CSCAMM-DAS13 – Lecture 2

- LETKF – No-Cost Smoothing
- Running in Place (RIP) (Kalnay & Yang, 2010)
  - Lorenz model (Yang et al, 2012)
  - Sinlaku typhoon (Yang et al, 2013)
  - 7 years of ocean assimilation (Penny et al, 2013)
- Proposed coupled ocean-atmosphere model
- A new type of hybrid (Penny, under review)

Shu-Chih Yang, Takemasa Miyoshi, Steve Penny,

and Eugenia Kalnay

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, Travis Sluka)

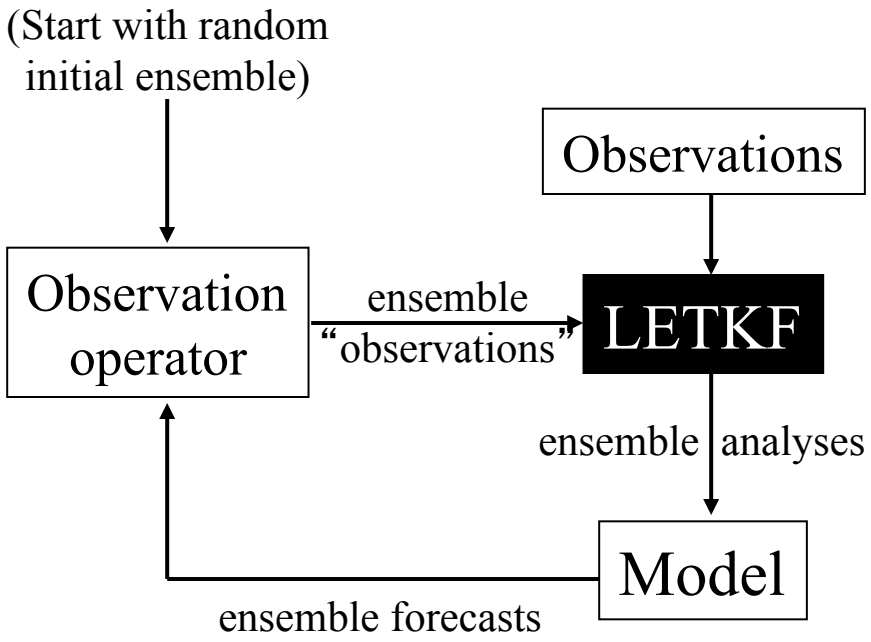
Also: **Y Ota**, Juan Ruiz, C Danforth, M Peña, M Corazza, A. Carrassi

Promising new tools for the LETKF

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!

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

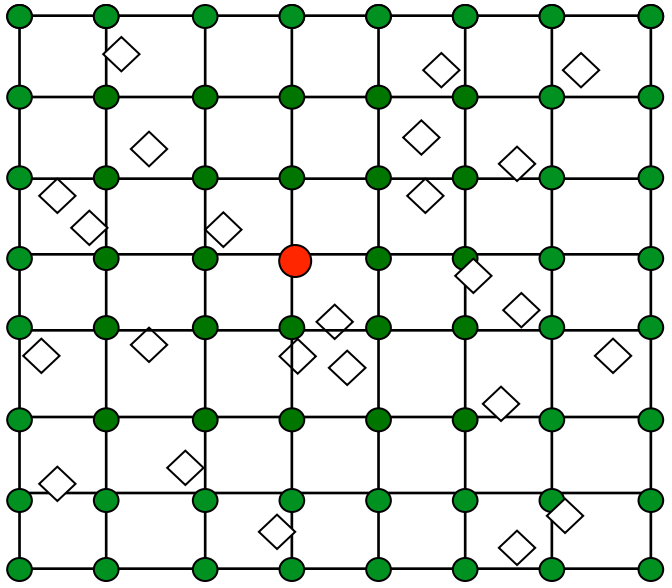


- Model independent (black box)
- Obs. assimilated simultaneously at each grid point
- 100% parallel
- No **adjoint** needed
- 4D LETKF extension
- Computes the **weights** for the ensemble forecasts explicitly

Localization based on observations

Perform data assimilation in a local volume, choosing observations

The state estimate is updated at the central grid red dot

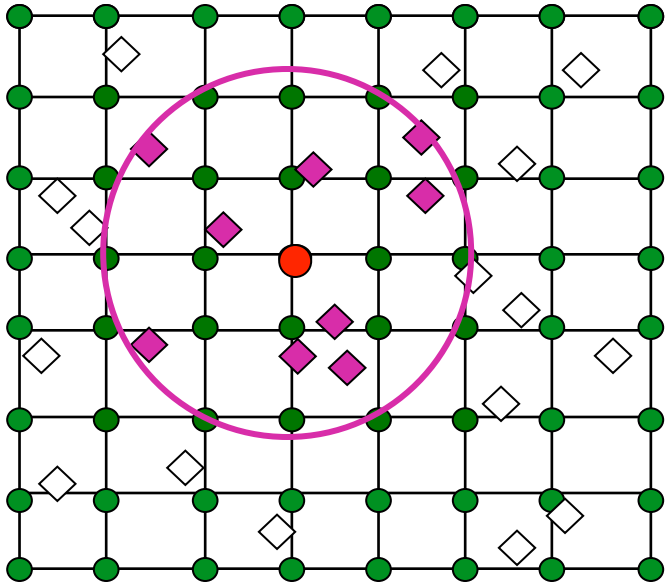


Localization based on observations

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 $\mathbf{X}^b = \left[\mathbf{x}^b_1 - \bar{\mathbf{x}}^b \,|\, \ldots \,|\, \mathbf{x}^b_K - \bar{\mathbf{x}}^b\right];$

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

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

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

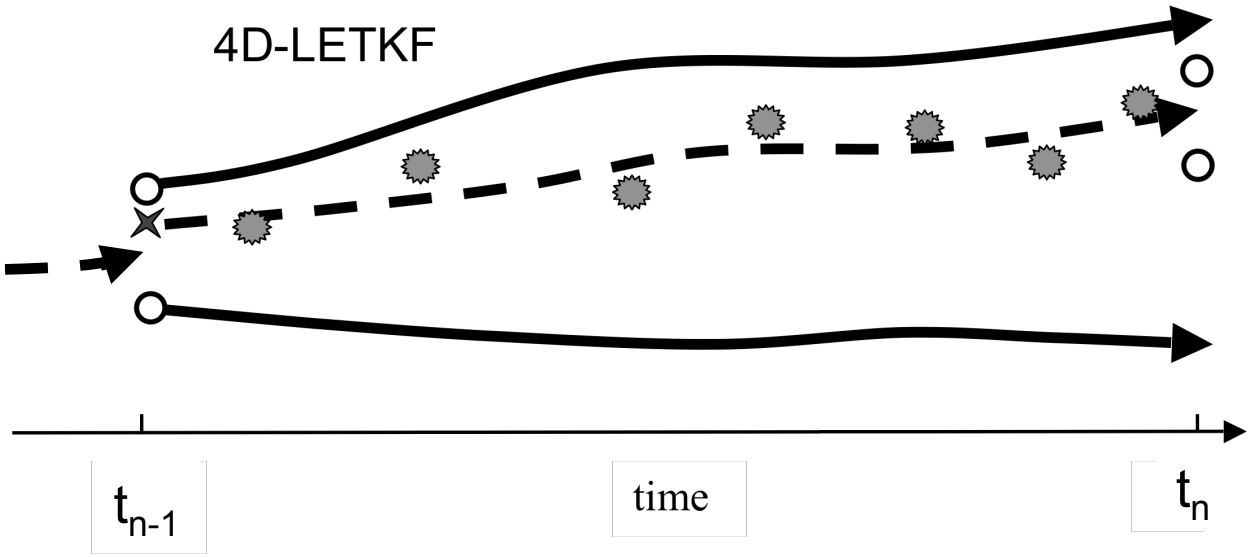
Analysis mean in ensemble space: $\bar{\mathbf{w}}^a = \tilde{\mathbf{P}}^a\mathbf{Y}^{bT}\mathbf{R}^{-1}(\mathbf{y}^o - \bar{\mathbf{y}}^b)$

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 $\mathbf{X}^a_n = \mathbf{X}^b_n\mathbf{W}^a + \bar{\mathbf{x}}^b$

Gathering the grid point analyses forms the new global analyses. Note that the the output of the LETKF are analysis weights $\bar{\mathbf{w}}^a$ 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

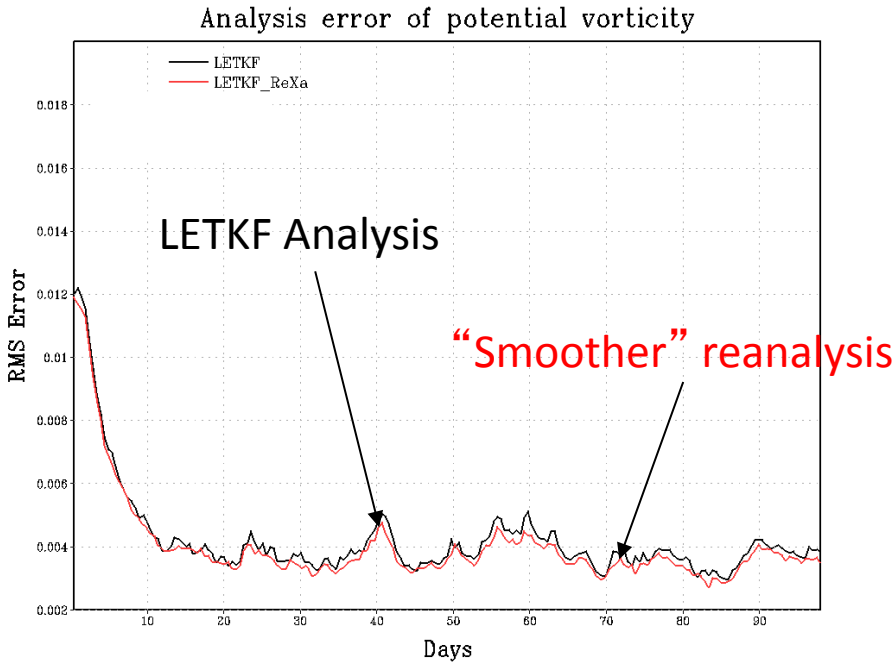
No-cost LETKF smoother first tested on a QG model: it works…

LETKF analysis at time n

$$\bar{\mathbf{x}}_n^a = \bar{\mathbf{x}}_n^f + \mathbf{X}_n^f \bar{\mathbf{w}}_n^a$$

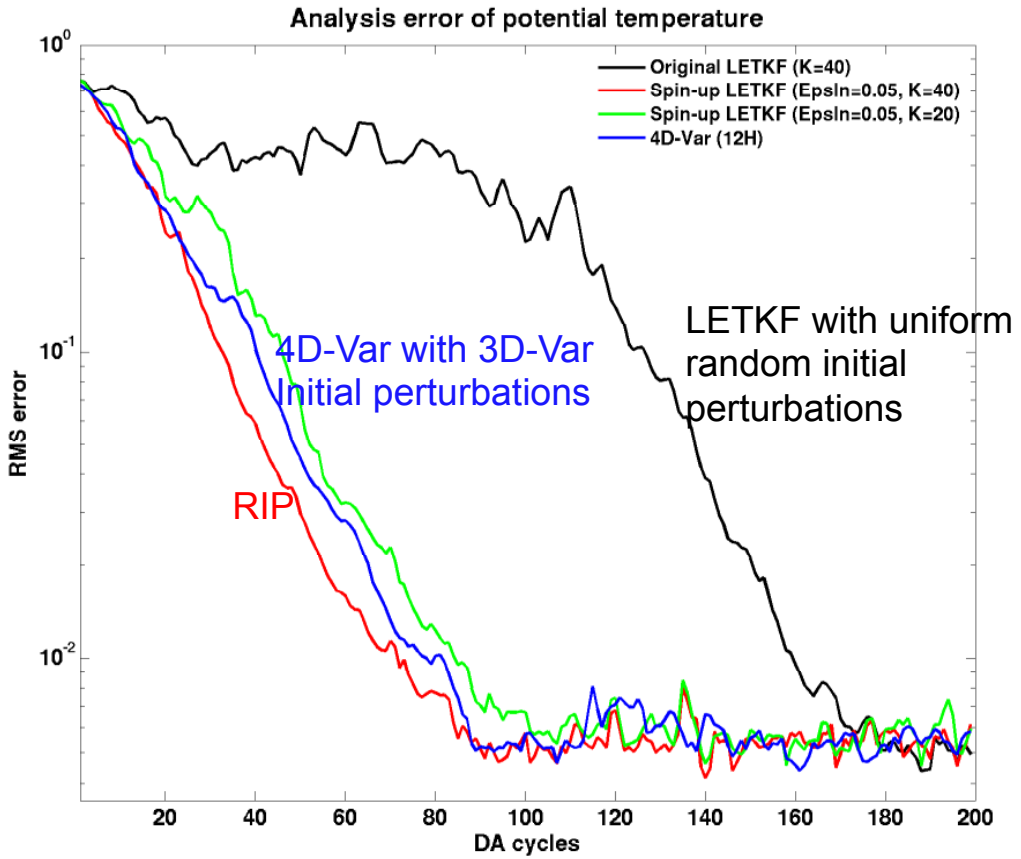
Smoother analysis at time n-1

$$\tilde{\mathbf{x}}_{n-1}^a = \bar{\mathbf{x}}_{n-1}^f + \mathbf{X}_{n-1}^f \bar{\mathbf{w}}_n^a$$



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.

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).

Nonlinearities: "Quasi Outer Loop" (QOL)

Quasi Outer Loop: use the final weights to correct only the mean 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	LETKF +RIP
Window=8 steps	0.31	0.30	**0.27**	0.27
Window=25 steps	0.53	0.66	**0.48**	**0.35**

Nonlinearities, "QOL" and "Running in Place"

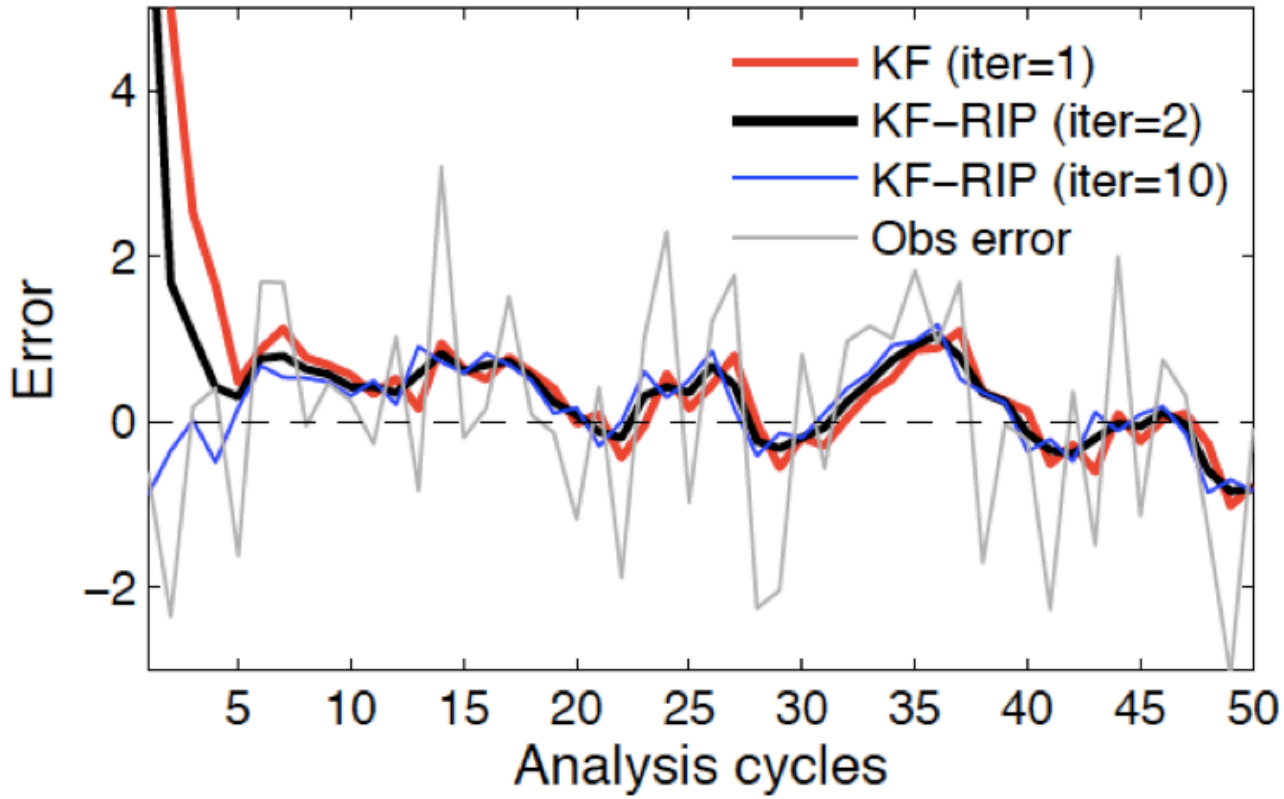
Quasi Outer Loop: similar to 4D-Var: use the final weights to correct only the mean initial analysis, keeping the initial perturbations. Repeat the analysis once or twice. It centers the ensemble on a more accurate nonlinear solution.

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…

Why RIP works: Results with a Linear model

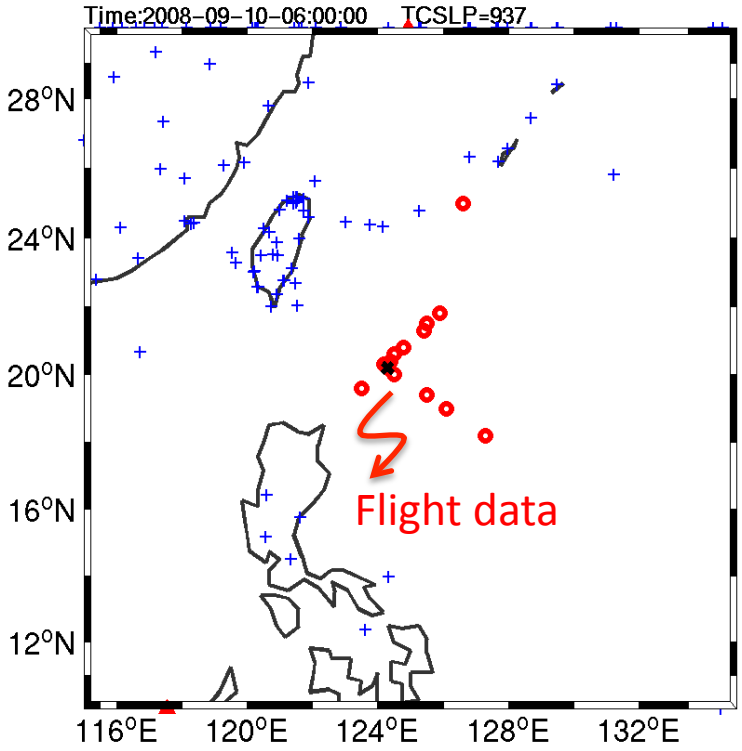


$$x_n = M(x_{n-1}) = x_{n-1} + \alpha$$

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

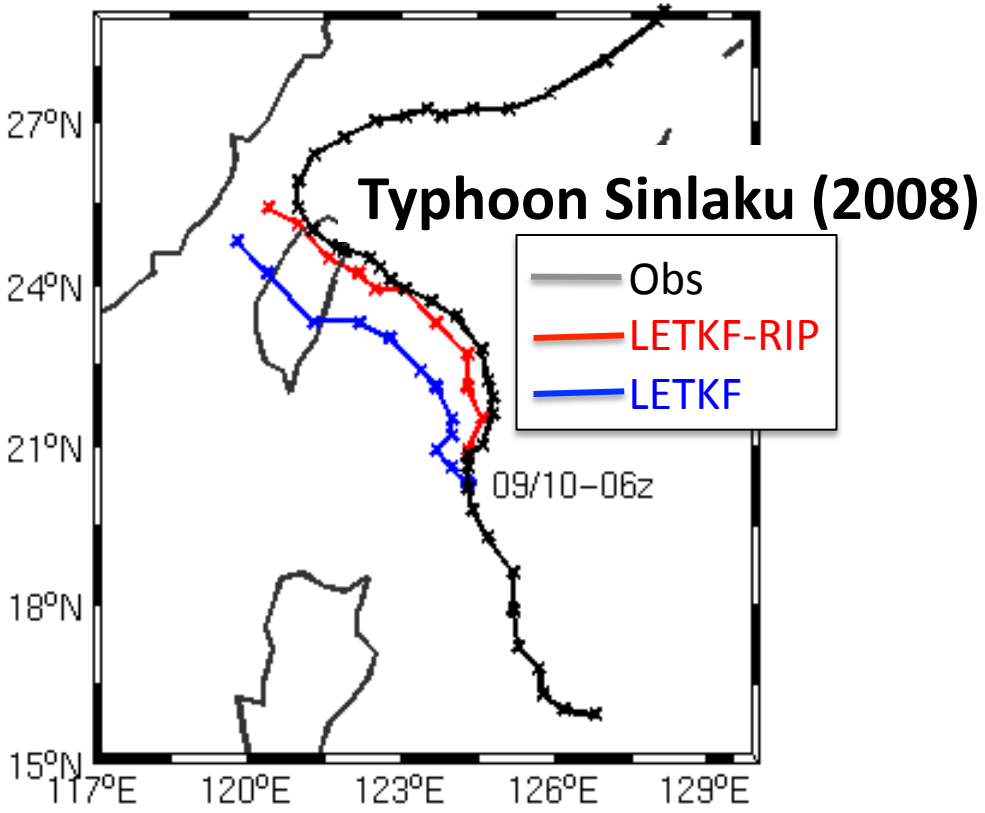
- RIP adapts to using an observation N-times by dividing the spread by N: RIP converges to the regular optimal KF solution.
- The spin-up is faster and the analysis update is “softer” (in small steps) rather than in large steps.

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



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

3-day forecast

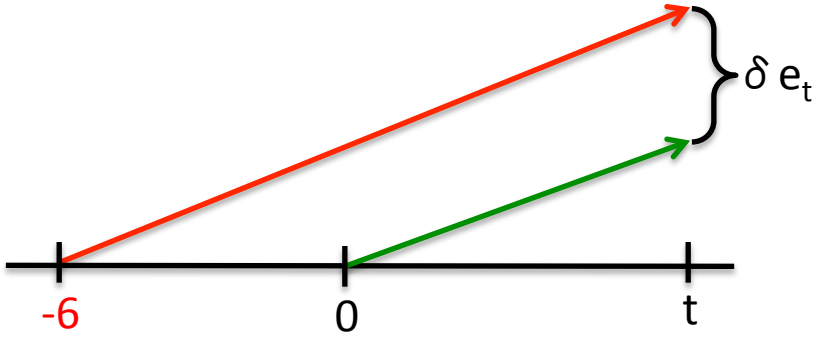


RIP uses better the "limited observations"!

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

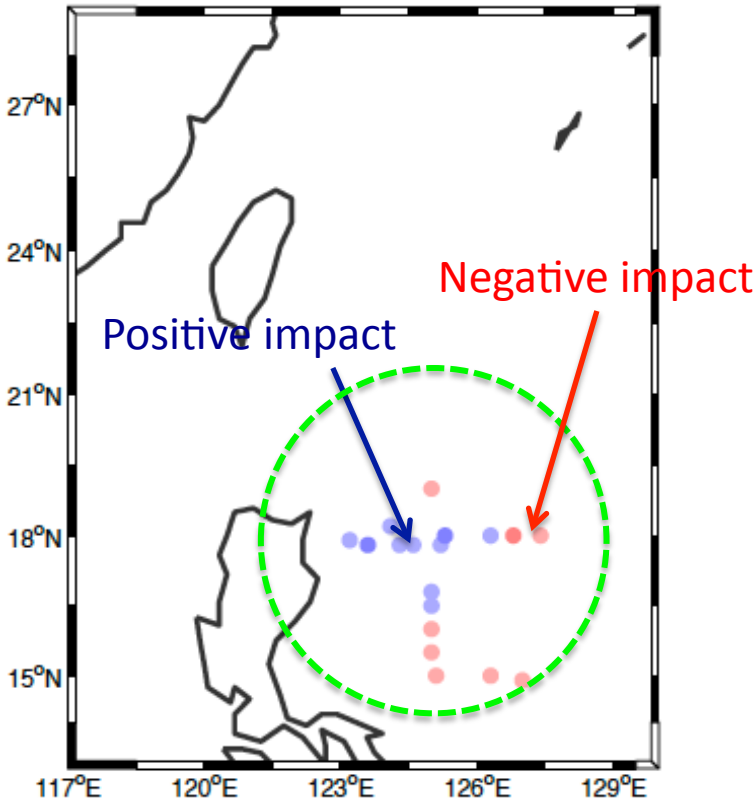
Observation impact on the forecast: Without RIP

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

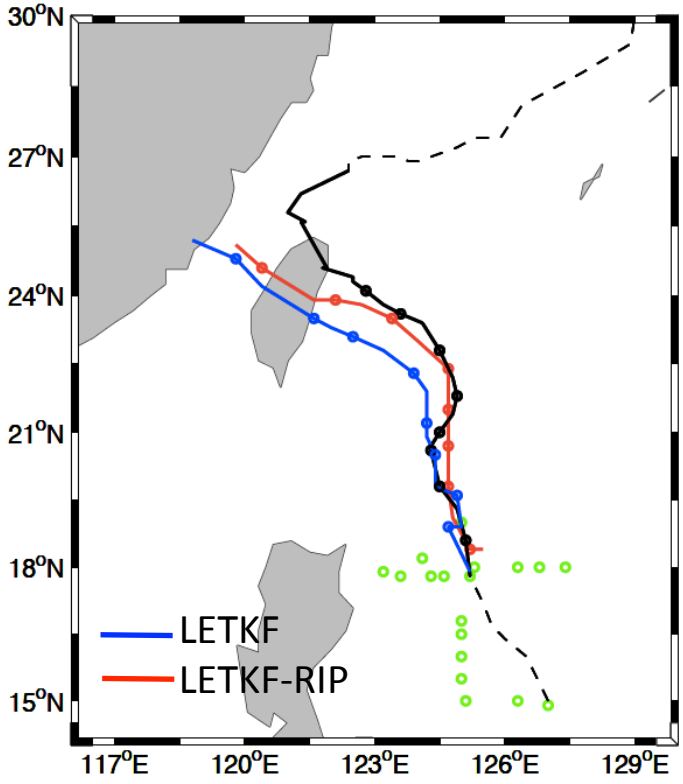
Observation impact with respect to **dropsondes** (standard LETKF)



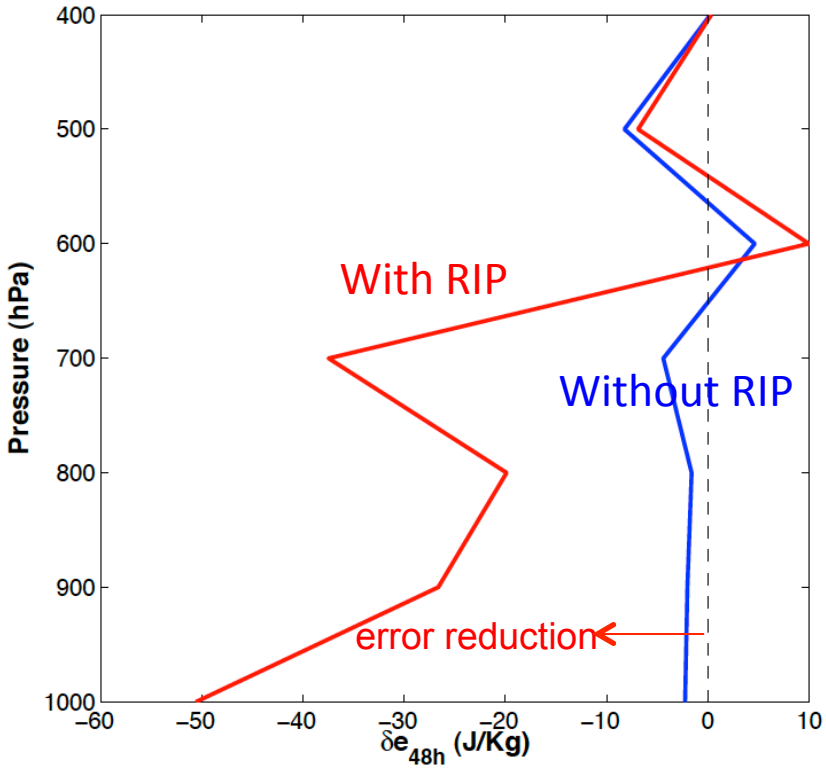
Targeted at TY's environmental condition

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.

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

Steve Penny's thesis defense

April 15, 2011

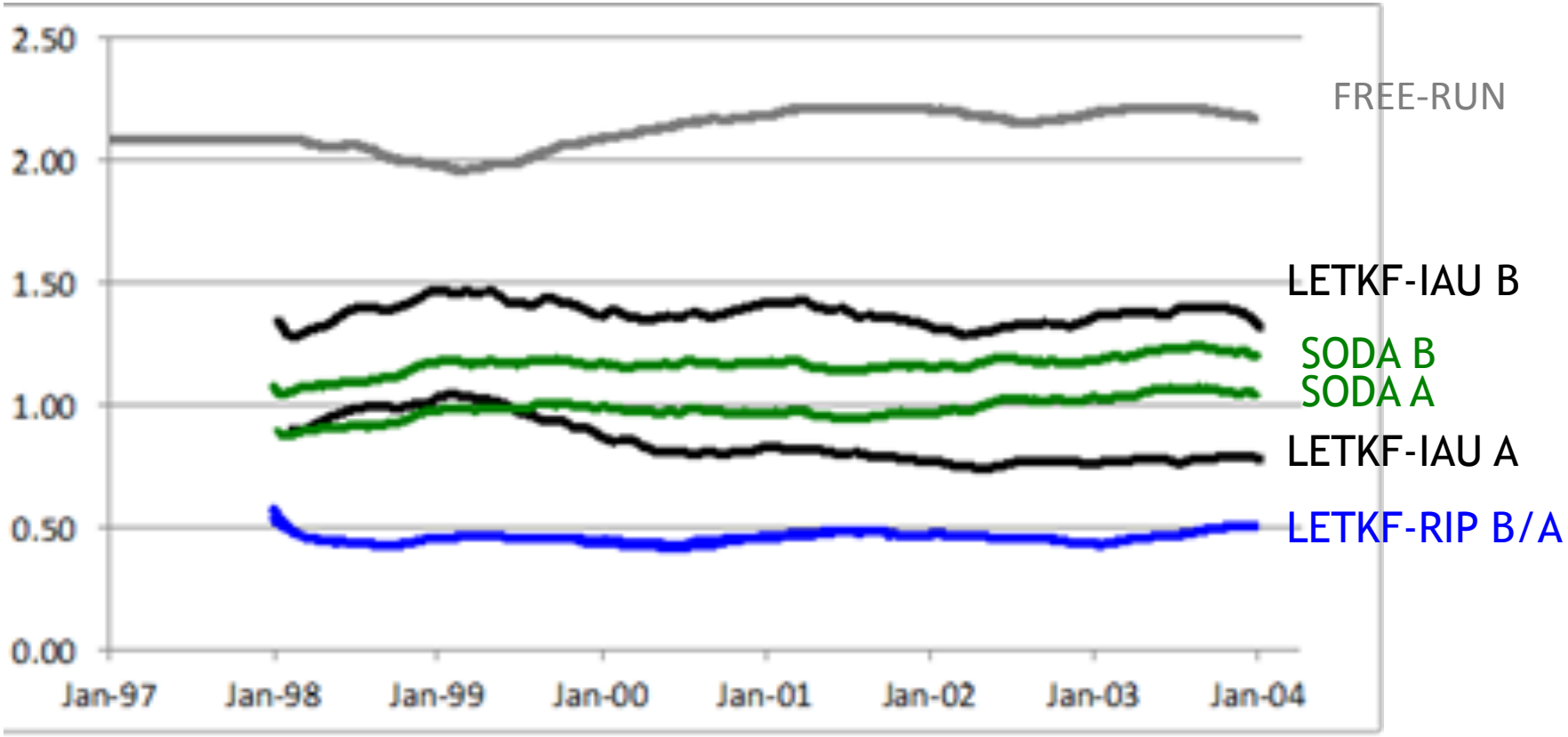
Advisors: E Kalnay, J Carton, K Ide, T Miyoshi, G Chepurin

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

RMSD (°C) (All vertical levels)

7 years of Ocean Reanalysis
Temperature

B: background
A: analysis

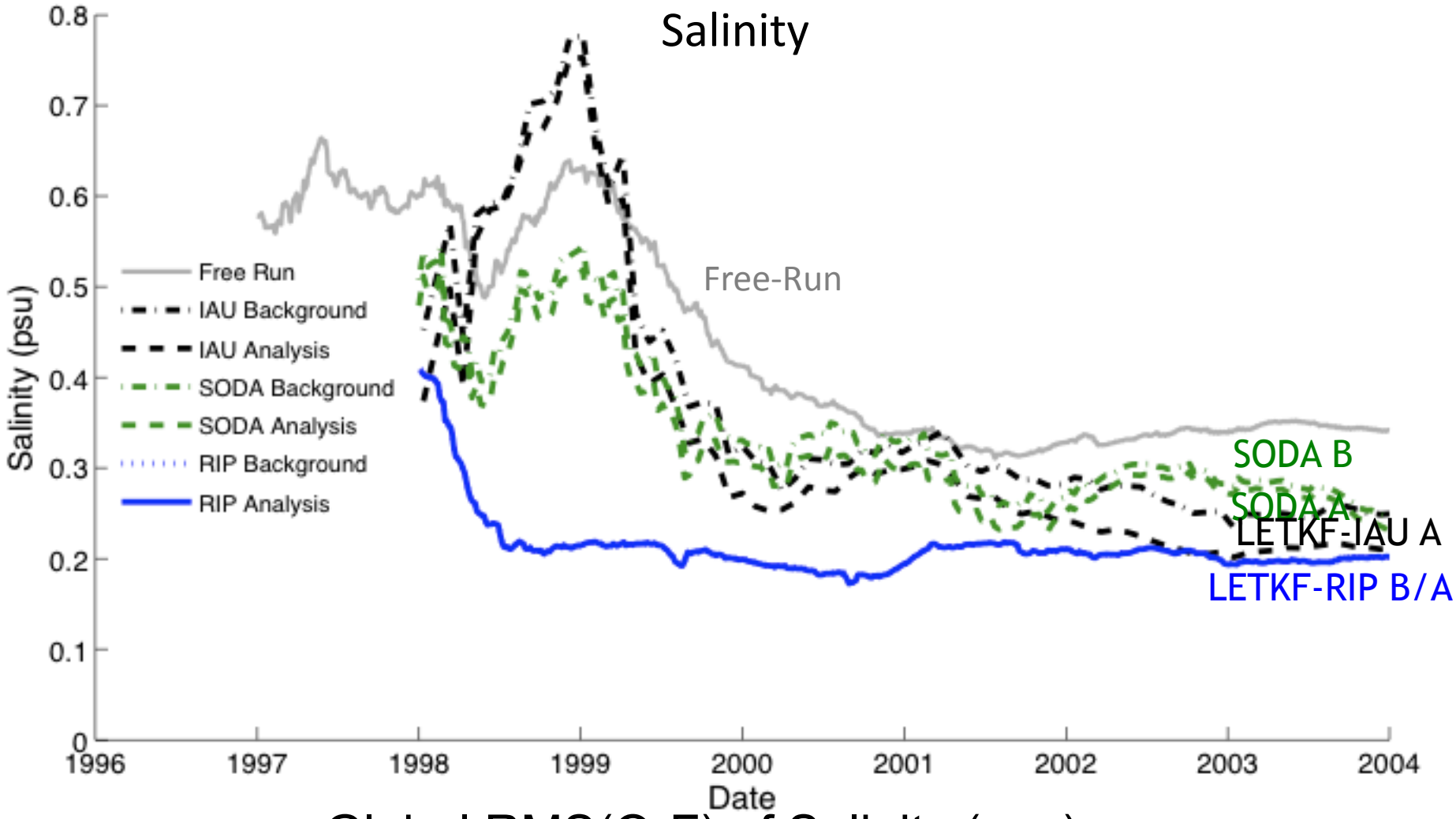


Global RMS(O-F) of Temperature (°C),
12-month moving average
LETKF (with **IAU**), **SODA** and LETKF with **RIP**

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, any 4D-EnKF) 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”.

GODAE Ocean View/WGNE Workshop 2013
19 March 2013

Data assimilation for the coupled ocean-atmosphere

Eugenia Kalnay, Tamara Singleton, Steve Penny, Takemasa Miyoshi, Jim Carton

Thanks to the UMD Weather-Chaos Group, to Daryl Kleist and to the India Monsoon Mission

Outline

- Traditional approaches.
- Thesis of Tamara Singleton (DA with toy coupled model).
- The LETKF and Running in Place.
- Steve Penny: 7 years ocean reanalysis.
- Steve Penny: New EnKF-based hybrid.
- Shaoqing Zhang: GFDL coupled EnKF.
- Our planned approach to coupled LETKF (India Monsoon Mission)
- Questions:
  - Can we do a robust coupled SST analysis? SSH? Scatterometer winds?
  - Should we do LETKF-RIP? Short windows for the ocean and atm.?
  - Should we do Gaussian Transformation? (Lien et al.)
  - Should we do Proactive QC with Ens. Fcst. Sens. to Obs. (EFSO)?
- Discussion

Traditional approaches

"In a typical coupling scheme for an ocean-atmosphere model, the ocean model passes SST to the atmosphere, while the atmosphere passes back heat flux components, freshwater flux, and horizontal momentum fluxes." (Neelin, Latif & Jin, 1994)

SST in the ocean model is frequently nudged from Reynolds SSTs, not assimilated from observations.

SSH may not be even be used.

The data assimilation windows are very different for the ocean and the atmosphere.

Tamara Singleton's thesis

Data Assimilation Experiments with a Simple Coupled Ocean-Atmosphere Model

Questions she addressed:

-- Which is more accurate: 4D-Var or EnKF?

-- Is it better to do an ocean reanalysis separately, or as a single coupled system?

-- ECCO is a version of 4D-Var where both the initial state and the surface fluxes are control variables. This allows ECCO to have very long windows (decades) and estimate the surface fluxes that give the best analysis.

Is ECCO the best approach for ocean reanalysis?

Simple Coupled Ocean-Atmosphere System

3 coupled Lorenz models: A slow "ocean" component strongly coupled with a fast "tropical atmosphere component", in turn weakly coupled with a fast "extratropical atmosphere" (Peña and Kalnay, 2004).

Model Parameter Definitions

Variables	Description	Values
c, c_z, c_e	Coupling coefficient	$c, c_z = 1$ $c_e = 0.08$
τ	time scale	$\tau = 0.1$
σ, b, and r	Lorenz parameters	$\sigma=10$, $b=8/3$, and $r=28$
k_1, k_2	Uncentering parameters	$k_1=10$ $k_2 = -11$

Extratropical atmosphere

$$\dot{x}_e = \sigma(y_e - x_e) - c_e(x_t + k_1)$$
$$\dot{y}_e = rx_e - y_e - x_e z_e - c_e(y_t + k_1)$$
$$\dot{z}_e = x_e y_e - bz_e$$

Tropical atmosphere

$$\dot{x}_t = \sigma(y_t - x_t) - c(X + k_2) - c_e(x_e + k_1)$$
$$\dot{y}_t = rx_t - y_t - x_t z_t + c(Y + k_2) + c_e(y_e + k_1)$$
$$\dot{z}_t = x_t y_t - bz_e + c_z Z$$

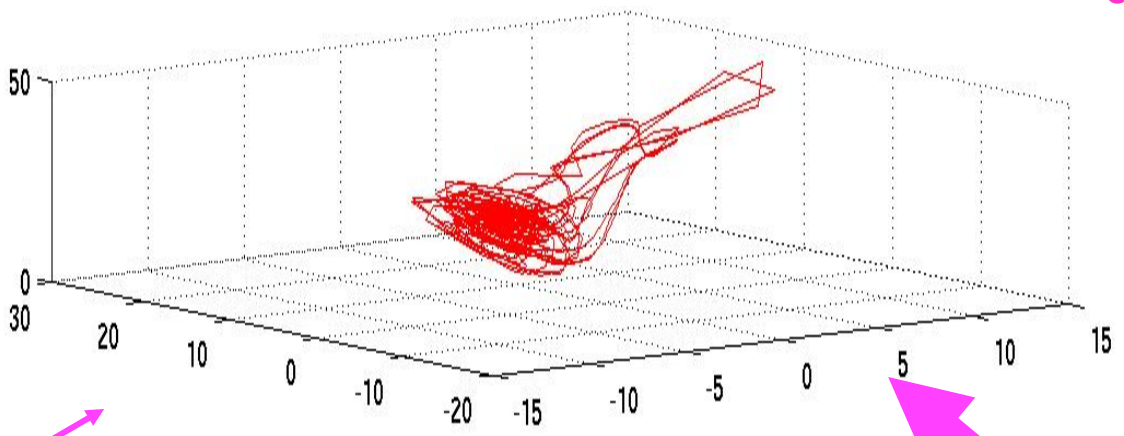
Ocean

$$\dot{X} = \tau\sigma(Y - X) - c(x_t + k_2)$$
$$\dot{Y} = \tau rX - \tau Y - \tau XZ + c(y_t + k_2)$$
$$\dot{Z} = \tau XY - \tau bZ + c_z z_t$$

Model State: $[x_e, y_e, z_e, x_t, y_t, z_t, X, Y, Z]^T$

Simple Coupled Ocean-Atmosphere Model (Peña and Kalnay, 2004)

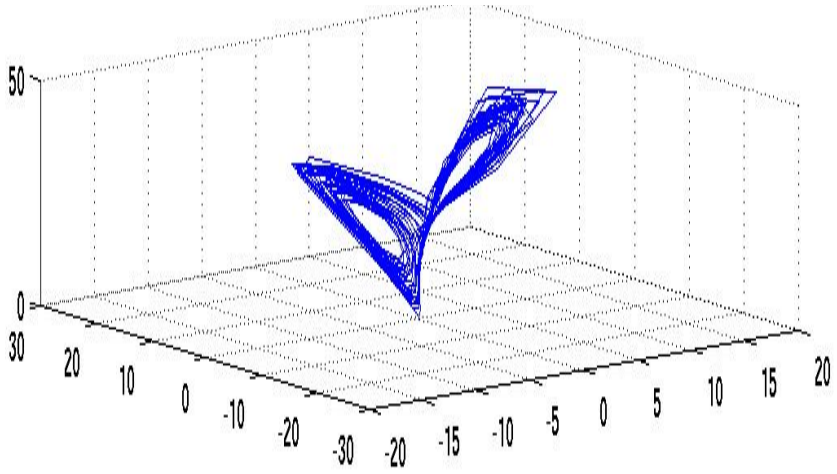
Tropical Atmosphere



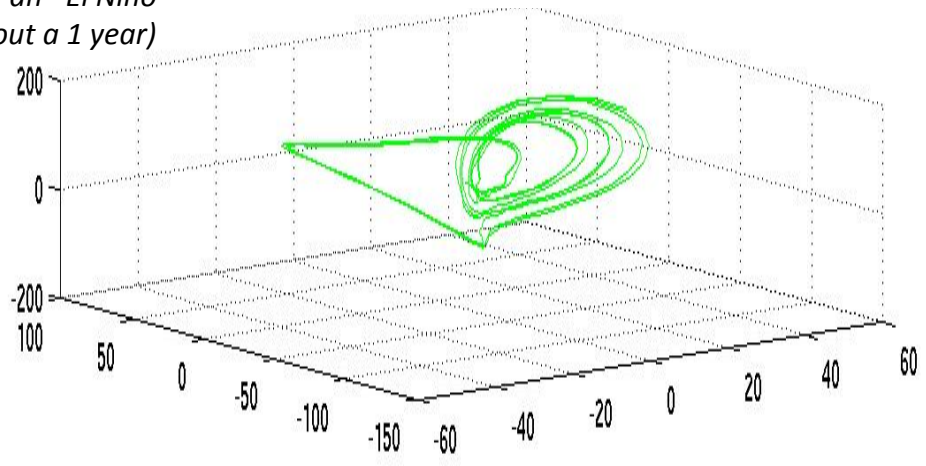
Coupling strength

Ocean is vacillating between a "normal year" (lasts from ~3-8 years) and an "El Nino" (lasts about a 1 year)

Extra-tropical Atmosphere



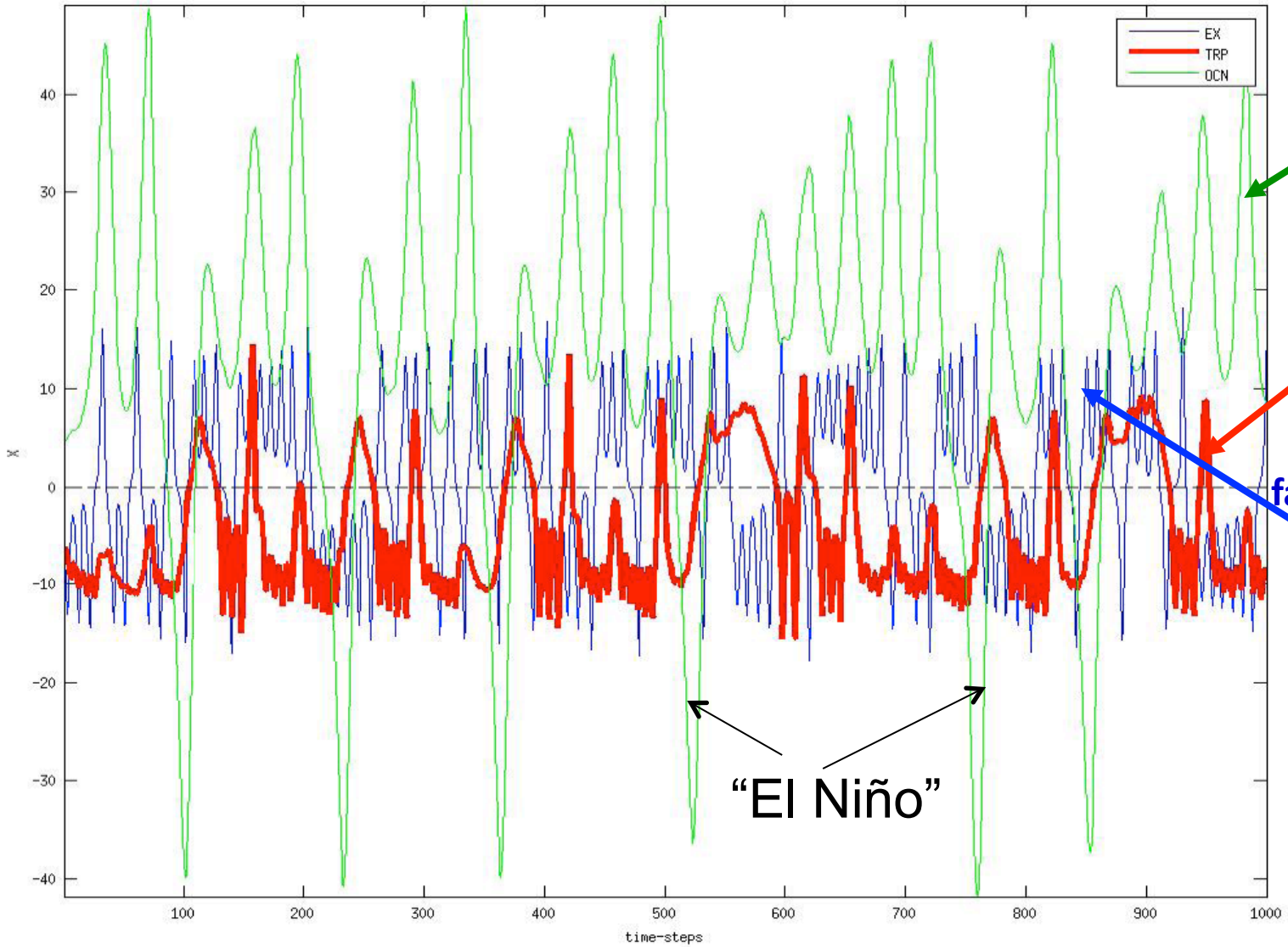
Tropical Ocean



We do OSSEs with this simple coupled model

Simple Coupled Ocean-Atmosphere Model (Peña and Kalnay, 2004)

Time series of the x-component



slow ocean

fast tropical atmosphere

fast extratropical atmosphere

“El Niño”

Δt=0.01

1 year~35Δt

We do OSSEs with this simple coupled model

4D-Var/ETKF Data Assimilation Results

- **We developed a 4D-Var data assimilation system for the simple coupled ocean-atmosphere model**
- We found that **lengthening the assimilation window** and **applying QVA improves the 4D-Var analysis**.
- **Tuning the amplitude of the background error covariance has an impact on the performance of the assimilation**.
- **EnKF-based methods (LETKF & ETKF-QOL) compete with 4D-Var analyses** for short and long assimilation windows.
- For much **longer assimilation windows, 4D-Var outperforms the EnKF-based methods**

- **Short windows are good for ETKF**
- **Long windows are good for 4D-Var**
- **Optimal accuracy similar for 4D-Var and ETKF**

ECCO-like 4D-Var

- The consortium for Estimating the Circulation and Climate of the Ocean (ECCO) is a collaboration of a group of scientists from the MIT, JPL, and the Scripps Institute of Oceanography

- The main characteristic of ECCO is that they include surface fluxes as control variables.
  - This allows them to have exceedingly long assimilation windows in 4D-Var (e.g. 10 years or even 50 years).
  - They used NCEP Reanalysis fluxes (Kalnay et al, 1996) as a first guess for the surface fluxes.

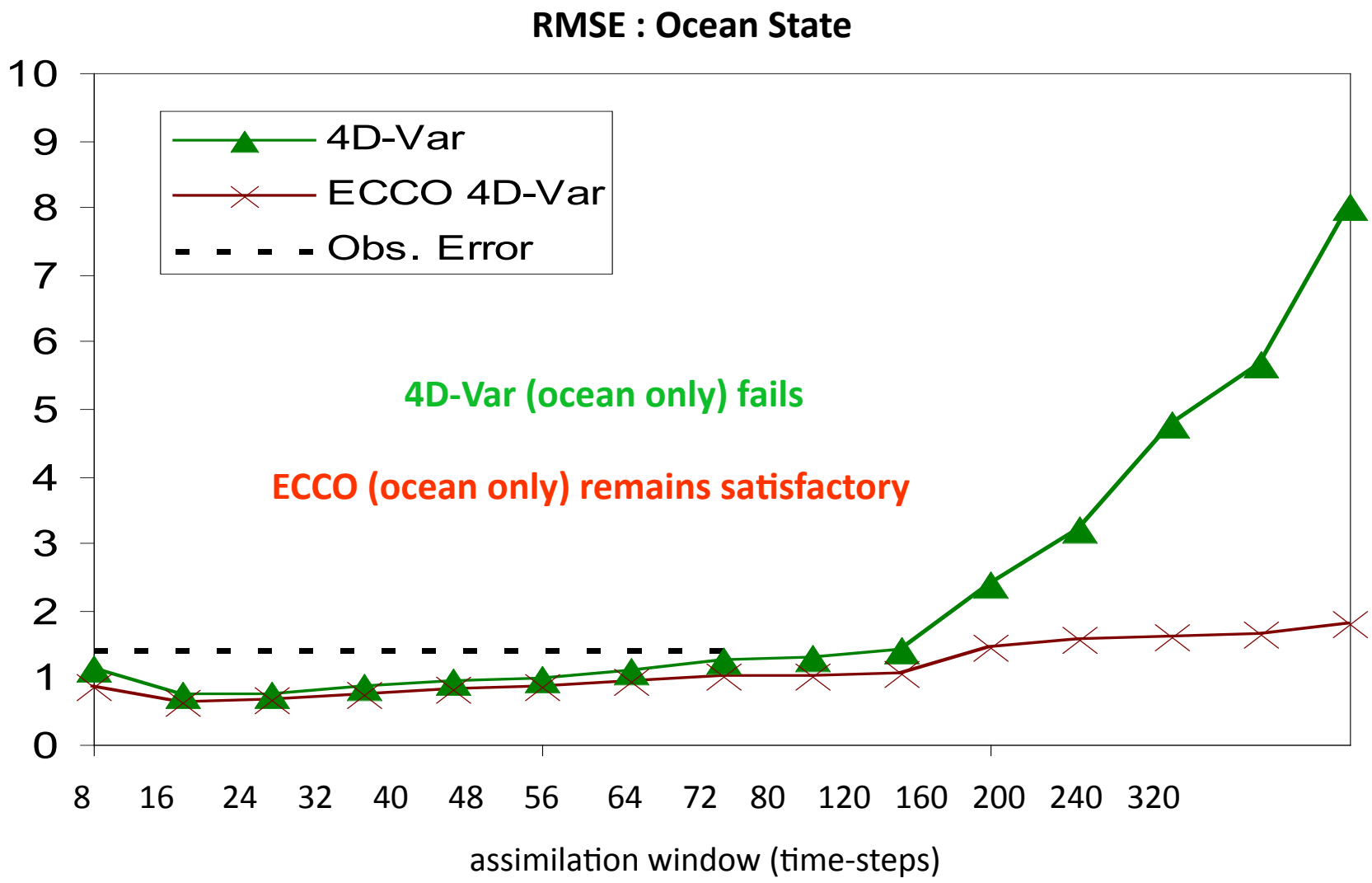
- ECCO used 4D-Var to estimate the initial ocean state and surface fluxes (Stammer et al., 2004; Kohl et al., 2007) in a 50-year reanalysis with a single assimilation window!

Comparison of ECCO-like & Ocean 4D-Var

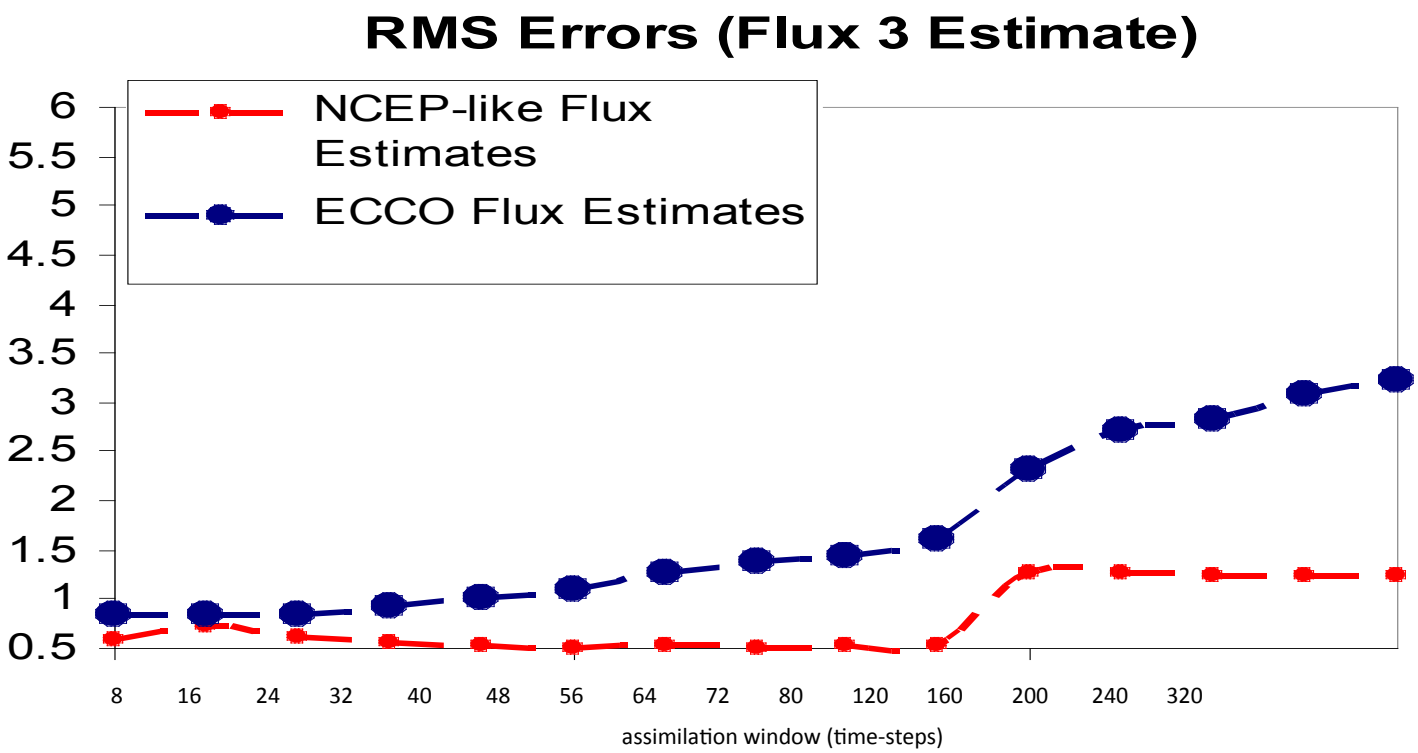
QVA APPLIED

OCEAN ONLY

Obs. s.d error = 1.41 for ocean



Are the ECCO fluxes more accurate?



ECCO does not improve the flux estimates

Answers to the Research Questions

Questions:

-- Which is more accurate: 4D-Var or EnKF?

Fully coupled EnKF (with short windows) and 4D-Var (with long windows) have about the same accuracy.

Answers to the Research Questions

Questions:

-- Which is more accurate: 4D-Var or EnKF?

Fully coupled EnKF (with short windows) and 4D-Var (with long windows) have about the same accuracy.

-- Is it better to do the ocean reanalysis separately, or as a single coupled system?

Both EnKF and 4D-Var are similar and most accurate when coupled, but uncoupled (ocean only) reanalyses are fairly good.

Answers to the Research Questions

Questions:

-- Which is more accurate: 4D-Var or EnKF?

Fully coupled EnKF (with short windows) and 4D-Var (with long windows) have about the same accuracy.

-- Is it better to do the ocean reanalysis separately, or as a single coupled system?

Both EnKF and 4D-Var are similar and most accurate when coupled, but uncoupled (ocean only) reanalyses are quite good.

-- Is ECCO 4D-Var with both the initial state and the surface fluxes as control variables the best approach?

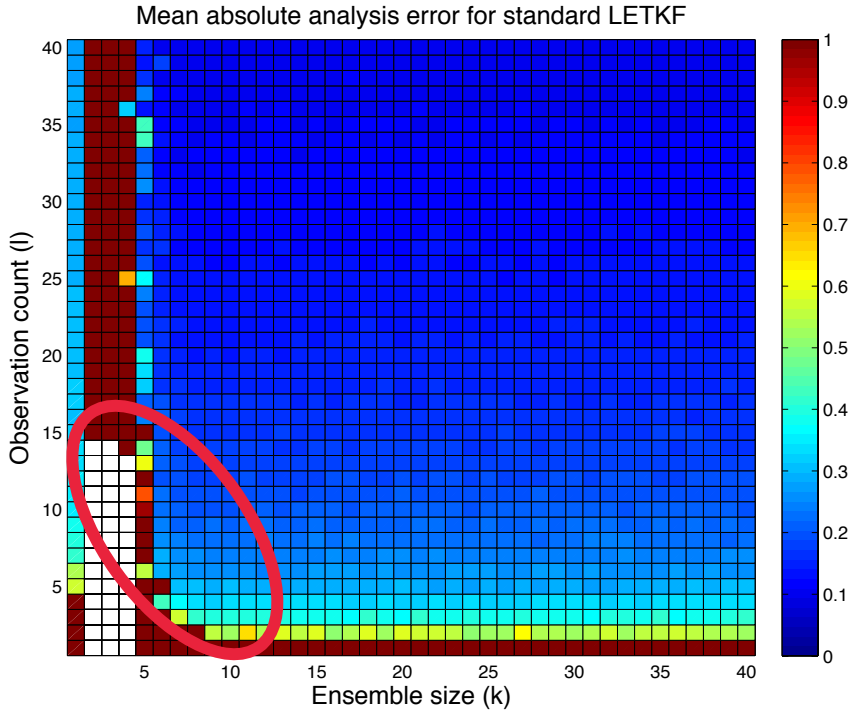
In our simple ocean model 4D-Var cannot remain accurate with very long windows. Our ECCO reanalysis remained satisfactory with very long windows but at the expense of less accurate fluxes.

How about hybrids between Var and EnKF?

- Hybrids have been very successful!!! (Kleist et al, 2013)
- They increase the rank of ***B*** subspace from K (ensemble size) to the size of the model.
- So far hybrids have been created combining an existing Var system with an ensemble to provide the flow dependence of the background error covariance.
- **We would like to start with a well-developed EnKF (like the LETKF) and add a simple local 3D-Var that provides the full rank that the ensemble lacks.**
- Steve Penny (under revision) developed a simple, locally Gaussian 3D-Var for this purpose, and tested the "hybrid/ mean" on the Lorenz-96, a 40 variable model.
- He plots the analysis error as a function of the number of ensemble members (2 to 40) and the number of observations (1 to 40). (3DVar errors are shown as k=1).

An ensemble based hybrid with a simple local 3D-Var (Steve Penny) applied to the Lorenz 96 model

Standard LETKF



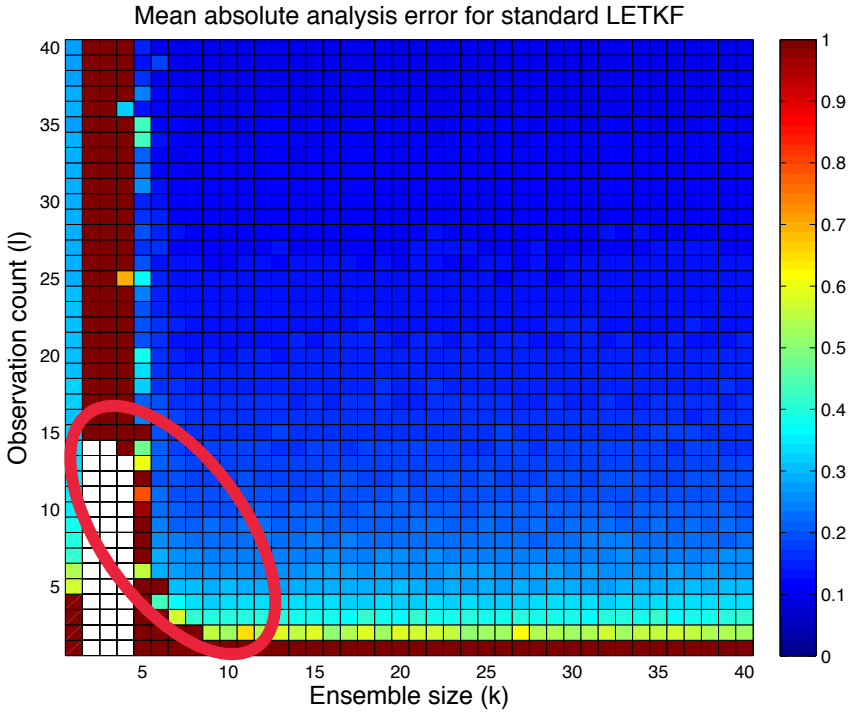
This is the corner where we are in ocean EnKF: too few obs, too few ensembles

The total model dimension is K=40

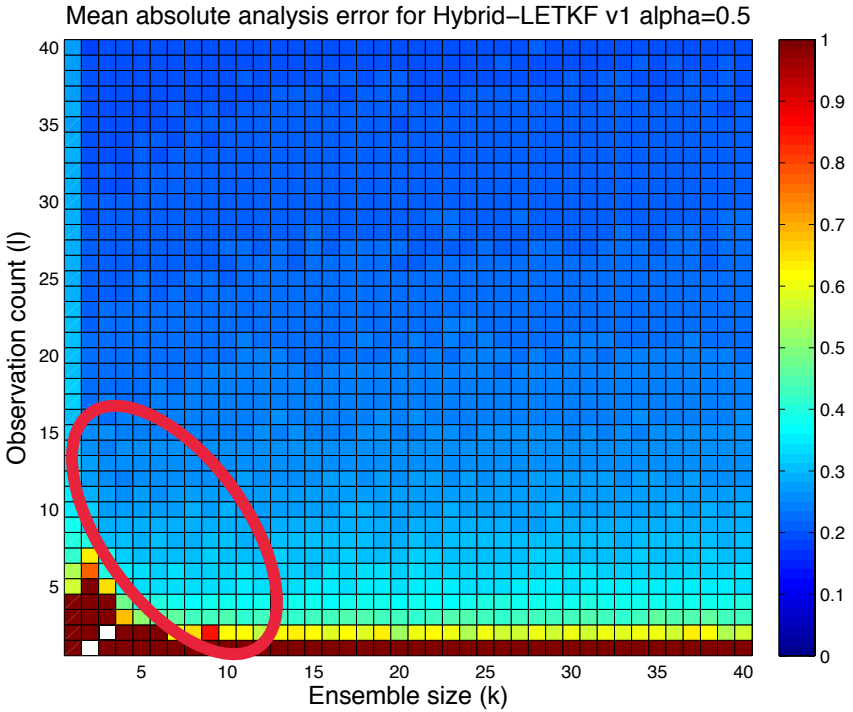
The LETKF is extremely accurate as long as k>7, number of obs>7.

An ensemble based hybrid with a simple local 3D-Var (Steve Penny) applied to the Lorenz 96 model (hybrid/mean LETKF)

Standard LETKF



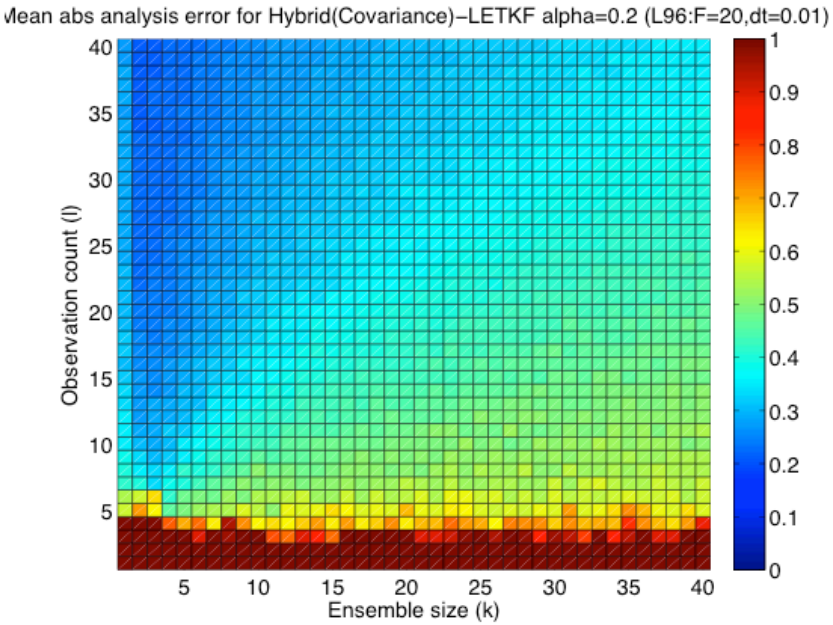
Add a simple 3D-Var to LETKF



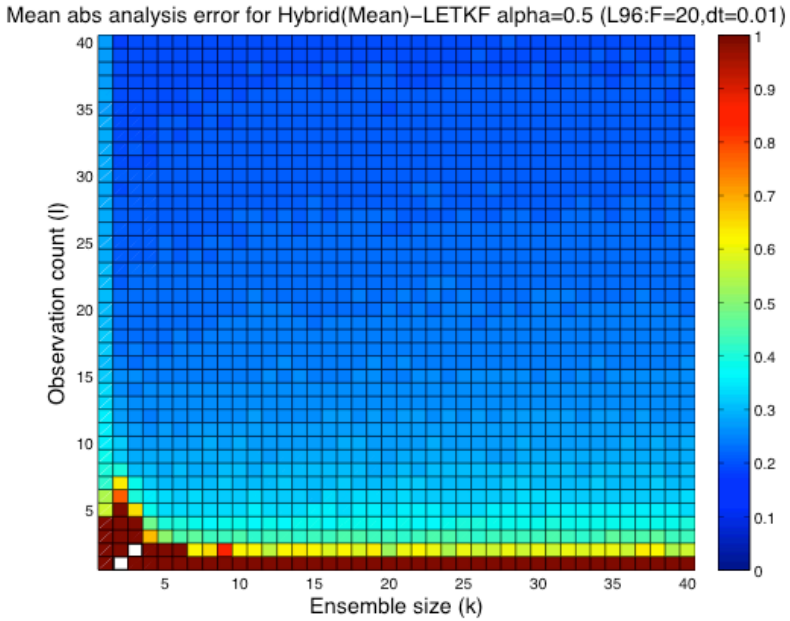
The hybrid/mean LETKF-simple 3D-Var is much more robust for few ensemble members and few observations, as they are for the ocean.

An ensemble based hybrid with a simple local 3D-Var (Steve Penny) applied to the Lorenz 96 model (hybrid/mean LETKF)

Hybrid/covariance LETKF



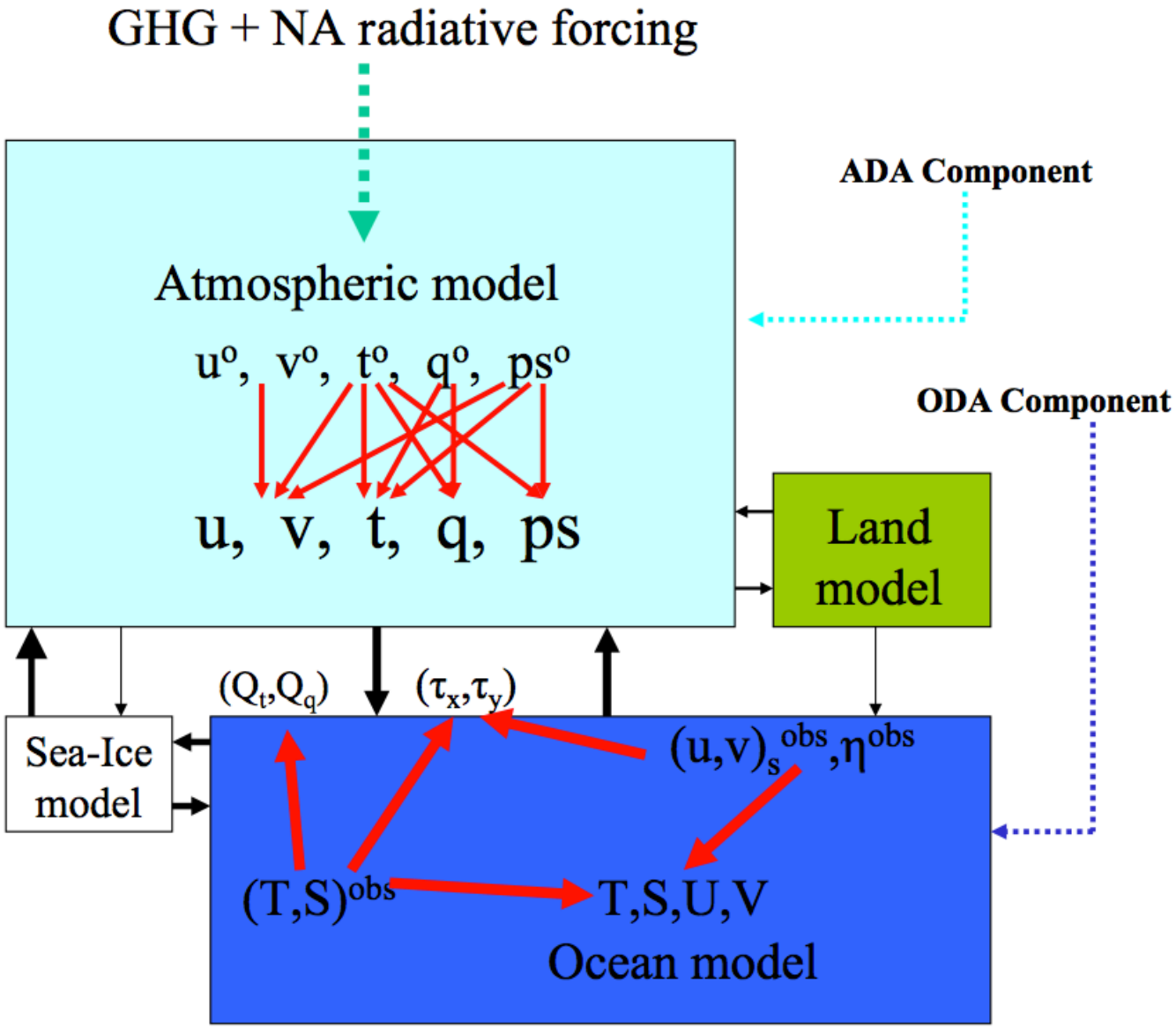
Hybrid/mean LETKF



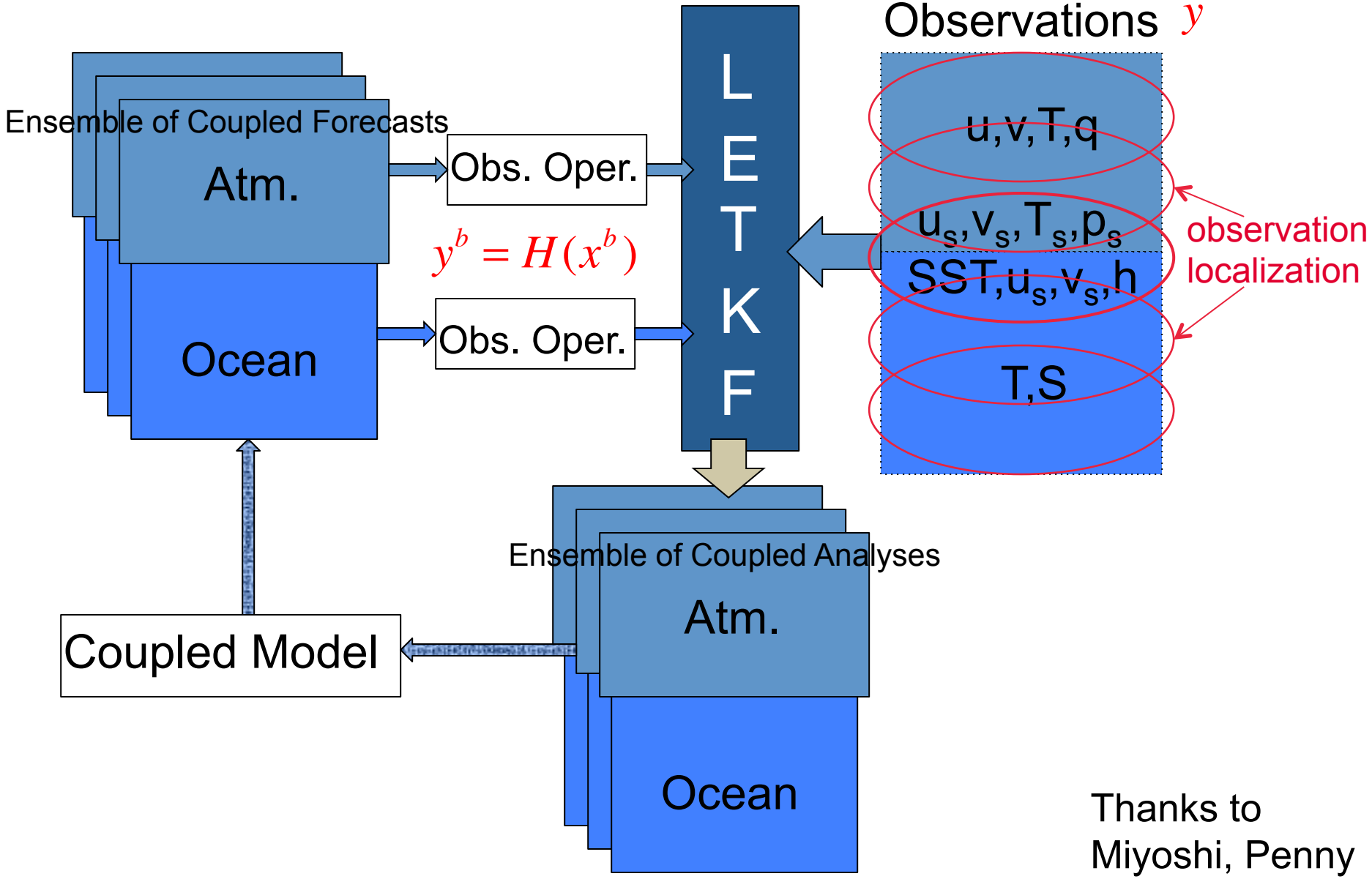
The hybrid/mean LETKF gives more accurate results than the standard hybrid/covariance LETKF.

When alpha=0.5, the accuracy in regime 1 (dark blue) decreases, but the total area of regime 1 increases.

S. Zhang et al.: GFDL Coupled Ocean-Atm EnKF Data Assimilation



Basic idea for our coupled LETKF assimilation



Thanks to
Miyoshi, Penny

Summary: ideas/questions for future coupled ocean-atmosphere EnKF

- Toy model: coupled assimilation and short windows are more accurate for LETKF even if ocean has longer time scales.
- Running in Place (RIP) extracts more information from the observations and allows the use of shorter windows.
- A new hybrid LETKF+simple 3D-Var would make the system more robust with fewer ensemble members and observations.
- For the coupled (India Monsoon Mission) CFS system, we will test the use of 6hr (short) windows for the ocean as well as the atmosphere assimilation.
- Assimilate SST and SSH observations directly.
- Localization of observations near the surface should allow for atm.-ocean interaction through the background error covariance

Summary: ideas/questions for future coupled ocean-atmosphere EnKF

- Toy model: coupled assimilation and short windows are more accurate for LETKF even if ocean has longer time scales.
- Running in Place (RIP) extracts more information from the observations and allows the use of shorter windows.
- A new hybrid LETKF+simple 3D-Var would make the system more robust with fewer ensemble members and observations.
- For the coupled (India Monsoon Mission) CFS system, we will test the use of 6hr (short) windows for the ocean as well as the atmosphere assimilation.
- Assimilate SST and SSH observations directly.
- Localization of observations near the surface should allow for atm.-ocean interaction through the background error covariance

Thanks!