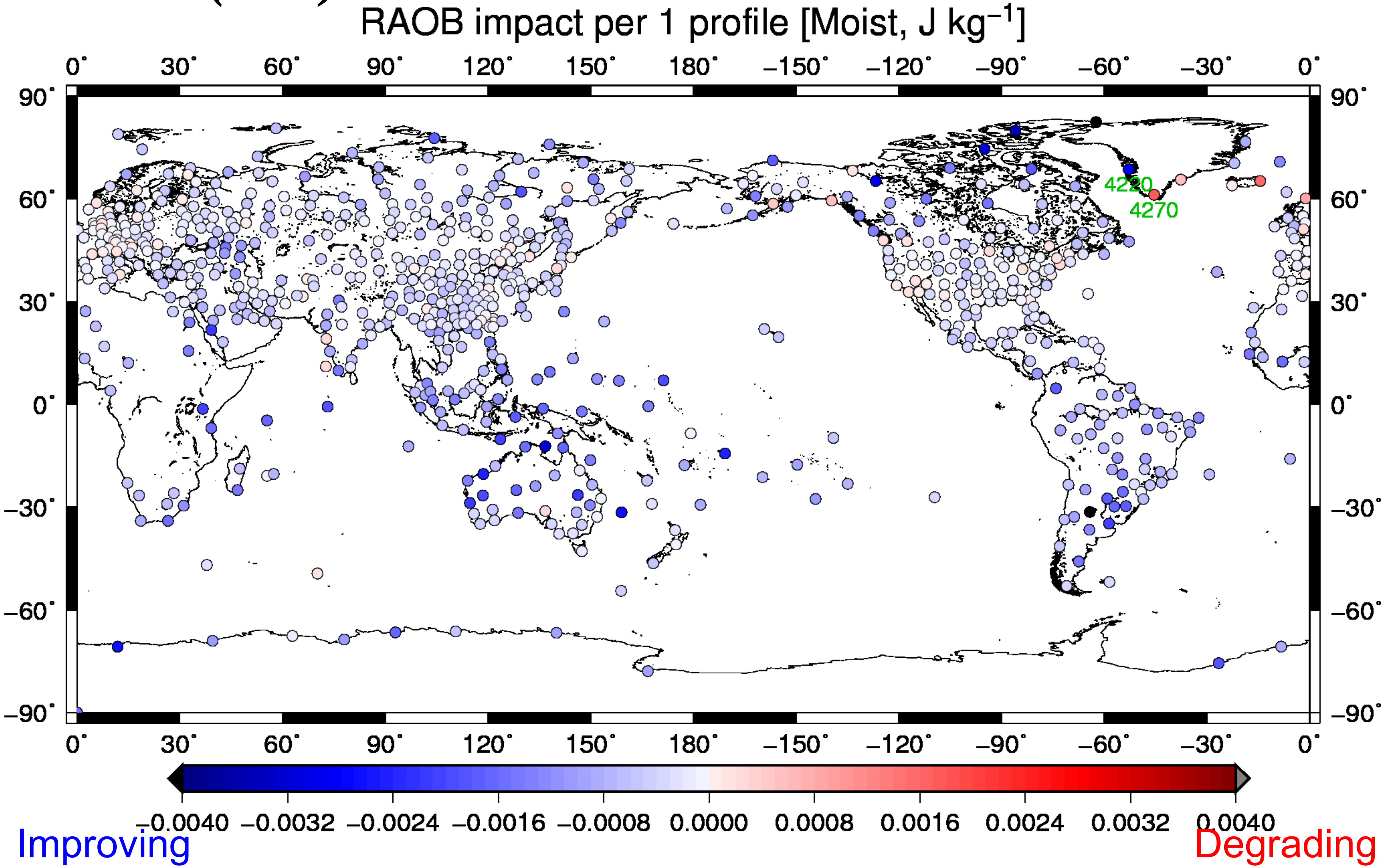
Improving

Impact of AIRS channels



RAOB impacts

Ota et al. (2012)



Future research ideas

- Regional ocean coupling
  - Considering flow-dependent SST perturbations.
- Higher-resolution runs, multi-scale considerations
  - We need more localization with higher resolution, but tight localization only allows using data for high-frequency components.
- Model parameter estimation, multi-model EnKF
- Expanding to WRF-Chem
  - Aerosols, air-quality, lidar data assimilation

A CHALLENGE: MULTI-SCALE TREATMENT

Miyoshi and Kondo (in preparation)

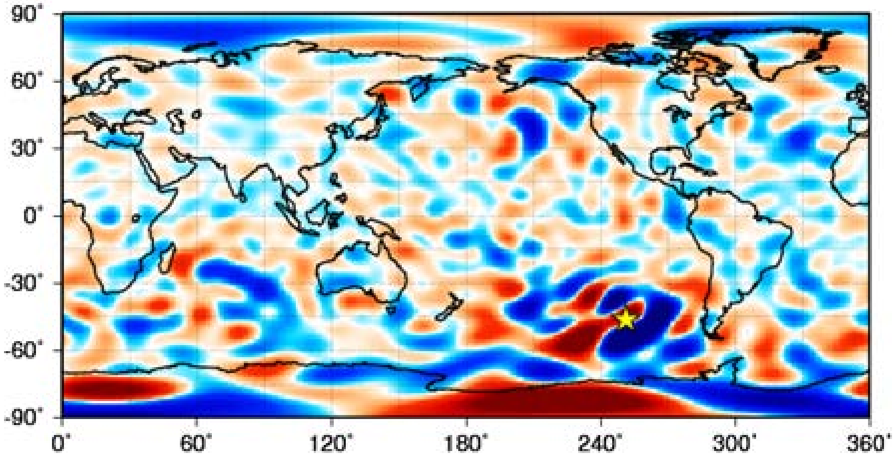
Motivation

Localization plays an essential role in an EnKF to cope with limited ensemble size.

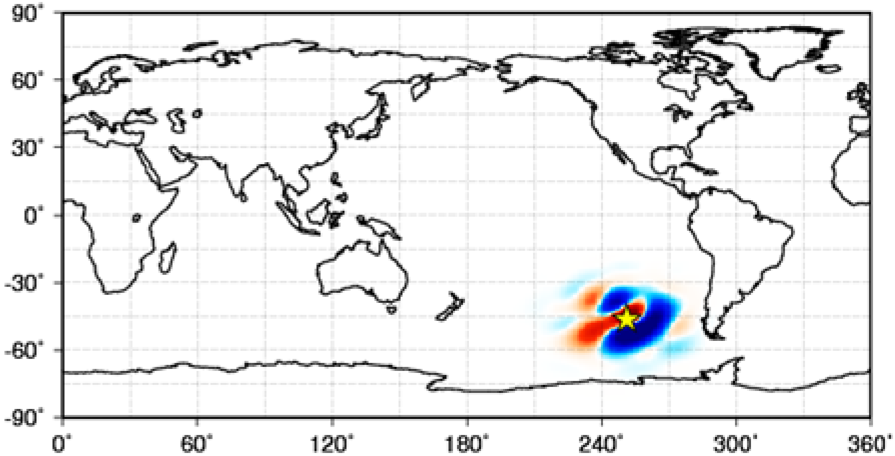
Higher resolution requires more localization, limiting the use of observations.

We look for better use of observations by separating the scales.

Raw covariance field estimated from 20 ensemble members



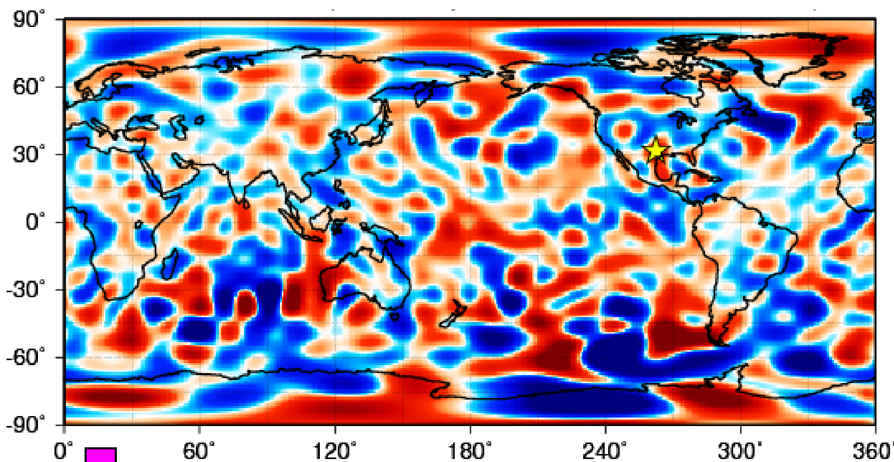
Localized covariance field



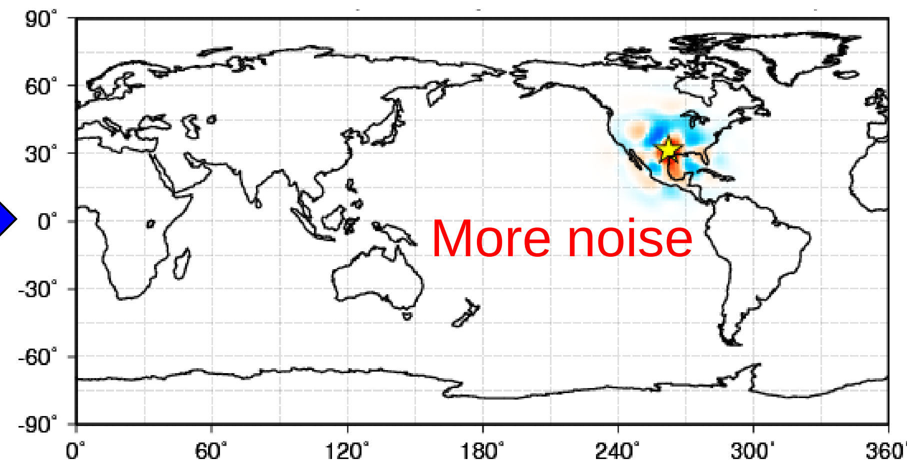
An idea of scale separation

✓ Extracting larger-scale covariance by spatial smoothing

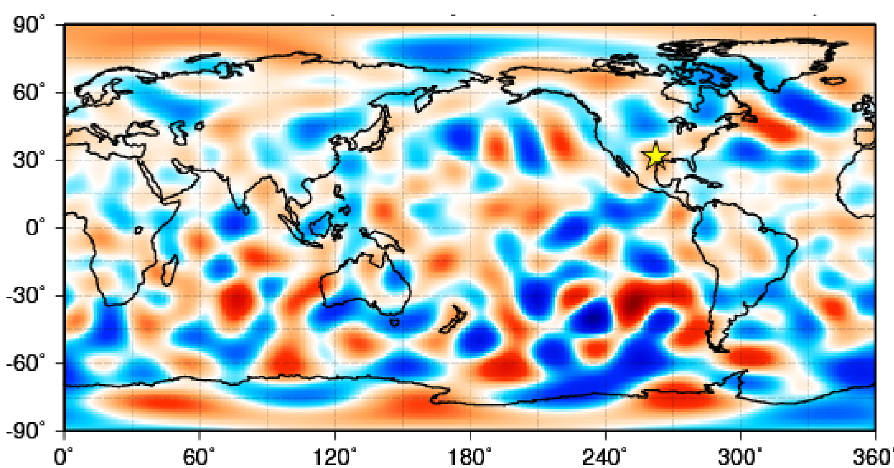
Raw covariance at T30



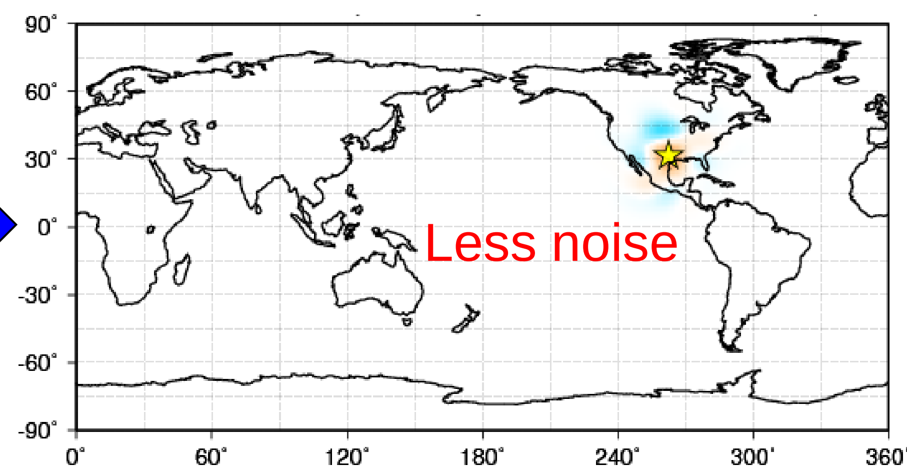
With 1000-km localization



Using smoothed perturbations



With 1000-km localization



Scale-separated analysis increments

We will construct analysis increments at high (*h*) and low (*l*) resolutions separately.

$$\delta x = \delta x_h + \delta x_l$$

δx_l is obtained by smoothed (low-resolution) forecast ensemble perturbations.

- ✓ Using larger localization

δx_h is obtained by the regular EnKF (smaller localization) minus $\delta x_l{}'$ (different from δx_l).

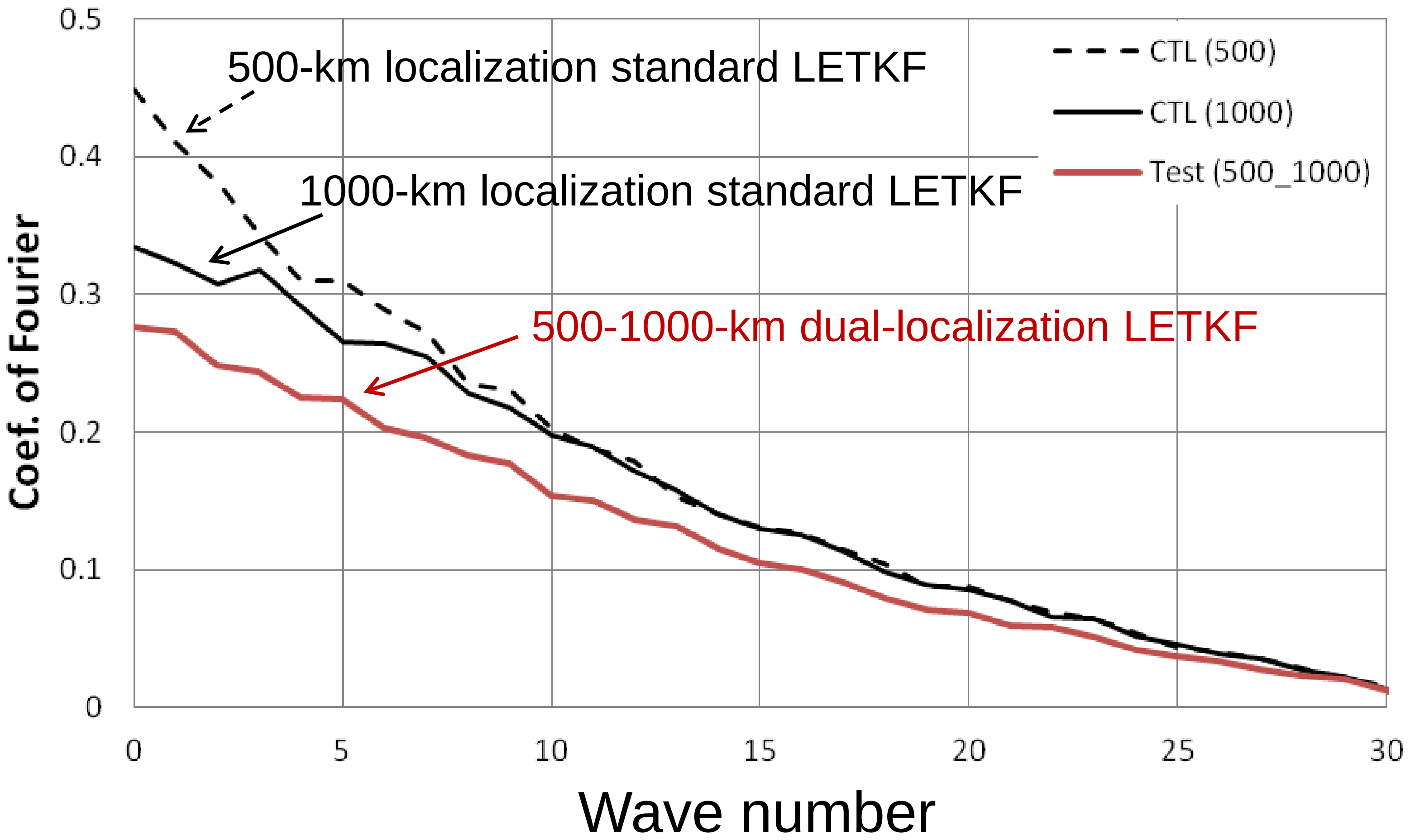
- ✓ $\delta x_l{}'$ is similar to δx_l, but with smaller localization.

In this study, we use the T30L7 SPEEDY model (Molteni 2003), with 500-km (smaller) and 1000-km (larger) localization settings.

Results are promising.

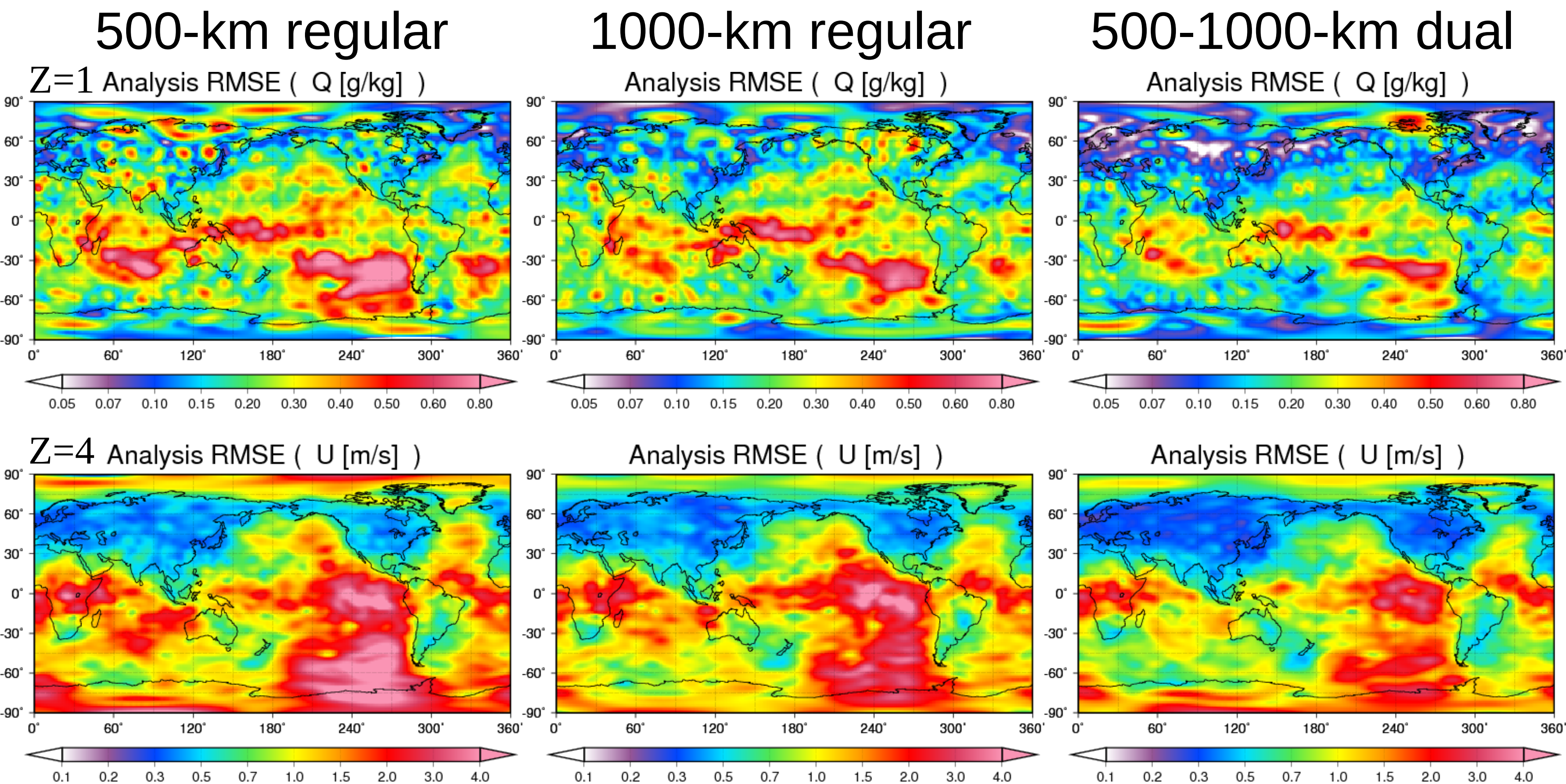
Successfully reducing the errors at almost all scales.

1-month average global analysis error power spectrum



Improvements are almost everywhere for all variables.

1-month average RMS errors



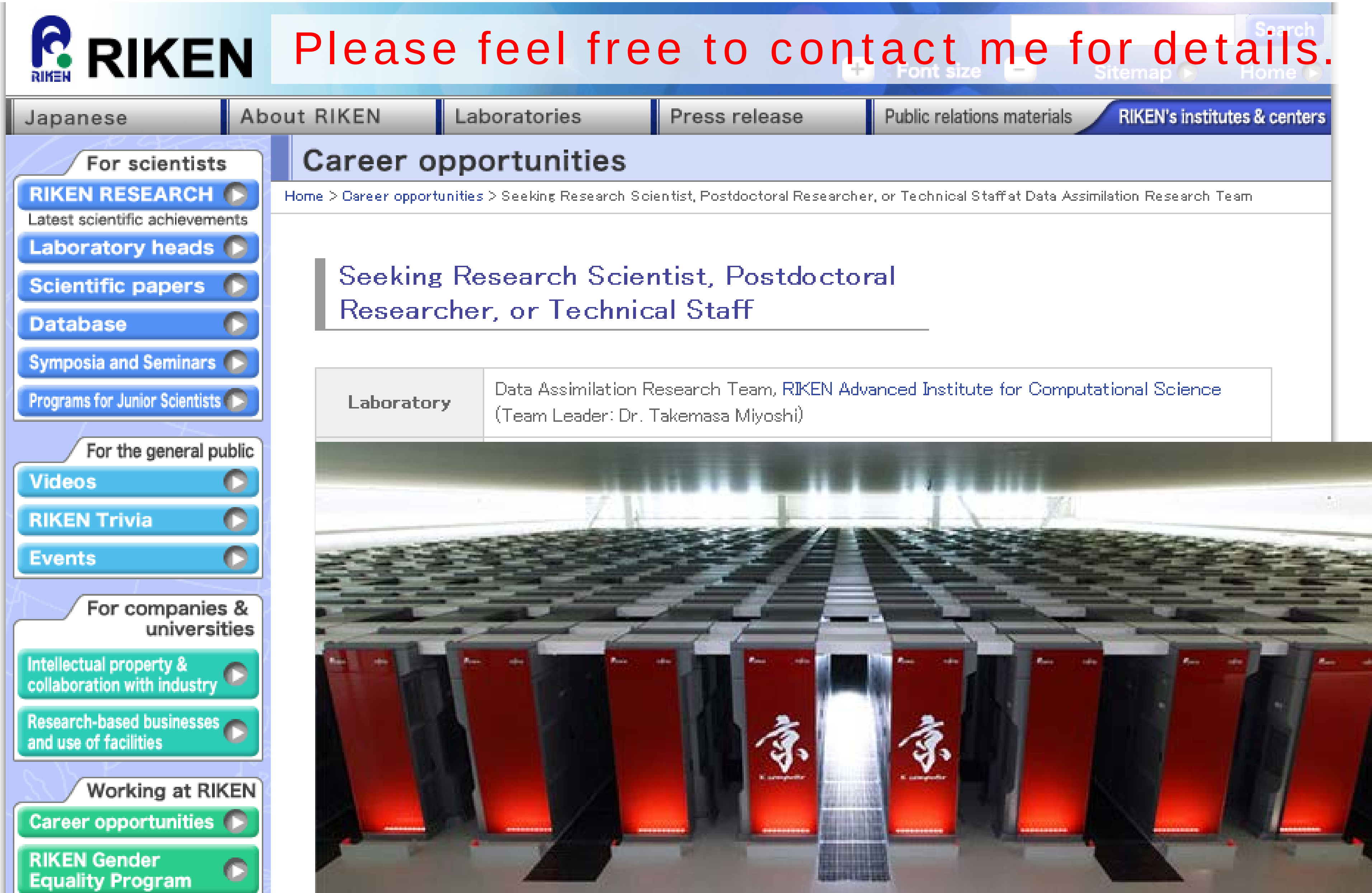
Summary

- EnKF was proposed by Evensen in 1994, and the Environment Canada first started using EnKF for operational NWP in 2005. We have achieved quite a lot to improve EnKF performances so far.

- We still have a lot more to improve. Challenges include…
  - Multi-scale treatment
  - Model errors
    - multi-model, model parameter optimization
  - Nonlinear, non-Gaussian filters/smoothers

We are hiring researchers.

Please feel free to contact me for details.




Thank you very much
for your kind attention!!