Recent Advances in EnKF: Running in Place, Assimilation of Rain, Ensemble Forecast Sensitivity to Observations

Shu-Chih Yang, Ji-Sun Kang, Takemasa Miyoshi, Steve Penny, Guo-Yuan Lien, Steve Greybush

Eugenia Kalnay

University of Maryland

UMD Weather-Chaos Group: **Kayo Ide**, **Brian Hunt**, Ed Ott, and students (Guo-Yuan Lien, Yan Zhou, Adrienne Norwood, Erin Lynch, Yongjing Zhao, Daisuke Hotta)

Also: **Y Ota**, Juan Ruiz, C Danforth, M Peña

NCU, 19 November 2012

Promising new tools for the LETKF(1)

1. Running in Place (Kalnay and Yang, QJ 2010, Yang, Kalnay and Hunt, MWR, 2012)

- It extracts more information from observations by using them more than once.
- Useful during spin-up (e.g., hurricanes and tornados).
- It uses the "no-cost smoother", Kalnay et al., Tellus, 2007b.
- Typhoon Sinlaku (Yang et al., 2012)
- 7-years of Ocean Reanalysis (Penny, 2011, Penny et al., 2012)
- Very good results!

Promising new tools for the LETKF(2)

2. **Effective assimilation of Precipitation** (Guo-Yuan Lien, Eugenia Kalnay and Takemasa Miyoshi, 2013)

- Assimilation of precipitation has generally failed to improve forecasts beyond a day.
- A new approach deals with non-Gaussianity, and assimilation of both zero and non-zero precipitation. The model now "remembers" the assimilation, so that that medium range forecasts are improved.

Promising new tools for the LETKF(3)

3. Forecast Sensitivity to Observations and "proactive QC"

(with Y Ota, T Miyoshi, J Liu, and J Derber)

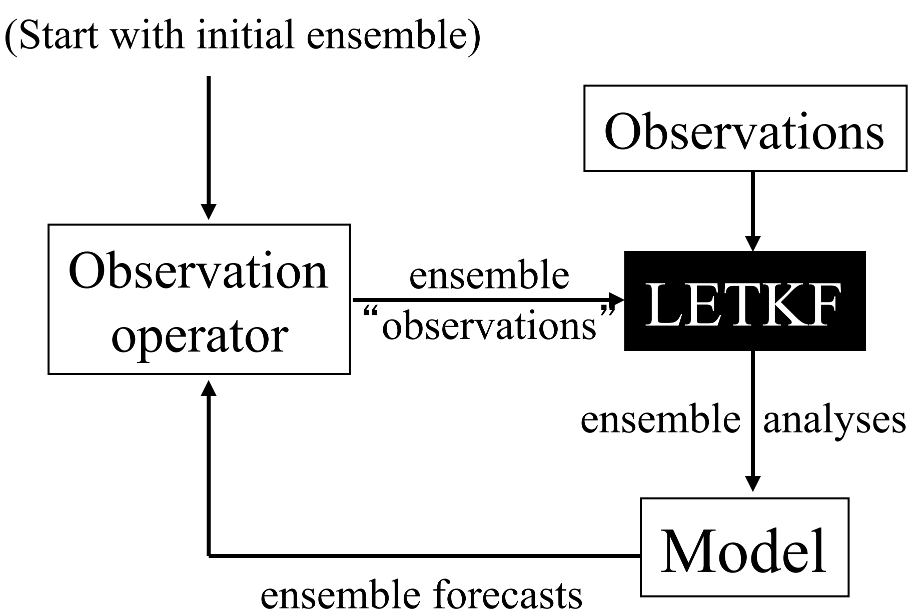
- A simpler, more accurate formulation for the Ensemble Forecast Sensitivity to Observations (EFSO, Kalnay et al., 2012, Tellus).
- Ota et al., 2012 tested it with the NCEP EnSRF-GFS operational system using all operational observations.
- Allows to identify "bad observations" after 12 or 24hr, and then repeat the data assimilation without them: "proactive QC".

4. Estimation of surface fluxes as evolving parameters

(Kang et al., 2011, Kang et al., 2012)

- Important for the carbon cycle, surface fluxes of heat, moisture and momentum (stress) and eventually for coupled data assimilation.

Local Ensemble Transform Kalman Filter (Ott et al, 2004, Hunt et al, 2004, 2007) (a square root filter)

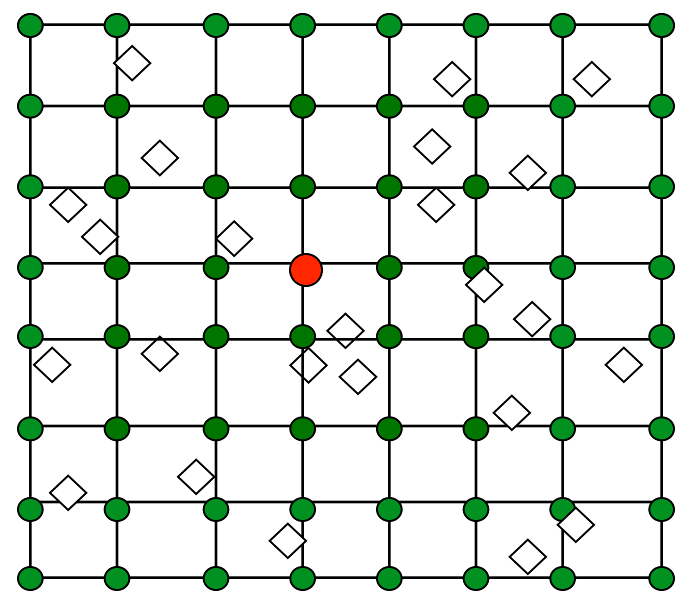


- Model independent (black box)
- Obs. assimilated simultaneously at each grid point
- 100% parallel
- No **adjoint** needed
- 4D LETKF extension
- Computes the **weights** for the ensemble forecasts explicitly

Localization based on observations

Perform data assimilation in a local volume, choosing observations

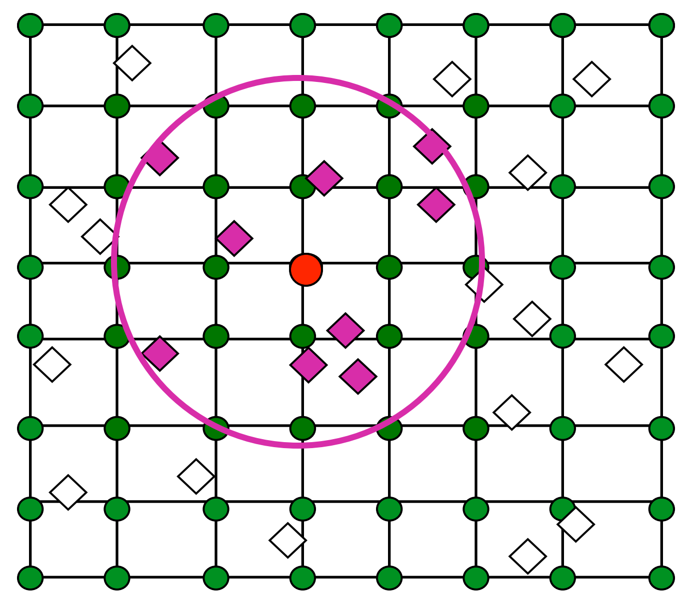
The state estimate is updated at the central grid red dot



Localization based on observations

Perform data assimilation in a local volume, choosing observations

The state estimate is updated at the central grid red dot



All observations (purple diamonds) within the local region are assimilated

The LETKF algorithm can be described **in a single slide**!

Local Ensemble Transform Kalman Filter (LETKF)

Globally:

Forecast step: $\mathbf{x}^b_{n,k} = M_n\left(\mathbf{x}^a_{n-1,k}\right)$

Analysis step: construct $\mathbf{X}^b = \left[\mathbf{x}^b_1 - \overline{\mathbf{x}}^b \,|\, ... \,|\, \mathbf{x}^b_K - \overline{\mathbf{x}}^b\right];$

$$\mathbf{y}^b_i = H(\mathbf{x}^b_i);\ \mathbf{Y}^b_n = \left[\mathbf{y}^b_1 - \overline{\mathbf{y}}^b \,|\, ... \,|\, \mathbf{y}^b_K - \overline{\mathbf{y}}^b\right]$$

Locally: Choose for each grid point the observations to be used, and compute the local analysis error covariance and perturbations in ensemble space:

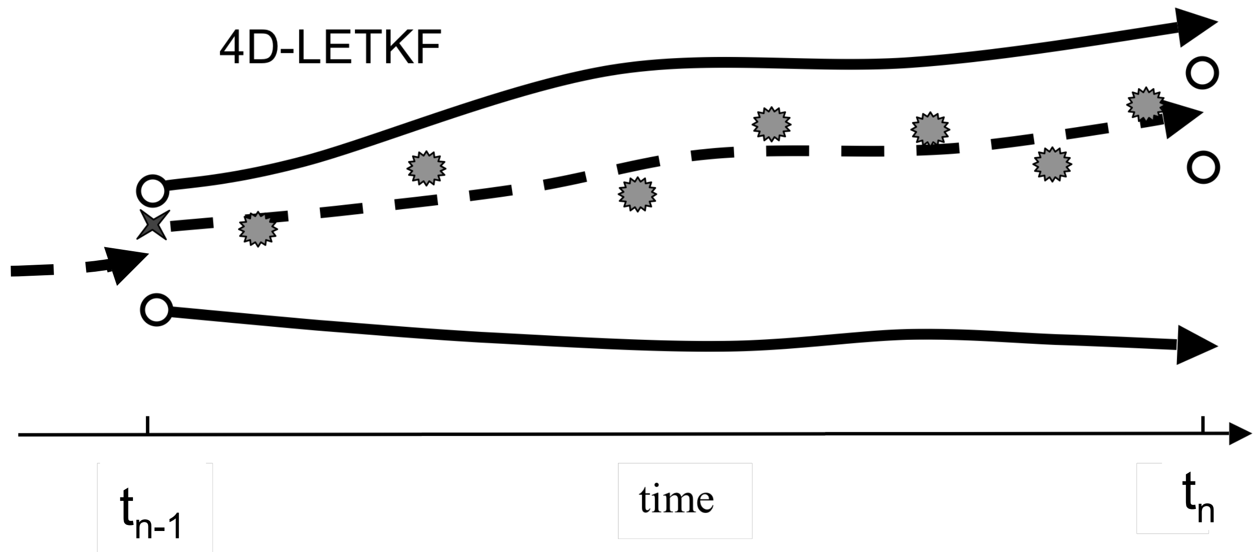
$$\tilde{\mathbf{P}}^a = \left[\left(K-1\right)\mathbf{I} + \mathbf{Y}^{bT}\mathbf{R}^{-1}\mathbf{Y}^b\right]^{-1};\ \mathbf{W}^a = [(K-1)\tilde{\mathbf{P}}^a]^{1/2}$$

Analysis mean in ensemble space: $\overline{\mathbf{w}}^a = \tilde{\mathbf{P}}^a\mathbf{Y}^{bT}\mathbf{R}^{-1}(\mathbf{y}^o - \overline{\mathbf{y}}^b)$

and add to $\mathbf{W}^a$ to get the analysis ensemble in ensemble space.

The new ensemble analyses in model space are the columns of $\mathbf{X}^a_n = \mathbf{X}^b_n\mathbf{W}^a + \overline{\mathbf{x}}^b$. Gathering the grid point analyses forms the new global analyses. Note that the the output of the LETKF are analysis weights $\overline{\mathbf{w}}^a$ and perturbation analysis matrices of weights $\mathbf{W}^a$. **These weights multiply the ensemble forecasts.**

No-cost LETKF smoother (×): apply at t_{n-1} the same weights found optimal at t_n. It works for 3D- or 4D-LETKF



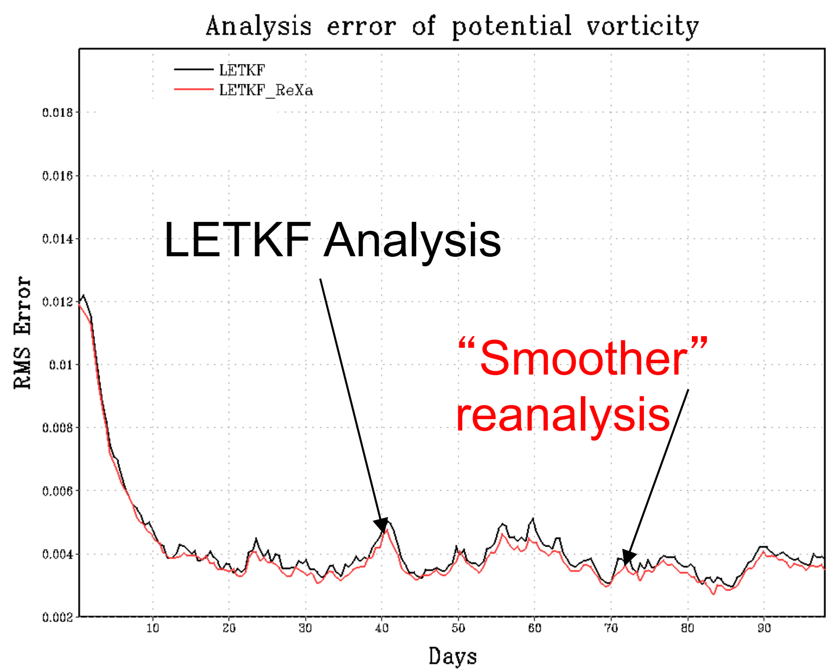
The no-cost smoother makes possible:

- ✓ Quasi Outer Loop (QOL)
- ✓ “Running in place” (RIP) for faster spin-up
- ✓ Use of future data in reanalysis
- ✓ Ability to use longer windows and nonlinear perturbations

No-cost LETKF smoother first tested on a QG model: it works…

LETKF analysis at time n $\quad \overline{\mathbf{x}}_n^a = \overline{\mathbf{x}}_n^f + \mathbf{X}_n^f \overline{\mathbf{w}}_n^a$

Smoother analysis at time n-1 $\quad \tilde{\mathbf{x}}_{n-1}^a = \overline{\mathbf{x}}_{n-1}^f + \mathbf{X}_{n-1}^f \overline{\mathbf{w}}_n^a$



Very simple smoother: apply the final weights at the beginning of the window. It allows assimilation of **future** data, and assimilating data more than once.

Nonlinearities: “Quasi Outer Loop” (QOL)

Quasi Outer Loop: use the final weights to correct only the <u>mean</u> initial analysis, keeping the initial perturbations. Repeat the analysis once or twice. It re-centers the ensemble on a more accurate nonlinear solution.

Lorenz -3 variable model RMS analysis error

	4D-Var	LETKF	LETKF +QOL
Window=8 steps	0.31	0.30	**0.27**
Window=25 steps	0.53	0.66	**0.48**

Nonlinearities, "QOL" and "Running in Place"

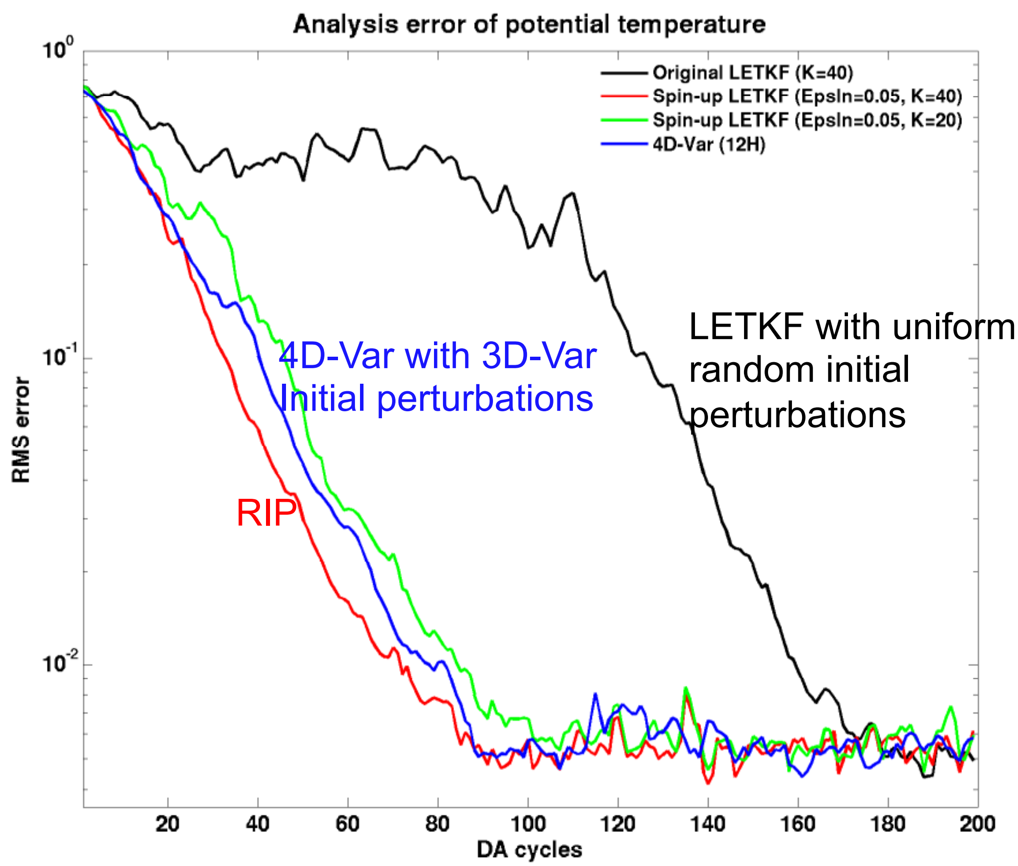
Quasi Outer Loop: similar to 4D-Var: use the final weights to correct only the mean initial analysis, keeping the initial perturbations. Repeat the analysis once or twice. It centers the ensemble on a more accurate nonlinear solution.

Lorenz -3 variable model RMS analysis error

	4D-Var	LETKF	LETKF +QOL	LETKF +RIP
Window=8 steps	0.31	0.30	**0.27**	0.27
Window=25 steps	0.53	0.68	**0.47**	**0.35**

"**Running in Place**" smoothes both the **analysis** and the **analysis error covariance** and iterates a few times…

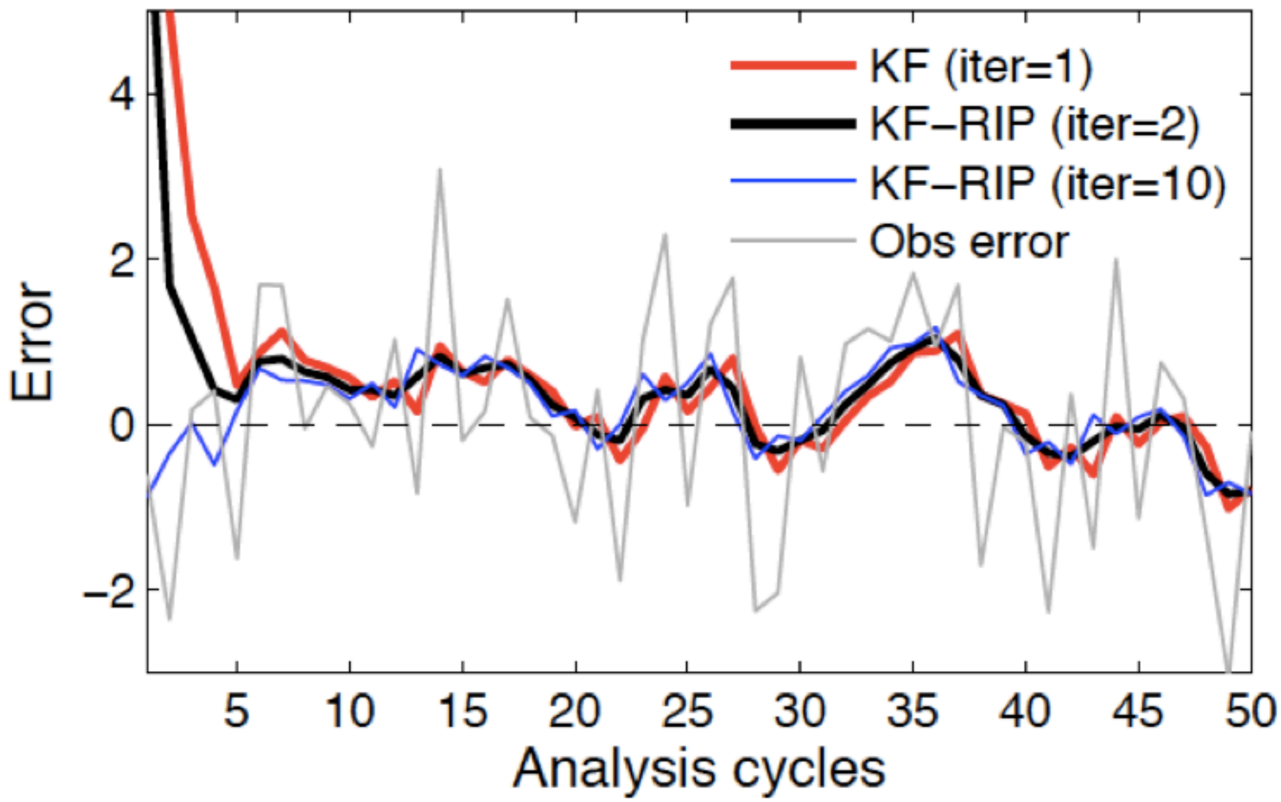
Running in Place: Spin-up with a QG model



RIP accelerates the EnKF spin-up (e.g., hurricanes, severe storms)

Spin-up depends on the initial perturbations, but RIP works well even with uniform random perturbations. RIP becomes even faster than 4D-Var (blue).

Why RIP works: Results with a Linear model

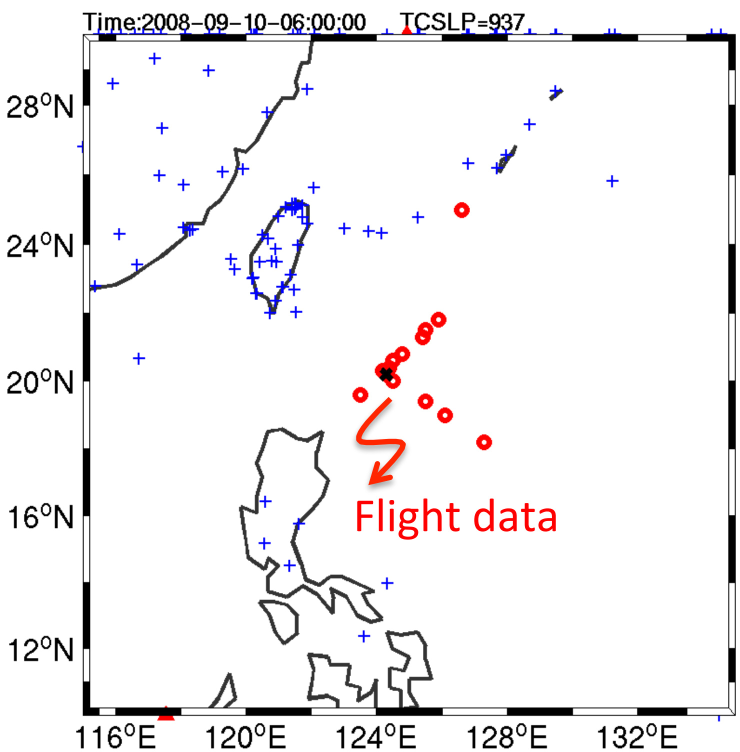


$$x_n = M(x_{n-1}) = x_{n-1} + \alpha$$

$$\sigma_n^2 = G(\sigma_{n-1}^2) = C\sigma_{n-1}^2$$

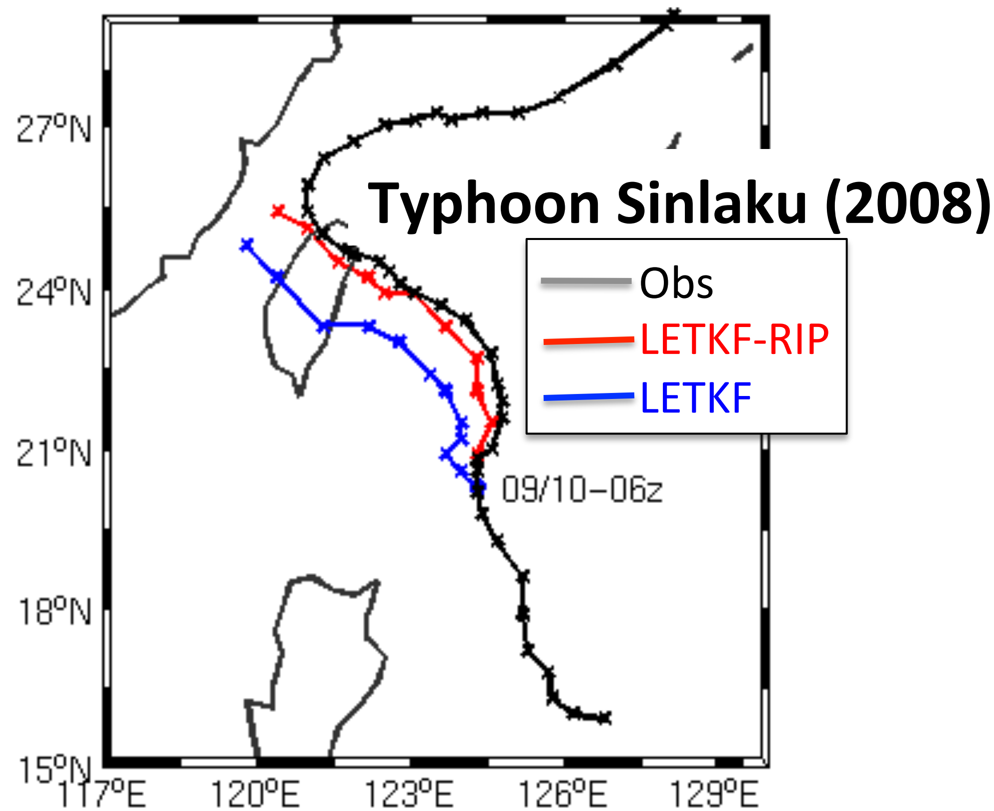
- RIP adapts to using an observation N-times by dividing the spread by N: RIP converges to the regular optimal KF solution.
- The spin-up is faster and the analysis update is "softer" (in small steps) rather than in large steps.

LETKF-RIP with real observations (Typhoon Sinlaku, 2008)



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

3-day forecast

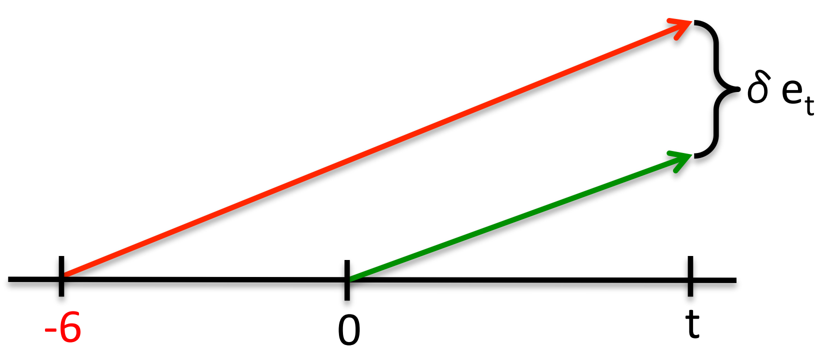


RIP uses better the "limited observations"!

Courtesy of Prof. Shu-Chih Yang (NCU, Taiwan)

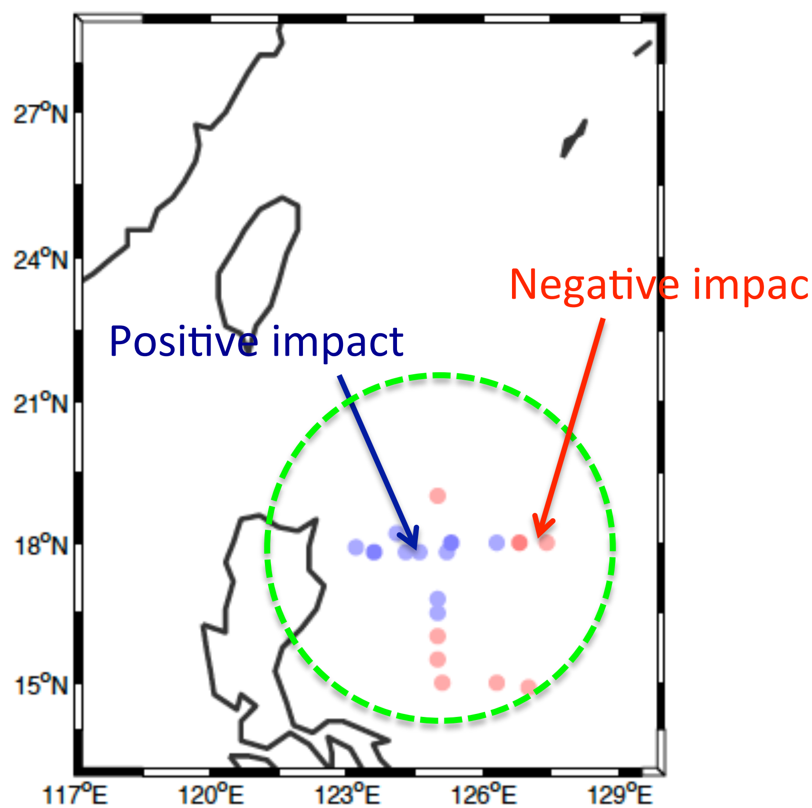
Observation impact on the forecast: Without RIP

Observations impact at t=0 on the forecast at time t
(Kalnay et al. 2012, Liu and Kalnay, 2008)



Forecast error at time t is reduced because of assimilating the observation at t=0

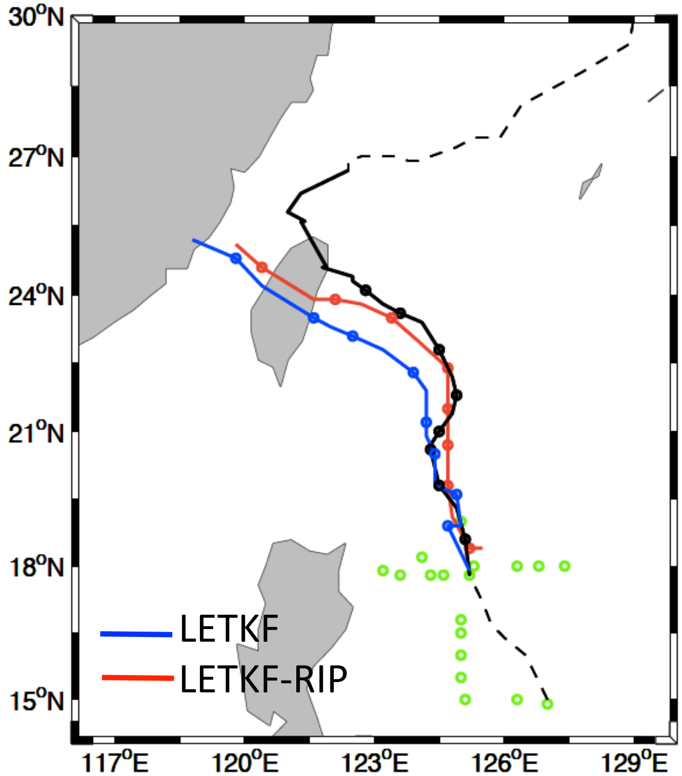
Observation impact with respect to **dropsondes** (standard LETKF)



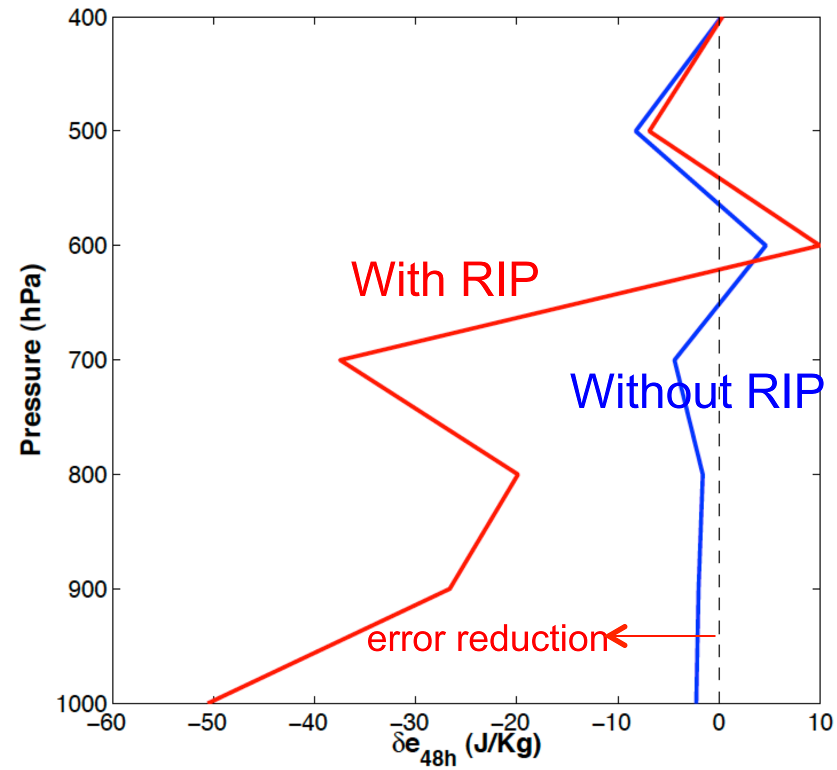
Targeted at TY's environmental condition

Observation Impact for the first set of dropsondes

4-Day track prediction initialized at 09/09 06Z



Mean observation impact of dropsondes



The effectiveness of the dropsonde data is greatly improved by RIP and the negative impact shown in the control LETKF is much reduced.

An application of LETKF-RIP to ocean data assimilation

Data Assimilation of the Global Ocean using 4D-LETKF, SODA(OI) and MOM2

Steve Penny's thesis defense

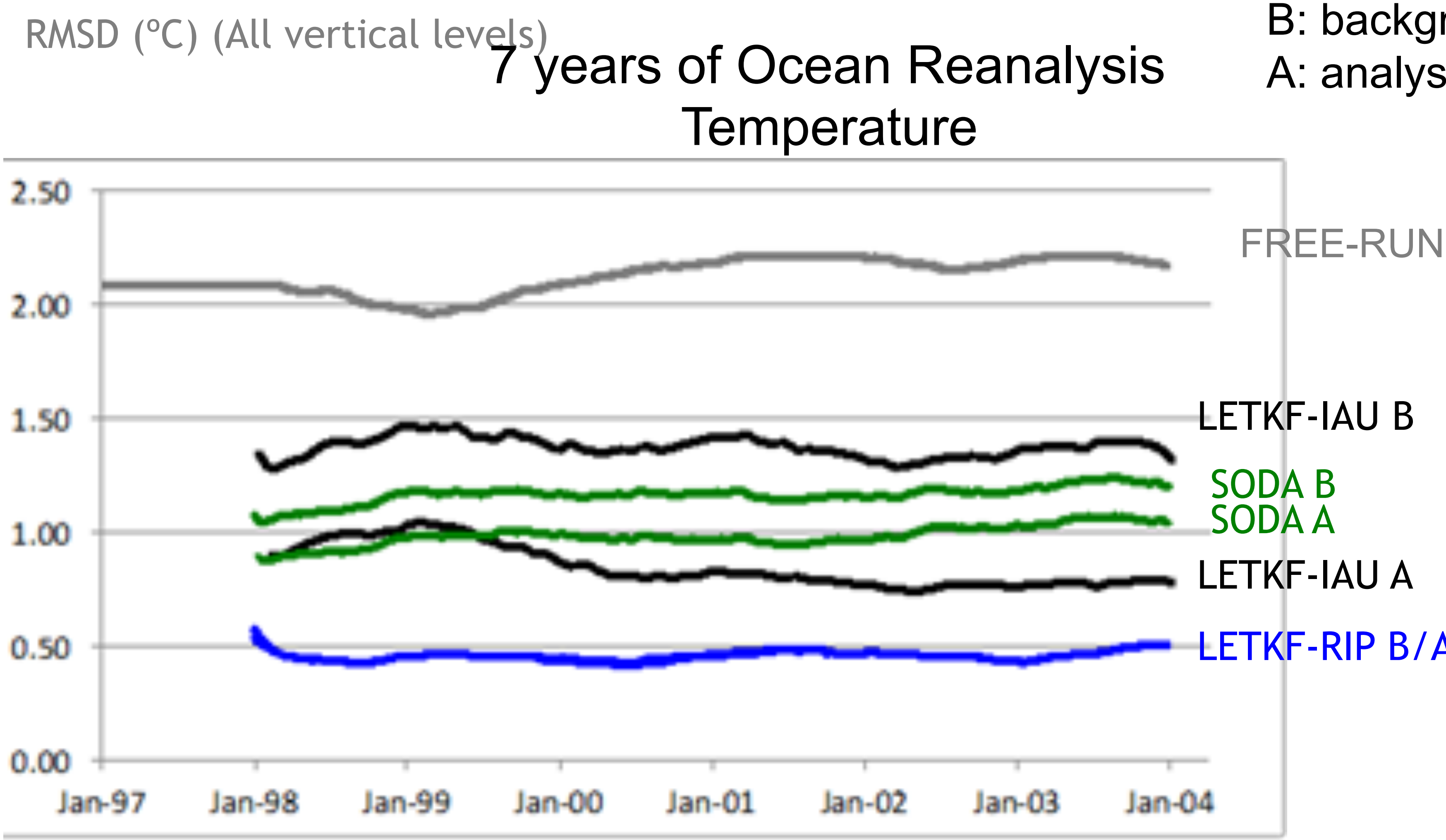
April 15, 2011

Advisors: E Kalnay, J Carton, K Ide, T Miyoshi, G Chepurin

Penny (now at UMD/NCEP) implemented the LETKF with either IAU or RIP and compared it with SODA (OI)

7 years of Ocean Reanalysis Temperature

B: background
A: analysis

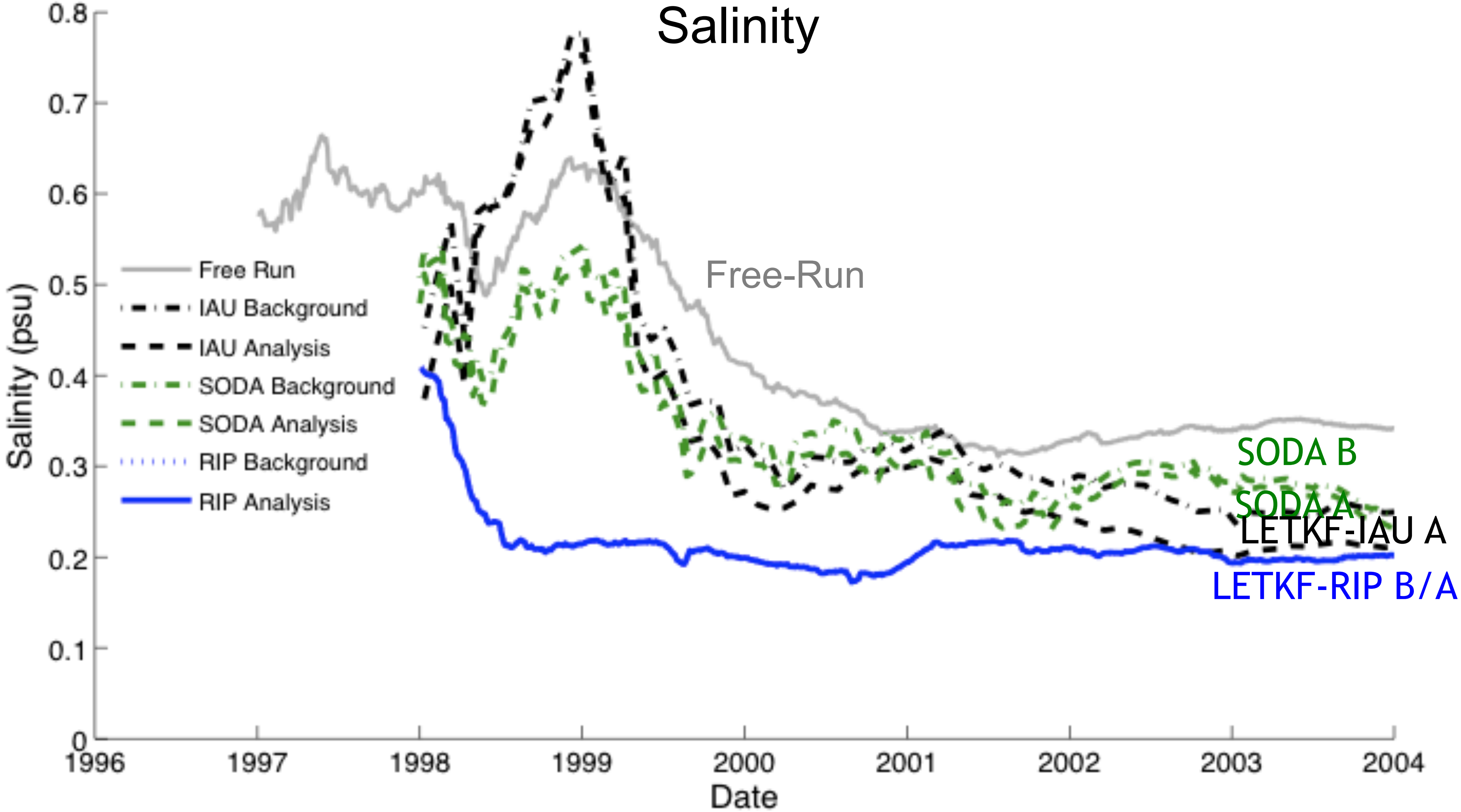


Global RMS(O-F) of Temperature (°C),
12-month moving average
LETKF (with **IAU**), **SODA** and LETKF with **RIP**

RMSD (psu) (All vertical levels)

B: background
A: analysis

7 years of Ocean Reanalysis Salinity



Global RMS(O-F) of Salinity (psu),
12-month moving average
LETKF (with **IAU**), **SODA** and LETKF with **RIP**

Why is LETKF-RIP so much better than SODA or LETKF-IAU for the ocean reanalysis?

- The ocean observations are too sparse for a standard EnKF, or even OI/3D-Var with a short (5-day) window.
- SODA and LETKF-IAU used a much longer window (30 days) in order to hammer the system with the available observations.
- LETKF-RIP uses a 5-day window but re-uses the observations in order to extract more information.

Summary for LETKF-RIP (or QOL)

- Kalman Filter is optimal for a linear, perfect model.
- During spin-up, or when the ensemble perturbations grow nonlinearly, EnKF is not optimal, since it does not extract enough information from the observations.
- The LETKF "no-cost" smoother (or, equivalently, the 4D-EnSRF) allows LETKF-RIP to use the observations more than once, and thus extract much more information.
- This shortens the spin-up and produces more accurate forecasts with the same observations.
- For linear models RIP converges to the same optimal KF solution but with spread reduced by ~ $\sqrt{N}$
- For long windows and nonlinear perturbations, RIP advances in smaller steps and approaches the true attractor more "softly".

(2) Effective Assimilation of Precipitation

(Guo-Yuan Lien, E. Kalnay and T Miyoshi)

- Assimilation of precipitation has been done by changing the moisture *Q* in order to make the model "rain as observed".
- Successful during the assimilation: e.g. the North American Regional Reanalysis had perfect precipitation!
- However the model **forgets** about the changes soon after the assimilation stops!
- The model **will remember** potential vorticity (PV).
- EnKF should modify PV efficiently, since the analysis weights will be larger for an ensemble member that is raining more correctly, because it has a better PV.

- However, 5 years ago, we had tried assimilating precipitation observations in a LETKF-SPEEDY model simulation but the results were POOR!
- Big problem: precipitation is not Gaussian.
- We tried a Gaussian transformation of precipitation and it worked!

How do we transform precipitation *y* to a Gaussian y_{transf}?

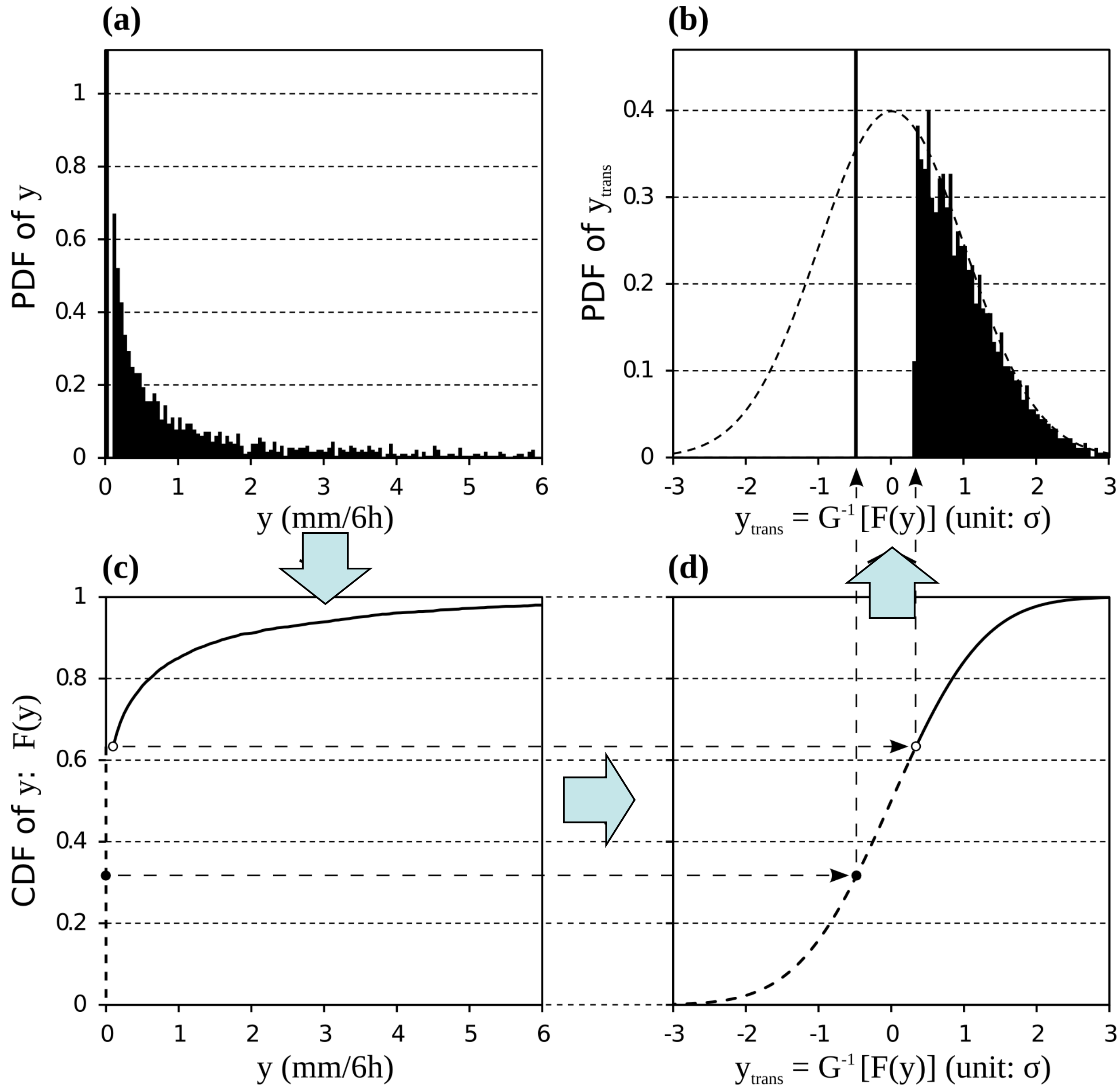
Start with pdf of *y*=rain at every grid point.

"No rain" is like a delta function that we cannot transform.

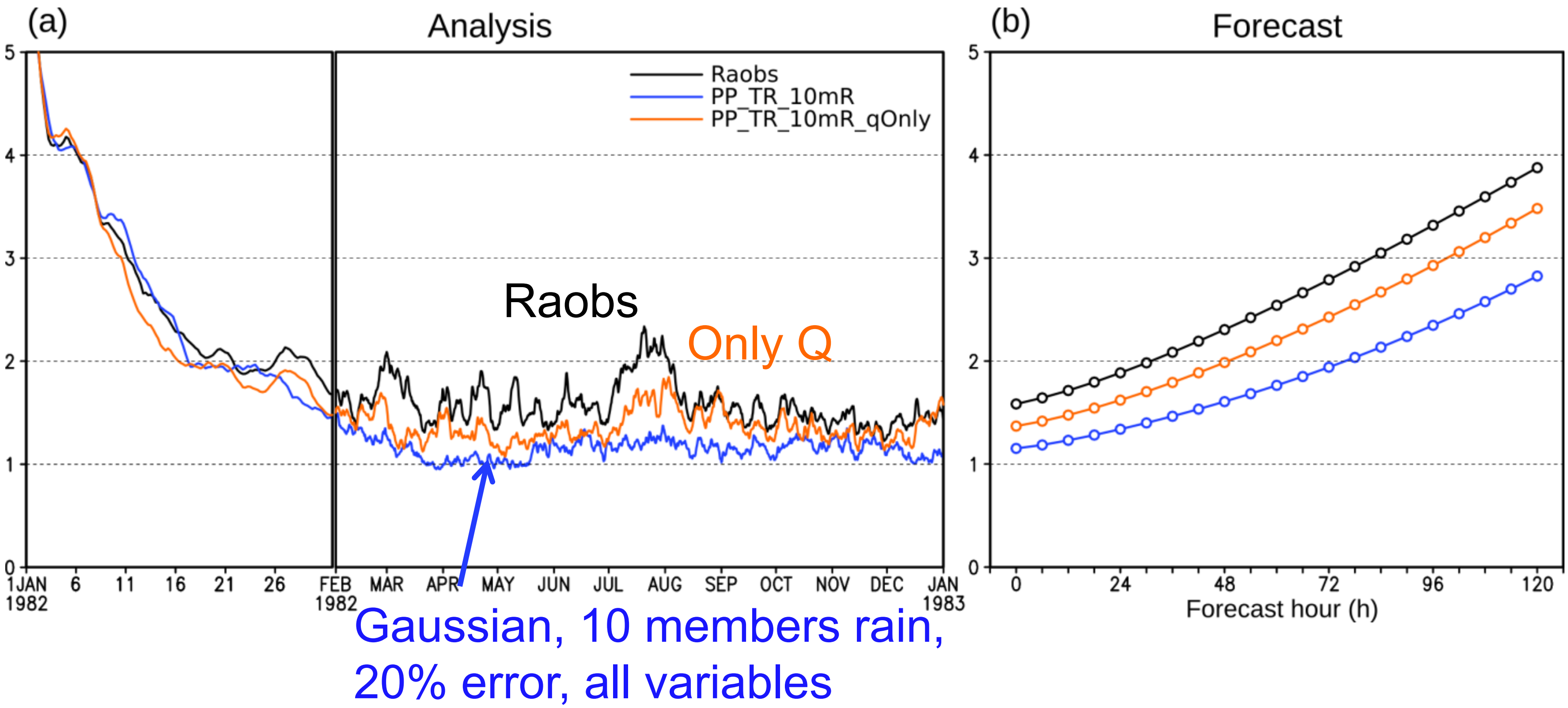
We assign all "no rain" to the **median** of the no rain CDF.

We found this works as well as more complicated procedures.

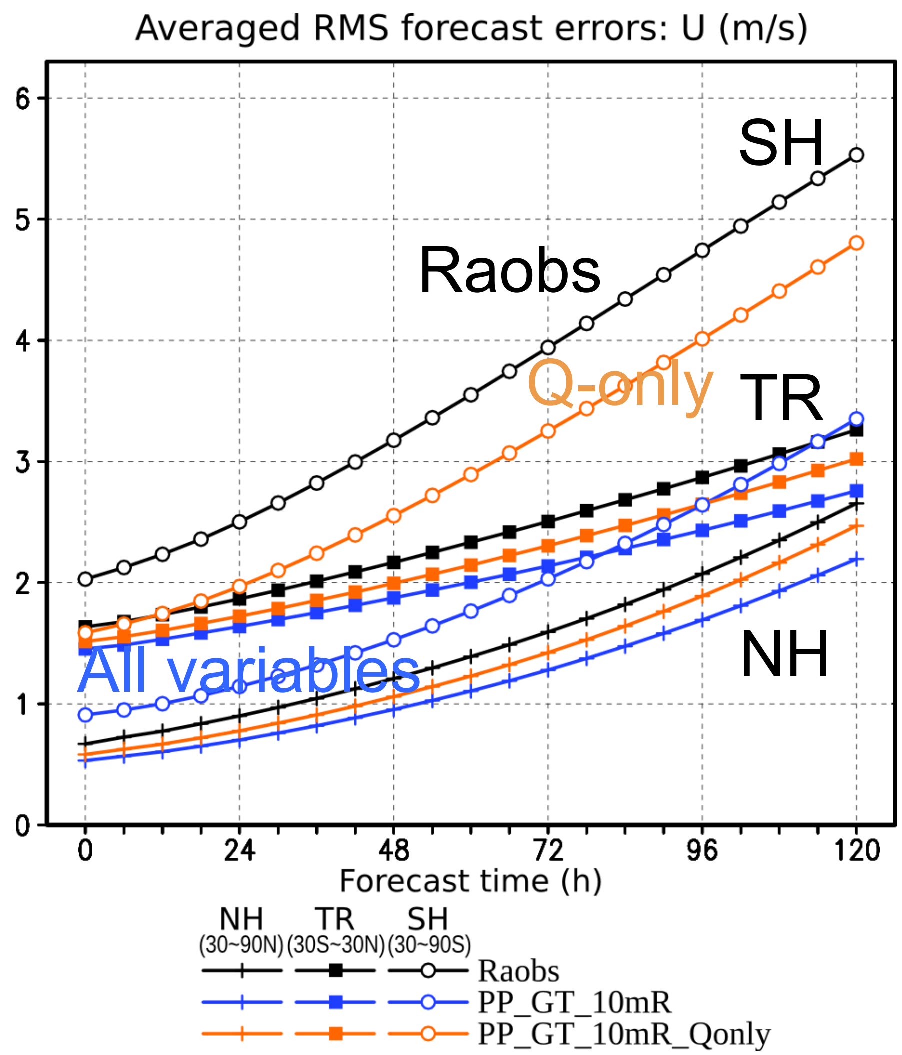
It allows to assimilate both rain and no rain.



$$G^{-1}(x) = \sqrt{2}\, erf^{-1}(2x-1)$$

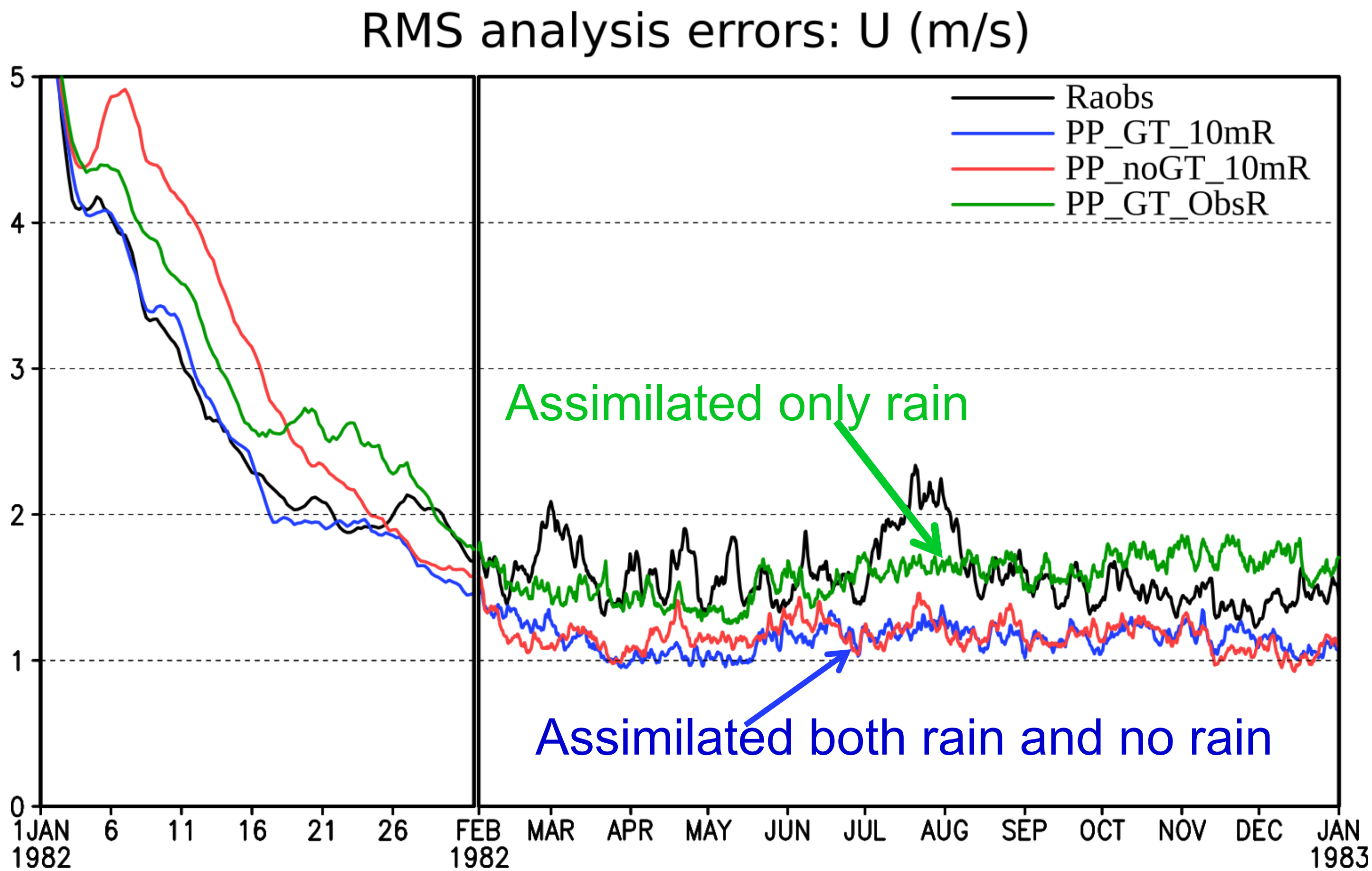


- Main result: with at least 10 ensemble members raining in order to assimilate an obs, updating all variables (including vorticity), with Gaussian transform, and rather accurate observations (20% errors), the analyses and forecasts are much improved!
- Updating only Q is much less effective.
- **The 5-day forecasts maintain the advantage.**

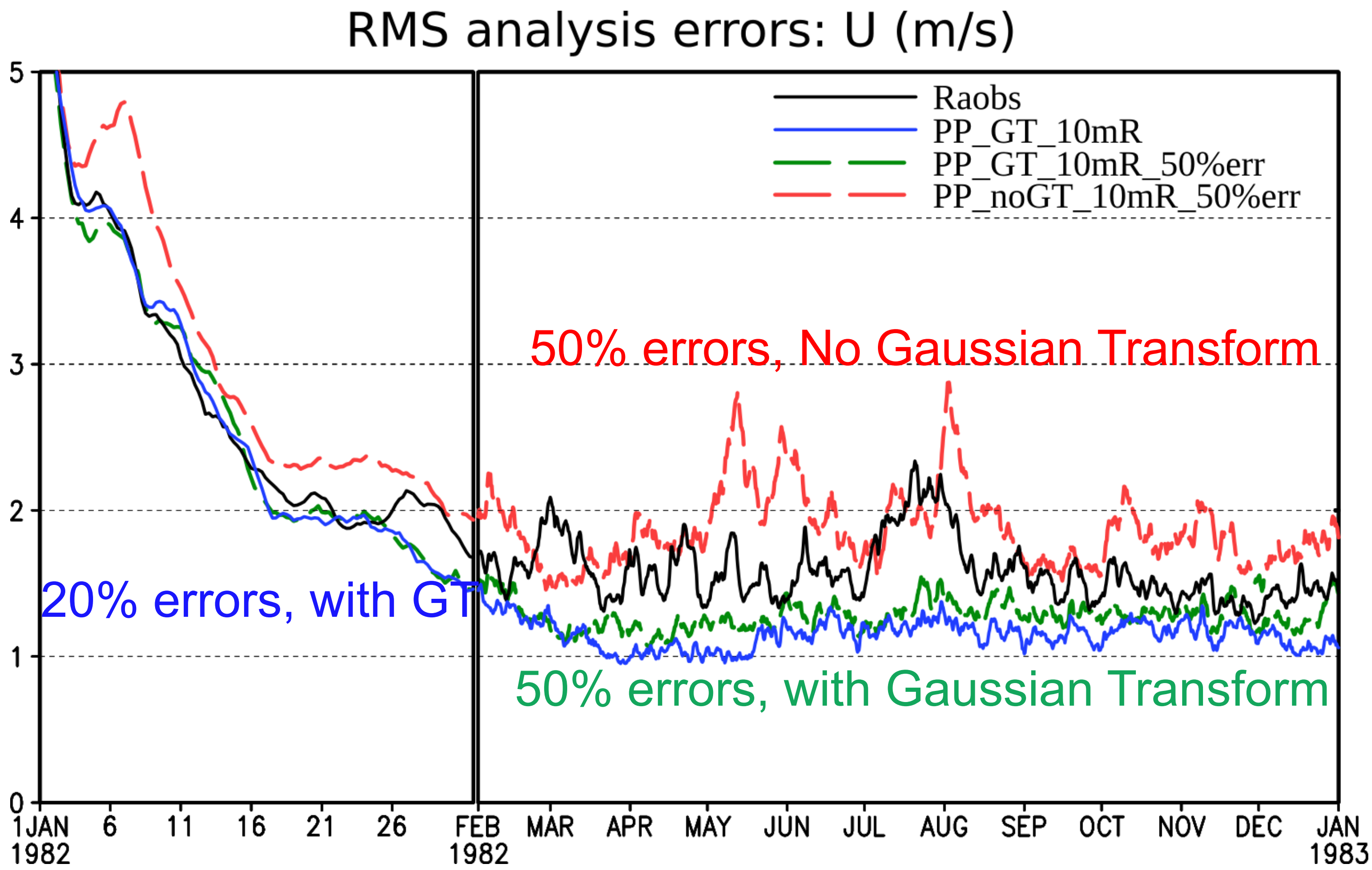


One year of 5-day forecasts

The model remembers the impact of pp assimilation in the SH, NH and tropics!

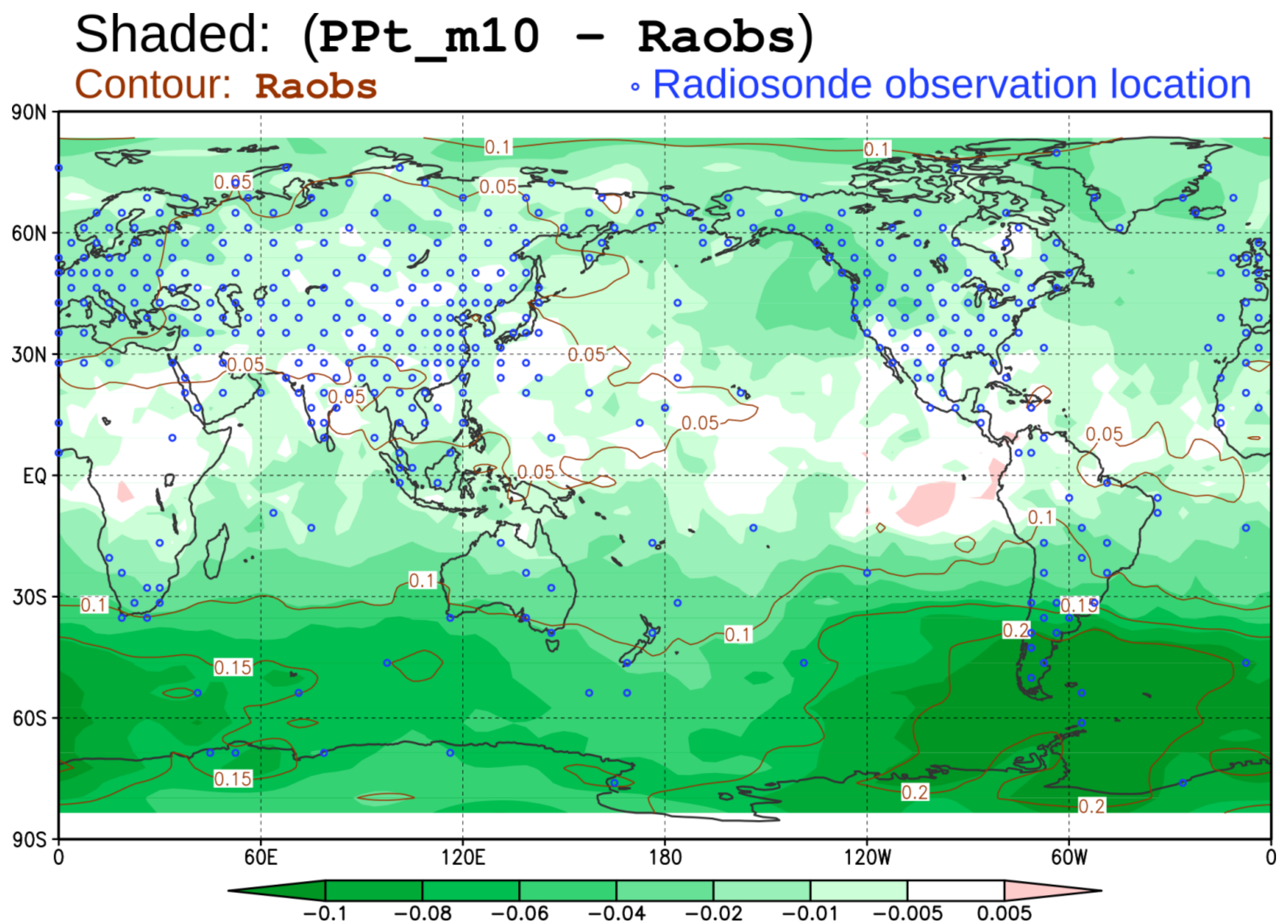


If we assimilate only rain the results are much worse!
We need to assimilate both rain and no rain!



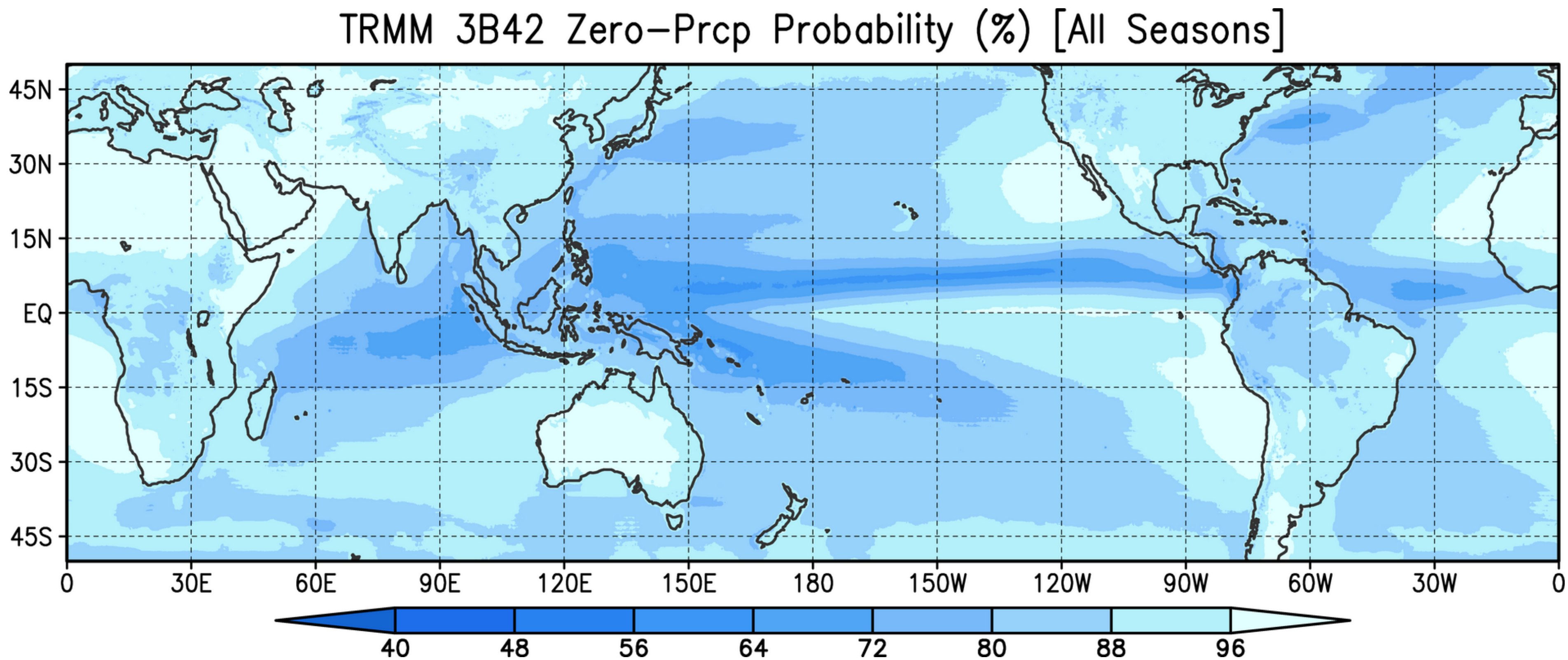
The impact of the Gaussian Transform is important with large observation errors (50% rather than 20%). The impact of GT50% is almost as good as GT20%.

Vorticity errors and corrections



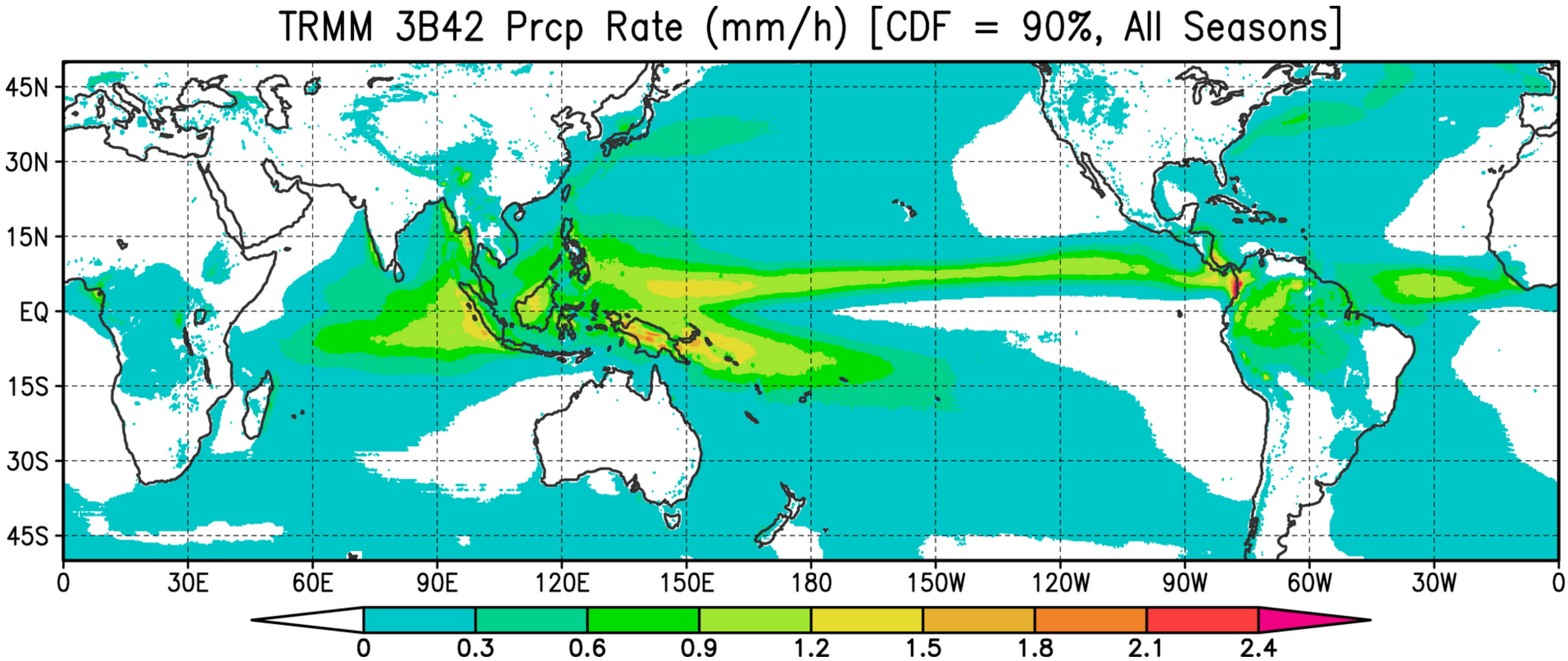
There is no vorticity information in the pp observations, but the LETKF clearly knows about the vorticity errors

How about real observations?
We will use TRMM/TMPA satellite estimates (from G. Huffman) with the NCEP GFS



TRMM/TMPA: 14 years of data, 50S-50N, 3hrs, 0.5 deg

TRMM/TMPA (data from G. Huffman)

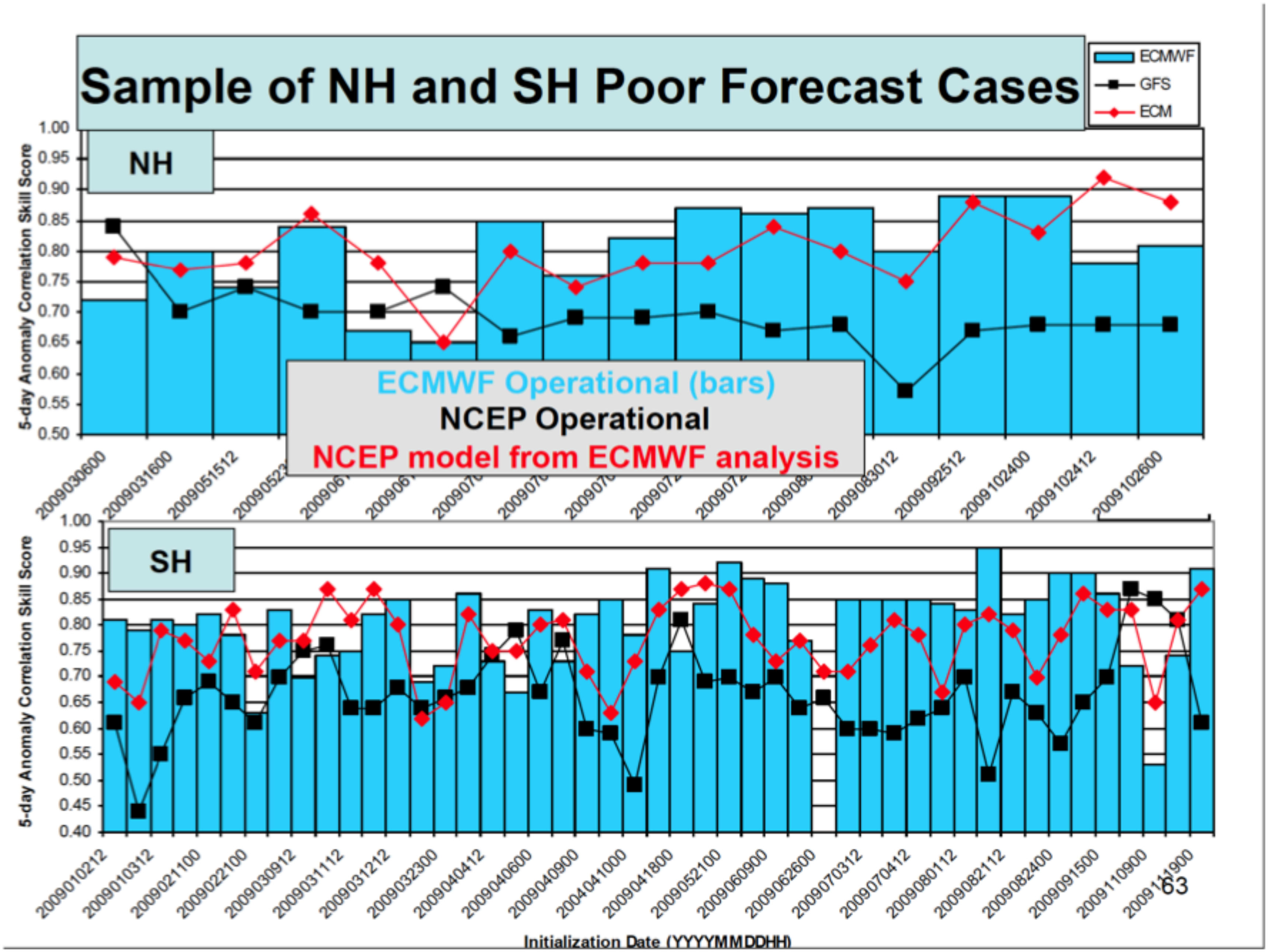


TRMM/TMPA: 14 years of data, 50S-50N, 3hrs, 0.5 deg

Summary for assimilation of precipitation

- The model remembers potential vorticity (dynamics), does not remember moisture changes, or even temperature.
- For this reason, when using nudging, or variational assimilation of precipitation to change Q and T, the model "forgets" this information and returns to the original forecast.
- EnKF has a better chance to assimilate potential vorticity by giving higher weights to ensemble members with good precip.
- In addition, EnKF has the advantage of not requiring model linearization, a problem for variational systems.
- We found that EnKF with a Gaussian transformation of precipitation assimilates rain info and remembers it during the forecast.
- Requiring at least several ensemble forecasts to have Rain>0 allows the effective assimilation of both rain and no rain.

The NCEP 5-day skill dropout problem

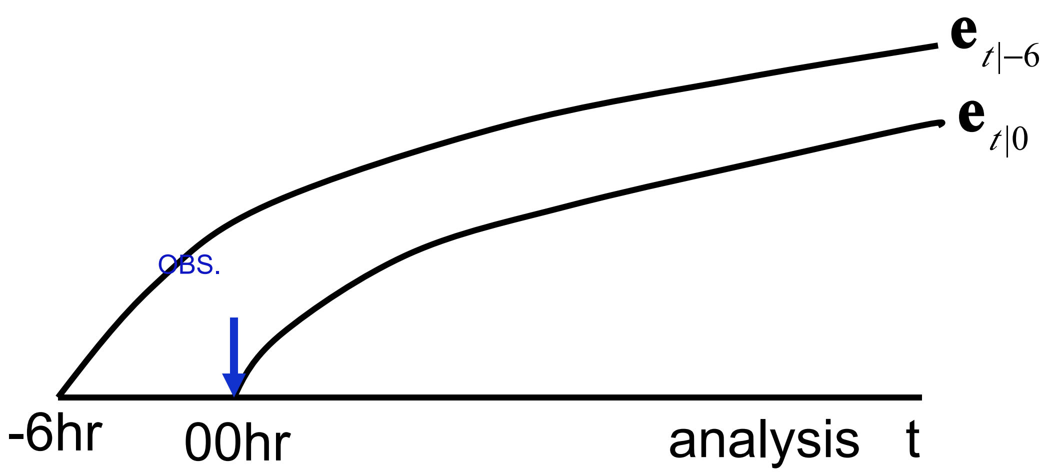


Ensemble Forecast Sensitivity to Observations

"Adjoint sensitivity without adjoint" (Liu and K, 2008, Li et al., 2010)

Here we show a simpler, more accurate formulation

(Kalnay, Ota, Miyoshi: Tellus, 2012)



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

(Adapted from Langland and Baker, 2004)

The only difference between $\mathbf{e}_{t|0}$ and $\mathbf{e}_{t|-6}$ is the assimilation of observations at 00hr:

$$(\overline{\mathbf{x}}^a_0 - \overline{\mathbf{x}}^b_{0|-6}) = \mathbf{K}(\mathbf{y} - H(\mathbf{x}^b_{0|-6}))$$

- Observation impact on the reduction of forecast error:

$$\Delta\mathbf{e}^2 = (\mathbf{e}^T_{t|0}\mathbf{e}_{t|0} - \mathbf{e}^T_{t|-6}\mathbf{e}_{t|-6}) = (\mathbf{e}^T_{t|0} - \mathbf{e}^T_{t|-6})(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6})$$

Ensemble Forecast Sensitivity to Observations

$$\Delta \mathbf{e}^2 = (\mathbf{e}_{t|0}^T \mathbf{e}_{t|0} - \mathbf{e}_{t|-6}^T \mathbf{e}_{t|-6}) = (\mathbf{e}_{t|0}^T - \mathbf{e}_{t|-6}^T)(\mathbf{e}_{t|0} + \mathbf{e}_{t|-6})$$

$$= (\overline{\mathbf{x}}_{t|0}^f - \overline{\mathbf{x}}_{t|-6}^f)^T (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6})$$

$$= \left[\mathbf{M}(\overline{\mathbf{x}}_0^a - \overline{\mathbf{x}}_{0|-6}^b)\right]^T (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}), \text{ so that}$$

$$\boxed{\Delta \mathbf{e}^2 = \left[\mathbf{MK}(\mathbf{y} - H(\mathbf{x}_{0|-6}^b))\right]^T (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6})}$$

Langland and Baker (2004), Gelaro, solve this with the adjoint:

$$\boxed{\Delta \mathbf{e}^2 = \left[(\mathbf{y} - H(\mathbf{x}_{0|-6}^b))\right]^T \mathbf{K}^T \mathbf{M}^T (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6})}$$

This requires the adjoint of the model $\mathbf{M}^T$ and of the data assimilation system $\mathbf{K}^T$ (Langland and Baker, 2004)

Ensemble Forecast Sensitivity to Observations

Langland and Baker (2004):

$$\Delta \mathbf{e}^2 = \left[\mathbf{MK}(\mathbf{y} - H(\mathbf{x}^b_{0|-6})) \right]^T (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6})$$
$$= \left[(\mathbf{y} - H(\mathbf{x}^b_{0|-6}) \right]^T \mathbf{K}^T \mathbf{M}^T (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6})$$

With EnKF we can use the original equation without "adjointing":

Recall that $\mathbf{K} = \mathbf{P}^a \mathbf{H}^T \mathbf{R}^{-1} = 1/(K-1)\mathbf{X}^a \mathbf{X}^{aT} \mathbf{H}^T \mathbf{R}^{-1}$ so that

$$\mathbf{MK} = \mathbf{MX}^a (\mathbf{X}^{aT} \mathbf{H}^T) \mathbf{R}^{-1} / (K-1) = \mathbf{X}^f_{t|0} \mathbf{Y}^{aT} \mathbf{R}^{-1} / (K-1)$$

Thus,

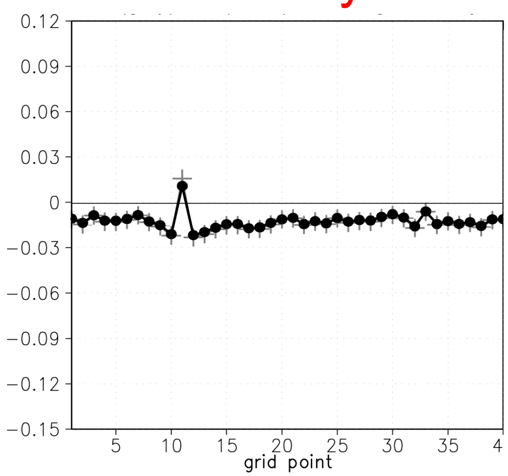
$$\Delta \mathbf{e}^2 = \left[\mathbf{MK}(\mathbf{y} - H(\mathbf{x}^b_{0|-6})) \right]^T (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6})$$
$$= \left[(\mathbf{y} - H(\mathbf{x}^b_{0|-6}) \right]^T \mathbf{R}^{-1} \mathbf{Y}^a_0 \mathbf{X}^{fT}_{t|0} (\mathbf{e}_{t|0} + \mathbf{e}_{t|-6}) / (K-1)$$

This uses the available **nonlinear** forecast ensemble products.

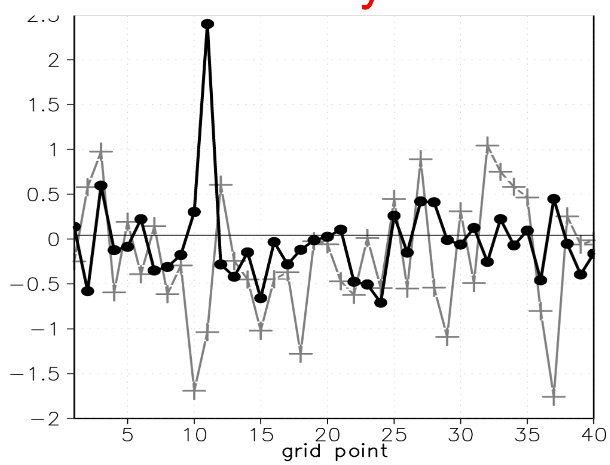
Tested ability to detect a poor quality ob impact on the forecast in the Lorenz 40 variable model

Observation impact from LB(+) and from ensemble sensitivity (•)

1 day



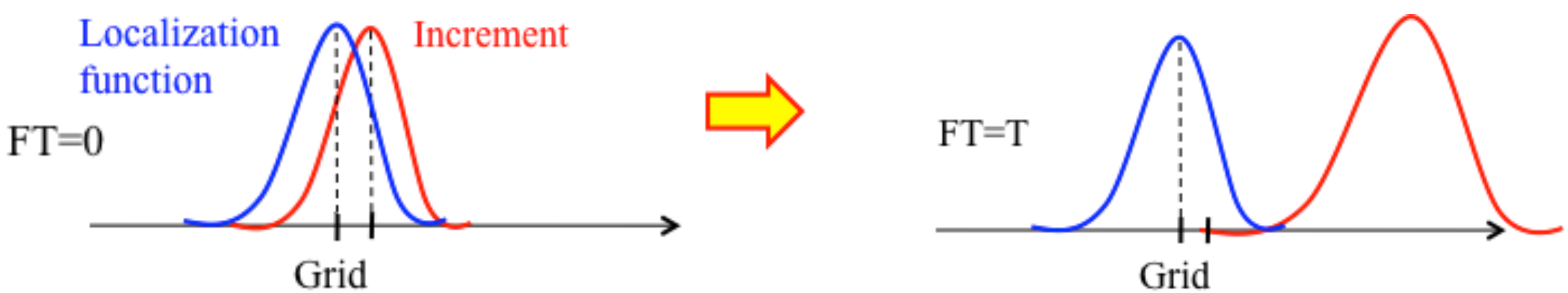
10 days



✓The adjoint and the ensemble sensitivity give similar observation impact on the 24 hr forecast.

✓The ensemble sensitivity is nonlinear and is able to detect bad obs for longer forecasts

✓This was done ignoring EnKF localization

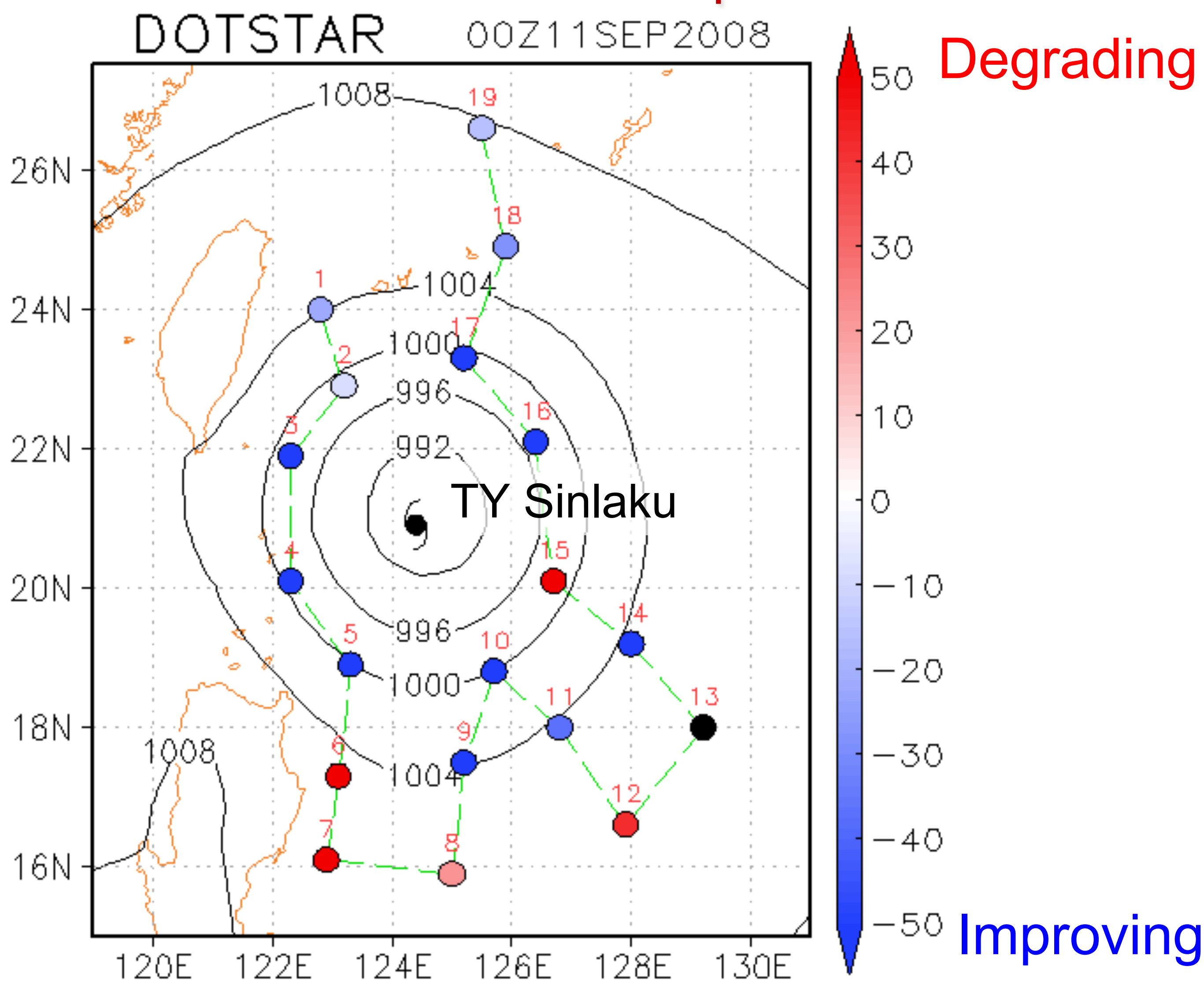


The localization center point for observation impact estimate is now moved with the horizontal wind: <u>an approximation</u>

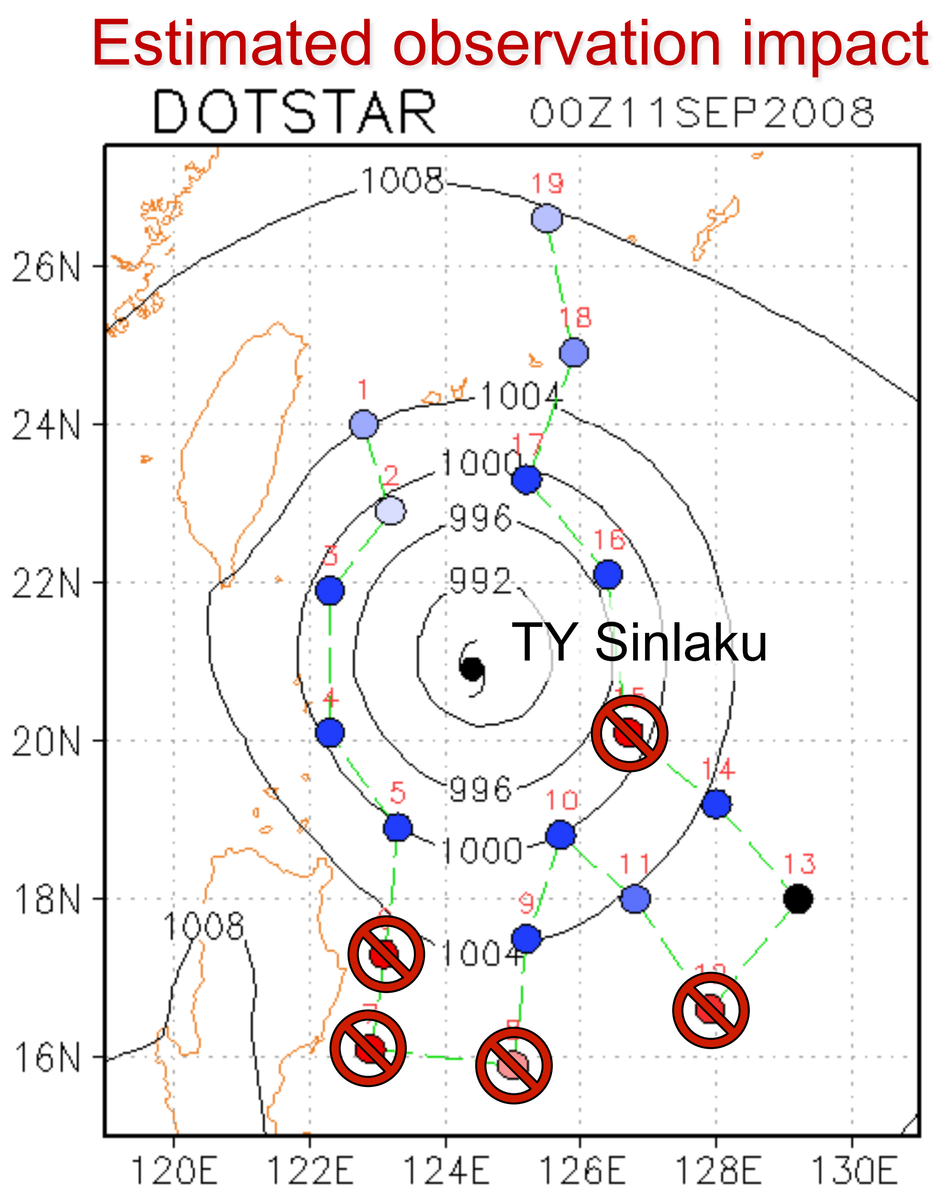
Impact of dropsondes on a Typhoon

(Kunii et al. 2012)

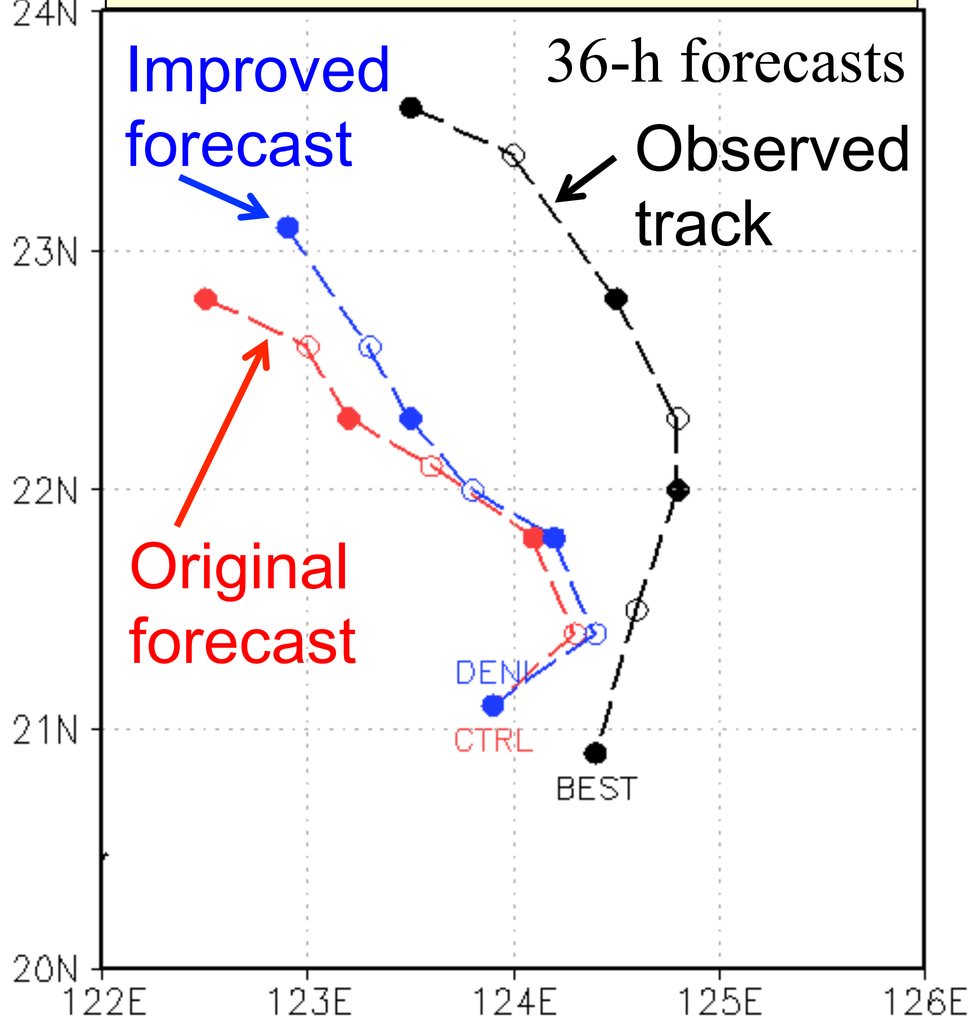
Estimated observation impact



Denying negative impact data improves forecast!



Typhoon track forecast is actually improved!!



Ota et al. 2012: Applied EFSO to NCEP GFS/ EnSRF using all operational observations. Determined regional 24hr "forecast failures"

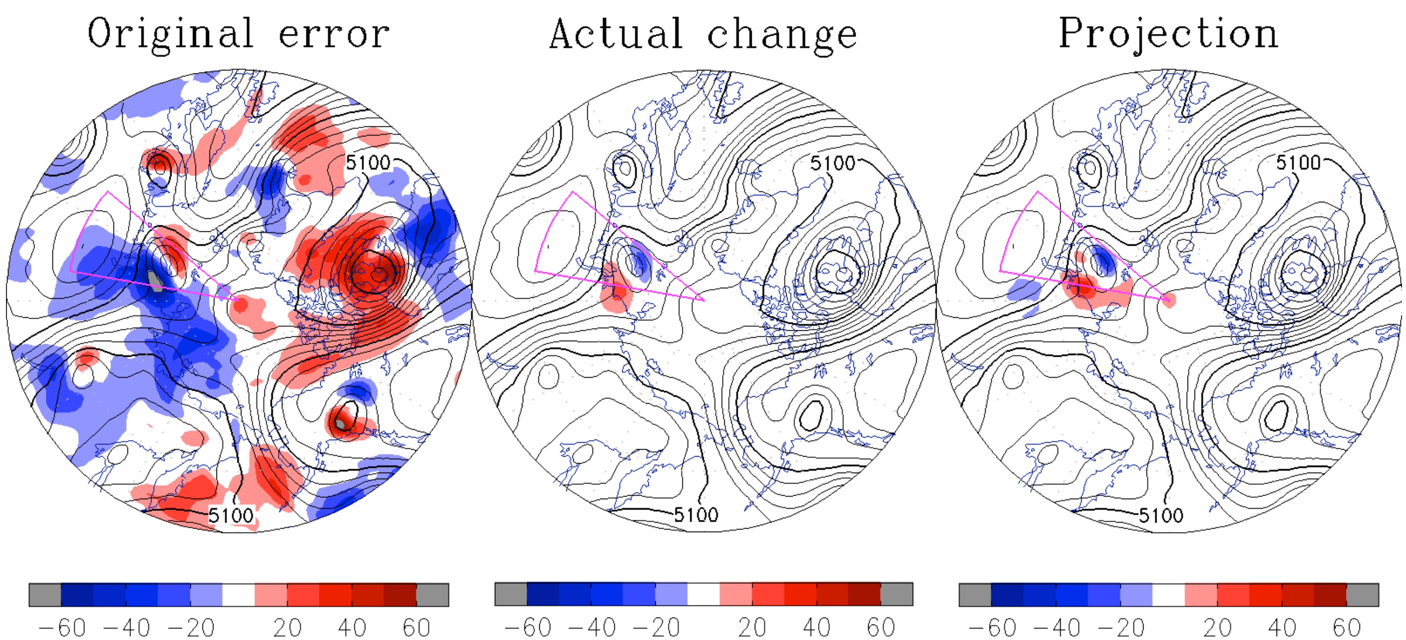
- Divide the globe into 30x30^{o} regions
- Find all cases where the 24hr regional forecast error is at least 20% larger than the 36hr forecast error verifying at the same time, and
- where the 24hr forecast has errors at least twice the time average.
- Identify the top observation type that has a negative impact on the forecast.
- Found 7 cases of 24hr forecast

24-hr forecast error correction (Ota et al.)

- identified 7 cases of large 30°x30° regional errors,
- rerun the forecasts denying bad obs.
- the forecast errors were substantially reduced
- this could be applied to improve the 5-day skill dropouts

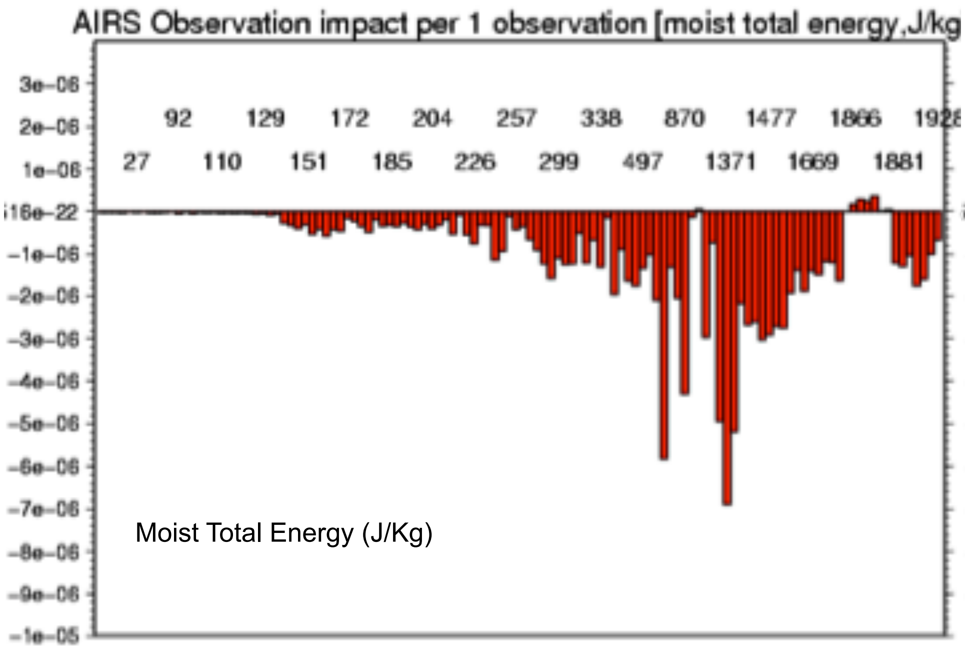
Initial	Area	Size	Rate	N	Denied observation	Change
12 UTC JAN 10	90S~60S 100E~130E	2.04	1.20	1	GPSRO (80S~60S, 90E~120E) ASCAT (60S~50S, 100E~120E)	-6.6%
06 UTC JAN 12	50N~80N 150E ~ 180	2.18	1.40	1	AMSUA (ch4: 45N~75N, 160E~170W, ch5:40N~55N, 155E~180, NOAA15 ch6: 50N~75N, 140E~170W, ch7: 70N~80N, 130E~170E)	-11.4%
00 UTC JAN 16	30N~60N 30W~0	2.13	1.31	2	Radiosonde wind (Valentia, Ireland), ASCAT (40N~47N, 20W~10W, 50N~55N, 35W~30W)	-1.0%
12 UTC JAN 22	90S~60S 130E~160E	2.34	1.22	2	AMSUA (ch5: 65S~50S, 90E~110E, 60S~50S, 120E~127E, ch6: 60S~45S, 110E~125E)	-2.2%
06 UTC FEB 2	50N~80N 150W~120W	3.10	1.32	4	IASI (35N~45N, 155W~150W) NEXRAD (55N~60N, 160W~135W)	-5.5%
MODIS 18 UTC FEB 6	60N~90N 50E~80E	2.06	1.71	2	MODIS_Wind (60N~90N, 30E~90E)	-39.0%
18 UTC FEB 6	90S~60S 20W~10E	3.56	1.22	1	MODIS_Wind (80S~50S, 30W~0)	-22.5%

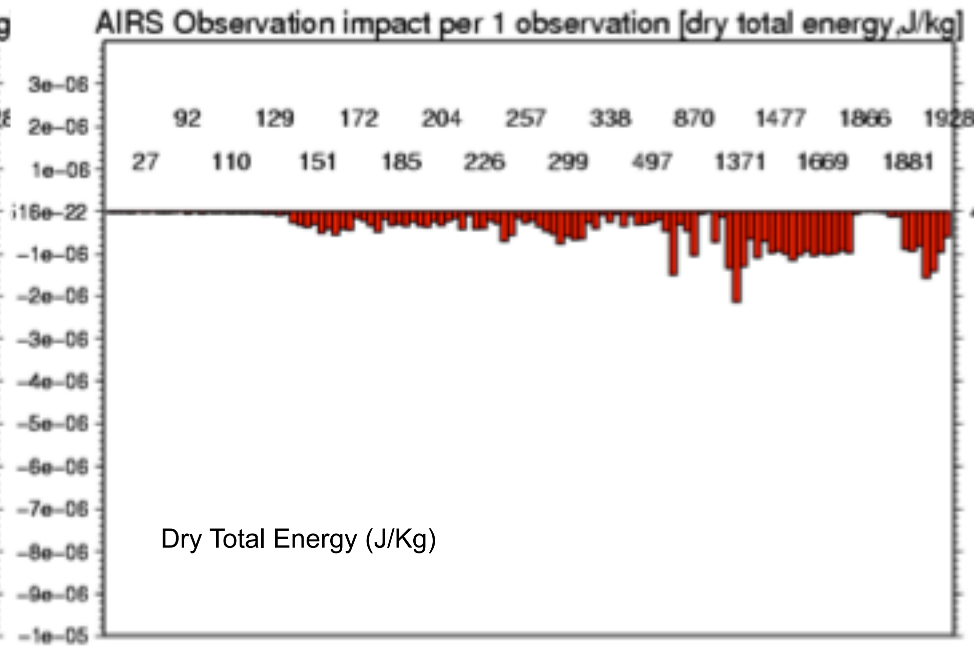
"Proactive" QC: Bad observations can be identified by EFSO and withdrawn from the data assimilation



After identifying MODIS polar winds producing bad 24 hr regional forecasts, the withdrawal of these winds reduced the forecast errors by 39%, as projected by EFSO.

Other applications: Impacts of Observing Systems





The EnKF formulation is nonlinear and thus allows computing Moist Total Energy and estimate more accurately the impact of the channels on the moisture forecast. Adjoint formulation needs TLM.

Summary and the future

- The new EFSO formulation works well and uses available EnKF products.
- It can be used to detect observations that give bad regional 12hr or 24hr forecasts.
- We can then repeat the data assimilation without the bad obs, a powerful tool for a "proactive" QC and monitoring.

- **EnKF is a newer, much simpler technology.**
- **There is much more potential not yet exploited or not even explored:**
  - **Estimation and correction of model errors and parameters (Ruiz et al, Danforth et al, Kang et al)**
  - **…**

Simultaneous data assimilation of CO_2 and meteorological variables within LETKF coupled with NCAR CAM model

*__Ji-Sun Kang__, *Eugenia Kalnay, +Junjie Liu, #Inez Fung, and *Takemasa Miyoshi

*University of Maryland, College Park,MD

+NASA/JPL, Pasadena, CA

University of California, Berkeley, CA

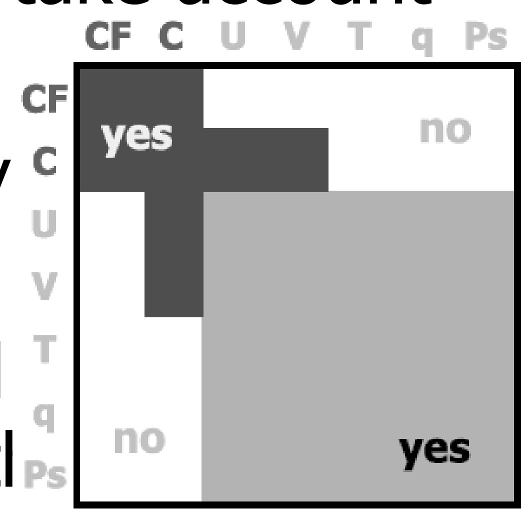
AGU fall meeting, 2012

Ensemble CO_2 Data Assimilation

- **Local Ensemble Transform Kalman Filter** (LETKF, Hunt et al. 2007) data assimilation system has been applied to analyze atmospheric CO_2 (C) and surface CO_2 fluxes (CF) in addition to meteorological variables (U, V, T, q, Ps)
  - **UMD-UCB LETKF-C** (Kang et al., 2011, 2012)
  - Unlike inverse methods, it assimilates atmospheric CO_2 as well as surface carbon fluxes
  - It assimilates observations of meteorological variables and atmospheric CO_2 simultaneously
  - Ensemble forecast within LETKF includes the uncertainty of surface CO_2 forcing as well as transport errors (forecast uncertainty of wind fields)
  - It uses an assimilation window of only 6 hrs

UMD-UCB LETKF-C

- Multivariate data assimilation with "**localization of variables**"
  - We zero out the error covariance between some variables, because CO_2 does not have a strong physical relation with every variables in the state vector, so that sampling errors are reduced
  - Analysis includes error covariance between atmospheric CO_2 and wind fields to take account for transport errors of CO
  - It has been tested very successfully experiments
  - Other configurations of background covariance matrix can be tested witl



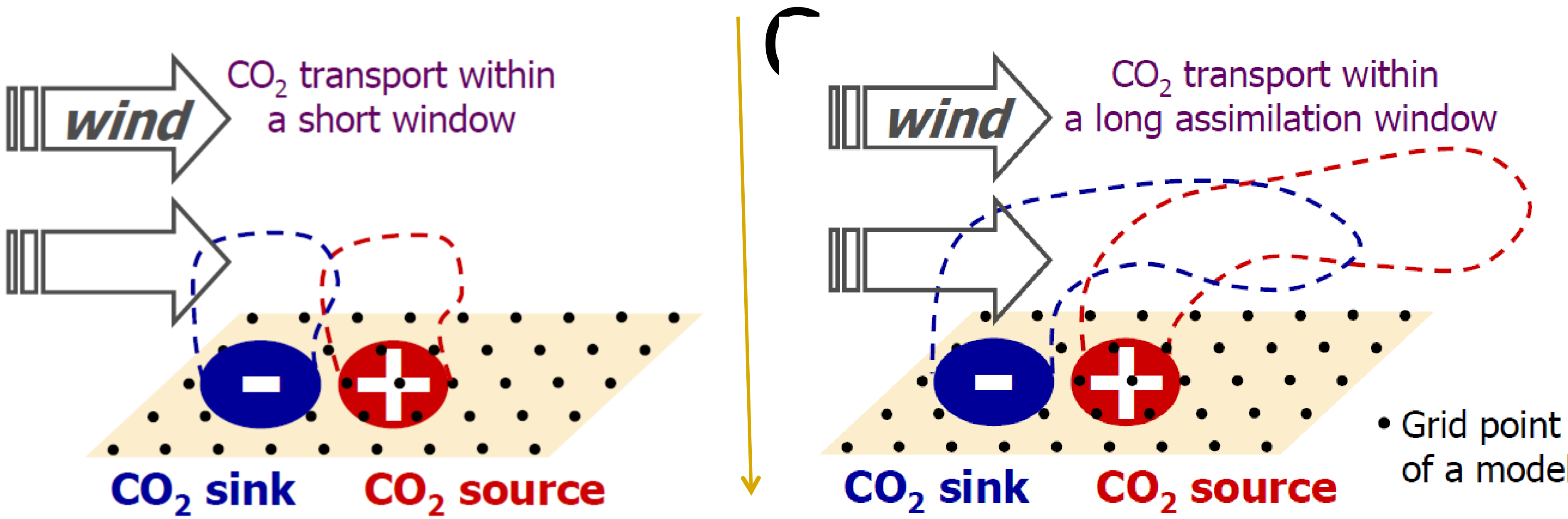
(Kang et al. 2011, JGR)

UMD-UCB LETKF-C

- Advanced inflation methods:
  - Adaptive multiplicative inflation (Miyoshi, 2011)
  - Additive inflation for parameters
- Vertical localization of column mixing CO_2 observations (GOSAT, OCO-2)
  - Emphasizes the lower levels of variability within the column, even though we use column observations whose sensitivity (averaging kernel) is fairly uniform in the vertical.
- **We use a short (6-hour) assimilation window**
- In contrast, CO_2 inversion methods adopt much **longer window lengths** (weeks to months)

(Kang et al. 2012, JGR)

Assimilation window in LETKF-C



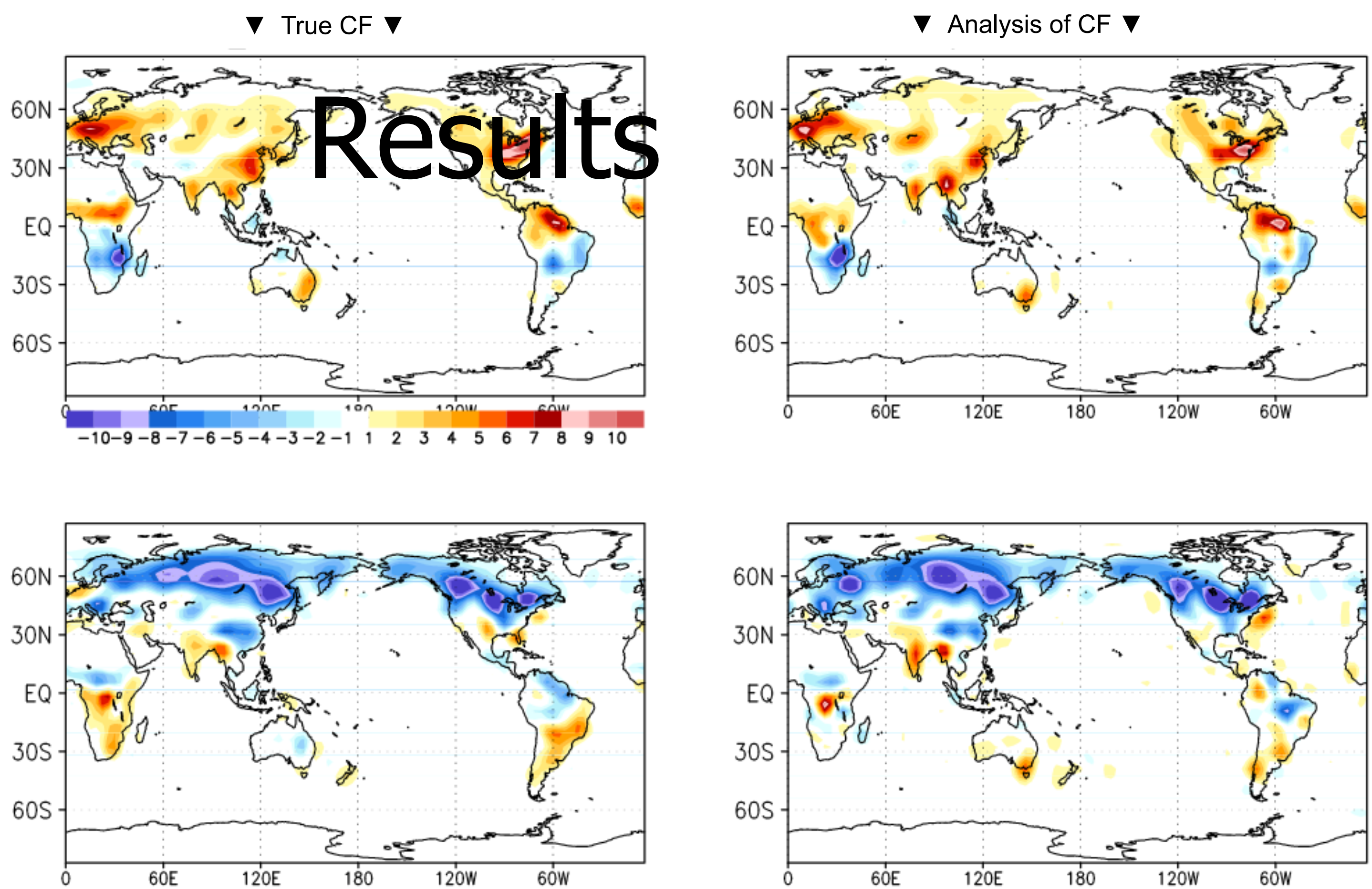
A short assimilation window reduces the attenuation of observed CO_2 information: the analysis system can use the strong correlation between C and CF before the transport of atmospheric CO_2 **blurs out** the essential information of surface CO_2 forcing.

A long assimilation window blurs this information and introduces sampling errors into the EnKF analysis

LETKF-C with SPEEDY-C

- Model: **SPEEDY-C** (Molteni, 2003; Kang, 2009)
  - Spectral AGCM model with T30L7
  - Prognostic variables: U, V, T, q, Ps, C
    - C (atmospheric CO_2): an inert tracer
  - Persistence forecast of CF
- Simulated observations
  - Rawinsonde observations of U, V, T, q, Ps
  - Ground-based observations of atmospheric CO_2
    - 18 hourly and 107 weekly data on the globe
  - Remote sensing data of column mixing CO_2
    - **AIRS** whose averaging kernel peaks at mid-troposphere

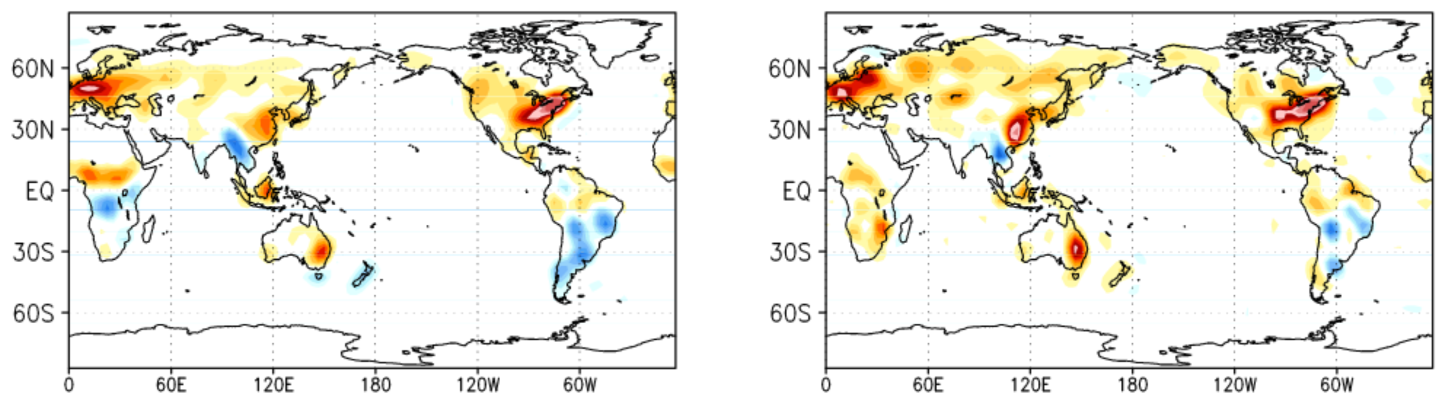
Results



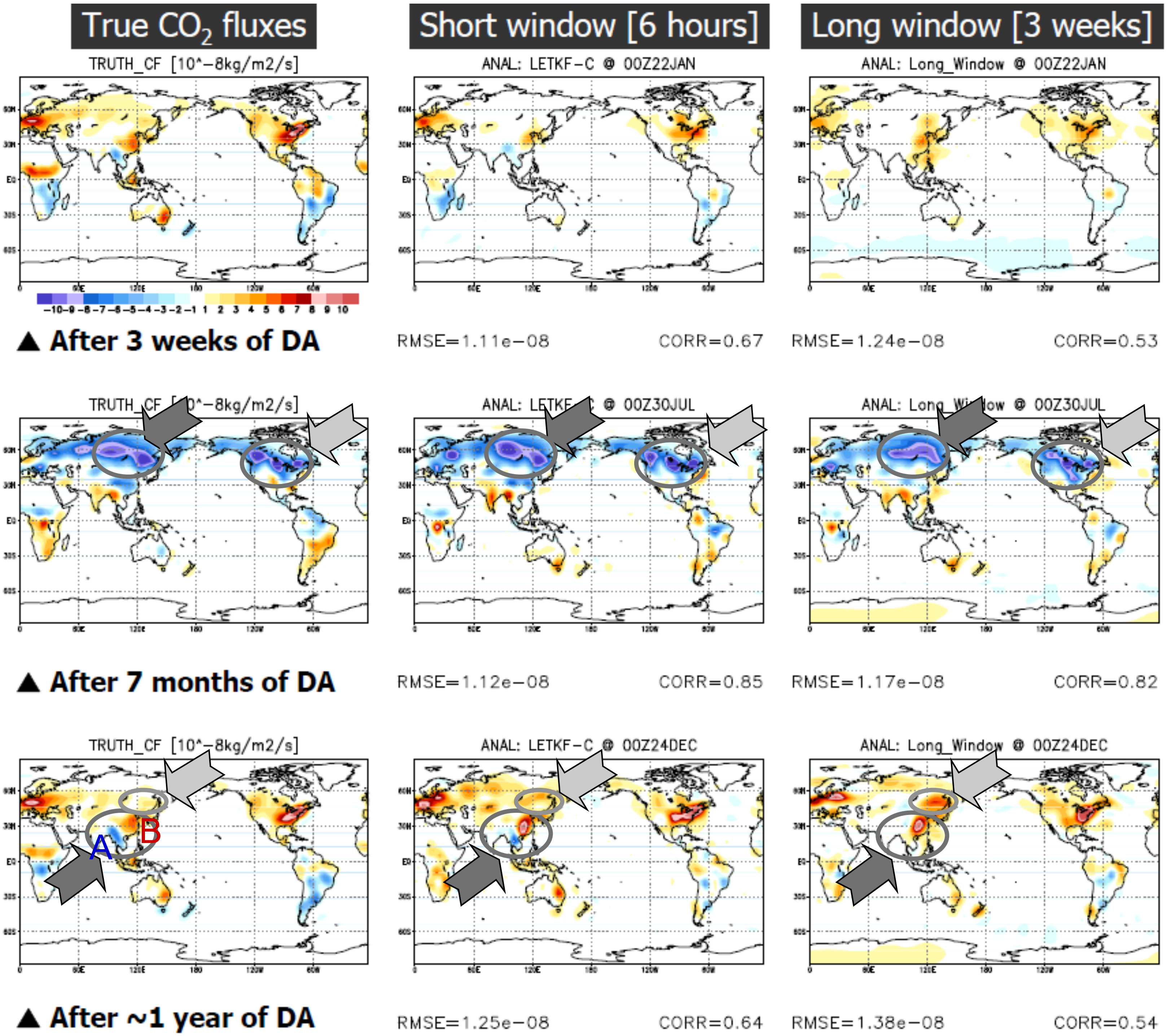
00Z01APR ▶
After three months of DA

00Z01AUG ▶
After seven months of DA

We succeed in estimating time-evolving CF at model-grid scale



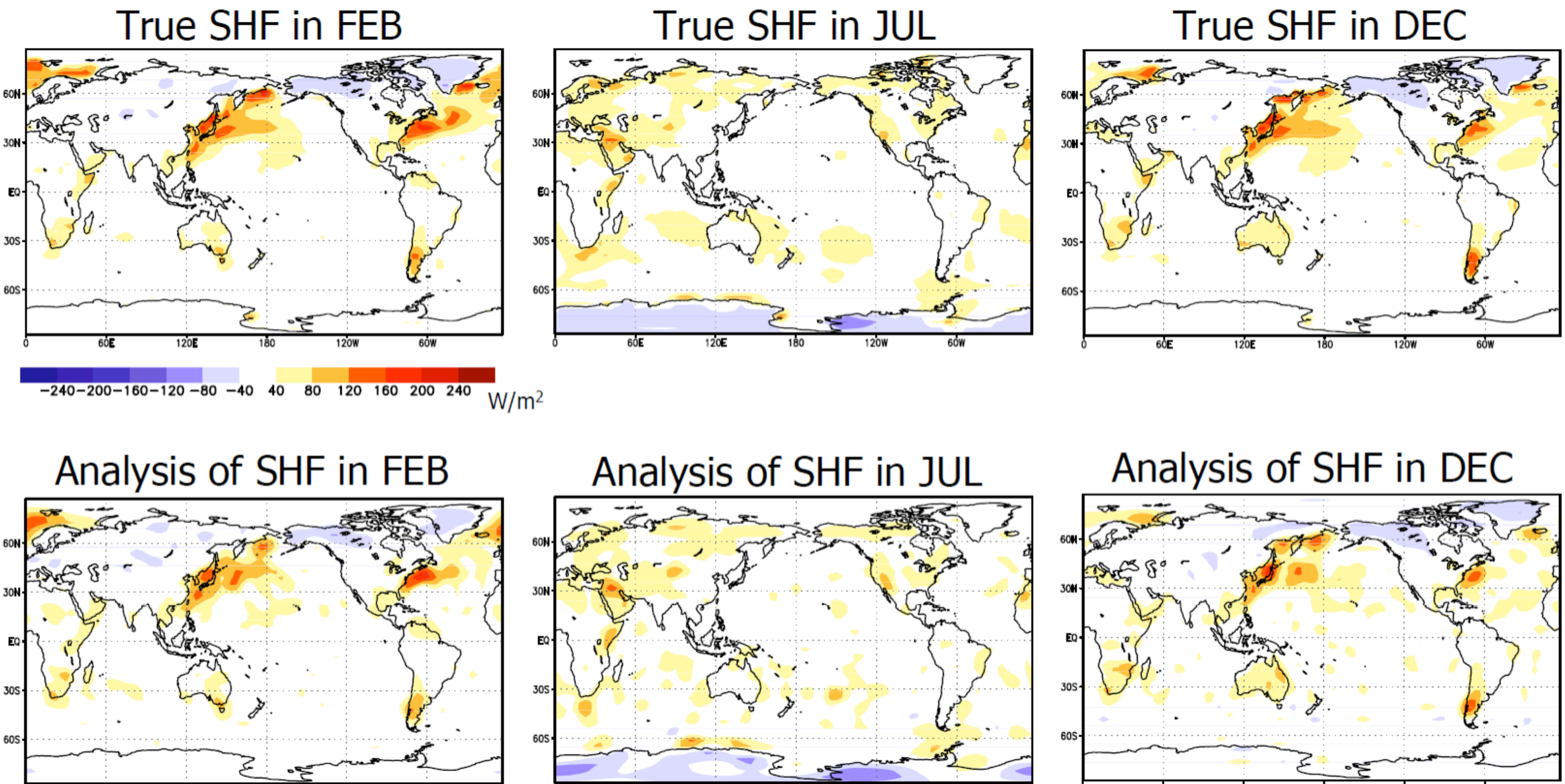
00Z01JAN ▶
After one year of DA


True CO2 fluxes
Short window [6 hours]
Long window [3 weeks]
TRUTH_CF [10^-8kg/m2/s]
ANAL: LETKF-C @ 00Z22JAN
ANAL: Long_Window @ 00Z22JAN
-10 -9 -8 -7 -6 -5 -4 -3 -2 -1 1 2 3 4 5 6 7 8 9 10
▲ After 3 weeks of DA
RMSE=1.11e-08
CORR=0.67
RMSE=1.24e-08
CORR=0.53
ANAL: LETKF-C @ 00Z30JUL
ANAL: Long_Window @ 00Z30JUL
▲ After 7 months of DA
RMSE=1.12e-08
CORR=0.85
RMSE=1.17e-08
CORR=0.82
ANAL: LETKF-C @ 00Z24DEC
ANAL: Long_Window @ 00Z24DEC
A
B
▲ After ~1 year of DA
RMSE=1.25e-08
CORR=0.64
RMSE=1.38e-08
CORR=0.54

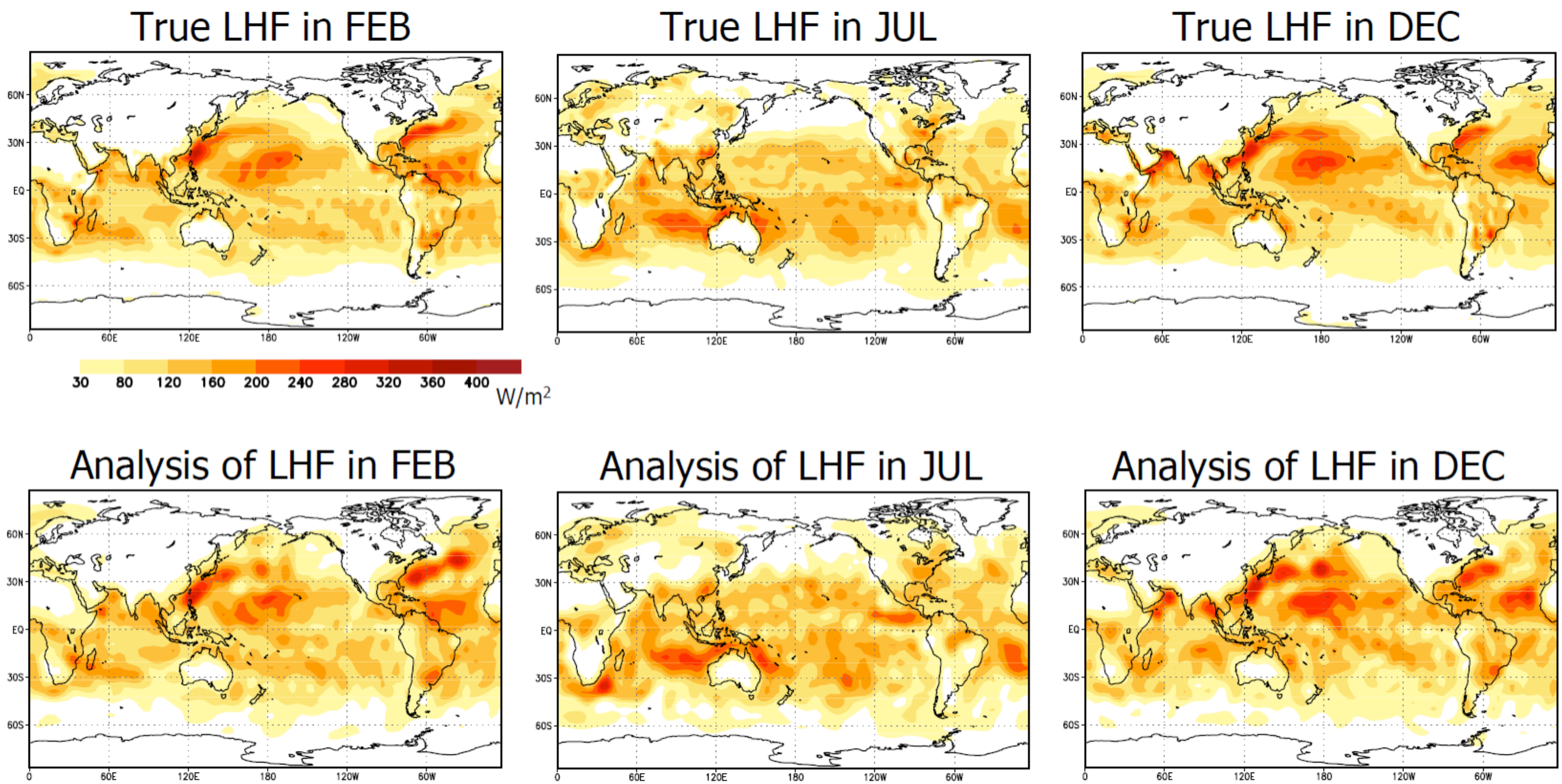
Summary

- We have shown the feasibility of simultaneous analysis of meteorological and carbon variables within LETKF framework through the simulation experiments.
- The system LETKF-C has been tested in a intermediate-complexity model SPEEDY-C with excellent results.
  - Multivariate data assimilation with "localization of the variables" (Kang et al. 2011)
  - Advanced data assimilation methods for CO_2 flux estimation have been explored (Kang et al. 2012)
- The implementation of the LETKF-C to NCAR CAM 3.5 model is now in progress
  - Analysis step shows very good performance in OSSE with real observation coverage
  - Analysis cycle with a forecast step will be operated soon
- The same methodology has been applied to estimating surface fluxes of heat, moisture and momentum, and the results are promising!

Result: Analysis of SHF



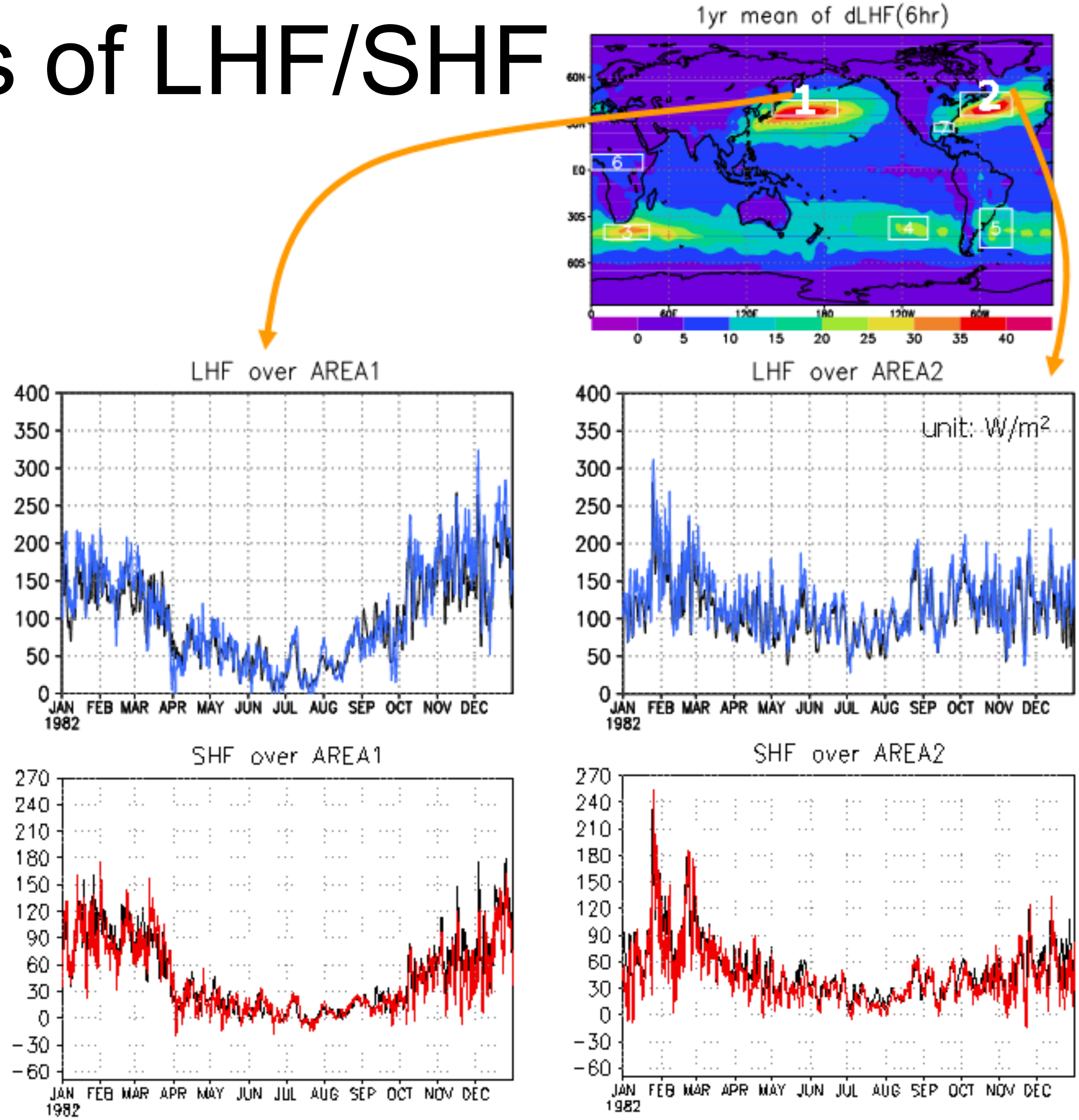
Result: Analysis of LHF



Time series of LHF/SHF

Recall that LHF & SHF are updated only by the data assimilation here!

Promising results from the estimation of "evolving parameters" with data assimilation



Summary and the future

- RIP can extract more information from observations and accelerate spin-up. Examples: Typhoon Sinlaku and Ocean.
- EnKF can be used to assimilate and remember precipitation information, using a Gaussian Transform.
- The Ensemble Forecast Sensitivity to Observations can be used to detect observations that give bad *regional* 6, 12 or 24hr forecasts. This allows repeating analysis without bad observations: "proactive QC" and monitoring.
- We can estimate surface fluxes of carbon, heat and moisture with the LETKF as evolving parameters.
- **EnKF is a newer, much simpler technology.**
- **There is much more potential not yet exploited or not even explored:**
  - **Estimation and correction of model errors and parameters (Ruiz et al, Danforth et al, Kang et al)…**

Comparison of 4D-Var and EnKF

Bold is the best option

Characteristics of DA system	**4D-Var**	**EnKF**
Scalable?	No	**Yes**
Needs TLM and Adjoint of Model	Yes	**No**
Needs TLM and Adjoint of Obs. Operator	Yes	**No**
Full rank B	**Yes**	No
Flow dependent B	No	**Yes**
Optimal Assimilation Window	Long	**Short**

Improvement for cross-track error due to RIP