Introduction to Data Assimilation Research at NRL

&

Flow Adaptive Error Covariance Localization

2 Aug, 2007

By

Craig H. Bishop

NRL, Marine Meteorology Division, Monterey, CA

Overview

- Intro to NRL
- NAVDAS and NAVDAS-AR
- Estimation of ob impact using adjoints
- Adaptive error covariance localization using Ensemble COrrelations Raised to A Power (ECO-RAP) and its use in the LETKF
- Conclusions

NRL/FNMOC Forecast Suite

- NOGAPS - Navy Operational Global Atmospheric Prediction System
  - Provides input/boundary conditions for
    - mesoscale, ocean, wave and ice prediction models,
    - ensemble forecasting system
    - Aircraft and ship routing programs
    - tropical cyclone forecast model (GFDN)
    - Aerosol forecasting model, NAAPs
    - Chemistry model, CHEM2D-OPP
- COAMPS®* - Coupled Ocean/ Atmosphere Mesoscale Prediction System
  - nonhydrostatic; globally relocatable, nested grids; explicit prediction of moisture variables
  - 5-10 different operational areas
  - drives ocean, wave, aerosol and EM propagation models
  - Ensemble forecasting system under development
- Both models used for basic research, predictability studies, adjoint sensitivity studies, adaptive observation-targeting

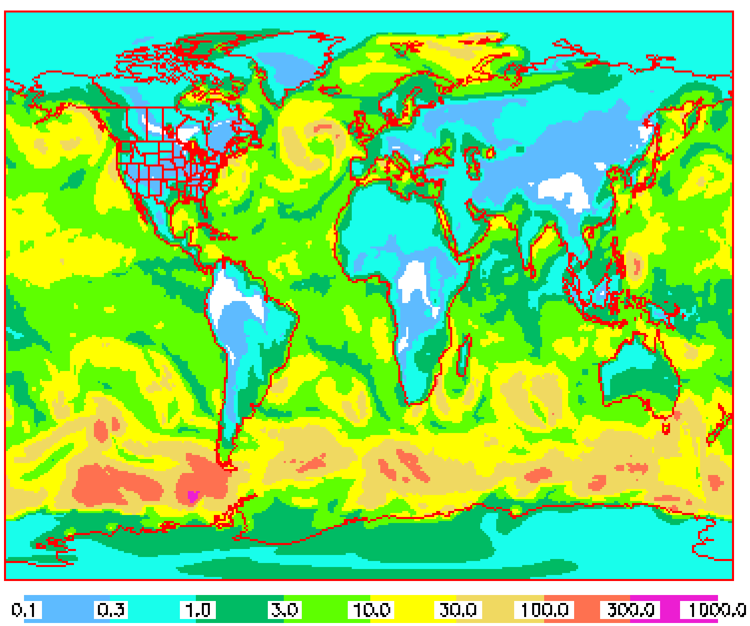
* COAMPS® is a registered trademark of the Naval Research Laboratory, Monterey CA

THE FUTURE OF NAAPS
Integrated into NOGAPS

- Predictive tropospheric and stratospheric aerosol fully embedded within NOGAPS.
- Fully interactive physics – aerosol, cloud formation, and radiative transfer.
- Aerosol coupled to wave model; salt production from WW3.
- Aerosol data assimilation integrated into NAVDAS-AR.
- Aerosol data assimilation will include UV and VIS radiances.
- Aerosol impacts will be included in radiance data assimilation for NWP.

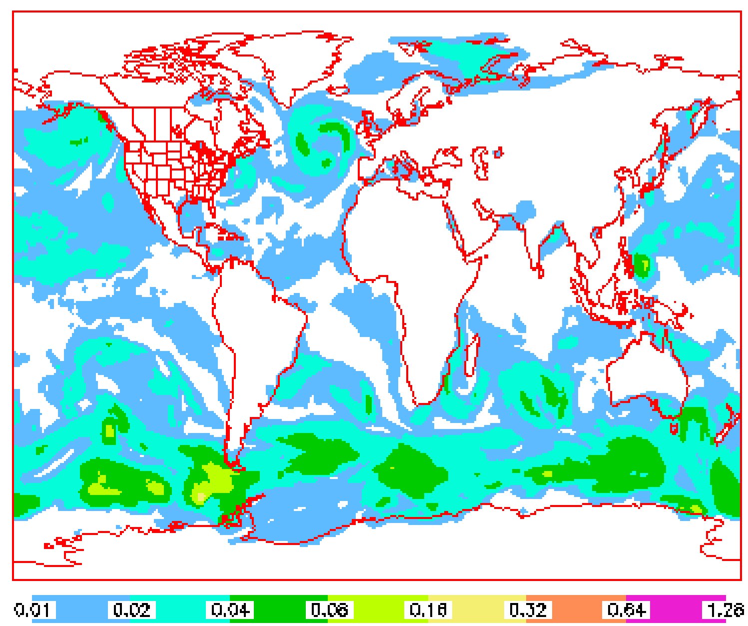
00:00Z 10 May 2006

NAAPS Salt Mass Concentration *(ug-m**3)*



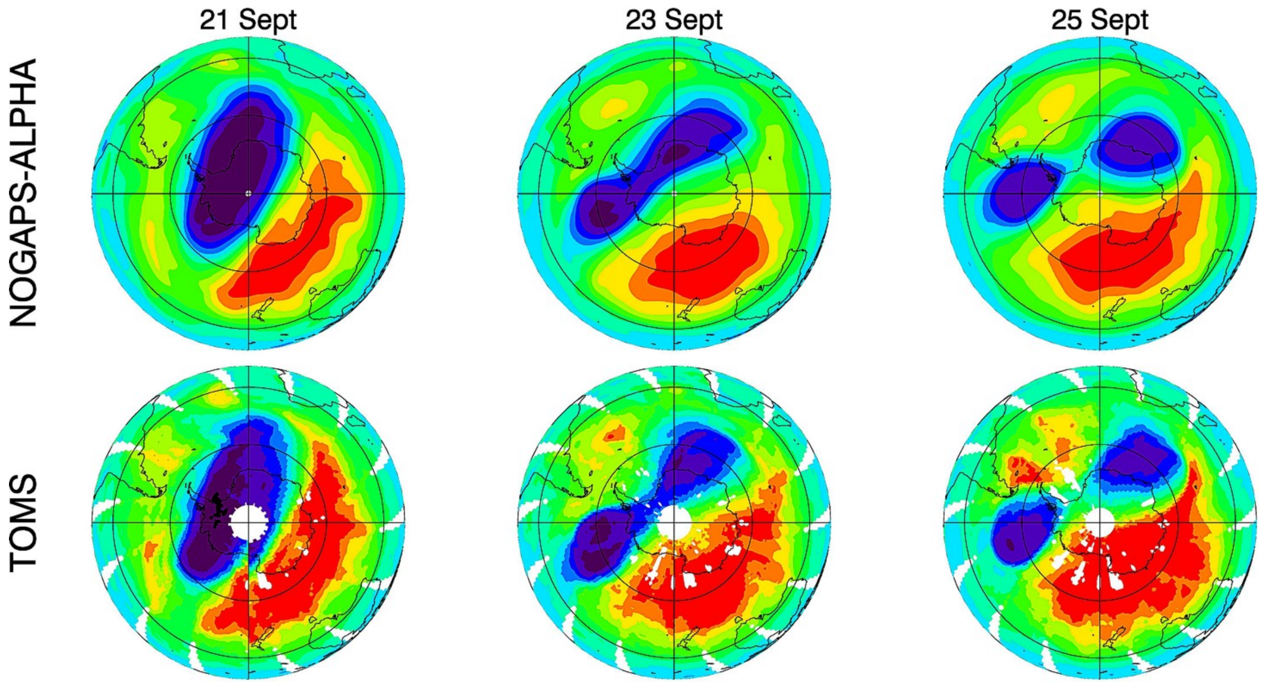
00:00Z 10 May 2006

NAAPS Salt Optical Depth



NOGAPS-ALPHA

Homogeneous Ozone Photochemistry Scheme



First Ever Splitting of Antarctic Ozone Hole: September 2002

NRL Space Science Division

"NOGAPS-ALPHA provides a state-of-the- art stratosphere for NWP applications"

"We get a much improved split vortex in the +5 day forecast by using

(a) new T239L54 NOGAPS-ALPHA

(b) new 3DVAR-based reanalysis (NAVDAS)

CHEM2D-OPP has to date proved superior to photochemistry schemes used in the ECMWF IFS, [former] NCEP GFS, & NASA GEOS5 & GISS models.

This work was funded by the JCSDA and was delivered to NCEP GFS and NASA GMAO as well as NOGAPS.

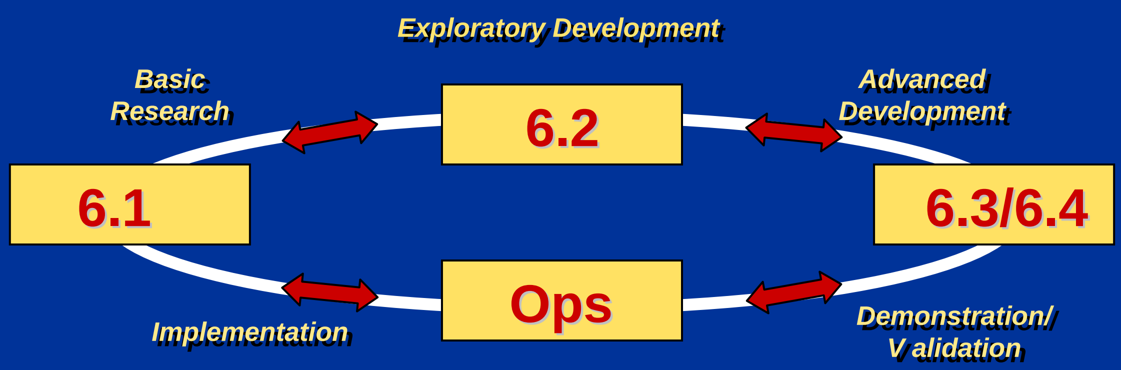
CHEM2D-OPP

Parameterized Ozone Photochemistry for Weather & Climate Models

MARINE METEOROLOGY DIVISION

Unique Program and Partnership

Special Partnership with Primary Customer FNMOC

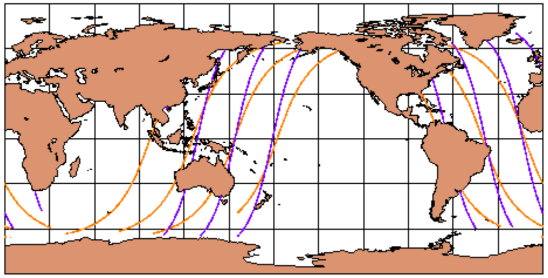


- Seamless transition from research to operations.
- Operational problems can be quickly addressed by NRL.
- Complex operational systems are used in basic research.
- Operational requirements can influence basic research.

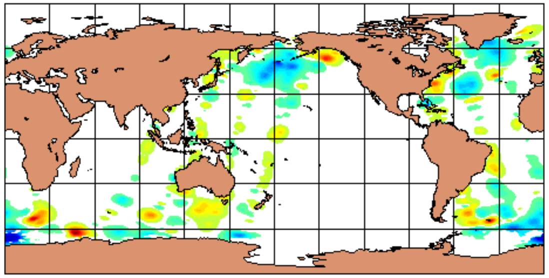
Wave Model Assimilation Validation

- Integrated wave model SWH assimilation and QC code in NCODA
- Completed QC of altimeter SWH data and free run of WW3 model as control
- Performing wave model assimilation runs for pre-beta validation
- Verification includes independent buoys and yet-to-be-assimilated altimeter data – SWH, mean wave period, and buoy spectra vs. model spectra

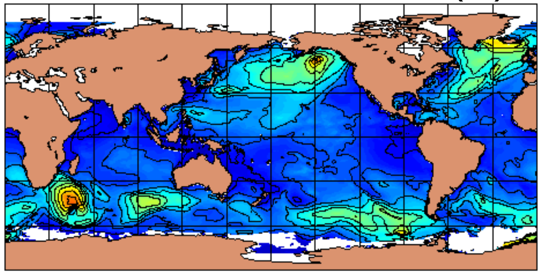
Altimeter SWH Observations



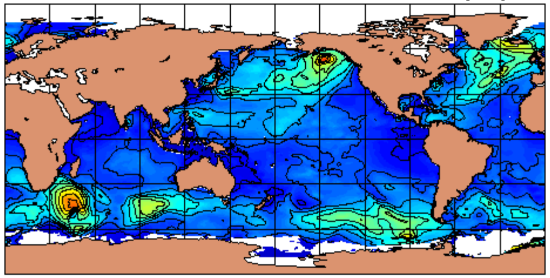
Analyzed Increment SWH (m)



Model Forecast SWH (m)



Corrected Model SWH (m)



Assimilation via 6-Hour Sequential Incremental Update Cycle

NAVDAS-AR*

$$\mathbf{x}_i^a = \mathbf{x}^f + \mathbf{P}_i^f \mathbf{H}_i^T \left(\mathbf{H}_i \mathbf{P}_i^f \mathbf{H}_i^T + \mathbf{R}\right)^{-1} \left(\mathbf{y} - \mathbf{H}_i \mathbf{x}^f\right)$$

Xu and Daley 2002, *Tellus*
Xu *et al.* 2005, *Tellus*

$$\mathbf{x}_i^a = \mathbf{x}^f + \mathbf{P}_o^b M_i^T \mathbf{H}_i^T \left(\mathbf{H}_i M_i \mathbf{P}_o^b M_i^T \mathbf{H}_i^T + \mathbf{R}\right)^{-1} \left(\mathbf{y} - \mathbf{H}_i \mathbf{x}^f\right), \text{ where } \mathbf{P}_i^f = M_i \mathbf{P}_o^b M_i^T$$

$$\mathbf{P}_i^f \approx \mathbf{P}^f = \left\langle \left(\mathbf{x}^f - \mathbf{x}^t\right)\left(\mathbf{x}^f - \mathbf{x}^t\right)^T \right\rangle, \ \mathbf{x}^a = \text{analysis}, \ \mathbf{x}^f = \text{forecast}, \ \mathbf{x}^t = \text{truth}, \ \mathbf{y} = \text{obs}$$

Step 1: Use CG to solve

$$\left(\mathbf{H}_i \mathbf{P}_i^f \mathbf{H}_i^T + \mathbf{R}\right)\mathbf{z}_i = \left(\mathbf{y} - \mathbf{H}_i \mathbf{x}^f\right)$$

so that

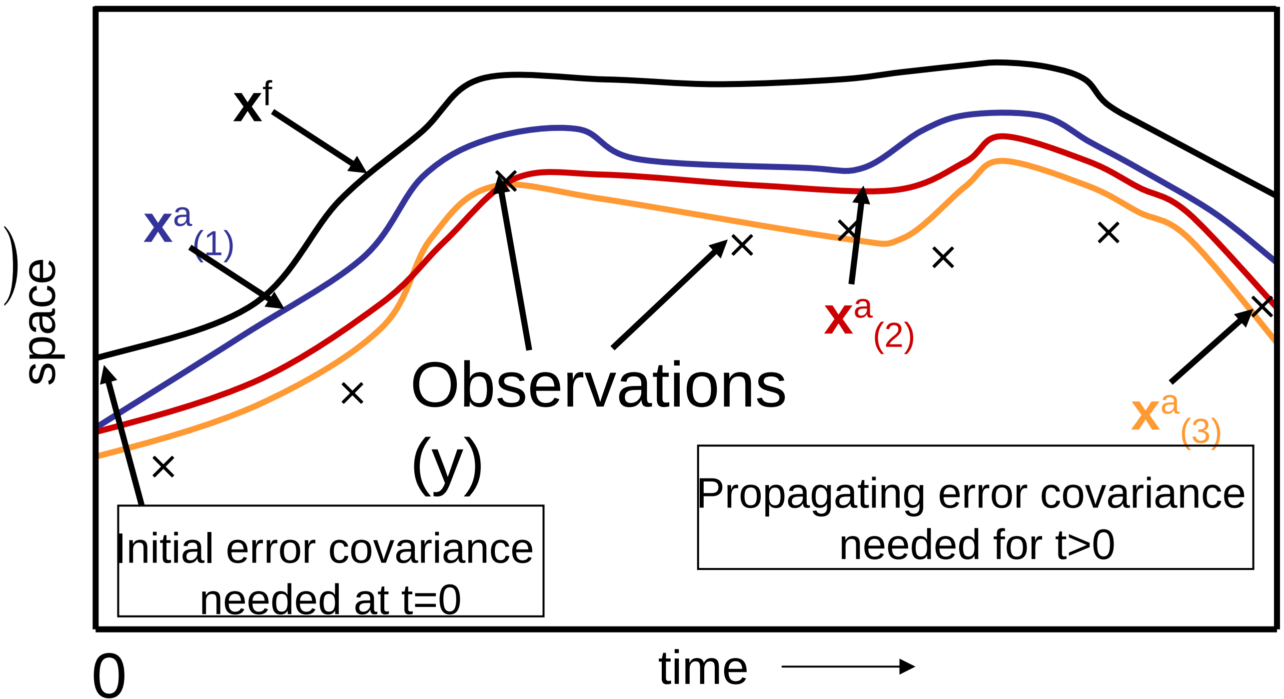
$$\mathbf{z}_i = \left(\mathbf{H}_i \mathbf{P}_i^f \mathbf{H}_i^T + \mathbf{R}\right)^{-1} \left(\mathbf{y} - \mathbf{H}_i \mathbf{x}^f\right)$$

Step 2: Post-multiply

$$\mathbf{x}_i^a = \mathbf{x}^f + \mathbf{P}_i^f \mathbf{H}_i^T \mathbf{z}$$

Step 3: Iterate outer loop

Set i=i+1 and go to Step 1.



NAVDAS-AR*

Let $\mathbf{L}_{ij}$ be the TLM around the ith trajectory that maps perturbations from time step j to time step j+1.

Let $\mathbf{Q}_u$ be the covariance of model errors associated with the time step that are uncorrelated with previous mosel errors.

Let $\mathbf{Q}_c$ be the covariance of model errors across the assimilation window.

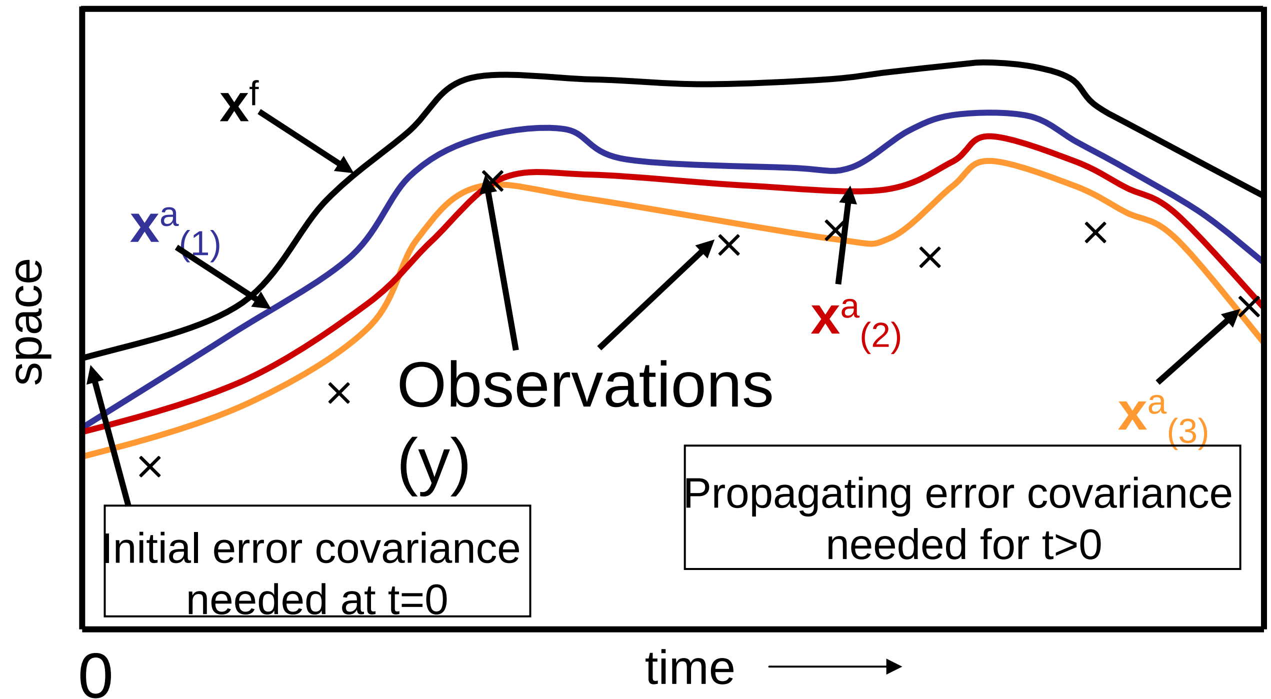
For, say, a 2-time step assimilation window, NAVDAS-AR incorporates this information using

$$\mathbf{P}_i^f = M_i \mathbf{P}_o^b M_i^T = \begin{bmatrix} \mathbf{P}_o^b & \mathbf{P}_o^b \mathbf{L}_{i0}^T & \mathbf{P}_o^b \mathbf{L}_{i0}^T \mathbf{L}_{i1}^T \\ \mathbf{L}_{i0} \mathbf{P}_o^b & \mathbf{L}_{i0} \mathbf{P}_o^b \mathbf{L}_{i0}^T + \mathbf{Q}_u & \left(\mathbf{L}_{i0} \mathbf{P}_o^b \mathbf{L}_{i0}^T + \mathbf{Q}_u \right) \mathbf{L}_{i1}^T \\ \mathbf{L}_{i1} \mathbf{L}_{i0} \mathbf{P}_o^b & \mathbf{L}_{i1} \left(\mathbf{L}_{i0} \mathbf{P}_o^b \mathbf{L}_{i0}^T + \mathbf{Q}_u \right) & \left(\mathbf{L}_{i1} \left(\mathbf{L}_{i0} \mathbf{P}_o^b \mathbf{L}_{i0}^T + \mathbf{Q}_u \right) \mathbf{L}_{i1}^T + \mathbf{Q}_u \right) \end{bmatrix} + \mathbf{Q}_c$$

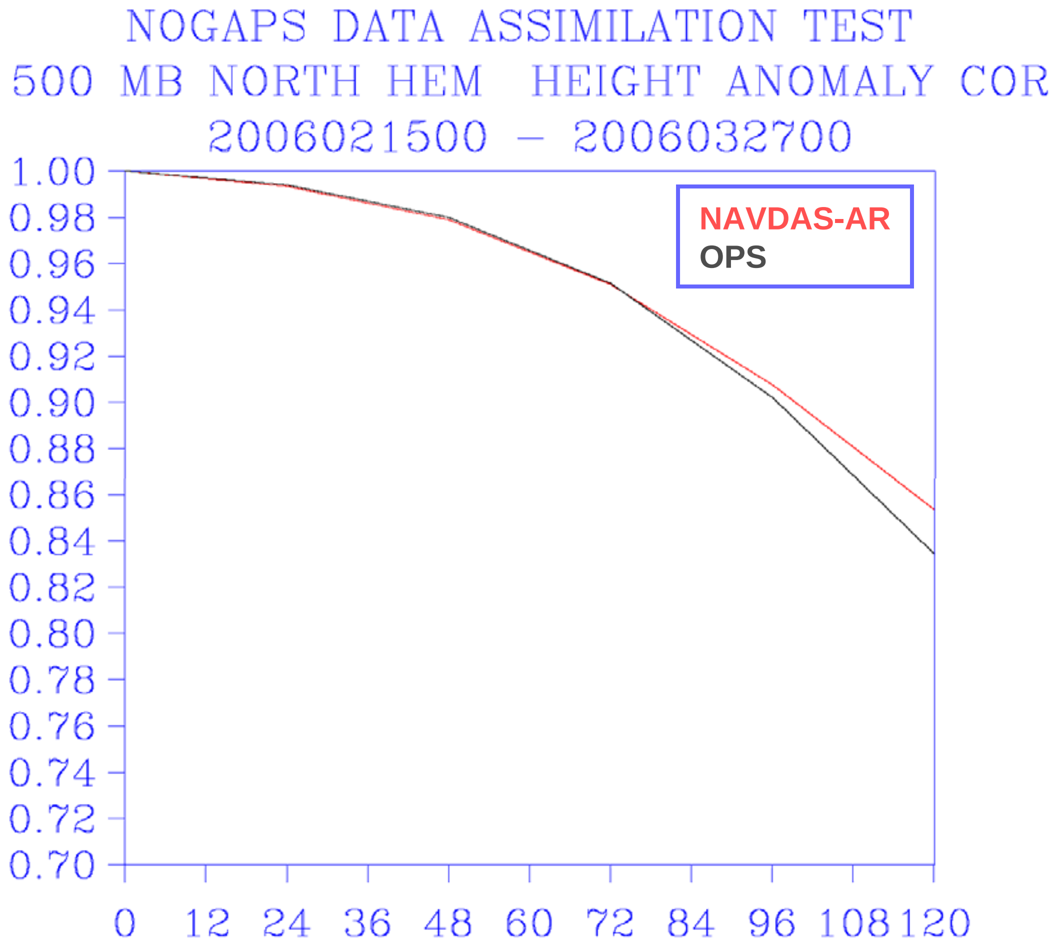
In other words, NAVDAS-AR is a comprehensive weak-constraint 4D-Var DA scheme.

Model Error, Outer Loop, Picard

- When model error included in 4D-Var, best estimate is *not* a model trajectory as it is in strong constraint 4D-VAR
- What to do?
- NRL's current approach is to simply linearize about the best state estimate and propagate error covariances about it. This approach to the outer loop is often called the "Picard" iteration.
- See Y. Tremolet's presentations for alternatives.



Preliminary Results for 06 Winter 'AR' vs. OPS



Forecasts produced with NAVDAS-AR are better than the ones produced with the FNMOC OPS for the winter of 2006.

Minimum Error Variance (MEV) versus (Tremolet's) Maximum Likelihood (ML) Formulation

MEV formulation:

$$J = \frac{1}{2}\mathbf{z}_i^T\left(\mathbf{H}_i M_i \mathbf{P}_0^b M_i^T \mathbf{H}_i^T + \mathbf{R}\right)\mathbf{z}_i - \mathbf{z}_i^T \mathbf{y}', \text{ where } \mathbf{y}' = \left(\mathbf{y} - \mathbf{H}_i \mathbf{x}^f\right)$$

$$\frac{\partial J}{\partial \mathbf{z}_i} = \left(\mathbf{H}_i M_i \mathbf{P}_0^b M_i^T \mathbf{H}_i^T + \mathbf{R}\right)\mathbf{z}_i - \mathbf{y}'$$

$$\mathbf{x}_i^a - \mathbf{x}^f = \mathbf{x}' = M_i \mathbf{P}_0^b M_i^T \mathbf{H}_i^T \mathbf{z}$$

ML formulation:

$$J = \frac{1}{2}\left[\mathbf{w}_i^T \mathbf{w}_i + \mathbf{w}_i^T \mathbf{U}^T M_i^T \mathbf{H}_i^T \mathbf{R}^{-1} \mathbf{H}_i M_i \mathbf{U}\mathbf{w}_i - 2\mathbf{y}'^T \mathbf{R}^{-1} \mathbf{H}_i M_i \mathbf{U}\mathbf{w}_i + \mathbf{y}'^T \mathbf{R}^{-1} \mathbf{y}\right], \text{ where } \mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{P}_0^b$$

$$\frac{\partial J}{\partial \mathbf{w}_i} = \left(\mathbf{U}^T M_i^T \mathbf{H}_i^T \mathbf{R}^{-1} \mathbf{H}_i M_i \mathbf{U} + \mathbf{I}\right)\mathbf{w}_i - \mathbf{U}^T M_i^T \mathbf{H}_i^T \mathbf{R}^{-1} \mathbf{y}'$$

$$\mathbf{x}_i^a - \mathbf{x}^f = \mathbf{x}' = M_i \mathbf{U}\mathbf{w}_i$$

Minimum Error Variance (MEV) versus (Tremolet's) Maximum Likelihood (ML) Formulation

- If error propagation is linear, error distributions are Gaussian and exact minima of cost-functions are found, then methods would be identical.
- Given that the outer loop in MEV is very similar to that in ML, can the outer loop introduce significant differences?
- Satellite observation errors have non-trivial correlations across space, time and channels.
- The MEV formulation does not require a precise inverse of the observation error covariance matrix.
- In order to rigorously handle these correlations, does the ML formulation require the exact inverse of the observation error covariance matrix?
- It is trivial to find the adjoint/gradient of the MEV formulation.
- Does the ML adjoint require line-by-line derivation?

Estimation of ob impact using adjoints

Langland and Baker (Tellus, 2004), Xu, Langland, Baker, and Rosmond (2006)

Assuming linear error propagation

$$\varepsilon^f_{24|-6} = \mathbf{M}\varepsilon^f_{0|-6},\ \ \varepsilon^f_{24|0} = \mathbf{M}\varepsilon^f_{0|-6} + \mathbf{MKv} \quad \left[\mathbf{K} = \mathbf{MP}_i^f\mathbf{H}_i^T\left(\mathbf{H}_i\mathbf{P}_i^f\mathbf{H}_i^T + \mathbf{R}\right)^{-1},\ \ \mathbf{v} = \left(\mathbf{y} - \mathbf{H}_i\mathbf{x}^f_{0|-6}\right)\right]$$

The difference in summed squared forecast error due to obs at 6 z is

$$J = \left(\varepsilon^{fT}_{24|0}\varepsilon^f_{24|0}\right) - \left(\varepsilon^{fT}_{24|-6}\varepsilon^f_{24|-6}\right) = e_{24} - e_{30} = \left(\varepsilon^{fT}_{24|0} + \varepsilon^{fT}_{24|-6}\right)\left(\varepsilon^f_{24|0} - \varepsilon^f_{24|-6}\right)$$

$$= \left(\varepsilon^{fT}_{24|0} + \varepsilon^{fT}_{24|-6}\right)\mathbf{M}(\varepsilon^f_{0|0} - \varepsilon^f_{0|-6})$$

$$= \left(\varepsilon^{fT}_{24|0} + \varepsilon^{fT}_{24|-6}\right)\mathbf{M}\left[(\mathbf{x}^a_{0|0} - \mathbf{x}^t_0) - (\mathbf{x}^f_{0|-6} - \mathbf{x}^t_0)\right]$$

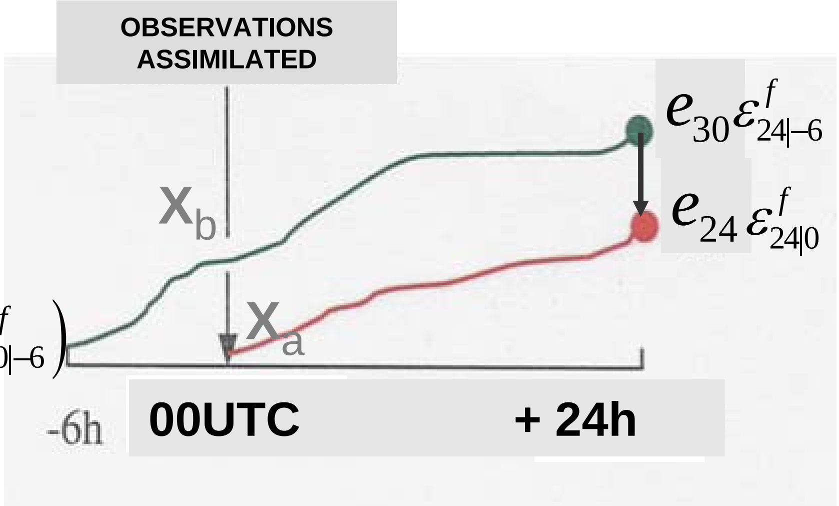
$$= \left(\varepsilon^{fT}_{24|0} + \varepsilon^{fT}_{24|-6}\right)\mathbf{M}\left[\mathbf{x}^a_{0|0} - \mathbf{x}^f_{0|-6}\right]$$

$$= \left(\varepsilon^{fT}_{24|0} + \varepsilon^{fT}_{24|-6}\right)\mathbf{MP}_i^f\mathbf{H}_i^T\left(\mathbf{H}_i\mathbf{P}_i^f\mathbf{H}_i^T + \mathbf{R}\right)^{-1}\left(\mathbf{y} - \mathbf{H}_i\mathbf{x}^f_{0|-6}\right)$$

$$= \left(\varepsilon^{fT}_{24|0} + \varepsilon^{fT}_{24|-6}\right)\mathbf{MKv}$$

$$= \left(2\varepsilon^{fT}_{0|-6}\mathbf{M}^T + \mathbf{v}^T\mathbf{K}^T\mathbf{M}^T\right)\mathbf{MKv}$$

$$= 2\varepsilon^{fT}_{0|-6}\mathbf{M}^T\mathbf{MKv} + \mathbf{v}^T\mathbf{K}^T\mathbf{M}^T\mathbf{MKv}$$



$$(\mathbf{x}^b = \mathbf{x}^f_{0|-6}, \mathbf{x}^a = \mathbf{x}^f_{0|0} = \mathbf{x}^a_{0|0})$$

Estimation of ob impact using adjoints

Since $\varepsilon_{0|-6}^{f}$ has the same value in the expressions for the 24 hr and 30 hr forecast errors, it follows that it is only variations in $\mathbf{v}=\mathbf{y}-\mathbf{H}_i\mathbf{x}^f=\varepsilon^o-\mathbf{H}\varepsilon_{0|-6}^{f}$ through variations in ε^o that can influence the reduction in squared forecast error. Taking the derivative of J with respect to $\mathbf{v}$ while holding $\varepsilon_{0|-6}^{f}$ constant yields

$$\left.\frac{\partial \mathrm{J}}{\partial \mathbf{v}}\right|_{\varepsilon_{0|-6}^{f}}=2\left[\mathbf{K}^T\mathbf{M}^T\mathbf{M}\varepsilon_{0|-6}^{f}+\mathbf{K}^T\mathbf{M}^T\mathbf{MKv}\right]=2\left[\mathbf{K}^T\mathbf{M}^T\varepsilon_{24|-6}^{f}+\mathbf{K}^T\mathbf{M}^T\mathbf{MKv}\right]$$

Since the variation δe_{30} due to a change $\delta\mathbf{x}_{0|-6}^{f}$ in $\mathbf{x}_{0|-6}^{f}$ is given by

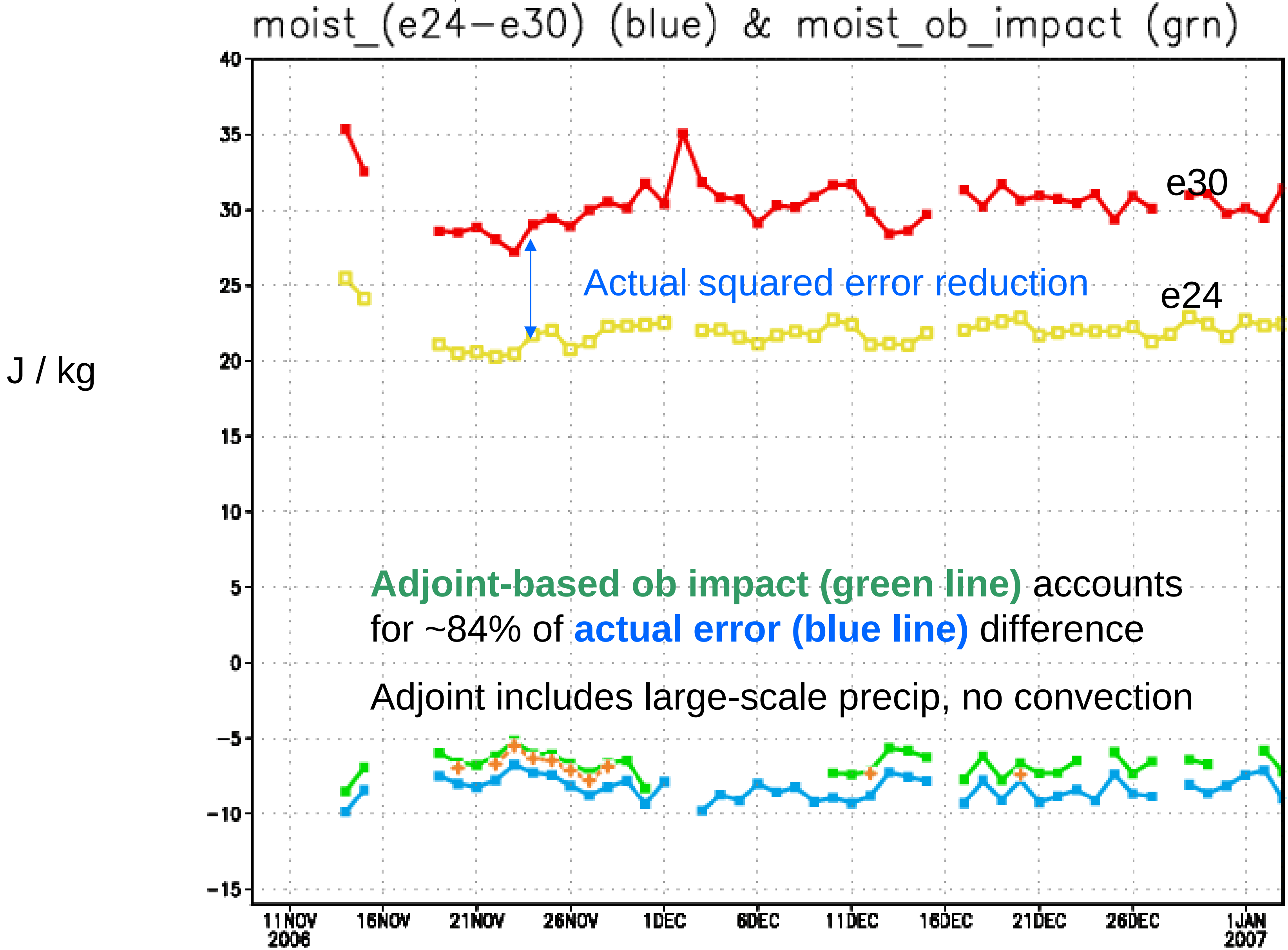
$$\delta e_{30}=\delta\mathbf{x}_{0|-6}^{fT}\mathbf{M}^T\mathbf{M}\delta\mathbf{x}_{0|-6}^{f}, \text{ it follows that } \frac{\partial e_{30}}{\partial\mathbf{x}_{0|-6}^{f}}=2\mathbf{M}^T\varepsilon_{24|-6}^{f}, \text{ when } \mathbf{M}\delta\mathbf{x}_{0|-6}^{f}=\varepsilon_{24|-6}^{f}$$

Similarly, since the variation δe_{24} due to a change $\delta\mathbf{x}_{0|0}^{a}$ in $\mathbf{x}_{0|0}^{a}$ is given by

$$e_{24}=\delta\mathbf{x}_{0|0}^{aT}\mathbf{M}^T\mathbf{M}\delta\mathbf{x}_{0|0}^{a}, \text{ it follows that } \frac{\partial e_{24}}{\partial\mathbf{x}_{0|0}^{a}}=2\mathbf{M}^T\mathbf{MKv} \text{ when } \delta\mathbf{x}_{0|0}^{a}=\mathbf{Kv}.$$

Thus, Langland and Baker (Tellus, 2004) find $\left.\frac{\partial \mathrm{J}}{\partial \mathbf{v}}\right|_{\varepsilon_{0|-6}^{f}}=\mathbf{K}^T\left[\frac{\partial e_{30}}{\partial\mathbf{x}_{0|-6}^{f}}+\frac{\partial e_{24}}{\partial\mathbf{x}_{0|0}^{a}}\right]$

How accurate is the $\delta J = \mathbf{v}^T \left.\frac{\partial \mathrm{J}}{\partial \mathbf{v}}\right|_{\varepsilon^f_{0|-6}}$ estimate using NOGAPS/NAVDAS adjoints?



The adjoint based ob impact is a good predictor of the true impact

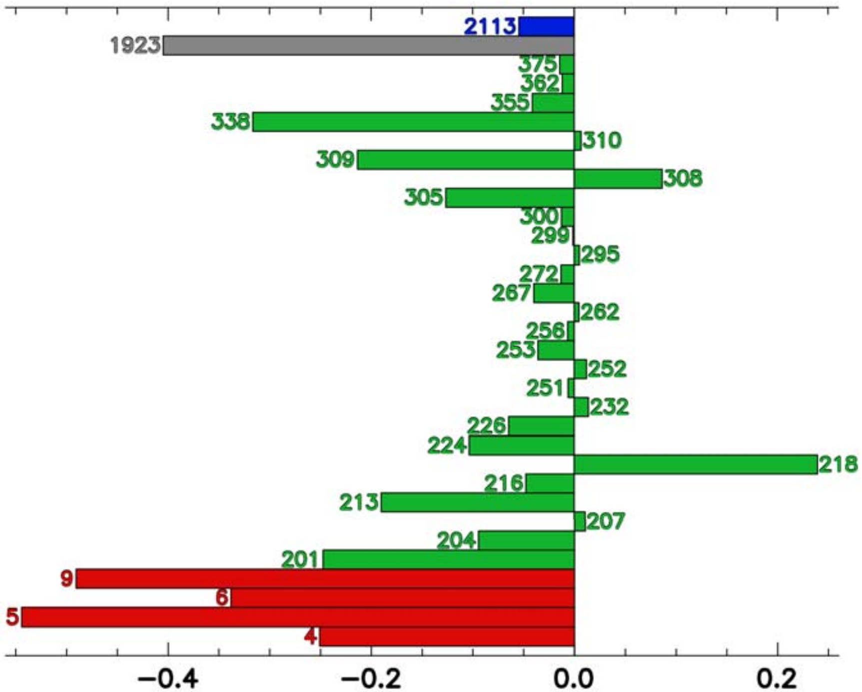
NAVDAS ADJOINT

Total Impact by Satellite Channel

Assessment of AQUA sensors
AMSU/A, AIRS longwave 14-13µm,
AIRS shortwave 4.474µm, AIRS shortwave 4.180µm

- **AIRS has 2378 spectral channels!**

AQUA sensitivity specified by channel number: Aug 15-26, 2006



Beneficial **Non-beneficial**

- **NRL pioneered methodology for quantifying reduction in forecast error for each individual satellite channel**
- **JCSDA partners will use methodology to optimally select satellite observations for maximum NWP impact.**
- Comparison of observation impact results between JSDCA partners will help identify problems with observing systems and assimilation systems.

Funded in part by JCSDA

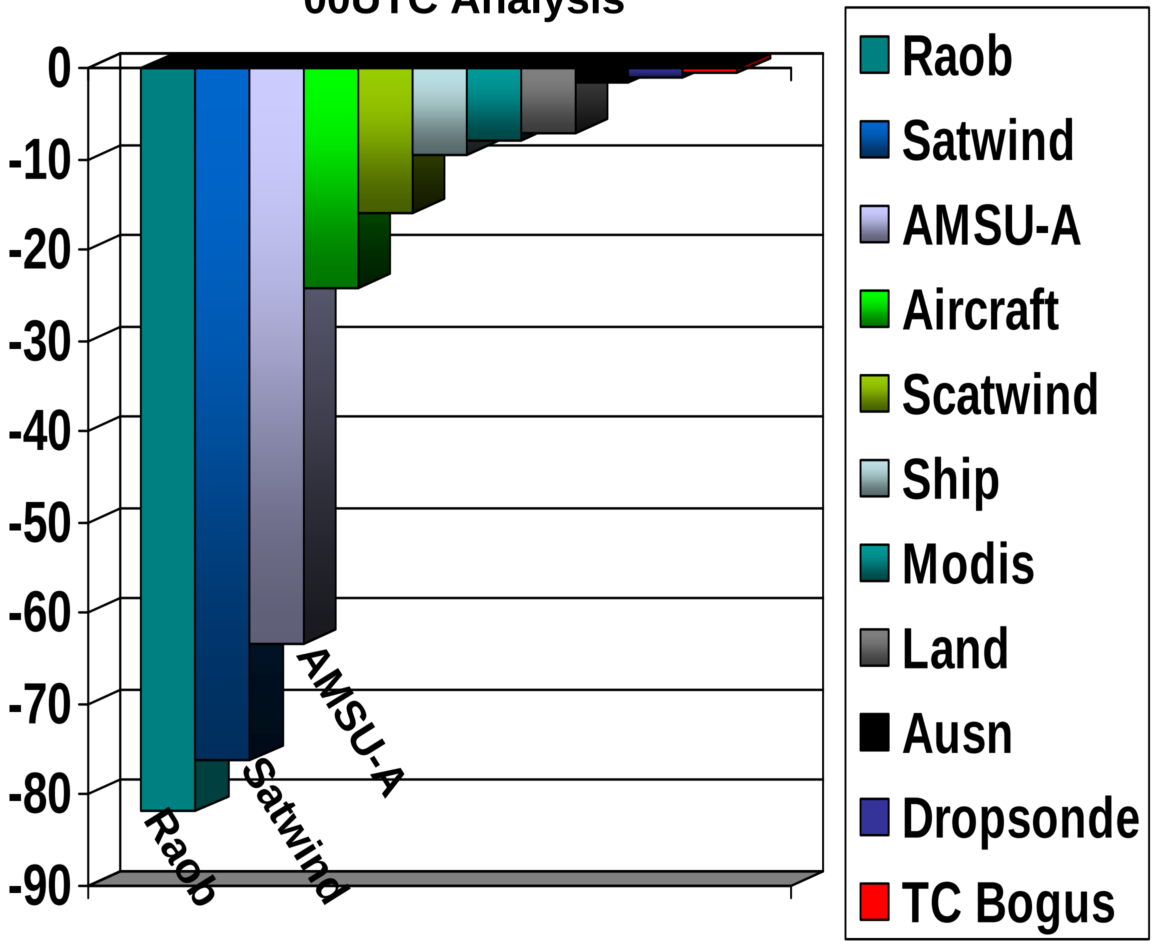
NAVDAS ADJOINT

Total Impact by Observation Type

Units of impact = J kg^{-1}

1 Jan – 28 Feb 2006
00UTC Analysis

Beneficial impact



ECO-RAP
A new adaptive error covariance localization tool for 4-dimensional ensemble data assimilation

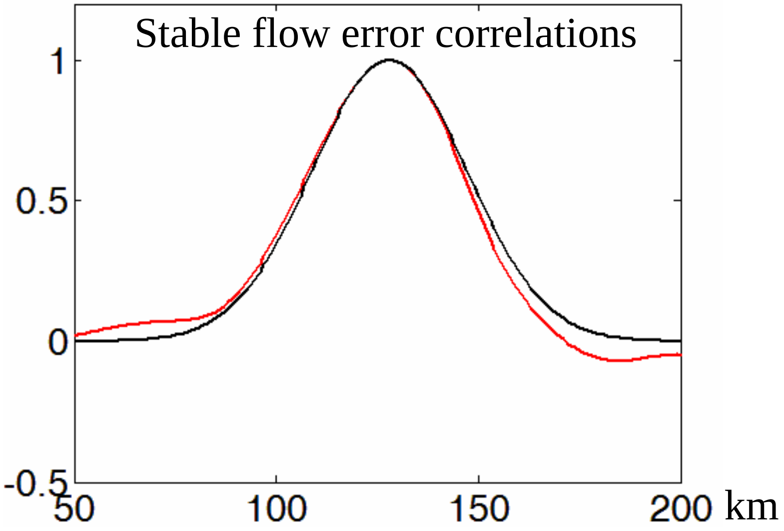
Craig H. Bishop, Daniel Hodyss,
William. F. Campbell, and Justin G. Mclay

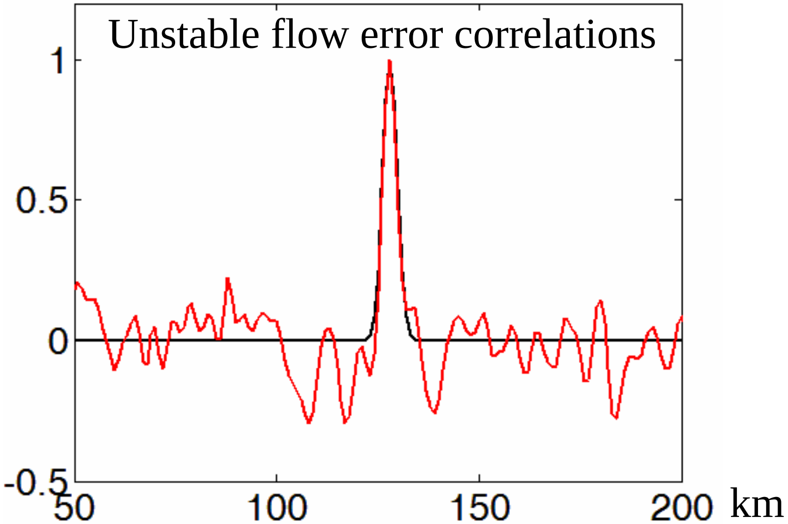
Naval Research Laboratory, Monterey, California

Outline

- Motivation
- How ECO-RAP works
- Idealized tests
- Review
- Computational considerations/speed-up
- Preliminary experiment with NWP model
- Conclusions

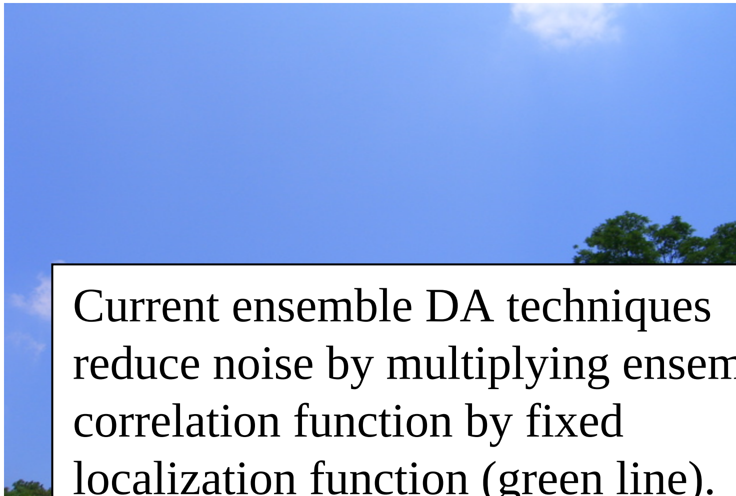
Small Ensembles and Spurious Correlations

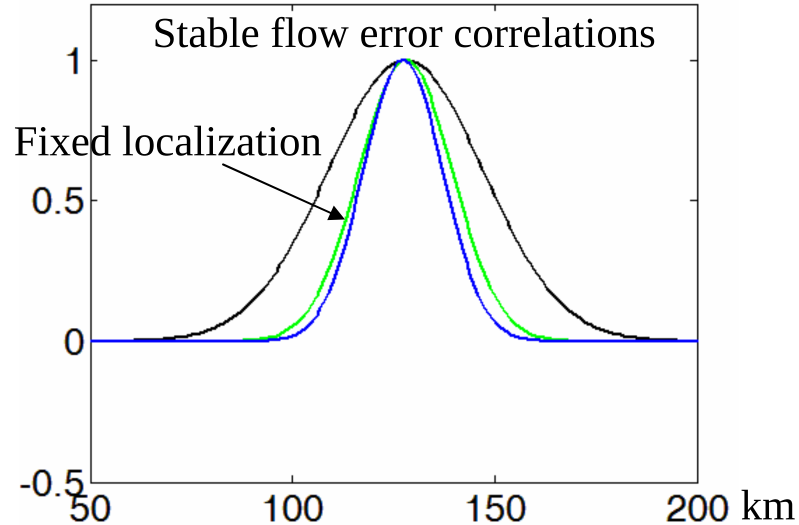




Ensembles give flow dependent, but ***noisy*** correlations

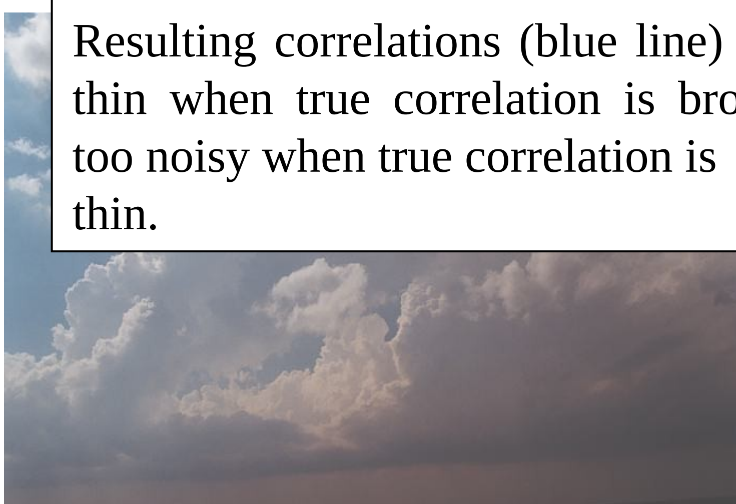
Small Ensembles and Spurious Correlations

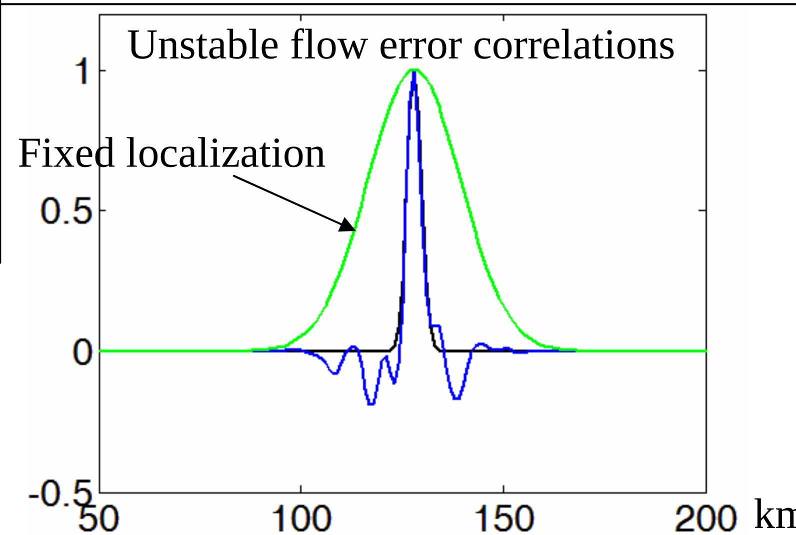




Current ensemble DA techniques reduce noise by multiplying ensemble correlation function by fixed localization function (green line).

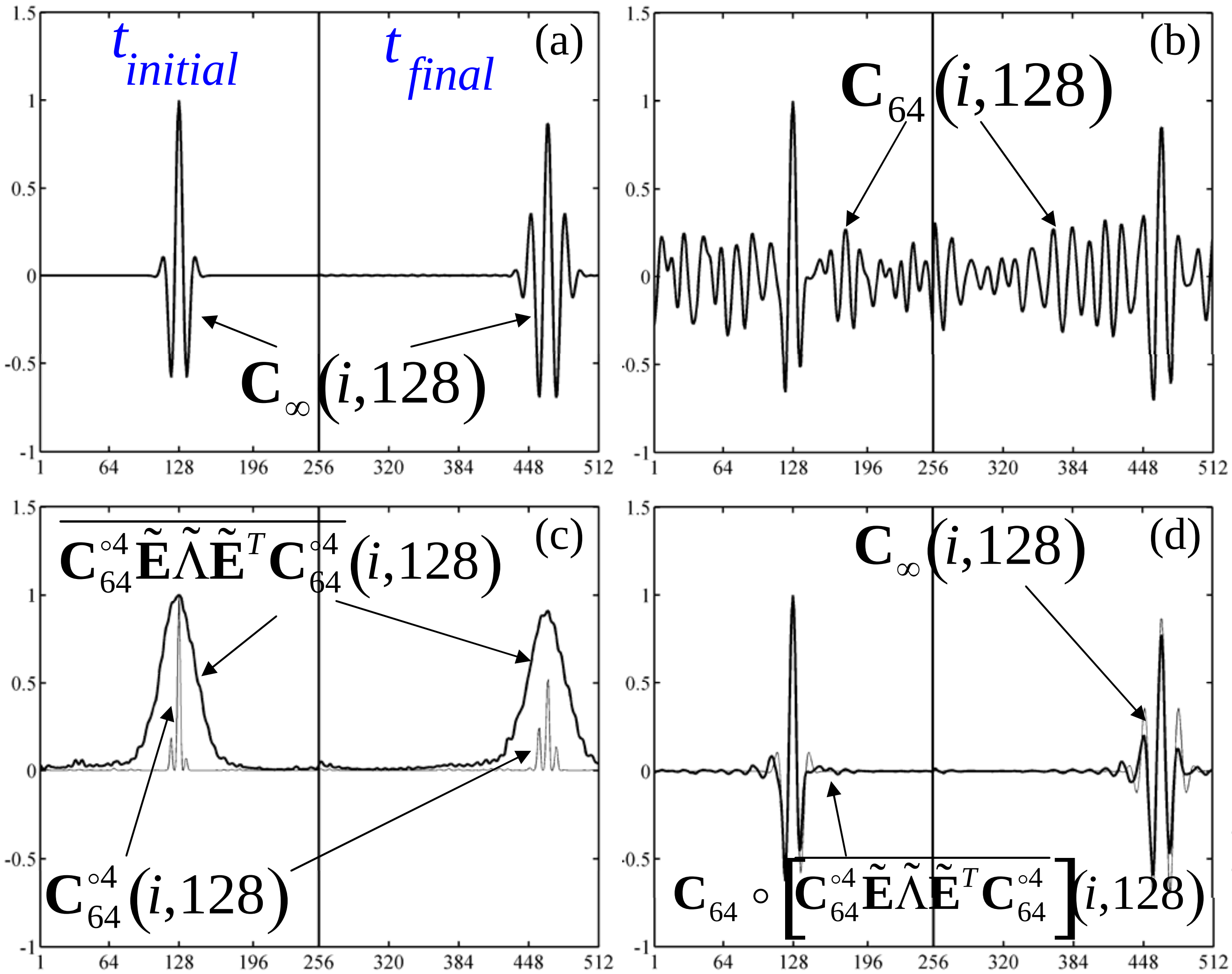
Resulting correlations (blue line) are too thin when true correlation is broad and too noisy when true correlation is thin.





Today's fixed localization functions limit adaptivity

ECO-RAP: Ensemble COrrelations RAised to a Power



$K = 64$ member ensemble

$\mathbf{C}_K =$ Ensemble correlation matrix

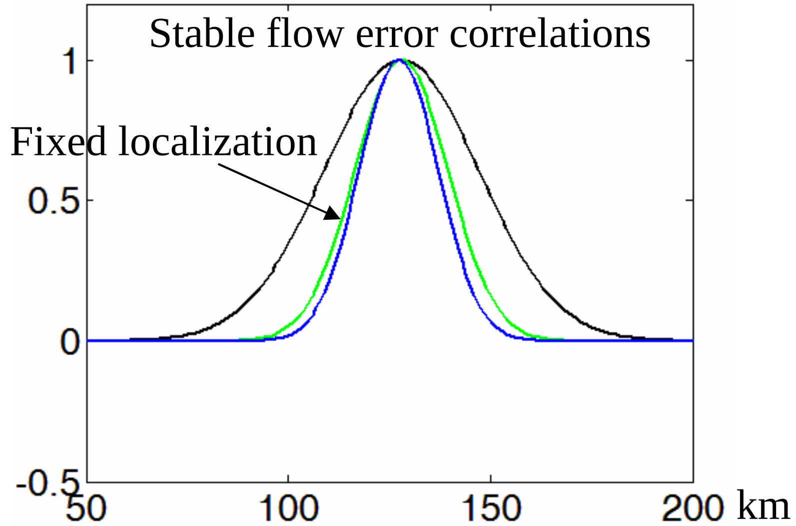
$\mathbf{C}_K^{\circ n} =$ n elementwise products of ensemble correlation

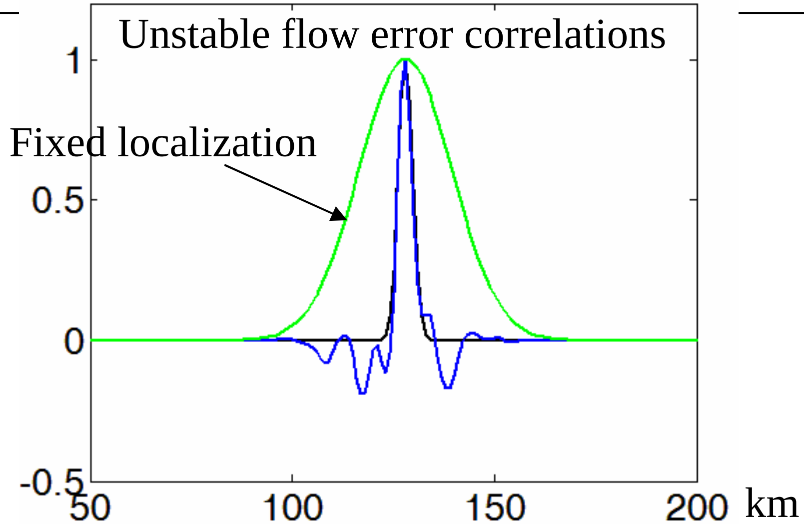
$\tilde{\mathbf{E}}\tilde{\Lambda}\tilde{\mathbf{E}}^T =$ Non-adaptive localization matrix

ECO-RAP: Ensemble COrrelations RAised to a Power

- Ensemble correlations contain propagation and length scale information.
- Ensemble correlations corresponding to large true correlations are bigger than those corresponding to true zero correlations. (Variance of spurious is 1/K).
- Raising ensemble correlations to a power attenuates small values more than large values.
- Sandwiching *non-adaptive* localization matrix between ensemble correlation matrices raised to a power yields *adaptive* localization matrix.

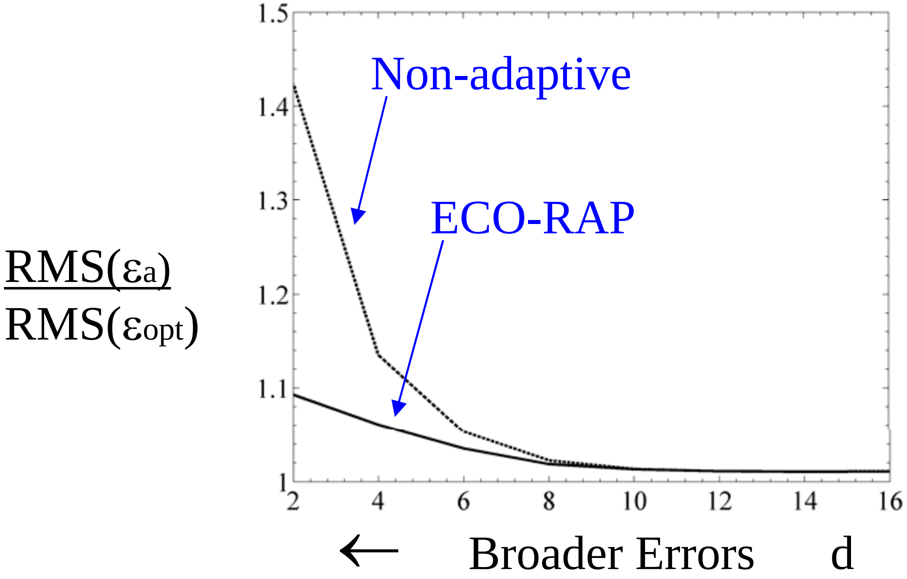
Length Scale Variability Experiment





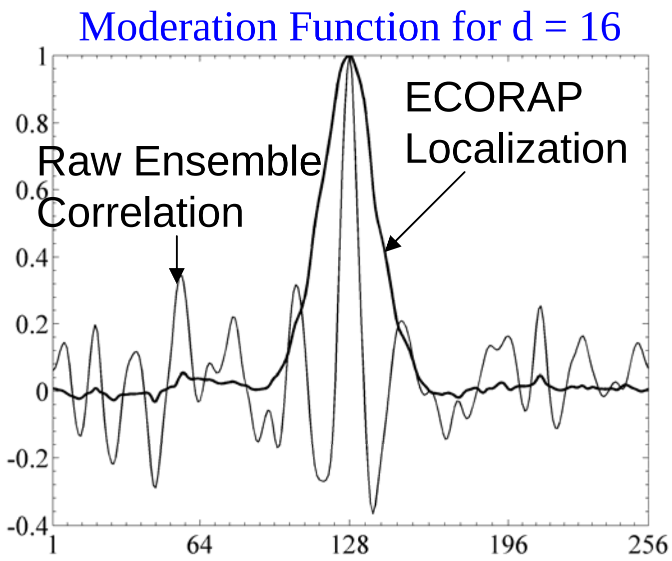
Tune for short error length scales then test on broad

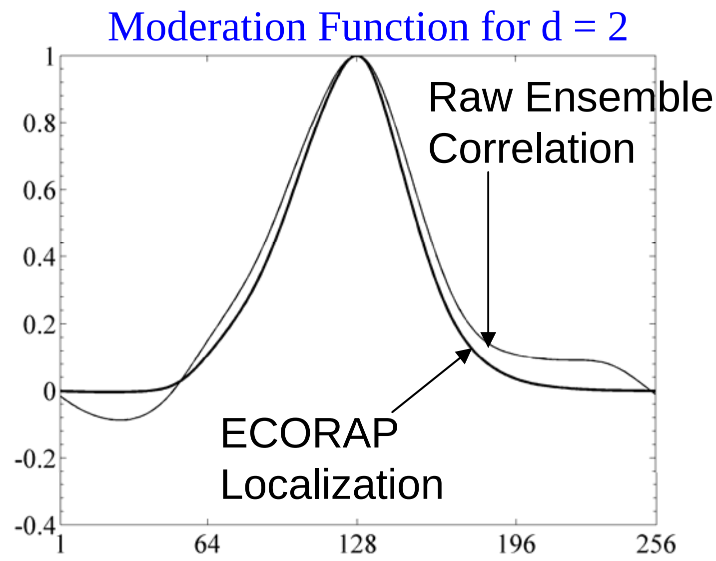
Length Scale Variability Experiment



- Tune ECO-RAP and Non-Adaptive localization methods for lowest analysis error at scale d = 16 with 156 obs

- Compare the performance of the two schemes when the true error correlation length scale is broader than that for d=16





Single ob. 4D Data Assimilation Test (16 members)

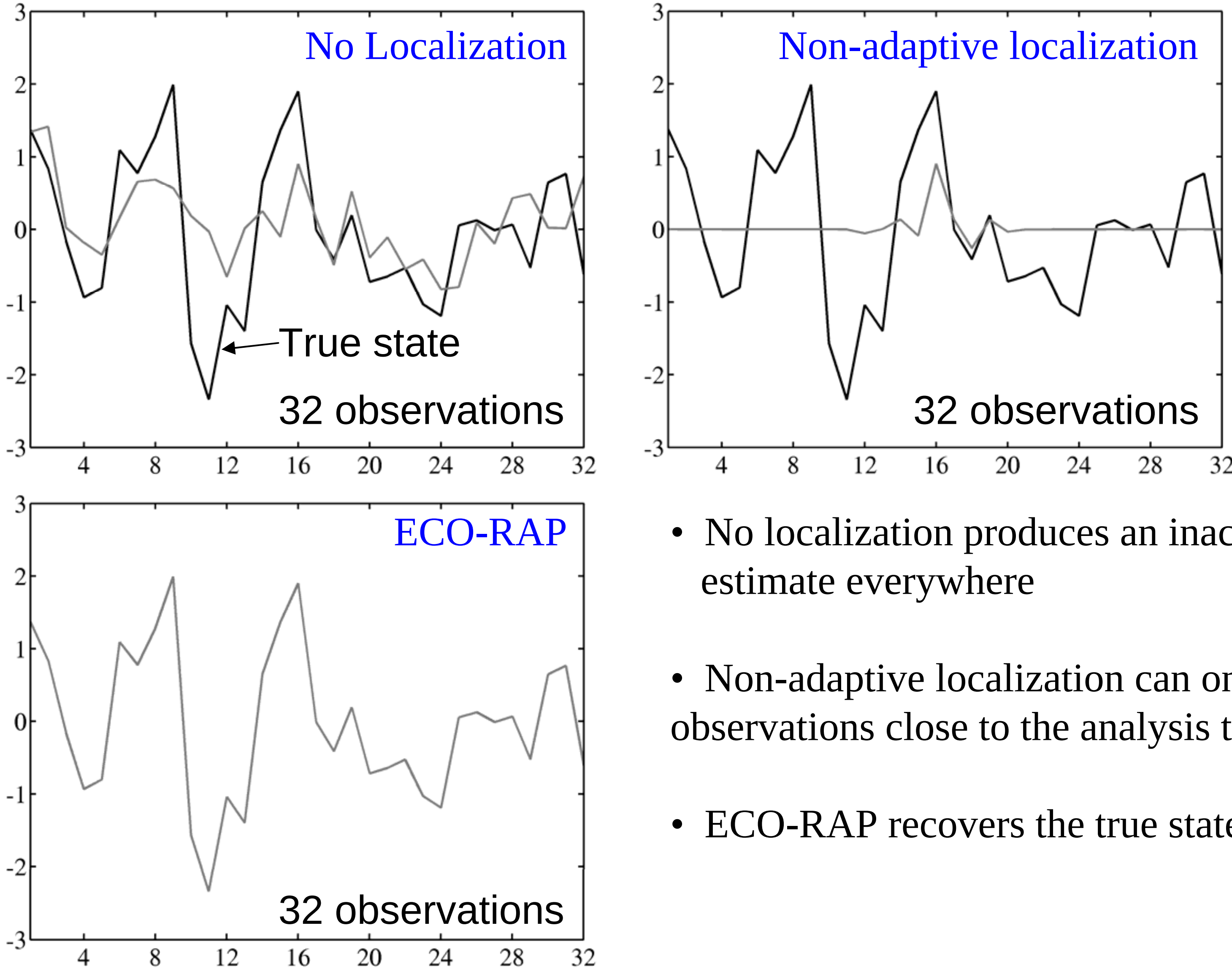
- 32 variables in periodic domain

- Truth moves to the right one grid point per time step

- One ob. per time step at variable 16 (very small ob error variance)

- After 32 time steps use all 32 collocated observations to estimate the initial state

Single ob. 4D Data Assimilation Test (16 members)

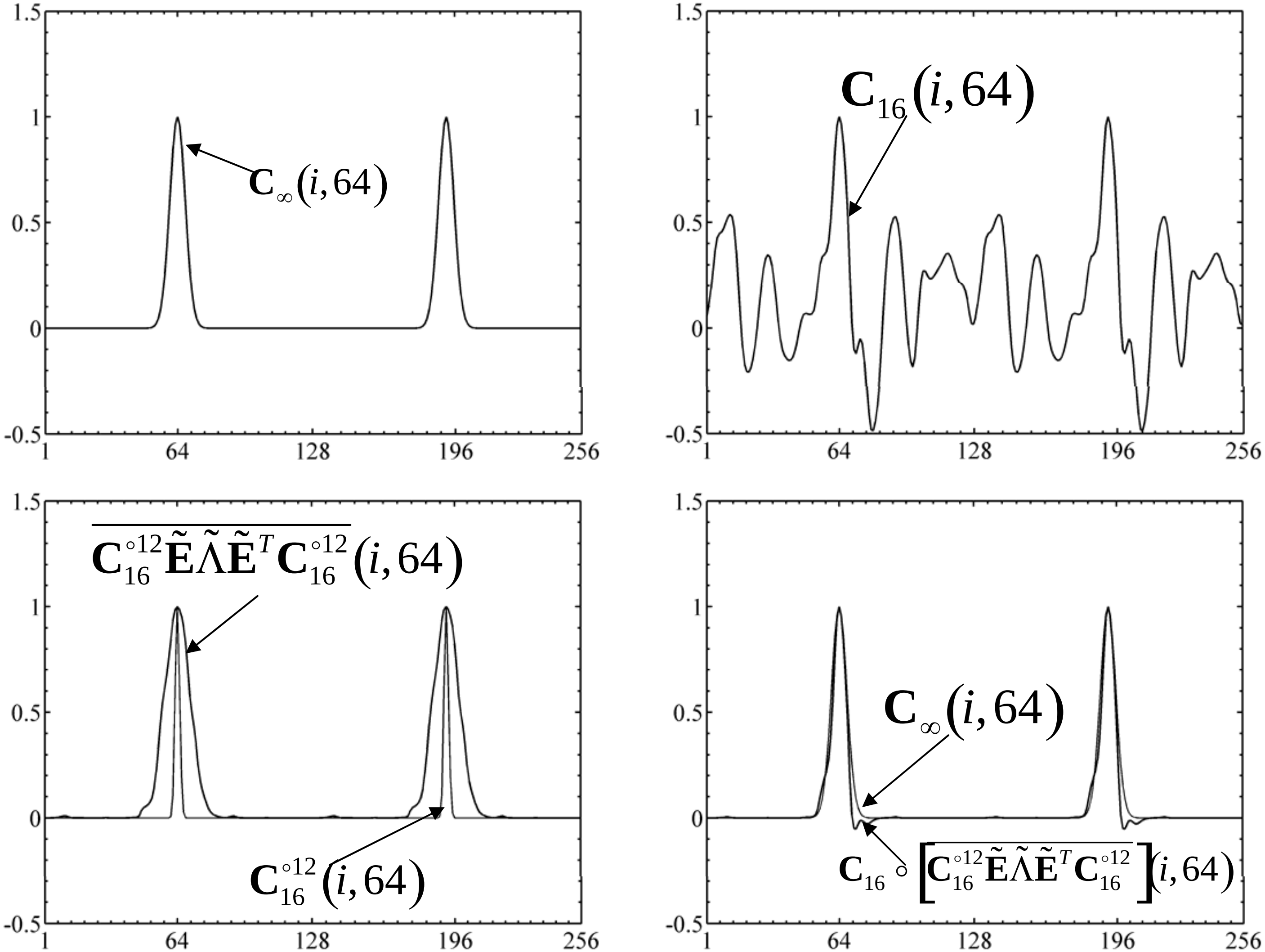


- No localization produces an inaccurate estimate everywhere
- Non-adaptive localization can only use observations close to the analysis time
- ECO-RAP recovers the true state

Localization or Moderation?

1) Are the variables whose forecast errors correlate with another variable confined to the geographic neighbourhood of that variable?
2) What is “local” about forecast errors due to a misspecification of the albedo of stratus clouds?
3) What is “local” about errors associated with a sudden stratospheric warming event?
4) ECO-RAP can moderate spurious correlations even when the answer to (1) is “No”.

Localization or Moderation?



$K = 16$ member ensemble

$\mathbf{C}_K =$ Ensemble correlation (K members)

$\mathbf{C}_K^{\circ n} =$ n Hadamard products of ensemble correlation

$\tilde{\mathbf{E}}\tilde{\Lambda}\tilde{\mathbf{E}}^{T} =$ a spectral smoother

Review

- ECO-RAP is a new flow-adaptive localization method for ensemble DA.
- It raises ensemble correlations to powers (Hadamard products) to selectively reduce spurious correlations.
- Broad localization functions are obtained by sandwiching non-adaptive localization matrices between correlation matrices raised to a power.
- ECO-RAP adapts to changes in the propagation and scale characteristics of errors.
- ECO-RAP is as good as non-adaptive localization when error distribution is invariant.

Computational Issues

- N=number of model variables
- $\mathbf{C}_{\text{ECO-RAP}} = \overline{\mathbf{C}_K^{\circ n}\tilde{\mathbf{E}}\tilde{\Lambda}\tilde{\mathbf{E}}^T\mathbf{C}_K^{\circ n}}$ has N^2 elements
- So does the Covariances Adaptively Localized with ECO-rap (CALECO) matrix

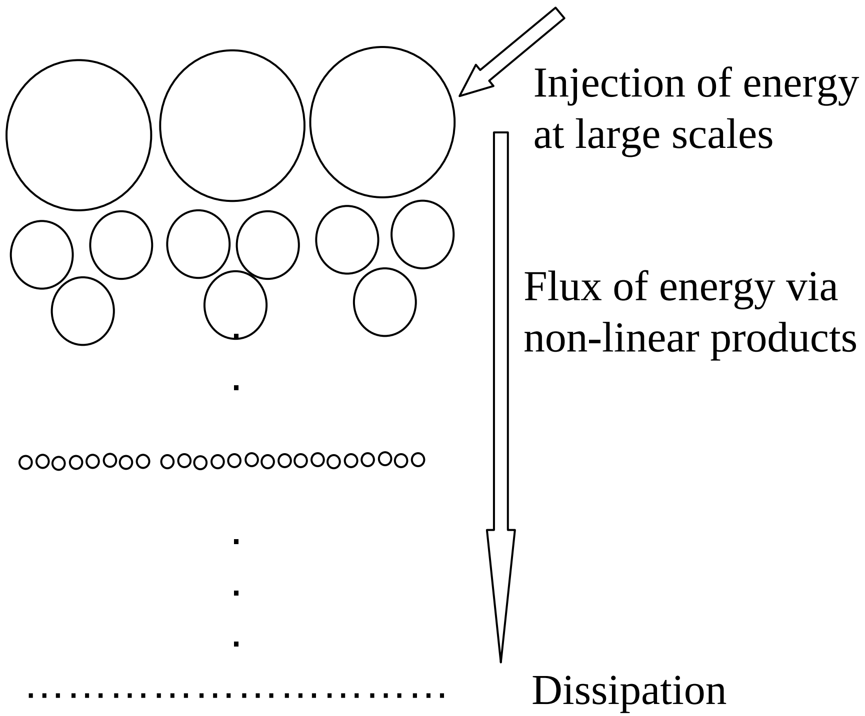
$$\mathbf{P}_{CALECO}^f = \mathbf{P}_K^f \circ \left[\overline{\mathbf{C}_K^{\circ n}\tilde{\mathbf{E}}\tilde{\Lambda}\tilde{\mathbf{E}}^T\mathbf{C}_K^{\circ n}}\right]$$

Local Ensemble Transform Kalman Filter

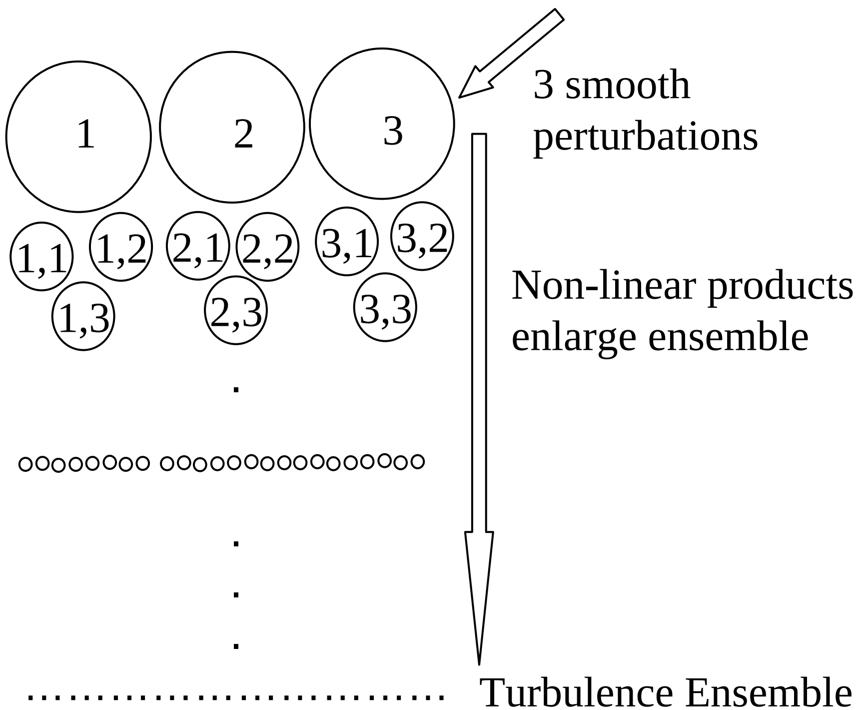
- Local Ensemble Transform Kalman Filter (LETKF) [Hunt et al (2007; Physica D)]
  - Each grid point is updated only with the observations lying within grid point’s observation volume.
  - Each grid point can be updated independently so algorithm is scalable.
  - Finite observation volume is needed to limit the effect of spurious long-distance correlations.
  - Problematic for observations of vertical integrals of model variables such as satellite obs.
  - Problematic for 4D assimilation when errors propagate further than the localization width over the time window of interest.
  - Redundancy in observation processing since there is a high degree of overlap between volumes.

Turbulence inspired recipe for huge 'turbulence' ensemble

Turbulent energy cascade following Richardson's (1922) ideas.



Recipe for creation of huge ensemble from small ensemble



If covariance of turbulence ensemble was CALECO then computer memory would only need to store "energy containing eddies".
Is there a turbulence ensemble whose covariance is CALECO?

Turbulence ensemble for CALECO

It may be shown that if

$$\mathbf{A} = \mathbf{U}\mathbf{U}^T \text{ and } \mathbf{B} = \mathbf{V}\mathbf{V}^T \text{ then } \mathbf{A}\circ\mathbf{B} = [\mathbf{U}\otimes\mathbf{V}][\mathbf{U}\otimes\mathbf{V}]^T,$$

where $\mathbf{U}\otimes\mathbf{V}$ indicates the matrix whose columns list all possible non-linear products of the columns of $\mathbf{U}$ and $\mathbf{V}$.

Consequently, covariance of turbulence ensemble one obtains by taking all possible non-linear products of "energy containing eddies" $\mathbf{U}$ and $\mathbf{V}$ is the element-wise product of the covariances of $\mathbf{U}$ and $\mathbf{V}$. Hence, since

$$\mathbf{P}_K^f = \mathbf{Z}_K\mathbf{Z}_K^T \text{ and } \mathbf{C}_{\text{ECO-RAP}} = \overline{\mathbf{C}_K^{\circ n}\tilde{\mathbf{E}}\tilde{\Lambda}\tilde{\mathbf{E}}^T\mathbf{C}_K^{\circ n}} = \left[\overline{\mathbf{C}_K^{\circ n}\tilde{\mathbf{E}}\tilde{\Lambda}^{1/2}}\right]\left[\overline{\mathbf{C}_K^{\circ n}\tilde{\mathbf{E}}\tilde{\Lambda}^{1/2}}\right]^T$$

It follows that the energy containing eddies for the turbulence ensemble whose covariance is CALECO are

$$\mathbf{Z}_K \text{ and } \left[\overline{\mathbf{C}_K^{\circ n}\tilde{\mathbf{E}}\tilde{\Lambda}^{1/2}}\right]$$

Since columns of $\left[\overline{\mathbf{C}_K^{\circ n}\tilde{\mathbf{E}}\tilde{\Lambda}^{1/2}}\right]$ are in spectral space, for broad localization functions can truncate leaving $\left[\overline{\mathbf{C}_K^{\circ n}\tilde{\mathbf{E}}\tilde{\Lambda}^{1/2}}\right]$ only has L columns $(L < N)$.

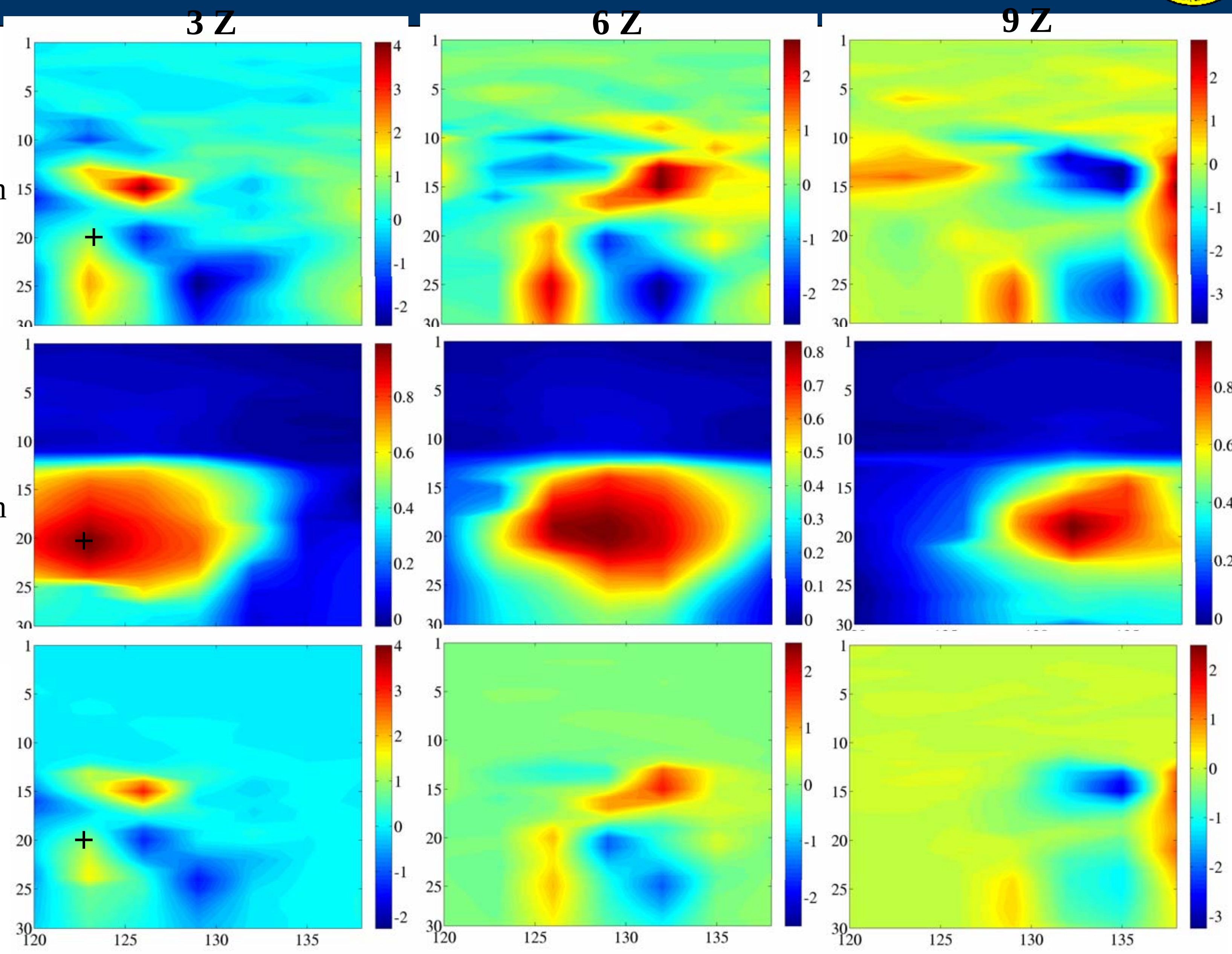
LETKF using CALECO turbulence

- When CALECO is used in LETKF size of observation volumes is unconstrained because localization is implicit in CALECO.
- Larger observation volumes are appropriate for satellite DA and 4D-DA
- Larger observation volumes enable entire grid columns (or indeed the entire globe) to be updated simultaneously and hence avoids redundancy in observation processing.
- Note that ensemble size is now given by the size of the turbulence ensemble.
- The size of the turbulence ensemble is an upper bound on the dimension of the error and is usually < number of obs.

Multi-variate or Uni-variate ECO-RAP

- ECO-RAP can provide multi-variate “localization”.
- However, in this experiment, to further increase the computational efficiency of ECO-RAP, we chose a single variable θ_e to localize u,v,T
- Future work will consider fully multivariate ECO-RAP together with alternative univariate formulations (e.g. using ϕ)

v Increment From a Single T Ob.

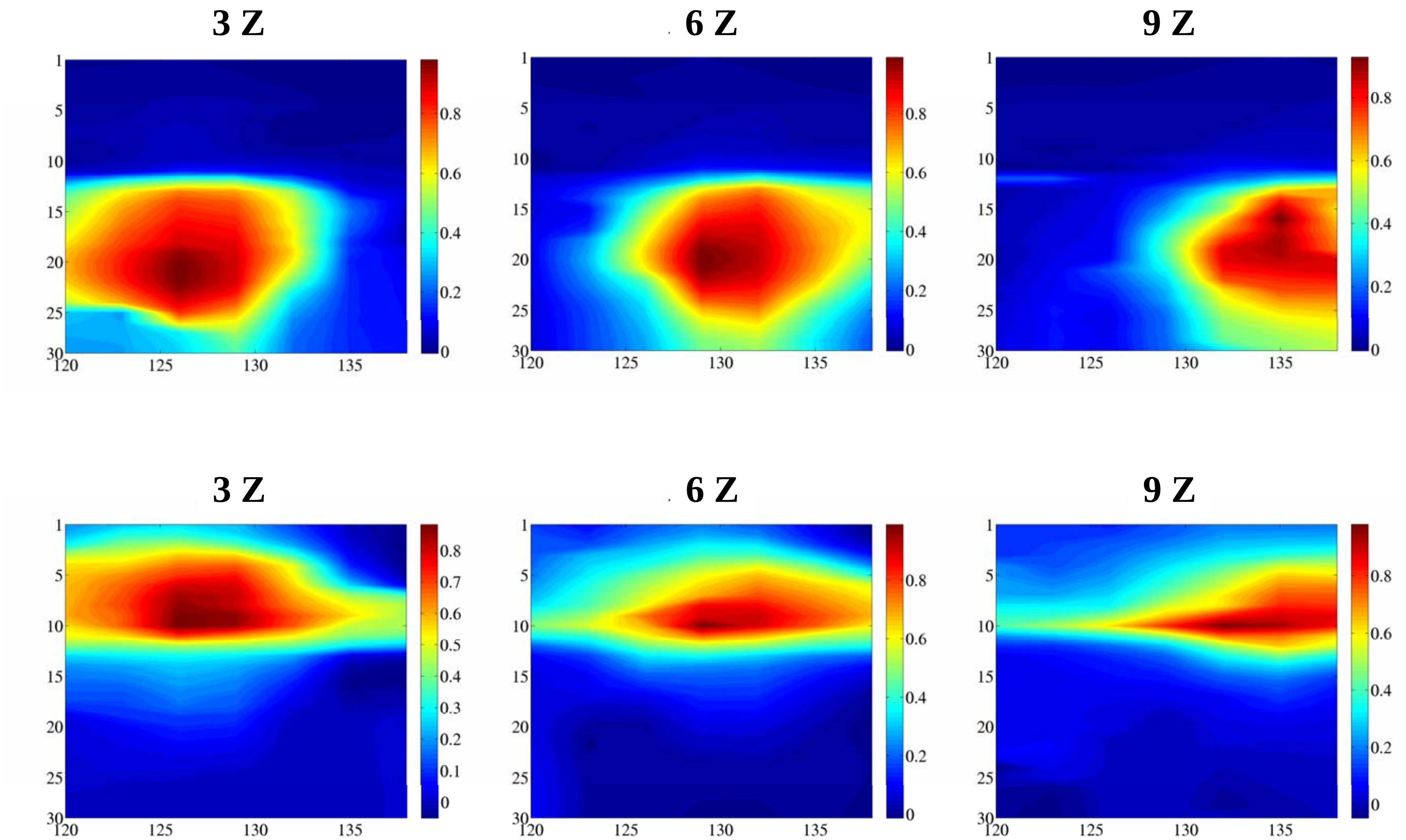


No Localization Increment

ECORAP Localization

ECORAP Increment

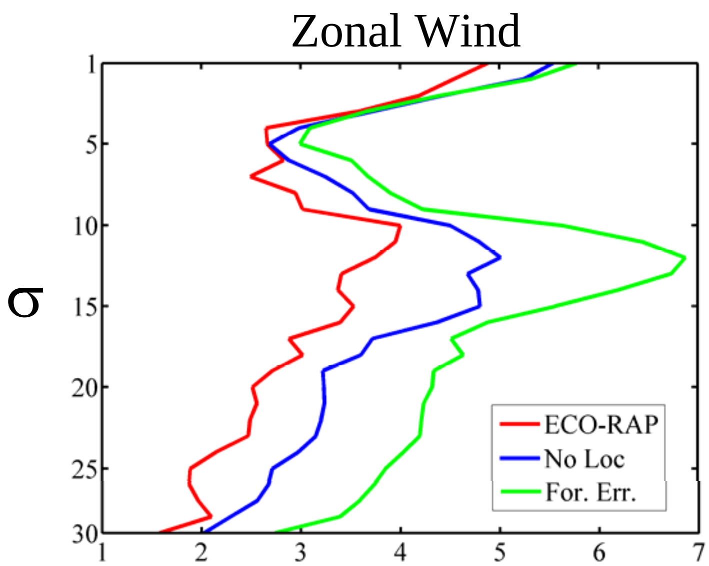
Example ECO-RAP Localization Functions

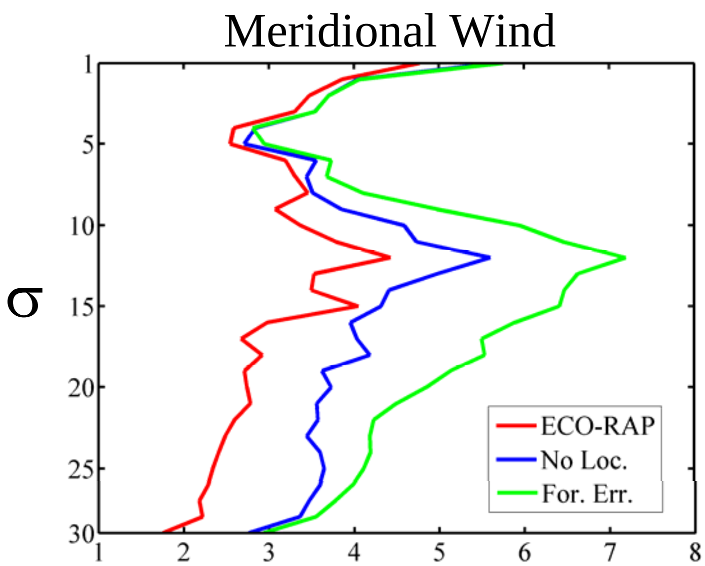


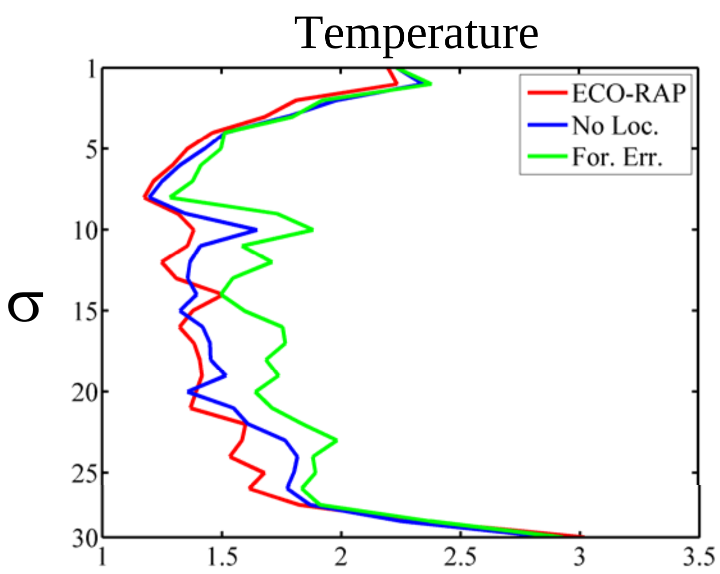
LETKF using CALECO: Preliminary Experiment

- K = 27 member ensemble, T119L30 NWP model (NOGAPS).
- 7x7x30 grid box size .
- 3^o grid resolution.
- We observe u,v,T at every point within the box at 3Z and 9Z, and attempt to estimate the state at 6Z.
- ‘Truth’ is assumed to be a 21-27 hour forecast.
- First guess/ensemble come from last 6 hrs of 9 hr forecast valid at the same time.
- Observations are the ‘truth’ plus random number
- Observation error variance is 1 m^2/s^2 and 1 K^2
- Number of obs = 8820, Number of variables=13230
- $K\_{\text{Turbulence}}$=1640
- Smoothed ensemble perturbations before applying ECO-RAP
- Correlations were raised to the 12^{th} power

Globally Averaged Results







Global RMS Error

	Forecast	ECORAP	No Loc.
u	4.6	3.1	3.8
v	4.8	3.1	3.9
T	1.8	1.6	1.7

Summary

- Ensemble localization is equivalent to running ensemble through a 1-step turbulent cascade where energy containing eddies are the raw ensemble and the columns of the square root of the localization covariance matrix.
- Turbulence analogy, separability, and spectral truncation enable computationally efficient DA algorithm – cost governed by error dimension.
- ECO-RAP allows larger observation volumes in LETKF - outperforms raw ensemble.