NOAA Earth System
Research Laboratory

What constrains spread growth in forecasts initialized from ensemble Kalman filters?

Tom Hamill & Jeff Whitaker

NOAA Earth System Research Lab
Boulder, Colorado, USA
tom.hamill@noaa.gov

a presentation at NCEP/EMC, December 2009; in review at MWR

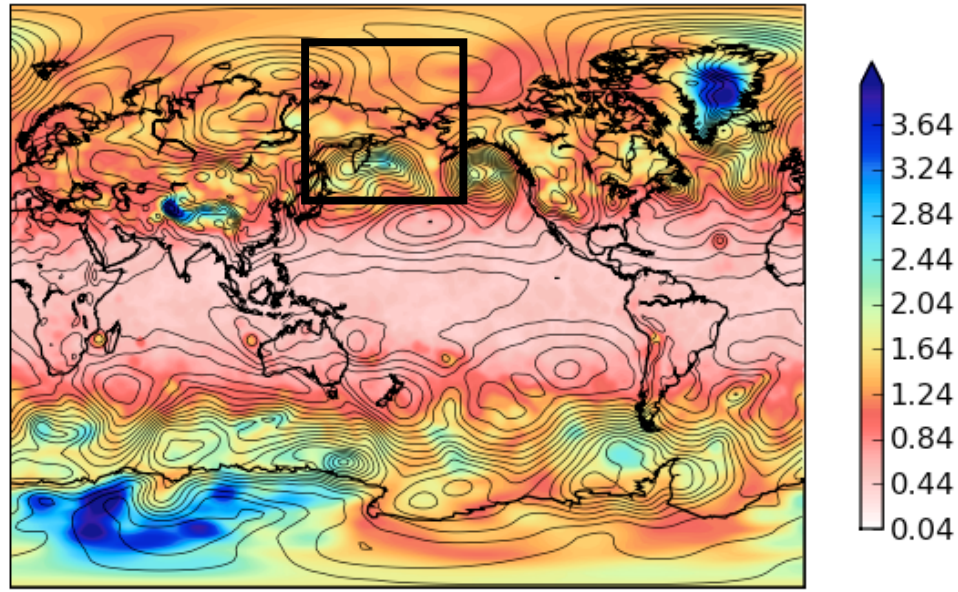
Two reasons to consider using ensemble Kalman filters

- To improve the accuracy of initial conditions .
  - >95% of research.

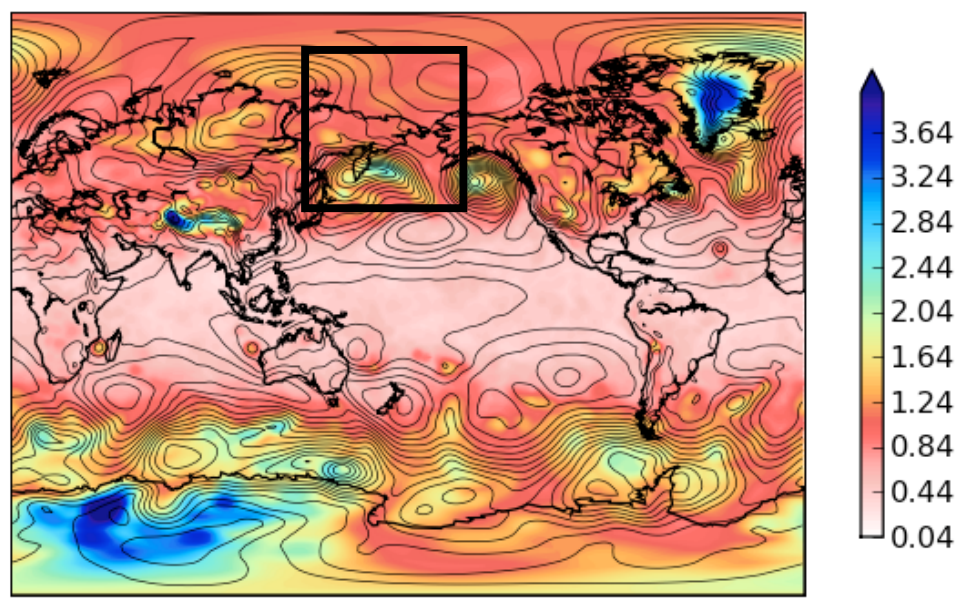
- To initialize ensembles in an optimal manner.
  - <5 % of research.

Example: lack of growth of spread in ensemble square-root filter using NCEP GFS

MSLP analysis spread, 2008-01-01 0600 UTC



First-guess spread 6 h later



Not much growth of spread in forecast, and decay in many locations. Why?

CMC has noticed problems, also

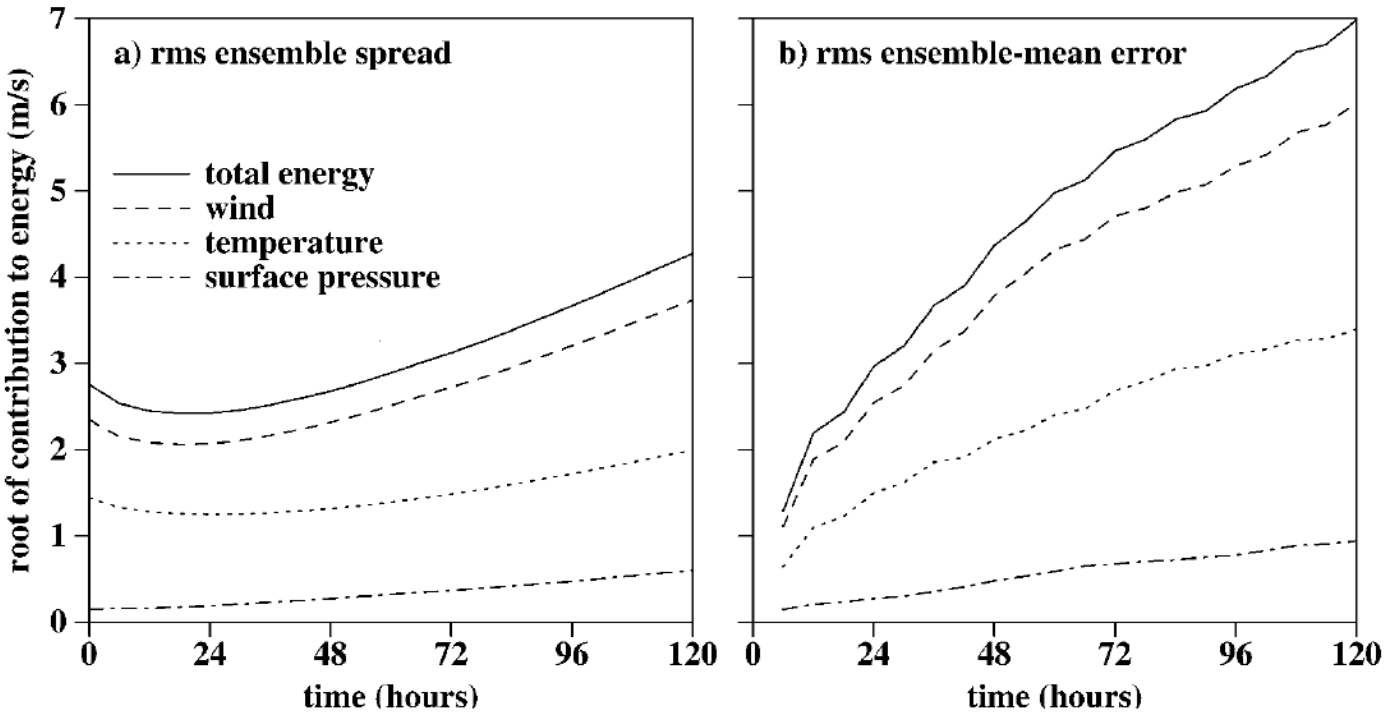


FIG. 8. The growth rate of (a) perturbations and (b) the ensemble mean error are shown in units of energy for an ensemble of sixty-four 120-h forecasts. The contributions from wind, temperature, and surface pressure are shown by the dashed, dotted, and dashed–dotted curves, respectively.

from Houtekamer et al. 2005 MWR; they attributed this to strong diffusive damping near model top, plus the additive noise they use to stabilize their EnKF.

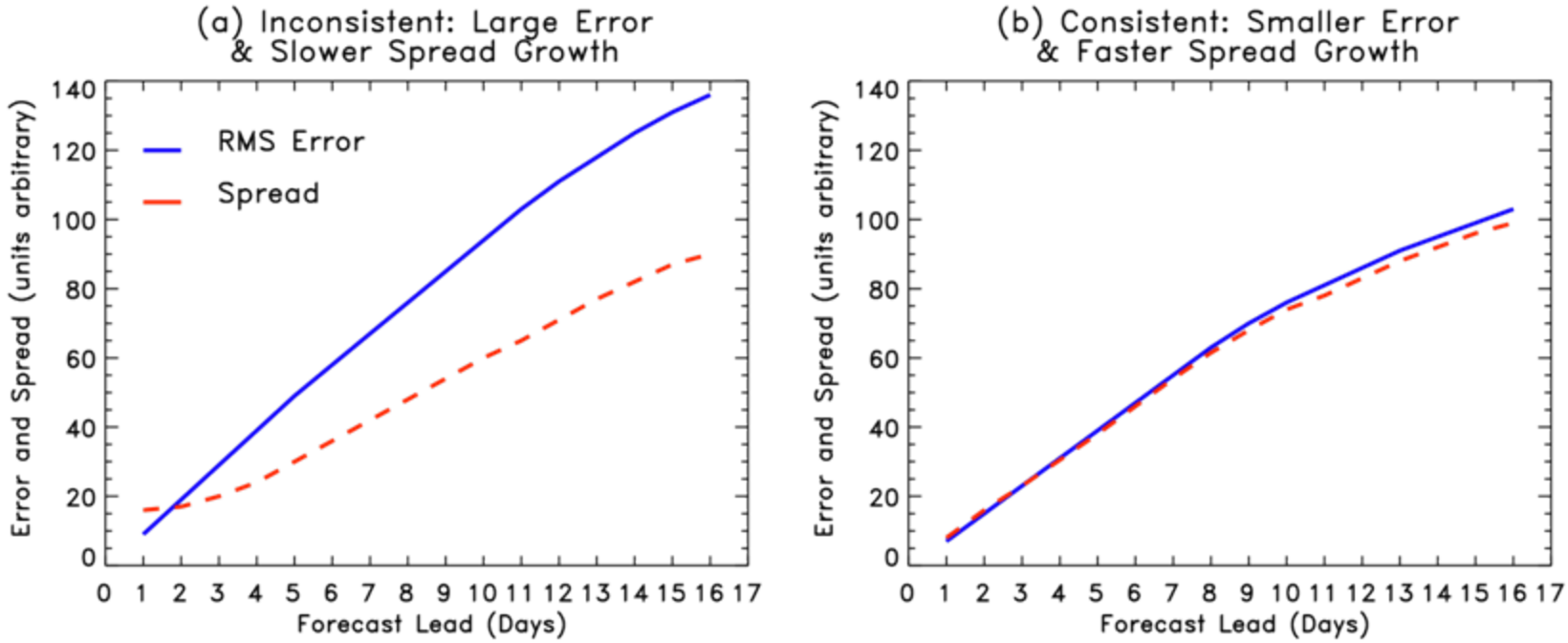
Why does this matter? "Spread-error consistency"

- Let ensemble member *X* and truth *T* be independent draws from the same distribution. Then can easily show that (Casella & Berger 1990, p. 57)

$$E\left(X - EX\right)^2 = \left(EX - T\right)^2$$

Spread^2 RMSE^2

Spread-error consistency



Spread should grow as quickly as error; part of spread growth from manner in which initial conditions are generated, some due to the model (e.g., stochastic physics, higher resolution increases spread growth). **If you don't have this consistency, your ensemble-based probability estimates will be inaccurate**.

Outline

- Review of ensemble Kalman filter and the variant we primarily use, the ensemble square-root filter (EnSRF).

- Understanding (and ameliorating) the mechanisms that limit spread growth in ensemble forecasts.

Part 1:

Review of the ensemble Kalman filter

Data assimilation terminology

- **y** : Observation vector (weather balloons, satellite radiances, etc.)
- $\mathbf{x}^b$: Background state vector ("prior")
- $\mathbf{x}^a$: Analysis state vector ("posterior")
- ***H*** : operator to convert model state → observation location & type
- **R** : Observation - error covariance matrix
- $\mathbf{P}^b$: Background - error covariance matrix
- $\mathbf{P}^a$: Analysis - error covariance matrix
- *M* : Forecast model operator
- **Q** : Model-error ("system error") covariance matrix

Canonical EnKF update equations (for time *t*)

$$\mathbf{x}_i^a = \mathbf{x}_i^b + \mathbf{K}\left(\mathbf{y}_i - H\mathbf{x}_i^b\right)$$

$$\mathbf{K} = \mathbf{P}^b H^{\mathbf{T}}\left(H\mathbf{P}^b H^{\mathbf{T}} + \mathbf{R}\right)^{-1}$$

$$\mathbf{P}^b = \mathbf{X}\mathbf{X}^{\mathbf{T}}$$

$$\mathbf{X} = \left(\mathbf{x}_1^b - \overline{\mathbf{x}^b}, \ldots, \mathbf{x}_n^b - \overline{\mathbf{x}^b}\right)$$

$$\mathbf{y}_i = \mathbf{y} + \mathbf{y}_i^{'}$$

$$\mathbf{y}_i^{'} \sim N(0, \mathbf{R})$$

Notes:

(1) An ensemble of *n* parallel data assimilation cycles is conducted, assimilating *perturbed observations* .

(2) Background-error covariances are estimated using the ensemble.

Propagation of state and error covariances in EnKF

$$\mathbf{P}^a(t) = \left\langle \left[\mathbf{x}_i^a(t) - \overline{\mathbf{x}}_i^a(t) \right] \left[\mathbf{x}_i^a(t) - \overline{\mathbf{x}}_i^a(t) \right]^T \right\rangle$$

($\mathbf{P}^a$ never explicitly formed)

$$\mathbf{x}_i^b(t+1) = M\mathbf{x}_i^a(t)$$

if forecast model is "perfect"

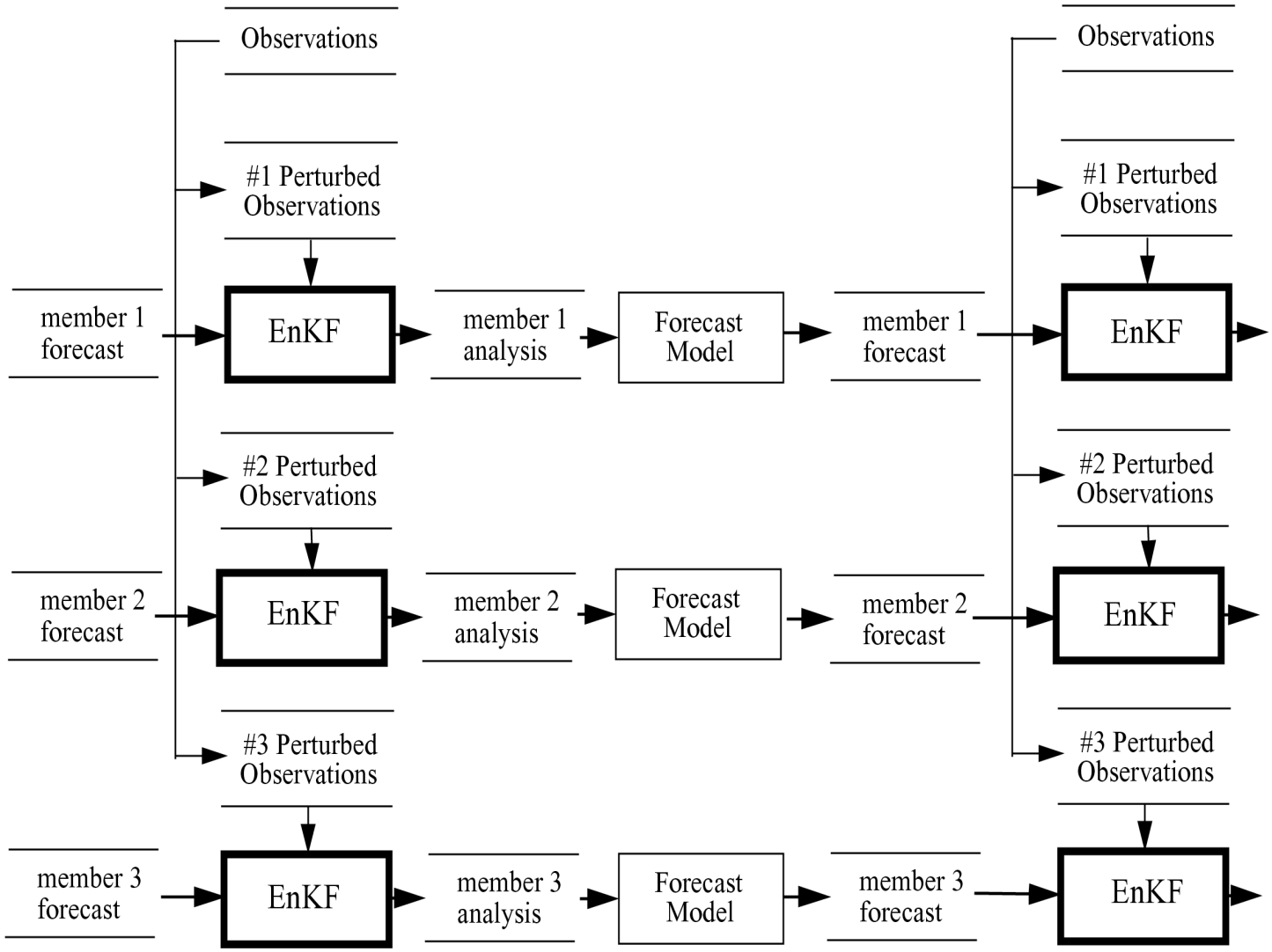
- or -

$$\mathbf{x}_i^b(t+1) = M\mathbf{x}_i^a(t) + \eta_i$$

$$\left\langle \eta_i \eta_i^T \right\rangle = \mathbf{Q}$$

…or something similar,
if forecast model imperfect.

Perfect-model EnKF schematic



(This schematic is a bit of an inappropriate simplification, for EnKF uses every member to estimate background-error covariances)

Propagation of state and error covariances in EnKF

$$\mathbf{P}^a(t) = \left\langle \left[\mathbf{x}_i^a(t) - \overline{\mathbf{x}}_i^a(t)\right]\left[\mathbf{x}_i^a(t) - \overline{\mathbf{x}}_i^a(t)\right]^T \right\rangle$$

($\mathbf{P}^a$ never explicitly formed)

$$\mathbf{x}_i^b(t+1) = M\mathbf{x}_i^a(t)$$

if forecast model is "perfect" and ensemble is very large

- or -

$$\mathbf{x}_i^b(t+1) = M\mathbf{x}_i^a(t) + \eta_i$$

$$\left\langle \eta_i \eta_i^T \right\rangle = \mathbf{Q}$$

if forecast model has model error and/or ensemble size small.

estimating this is key to EnKF performance in real-world scenarios

Two common ways of estimating system error

Covariance inflation: $\mathbf{x}_i^b \leftarrow r\left(\mathbf{x}_i^b - \overline{\mathbf{x}}_i^b\right) + \overline{\mathbf{x}}_i^b$

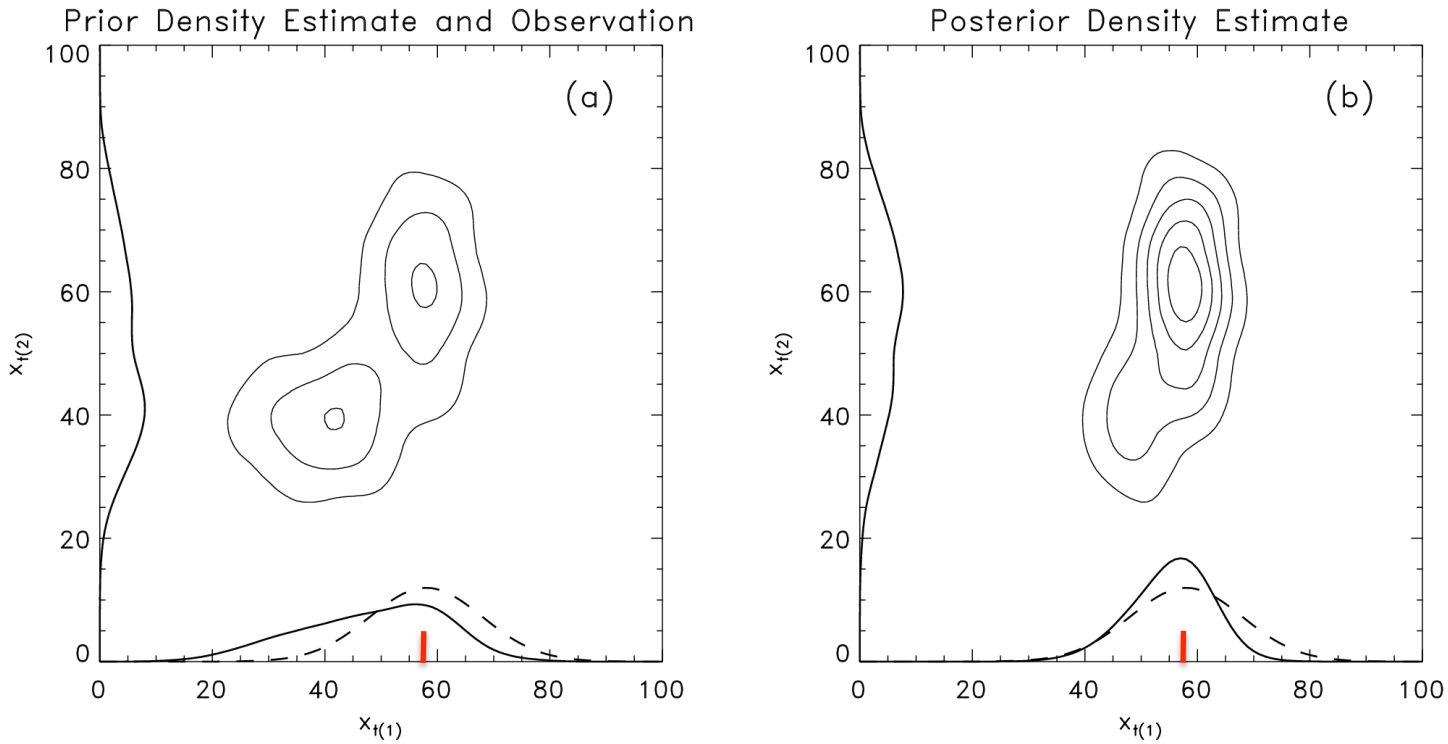
before update, pump up the spread around the ensemble mean by some factor *r > 1.0*

Additive noise: $\mathbf{x}_i^a \leftarrow \mathbf{x}_i^a + \alpha\mathbf{x}_i^n, \quad \alpha\mathbf{x}_i^n \sim N\left(0,\mathbf{Q}\right)$

add some differently structured perturbations that hopefully sample the model-error covariance **Q**

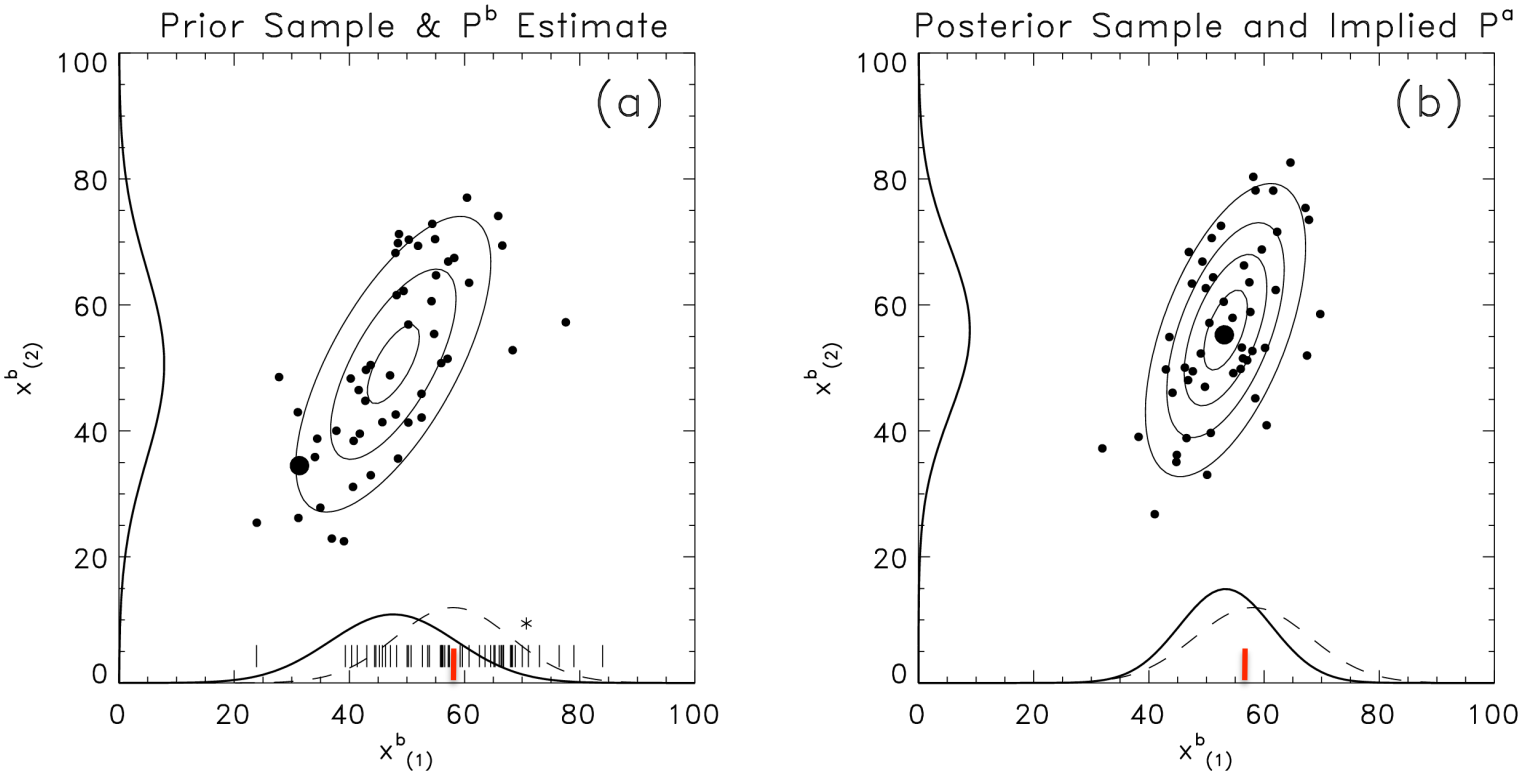
Note: this is an active area of research for us in THORPEX. There must be better ways.

Bayesian data assimilation: 2-D example as prelude to EnKF



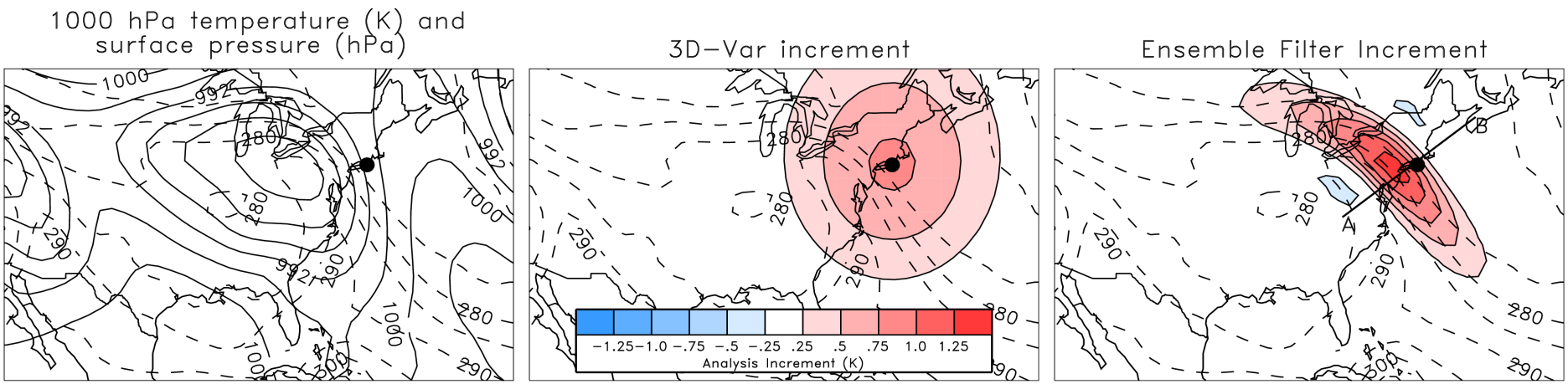
Computationally expensive when highly dimensional! Here, probabilities explicitly updated on 100x100 grid; costs multiply geometrically with the number of dimensions of model state. Also: "curse of dimensionality"

How the EnKF update works: 2-D example



Start with a random sample from bimodal distribution used in previous Bayesian data assimilation example. Contours reflect the Gaussian distribution fitted to ensemble data.

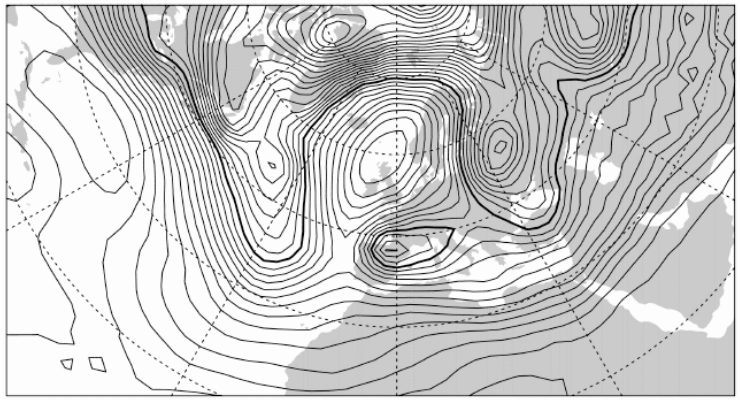
Potential advantage of EnKF: flow-dependent background-error covariances



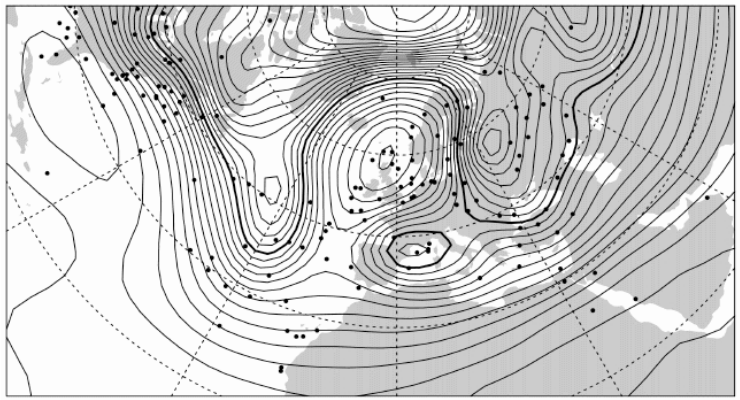
Output from a "single-observation" experiment. The EnKF is cycled for a long time. The cycle is interrupted and a single observation 1K greater than the mean prior is assimilated. Maps of the analysis minus first guess are plotted. These "analysis increments" are proportional to the background-error covariances between every other model grid point and the background at the observation location.

Flow-dependent covariances permit meteorologically reasonable adjustments to be made to non-observed fields from what's available

Full NCEP-NCAR Reanalysis (3D-Var) (120,000+ obs)



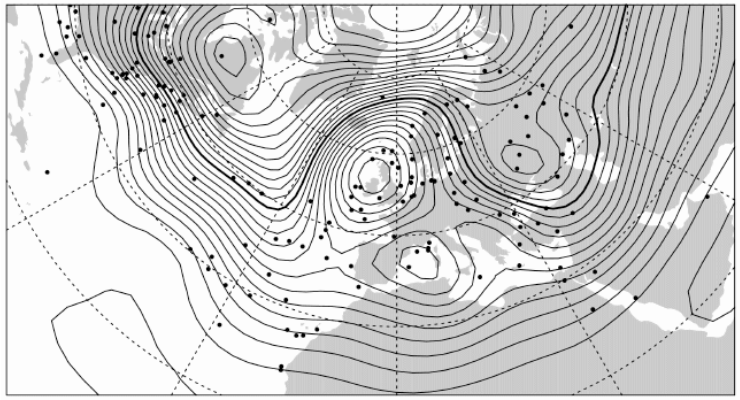
Ensemble Kalman Filter (214 surface pressure obs)



Black dots show surface pressure ob observation locations

RMS = 39.8 m

Older OI method, similar to 3D-Var (214 surface pressure obs)



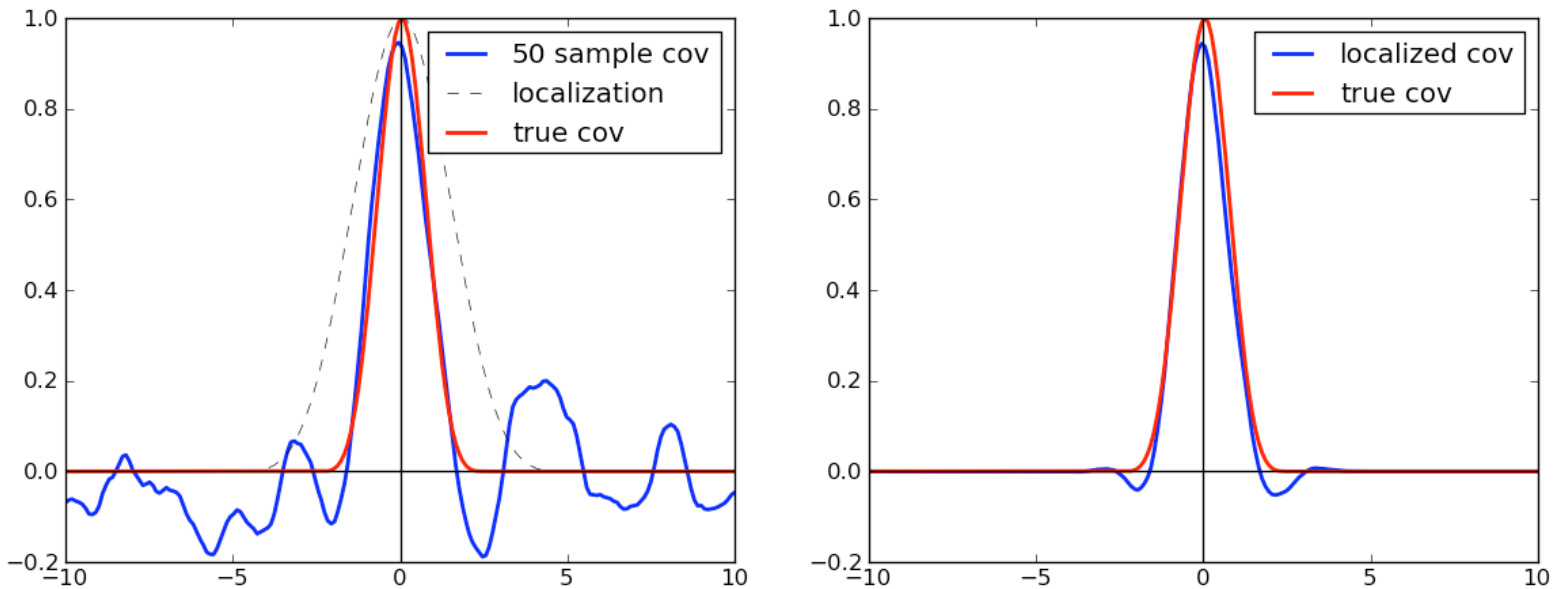
RMS = 82.4 m

fig. from Jeff Whitaker

Some of our major algorithmic modifications to basic EnKF

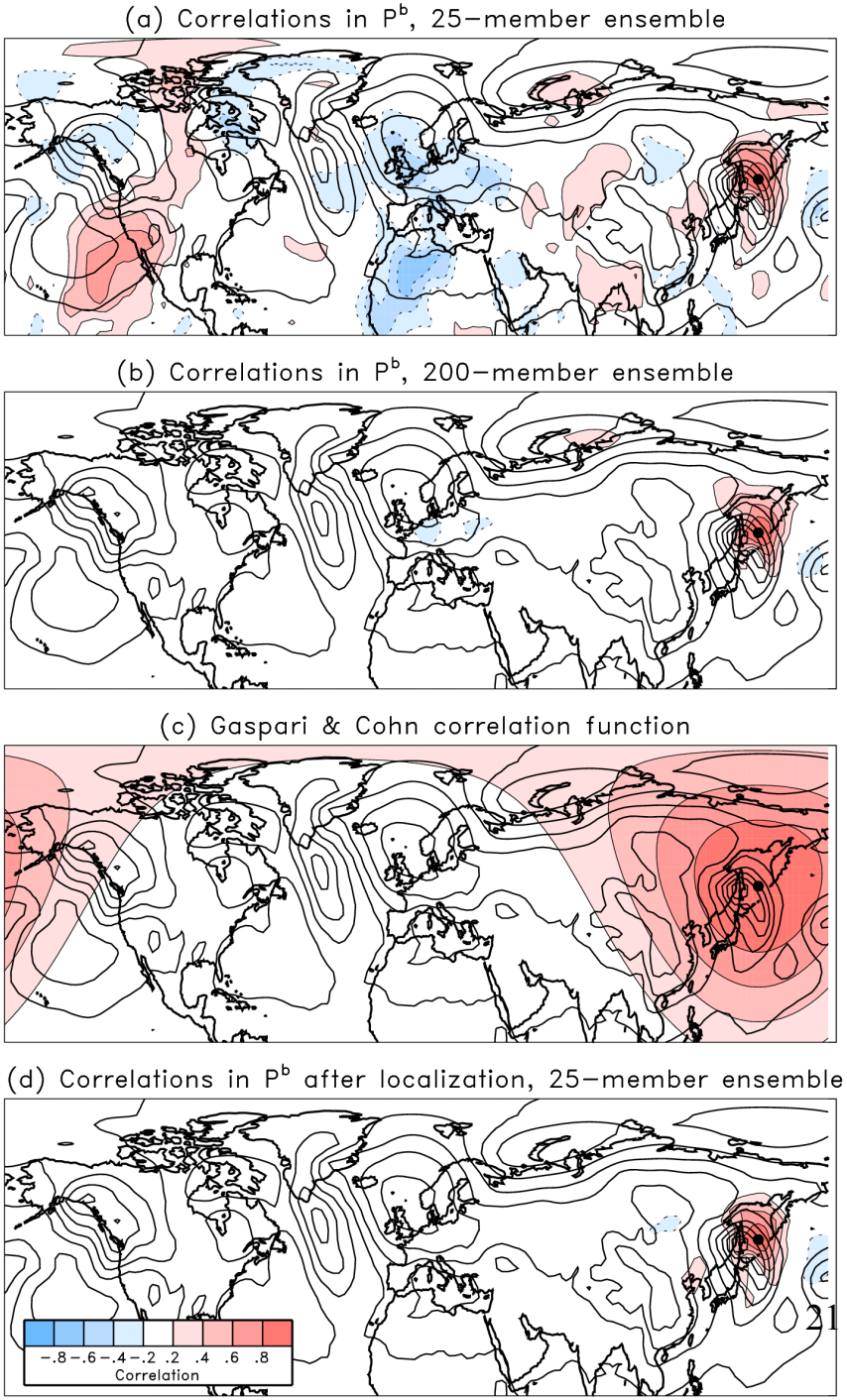
1. Covariance localization
2. Serial processing of observations
3. Simplification of Kalman-gain calculations
4. Change formulation to “ensemble square-root filter”

“Covariance localization”



Estimates of covariances from a small ensemble will be noisy, with signal-to-noise small especially when covariance is small

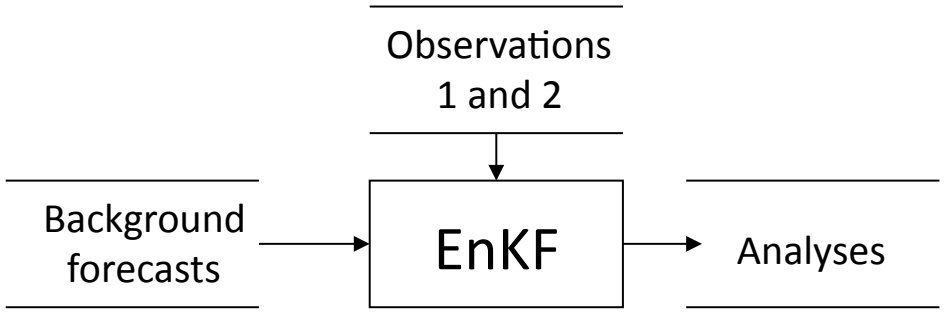
Covariance localization in practice



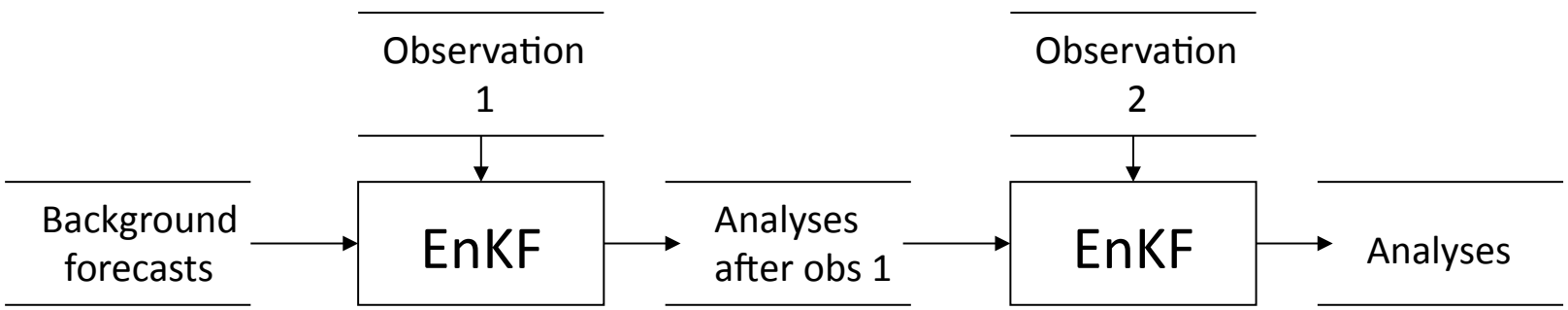
from Hamill review paper in "*Predictability of Weather and Climate*" (Cambridge Press), 2006

Serial processing of observations

Method 1



Method 2



Equivalent results, at least in absence of covariance localization

Simplifying Kalman-gain calculation

$$\mathbf{K} = \mathbf{P}^b H^{\mathrm{T}} \left(H\mathbf{P}^b H^{\mathrm{T}} + \mathbf{R} \right)^{-1}$$

$$define \quad \overline{H\mathbf{x}^b} = \frac{1}{m} \sum_{i=1}^{m} H\mathbf{x}_i^b$$

$$\mathbf{P}^b H^{\mathrm{T}} = \frac{1}{m-1} \sum_{i=1}^{m} \left(\mathbf{x}_i^b - \overline{\mathbf{x}^b} \right) \left(H\mathbf{x}_i^b - \overline{H\mathbf{x}^b} \right)^{\mathrm{T}}$$

$$H\mathbf{P}^b H^{\mathrm{T}} = \frac{1}{m-1} \sum_{i=1}^{m} \left(H\mathbf{x}_i^b - \overline{H\mathbf{x}^b} \right) \left(H\mathbf{x}_i^b - \overline{H\mathbf{x}^b} \right)^{\mathrm{T}}$$

The key here is that the huge matrix $\mathbf{P}^b$ is never explicitly formed

Different implementations of ensemble filters

- Double EnKF (Houtekamer and Mitchell, *MWR*, March 1998); more recently Quad EnKF.
- Ensemble adjustment filter (EnAF; Anderson, *MWR*, Dec 2001)
- Ensemble square-root filter (EnSRF; Whitaker and Hamill, *MWR*, July 2002)
- Ensemble transform Kalman filter (ETKF; Bishop et al, *MWR*, March 2001)
- Local ETKF (Hunt et al., *Physica D*, 2007)
- Others as well (Lermusiaux, Pham, Keppenne, Heemink, etc.)

Serial ensemble square-root filter (EnSRF); different update

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^b + \mathbf{K}\left(\mathbf{y}^o - \mathbf{H}\overline{\mathbf{x}}^b\right)$$

$$\mathbf{K} = \mathbf{P}^b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}\right)^{-1}$$

$$\mathbf{x}_i^{a'} = \mathbf{x}_i^{b'} - \tilde{\mathbf{K}}\mathbf{H}\mathbf{x}_i^{b'}$$

$$\tilde{\mathbf{K}} = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}}}\right)^{-1}$$

No perturbed obs; instead, updates to the mean and perturbations around the mean are handled separately, with "reduced" Kalman gain $\tilde{\mathbf{K}}$ used for perturbations. Rationale in Whitaker and Hamill, 2002 MWR

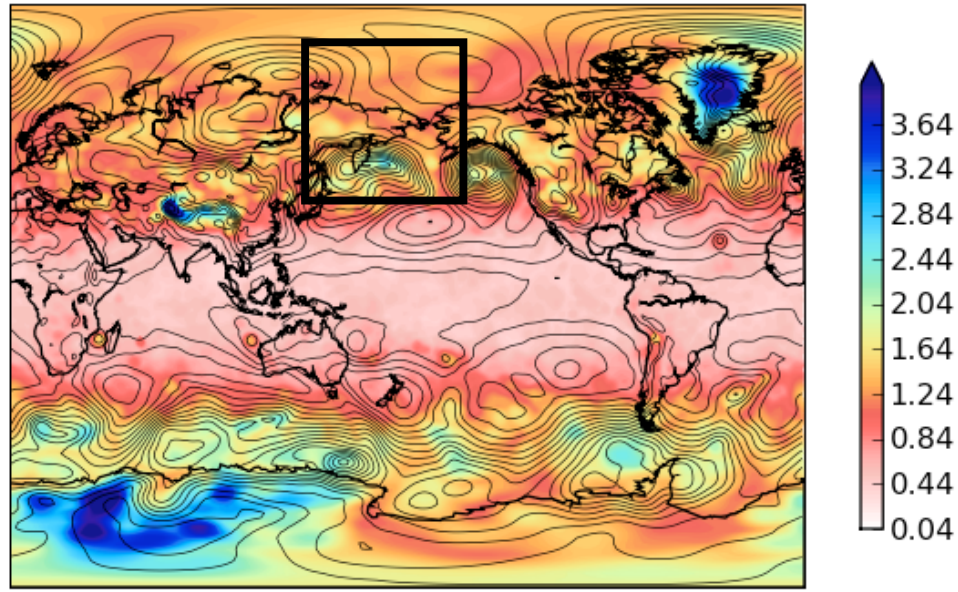
Part 2:

Understanding the mechanisms that limit spread growth in ensemble Kalman filters

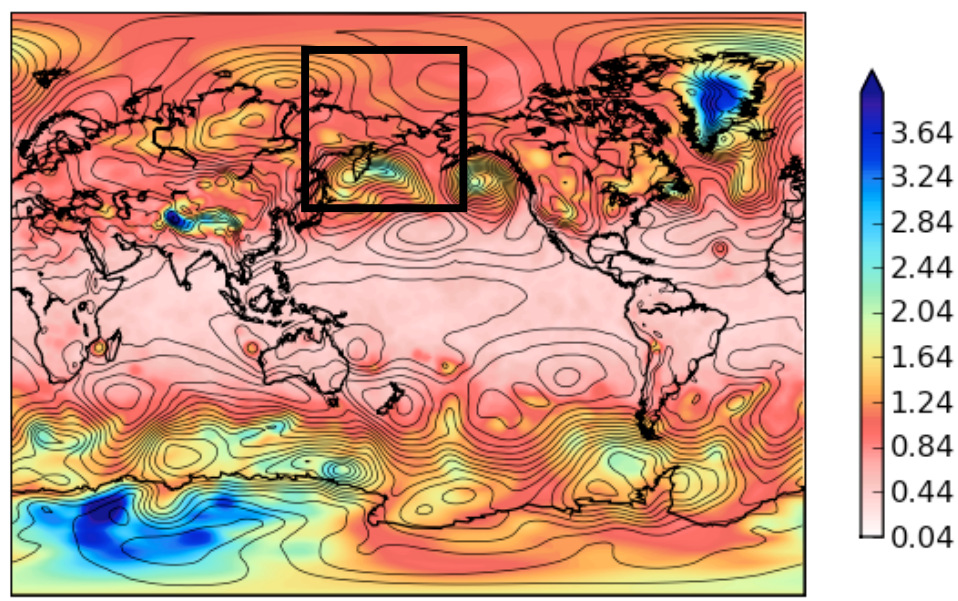
(and a proposed remedy)

Example: lack of growth of spread in ensemble square-root filter using NCEP GFS

MSLP analysis spread, 2008-01-01 0600 UTC



First-guess spread 6 h later



Not much growth of spread in forecast, and decay in many locations. Why?

Mechanisms that may limit spread growth from ensemble-filter ICs

- Covariance localization introduces imbalances.
- Additive noise used to treat system error in EnKF projects onto non-growing structures.
- Model attractor different from nature's attractor; assimilation kicks model from own attractor, transient adjustment process.
- Assumption that observation errors are independent when they are spatially correlated introduces unrealistic, small-scale increments, requiring adjustment.
- Non EnKF issues, such as neglect or improper treatment of model-related uncertainties.

(we'll consider only the first three)

Covariance localization & imbalance

$$\mathbf{P}^b = \begin{bmatrix} \sigma^2(u_1) & & Cov(u_1,u_n) & & Cov(u_1,t_1) & & Cov(u_1,t_n) \\ & \cdots & & & & & \\ Cov(u_1,u_n) & & \sigma^2(u_n) & & Cov(u_n,t_1) & & Cov(u_n,t_n) \\ & & & \cdots & & & \\ Cov(u_1,t_1) & & Cov(u_n,t_1) & & \sigma^2(t_1) & & Cov(t_1,t_n) \\ & & & & & \cdots & \\ Cov(u_1,t_n) & & Cov(u_n,t_n) & & Cov(t_1,t_n) & & \sigma^2(t_n) \end{bmatrix}$$

envision a covariance matrix, here with winds and temperatures at *n* grid points

$$\rho = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

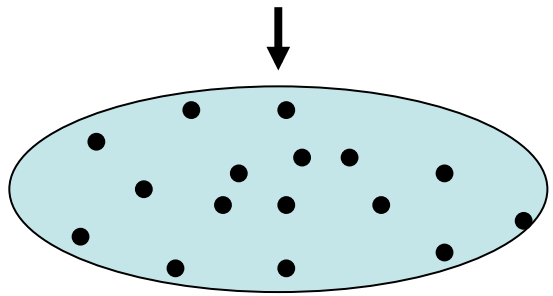
envision a covariance localization at its most extreme, a Dirac delta function, i.e., the identity matrix.

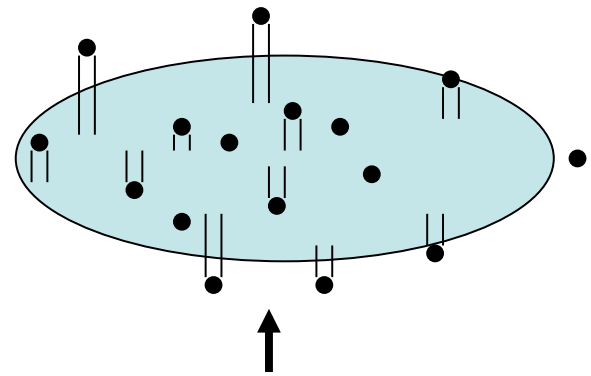
$$\rho \circ \mathbf{P}^b = \begin{bmatrix} \sigma^2(u_1) & & 0 & & 0 & & 0 \\ & \cdots & & & & & \\ 0 & & \sigma^2(u_n) & & 0 & & 0 \\ & & & \cdots & & & \\ 0 & & 0 & & \sigma^2(t_1) & & 0 \\ & & & & & \cdots & \\ 0 & & 0 & & 0 & & \sigma^2(t_n) \end{bmatrix}$$

The localized covariance matrix has totally decoupled any initial balances between winds and temperature

Additive noise

Before additive noise:
ensembles may tend to lie on lower-dimensional attractor

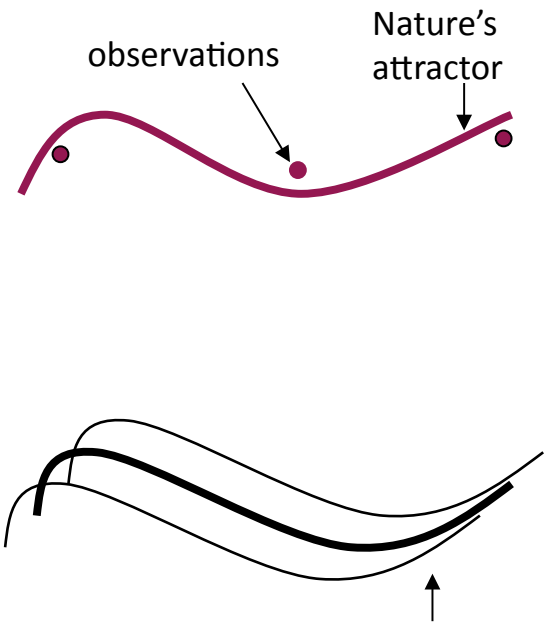




After additive noise:
some of the noise added takes model states off attractor; resulting transient adjustment & spread decay

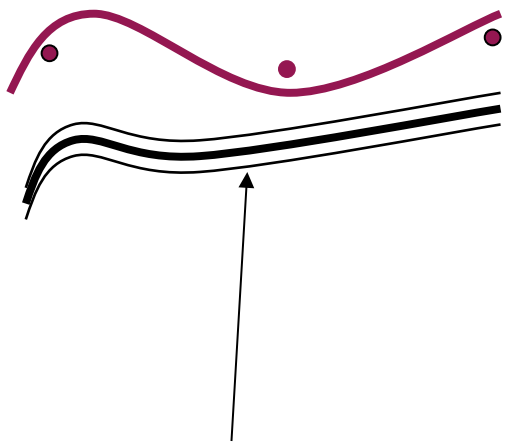
Model error

before data assimilation



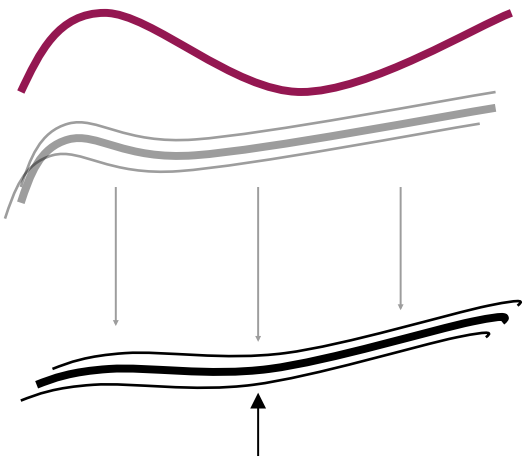
forecast mean background and ensemble members, ~ on model attractor

after data assimilation



analyzed state, drawn toward obs; ensemble (with smaller spread) off model attractor

after short-range forecasts



forecast states snap back toward model attractor; perturbations between ensemble members fail to grow.

Testing the effects on spread growth of covariance localization, additive noise, & model error

- Apply EnSRF in 2-level primitive equation "toy" model.
  - Perfect-model experiments
    - Determining maximum growth rate? Use very large ensemble, no localization.
    - Determine how much covariance localization, additive noise reduce growth rate of perturbations in moderately (n=50) sized ensemble.
  - Imperfect-model experiments
    - How do results differ from perfect-model? Similar conclusions for relative effect of localization, additive noise?
    - Apply corrective effect for additive noise (discussed later)
    - Diagnose the relative contribution of model error from how much dimunition of spread growth exists even after corrective effect applied.
- Test corrective effect in full T62 NCEP GFS EnSRF, real observations.

Assimilation & toy-model details

- Assimilation:
  - EnSRF of Whitaker and Hamill (2002) *MWR*. 50 members unless otherwise specified.
  - Ensemble forecasts at T31 resolution.
  - Observations: *u*,*v* at 2 levels every 12 h, plus potential temperature at 490 ~ equally spaced locations on geodesic grid. 1.0 m/s and 1.0 K observation errors σ.
- Model: 2-level GCM following Lee and Held (1993) *JAS*
  - State: vorticity at two levels, baroclinic divergence, barotropic potential temperature.
  - Forced by relaxation to radiative equilibrium state with pole-to-equator temperature difference of 80K, with 20-day timescale.
  - Lower-level winds damped at 4-day timescale.
  - ∇^8 diffusion, smallest resolvable scale damped with 3-h timescale.
  - T31 error-doubling time of 2.4 days
  - For imperfect model experiments, T42, with 74K pole-to-equator temperature difference, wind damping timescale of 4.5 days

Definitions

- **Covariance inflation**: $\mathbf{x}_i^b \leftarrow r\left(\mathbf{x}_i^b - \overline{\mathbf{x}}_i^b\right) + \overline{\mathbf{x}}_i^b$

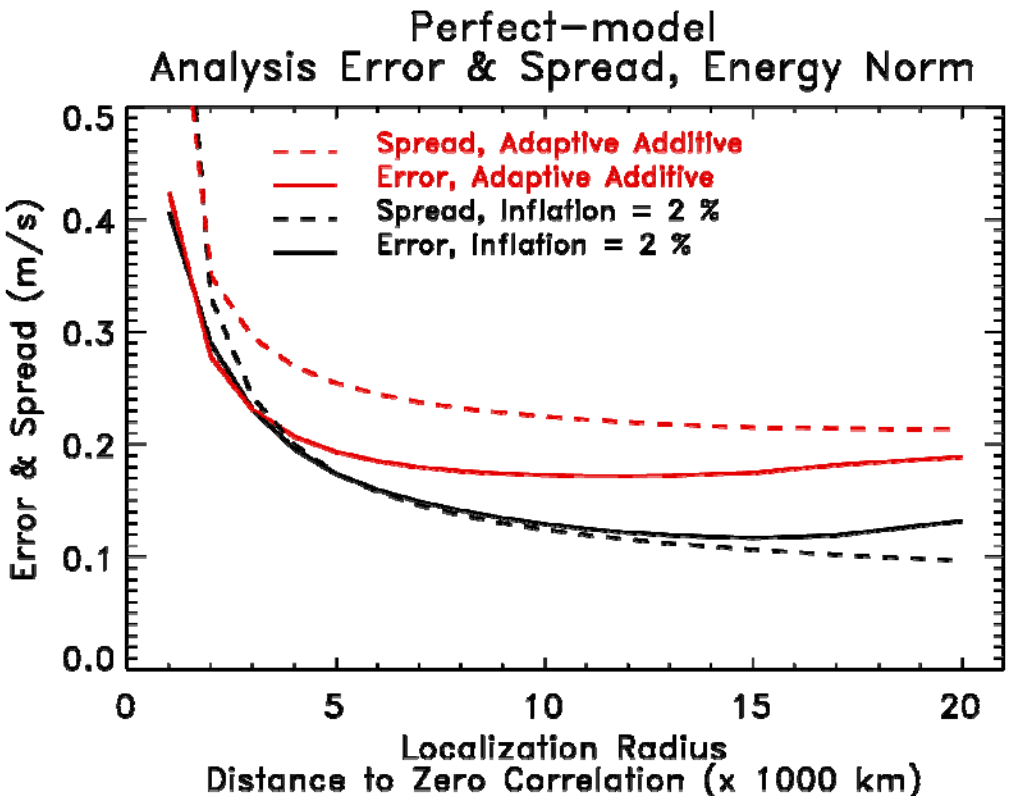
- **Additive noise**: $\mathbf{x}_i^a \leftarrow \mathbf{x}_i^a + \alpha\mathbf{x}_i^n, \quad \alpha\mathbf{x}_i^n \sim N\left(0, \mathbf{Q}\right)$

  - noise added after analyses, not prior to them.
  - 0-24h tendencies are used to generate $\mathbf{x}_i^n$ for perfect-model experiments; zero mean enforced.
  - Random samples of model states using perturbed models for imperfect-model experiments. Again, zero mean enforced.

- **Energy norm**:

$$\|\cdot\| = \frac{\sqrt{\frac{1}{2}\int_A \left[u^2 + v^2 + \frac{c_p}{T_{ref}} T^2\right] dA}}{\int_A dA}$$

Error/spread as functions of localization length scale, T31 perfect model



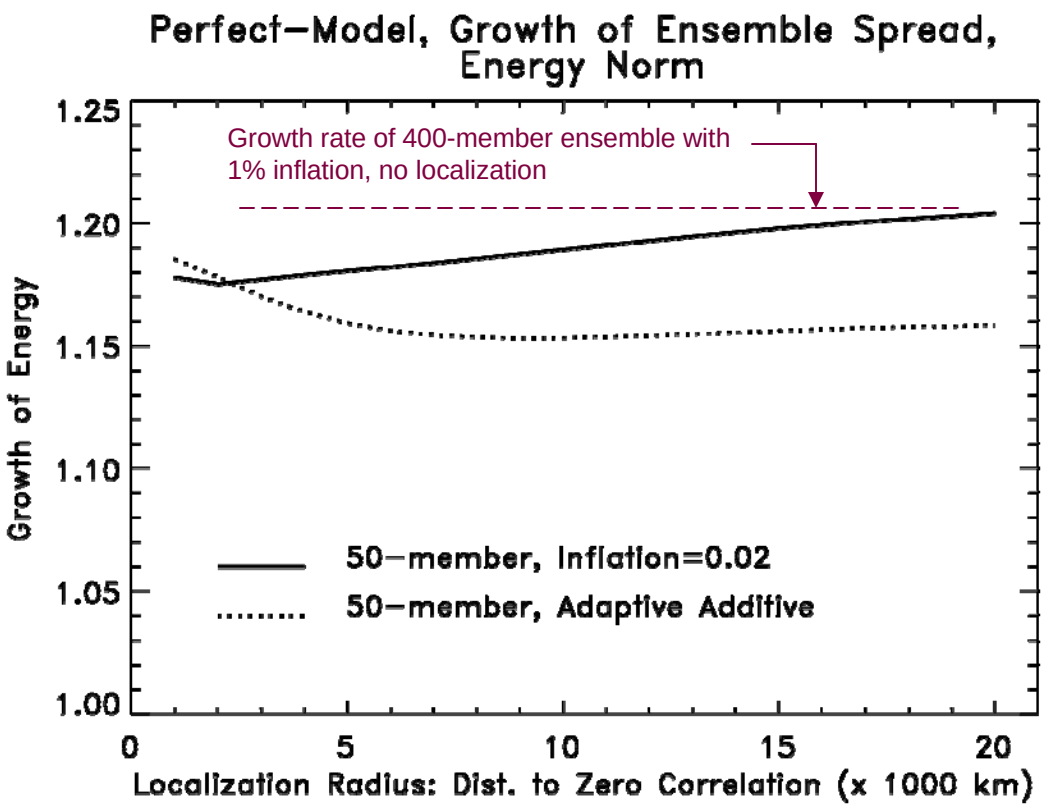
Proper data assimilation provides this:

Given some additive noise vector ς_i, adaptive additive finds α such that

at *observation locations* (whereas analysis errors & spread computed over globe).

Bottom line on errors: for *perfect-model simulation*, covariance inflation is more accurate; deleterious effect of additive random noise.

How does spread **growth** change due to localization? (perfect model)



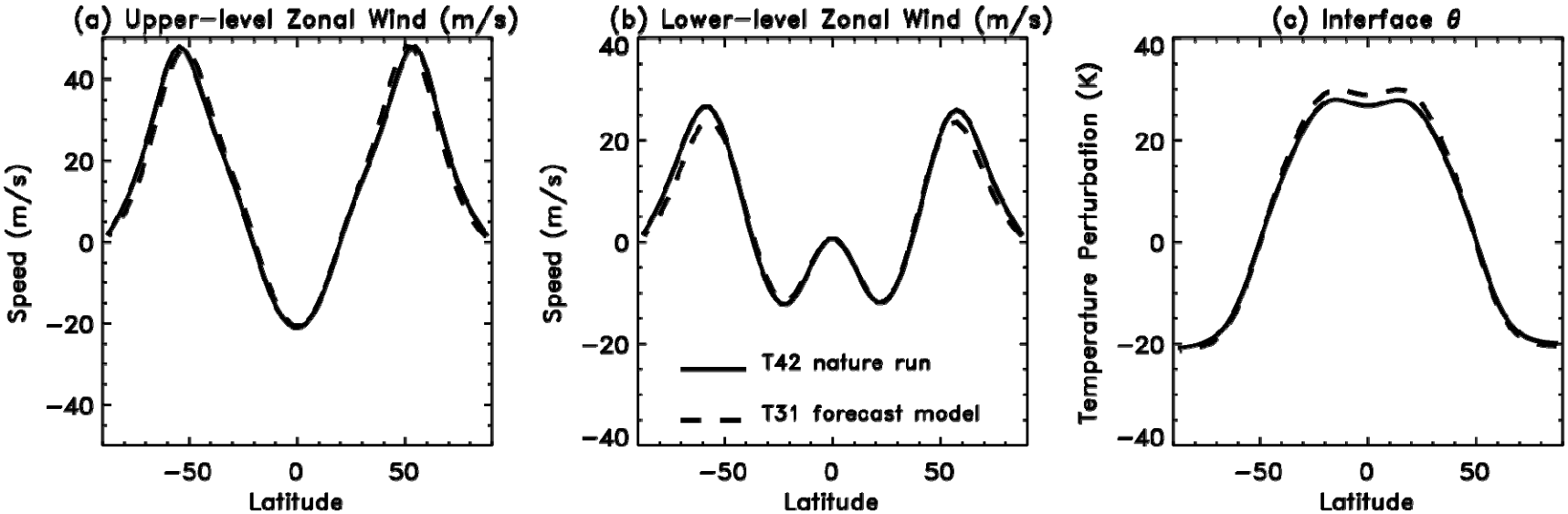
Notes:

(1) Growth rate of 50-member ensemble over 12-h period with large localization radius is close to "optimal"

(2) Increasing the localization radius with constant inflation factor has relatively minor effect on growth of spread. Suggests that in this model, covariance localization is secondary factor in limiting spread growth.

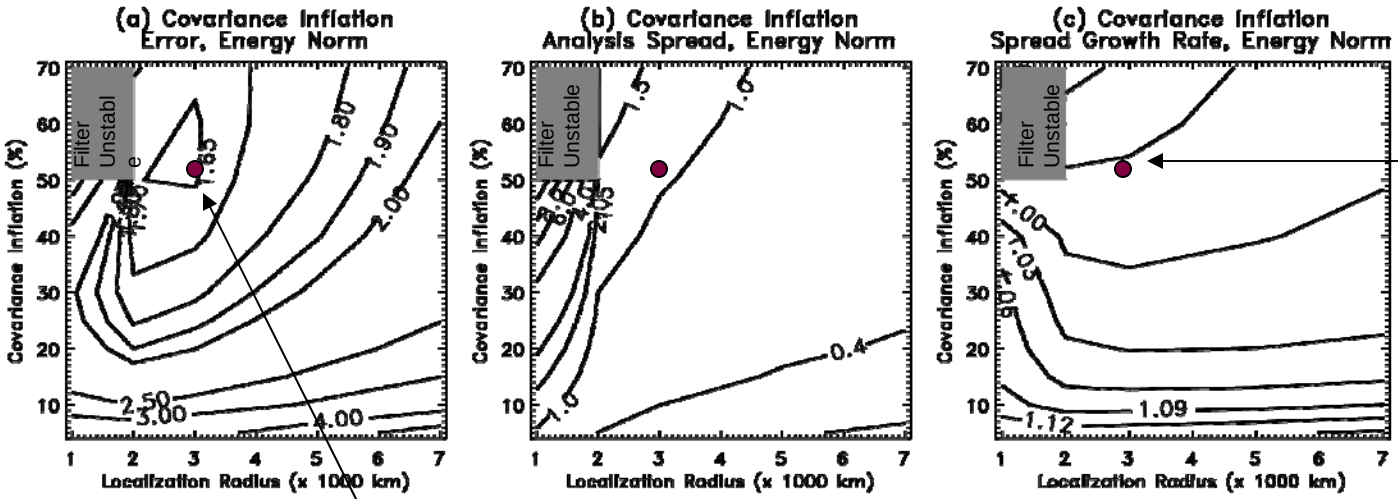
(3) Additive noise reduces spread growth somewhat more than does localization. Adaptive algorithm added virtually no additive noise at small localization radii, then more and more as localization radius increased. Hence, adaptive additive spread doesn't grow as much as localization radius increases because the diminishing imbalances from localization are offset by increasing imbalances from more additive noise.

Imperfect-model results:
nature run & imperfect model climatologies



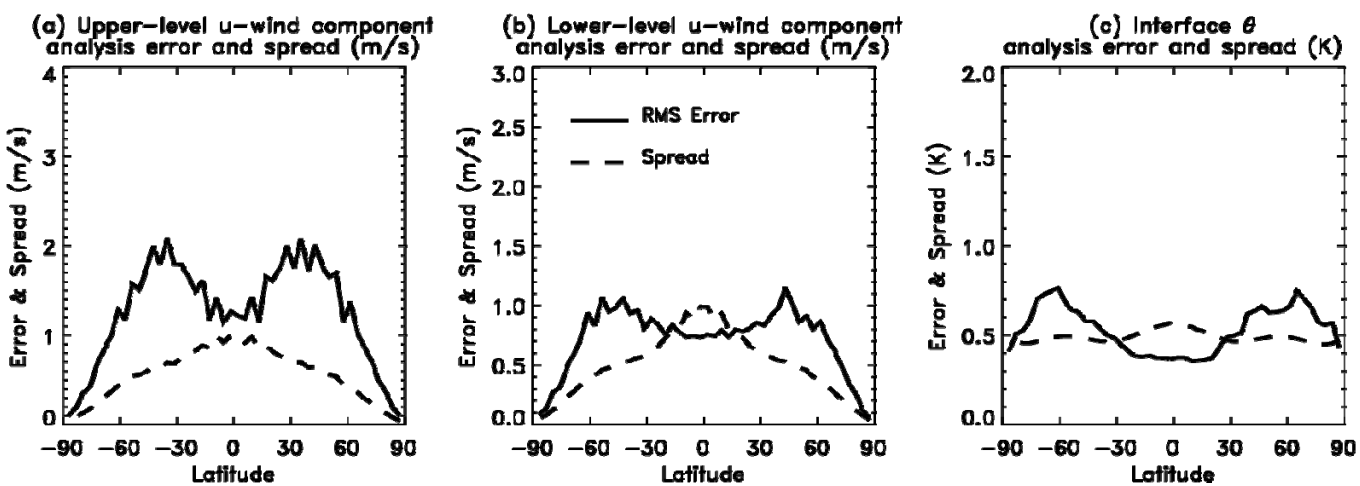
- 6 K less difference in pole-to-equator temperature difference in T42 nature run
- Less surface drag in T42 nature run results in more barotropic jet structure.

Covariance inflation, imperfect model



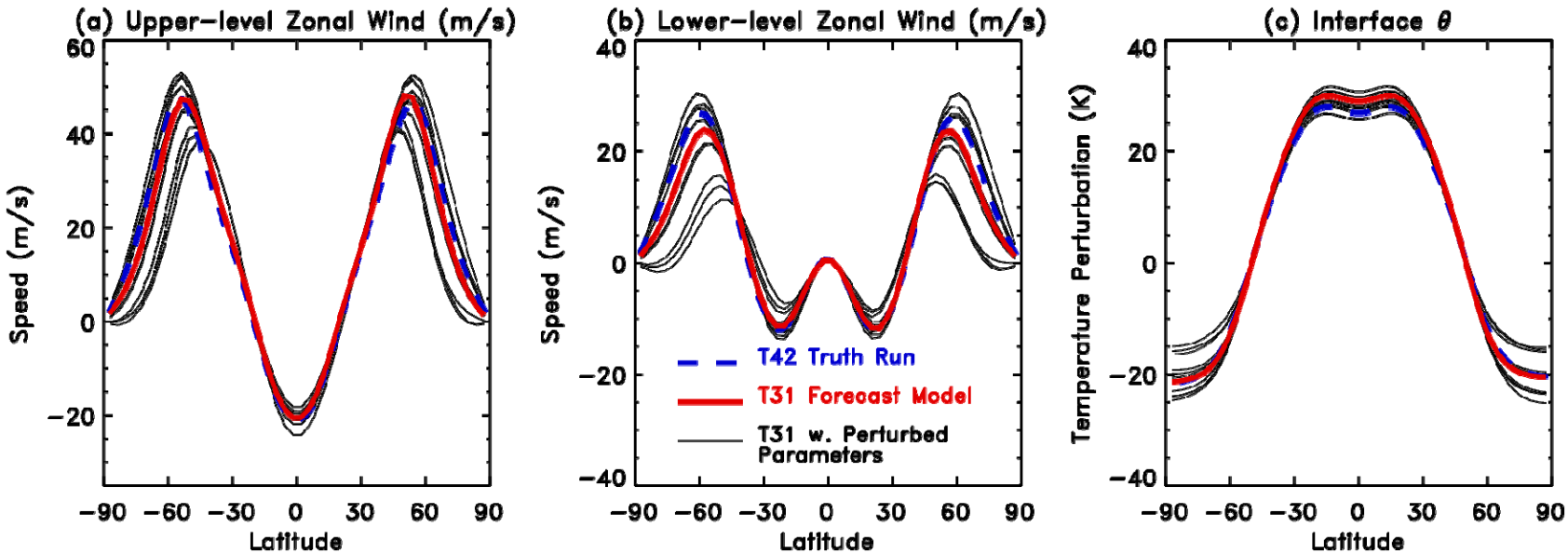
Spread decays in region of parameter space where analysis error is near its minimum.

3000 km localization 50 % inflation



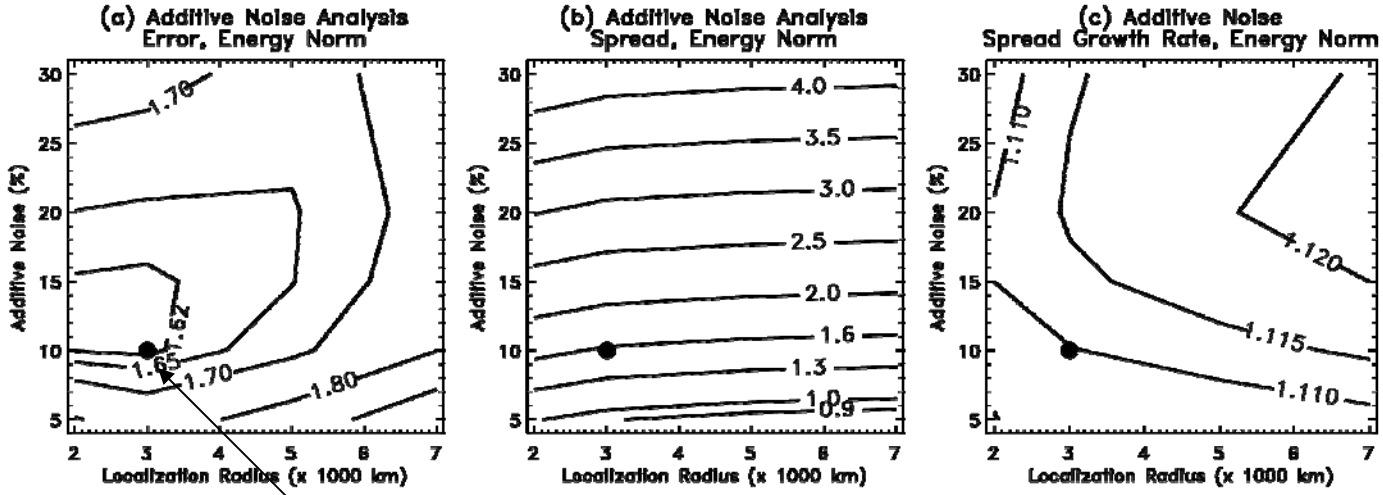
Differential growth rates of model error result in difficulties in tuning a globally constant inflation factor (see also Hamill and Whitaker, *MWR*, November 2005)

Model error additive noise zonal structure



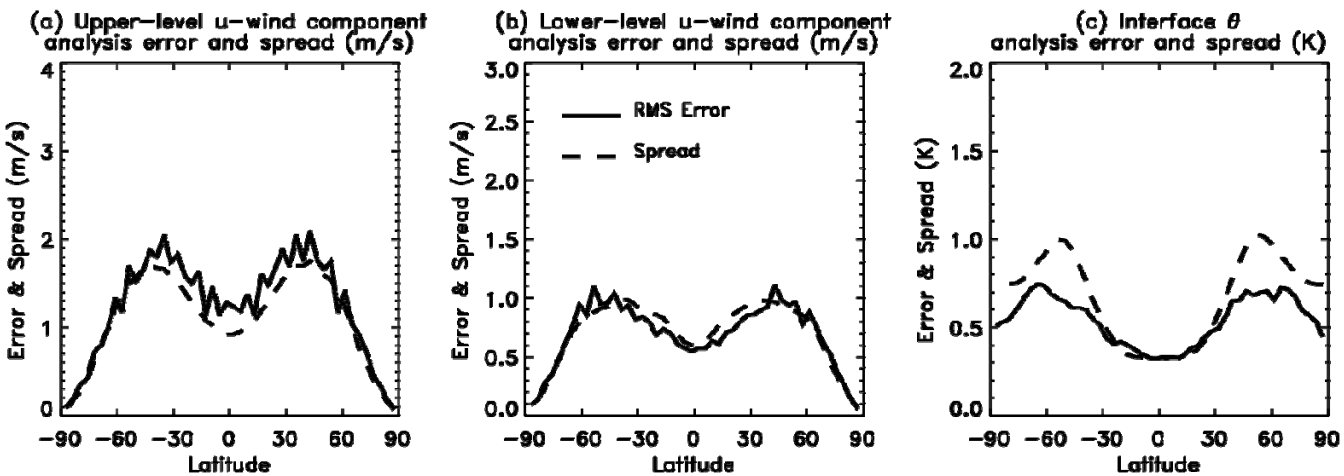
- Plots show the zonal-mean states of the various perturbed model integrations that were used to generate the additive noise for the imperfect-model simulations.
- Additive noise for imperfect model simulations consisted of 50 random samples from nature runs from perturbed models; zero-mean perturbation enforced. 0-24 h tendencies as with perfect model did not work well given substantial model error.

Additive noise, imperfect model



Spread growth is smaller than in perfect-model experiments, but is ~ constant over the parameter space. Decrease in spread growth should be attributable largely to imperfect vs. perfect model.

3000 km localization, 10% additive

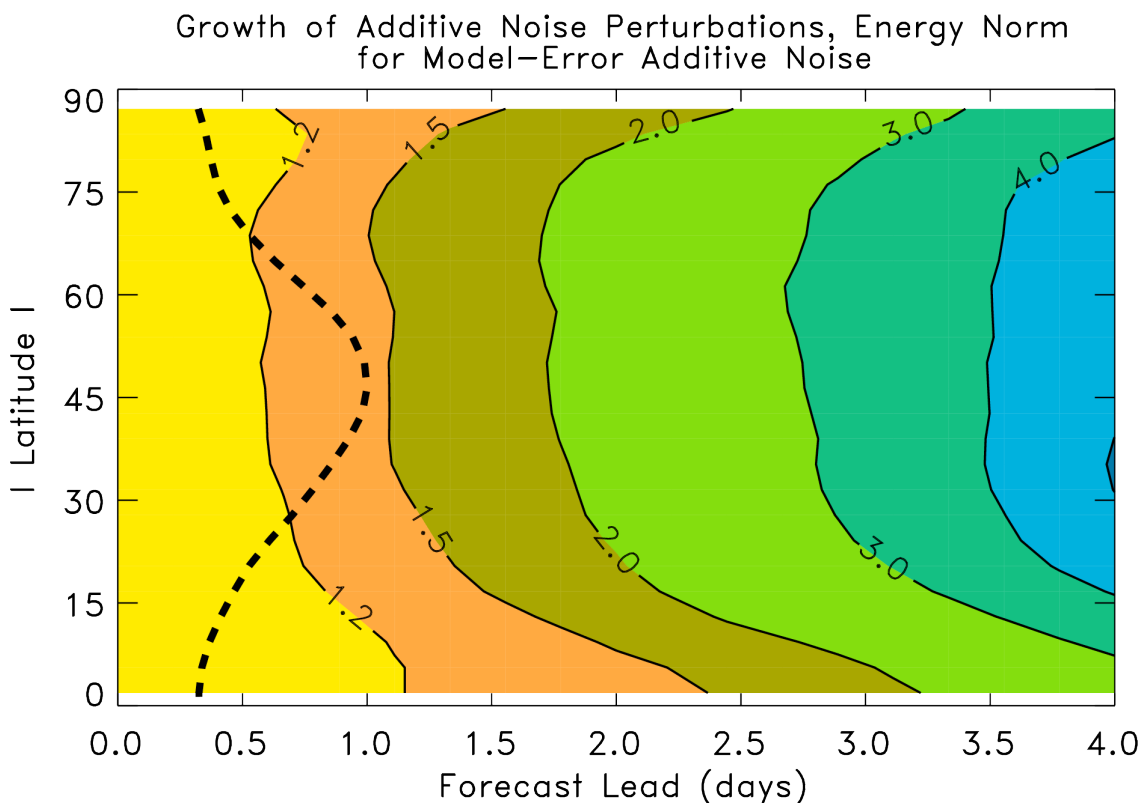


There is more consistency in spread and error than with the covariance inflation.

Synthesis (model-dependent result) rate of spread growth "g"

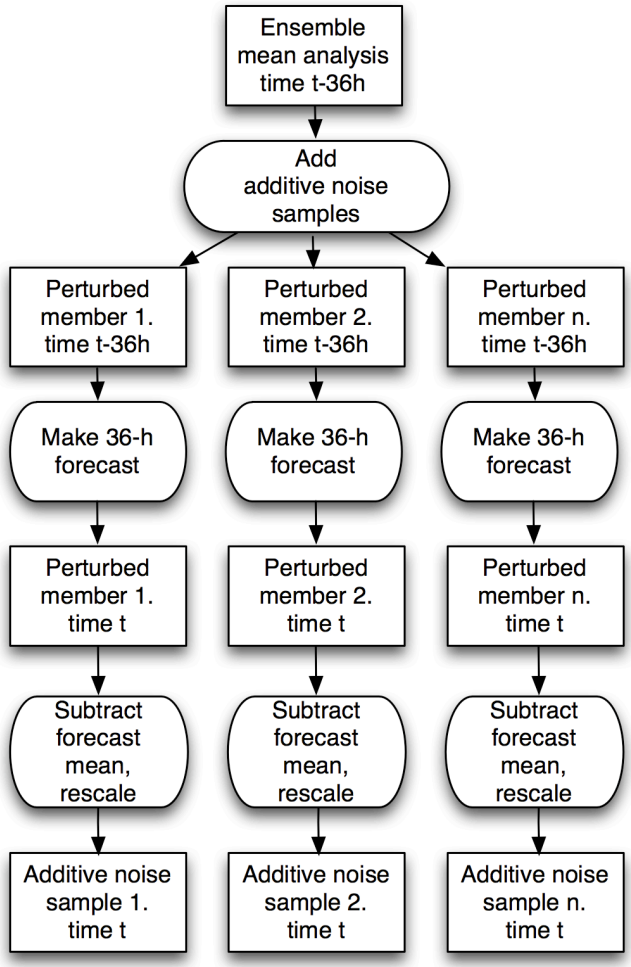
- Perfect model, 400 members, covariance inflation, no localization: g=1.2
- Perfect model: 50 members,
  - Covariance inflation + localization: g = 1.175 to 1.2 ; virtually no loss of potential spread growth lost due to use of covariance localization with large radii.
  - Additive noise + localization: g = 1.15; noise reduces spread by ~5 percent, introduces perturbations that don't project as highly onto growing forecast structures.
- Imperfect model, 50 members:
  - Globally constant covariance inflation doesn't work properly.
  - Additive noise (type of noise changed relative to perfect-model experiment) g = 1.11, and tighter localization needed.
- Implications:
  - Perfect vs. imperfect: the better the forecast model fits the observations, the less spread growth should be a problem in ensemble filters, for the less additive noise.
  - We need additive noise that have growing structures.

Average growth of additive noise perturbations around nature run

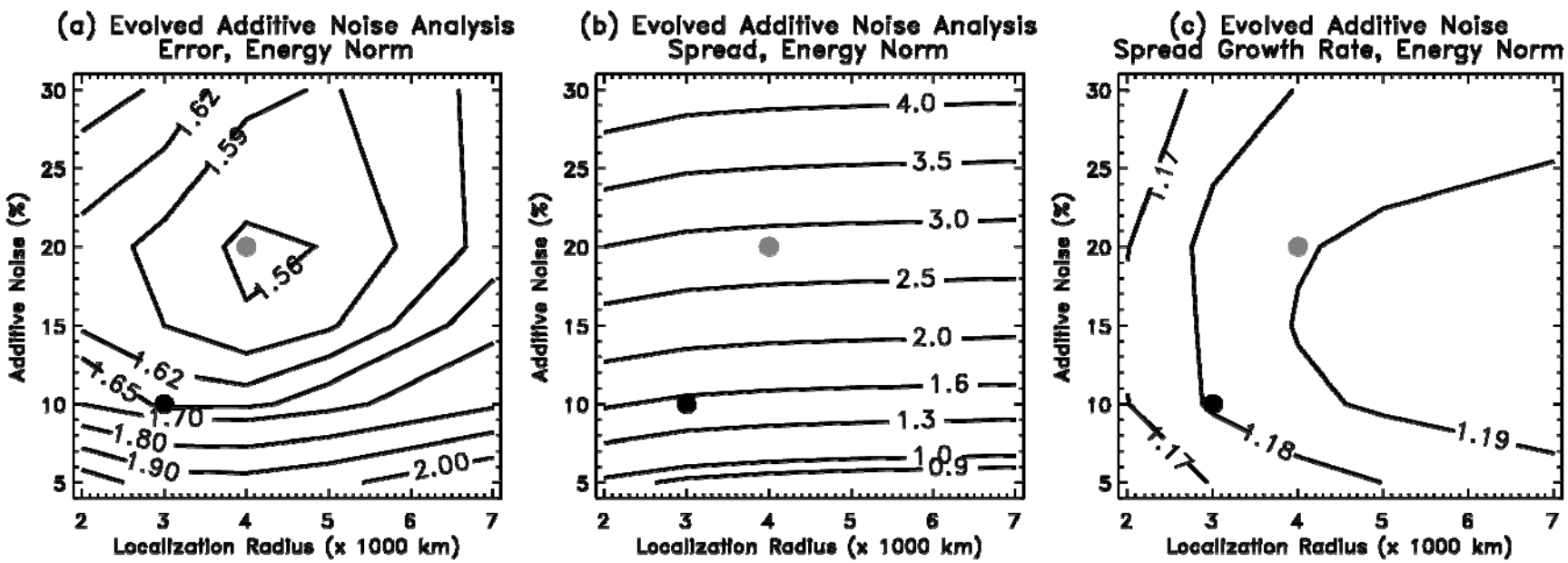


dashed line shows magnitude of initial perturbation

Suppose we evolve the additive noise for 36 h before adding to posterior?

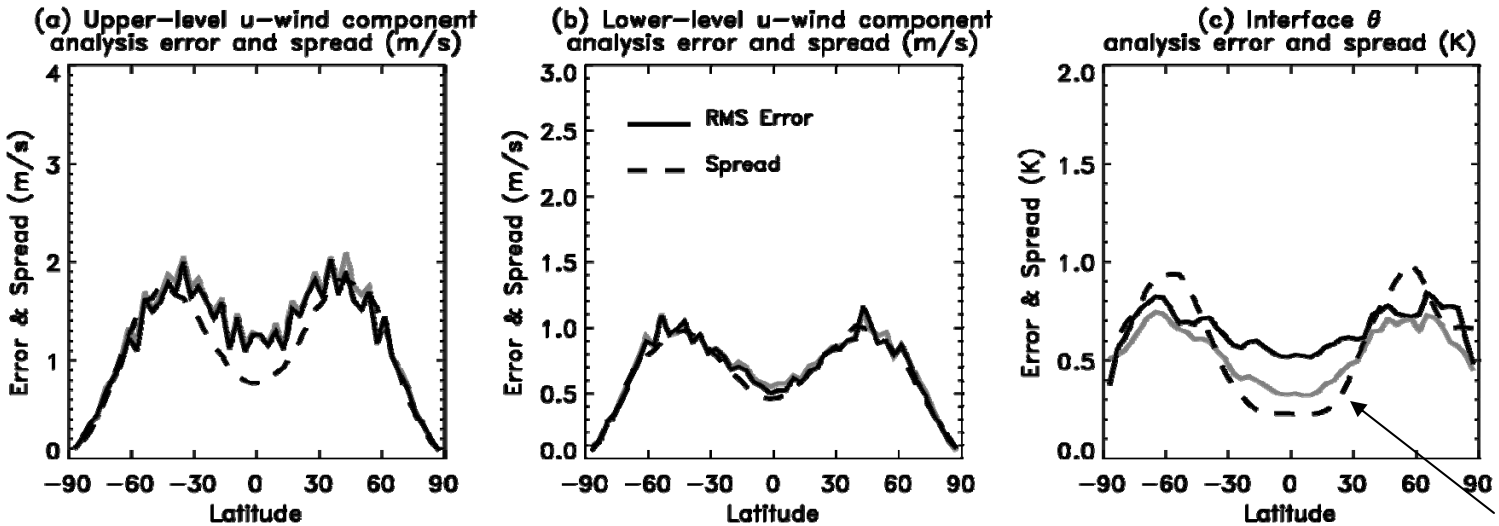


Suppose we evolve the additive noise for 36 h before adding to posterior?



Now the ensemble mean error is lower, but at a different optimum localization radius and additive noise amount. Note spread growth is much larger.

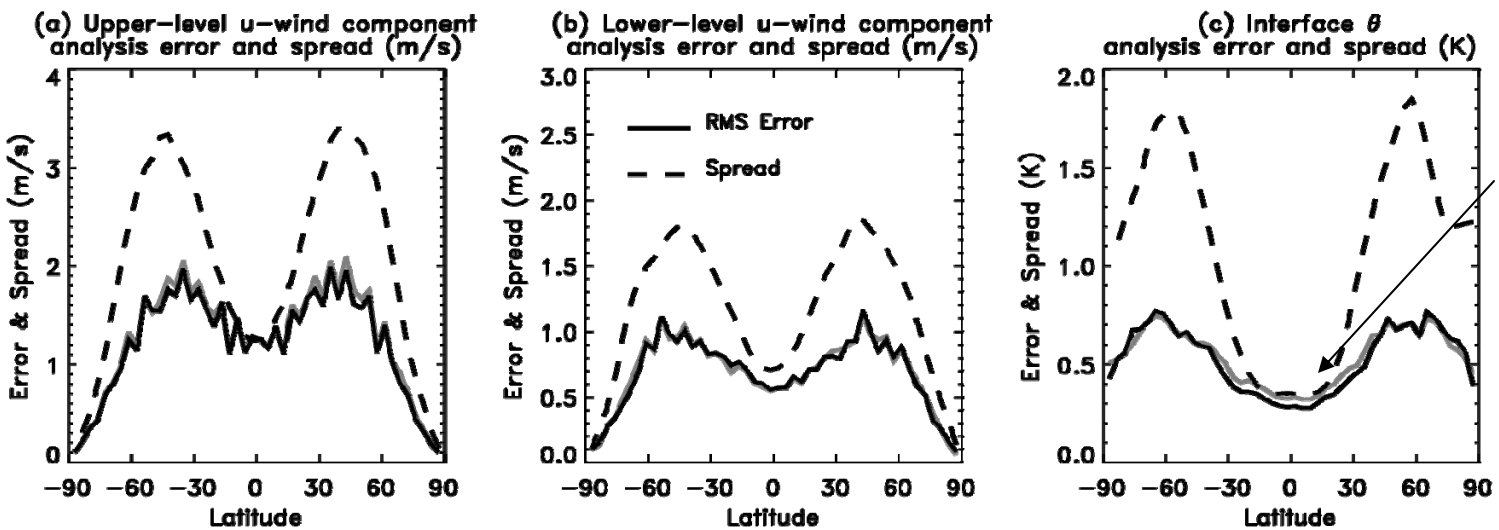
- Evolved, 3000 km localization, 10% inflation



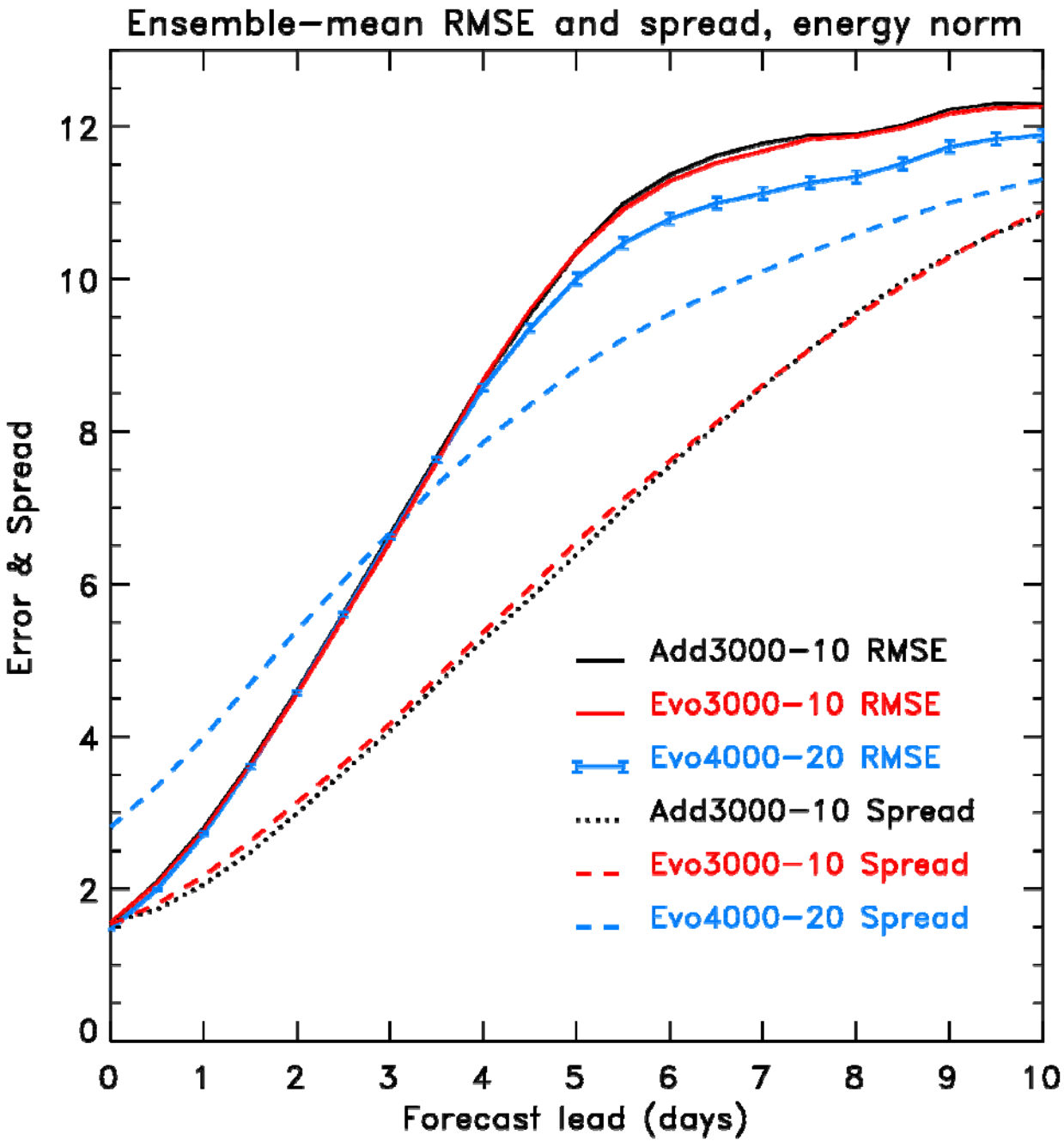
grey line is error result from non-evolved additive noise (replicated from slide 12)

higher error in tropics, less spread than error.

- Evolved, 4000 km localization, 20% inflation



now slightly reduced error in tropics, much greater spread than error.

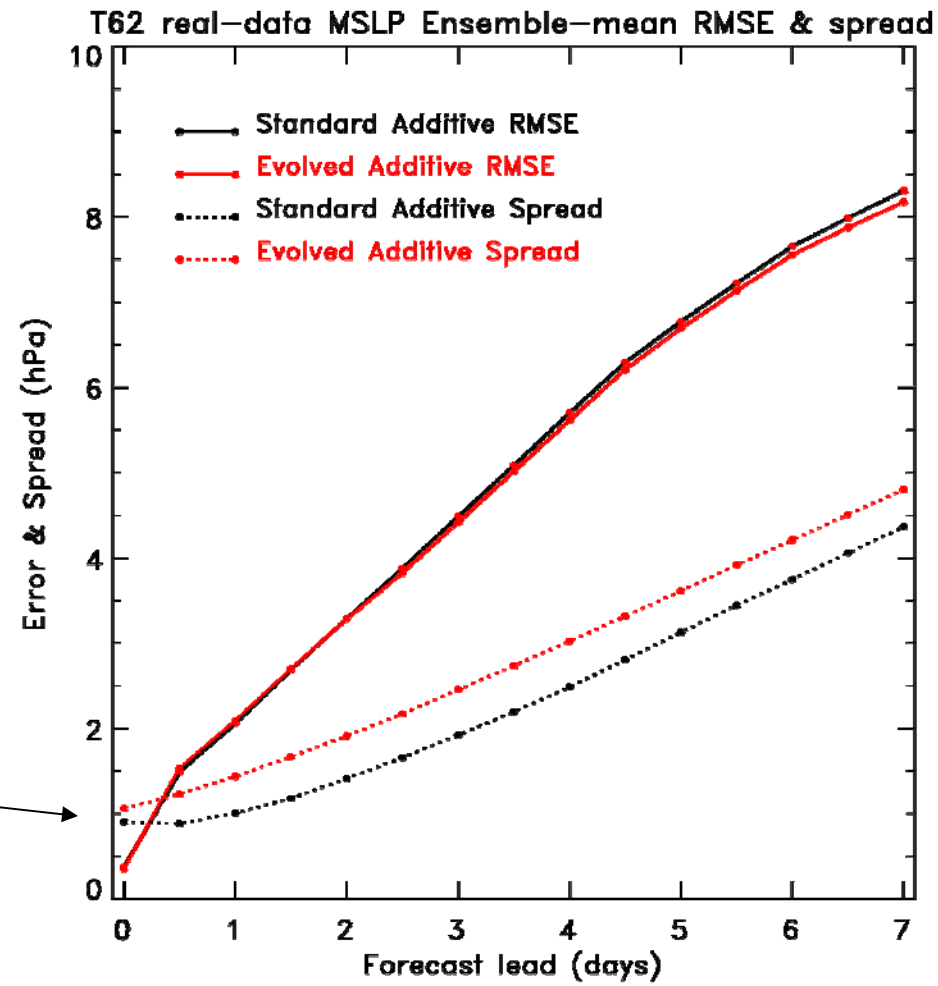


What is the effect on longer-lead ensemble forecasts?

- Not much difference, evolved vs. additive, with same localization / additive noise size.
- An improvement in error, more spread, bigger spread growth with longer localization, more evolved additive noise.

Will results hold with real model, real observations?

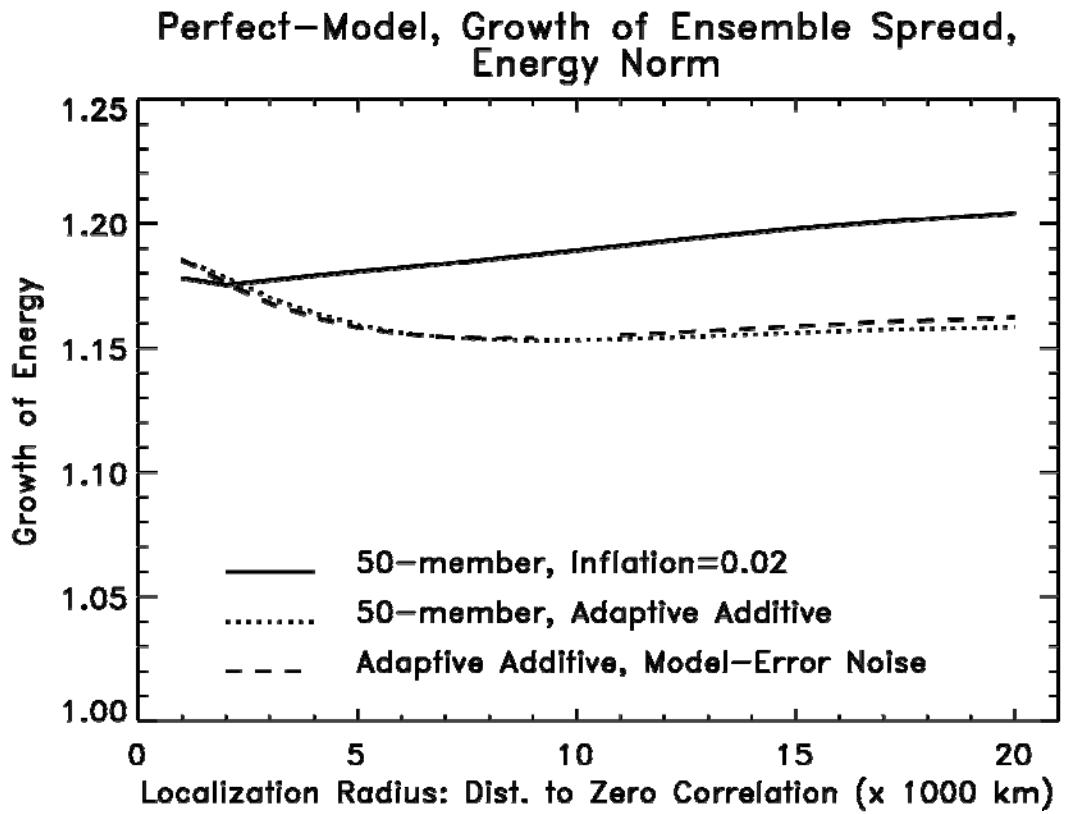
- EnKF with T62 NCEP GFS, 10 Dec 2007 to 10 Jan 2008. Nearly full operational data stream.
- 24-h evolved additive error using NMC method (48-24h forecasts) multiplied by 0.5.
- 10-member forecasts 1x daily, from 00Z.
- Main result: slightly higher spread growth at beginning of forecast.
- Other results (T190L64) less encouraging, still being analyzed.



Conclusions

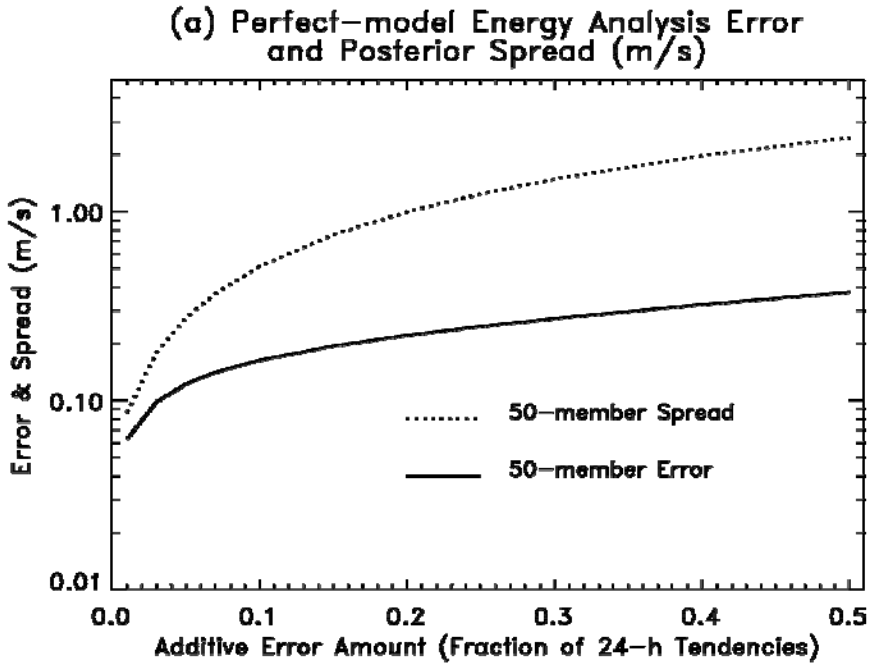
- The non-flow dependent structure of additive noise may be a primary culprit in the lack of spread growth in forecasts from EnKFs
  - Model-dependent result?
- Pre-evolving the additive noise used to stabilize the EnKF results in improved spread in the short-term forecasts, and possibly a reduction in ensemble mean error at longer leads.
  - operationally this would increase the cost of the EnKF, but perhaps the evolved additive noise could be done with a lower-resolution model.
- More generally, **the methods to treat system error will affect performance of EnKF for assimilation, ensemble forecasting; require more thought & research**.

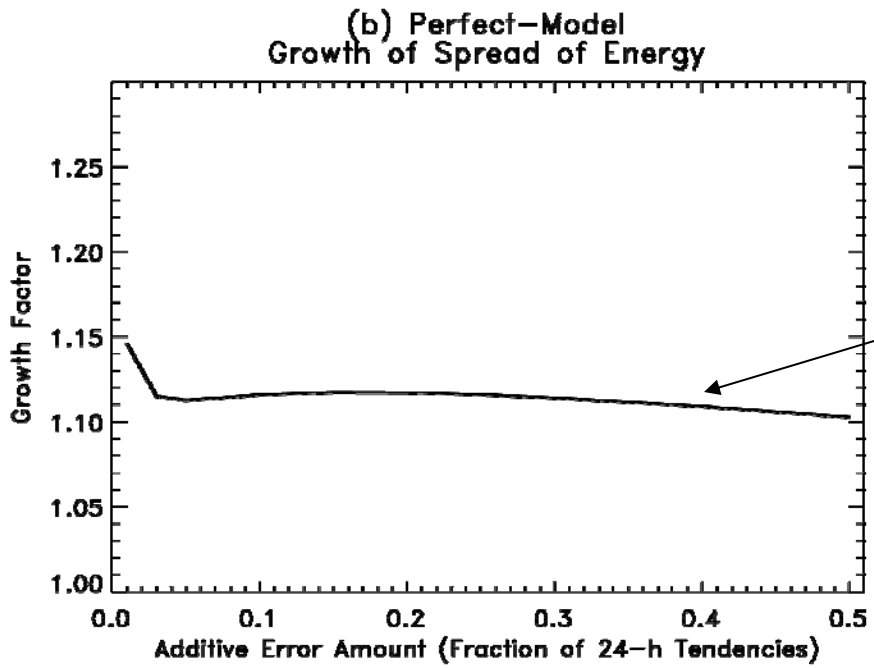
Is it the structure of this new type of additive error responsible for the lesser spread growth? NO.



- used model-error additive noise back in perfect-model data assimilation experiment. Little change in growth of energy.

[more]





Does more and more additive noise decrease the spread growth?

(a test with fixed 10 000 km localization radius)

- Answer:slightly. Moderate detrimental effect of on spread growth from increasing amounts of additive noise when localization radius is fixed.

Covariance localization and size of the ensemble

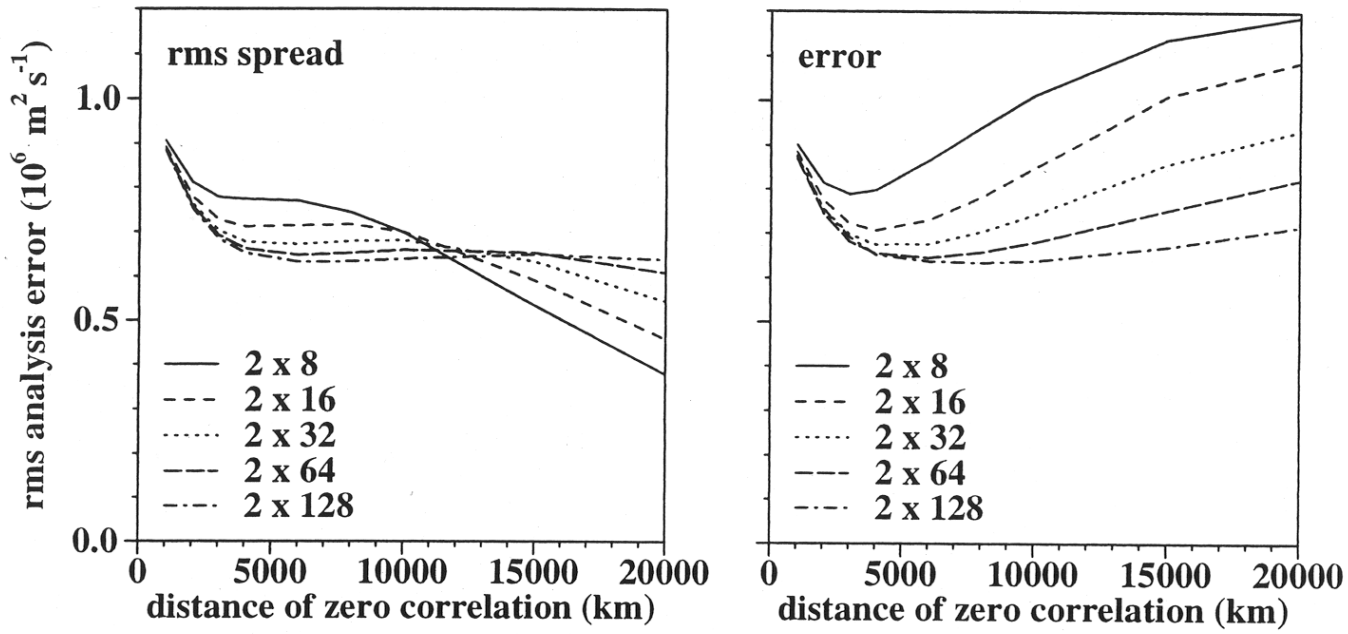
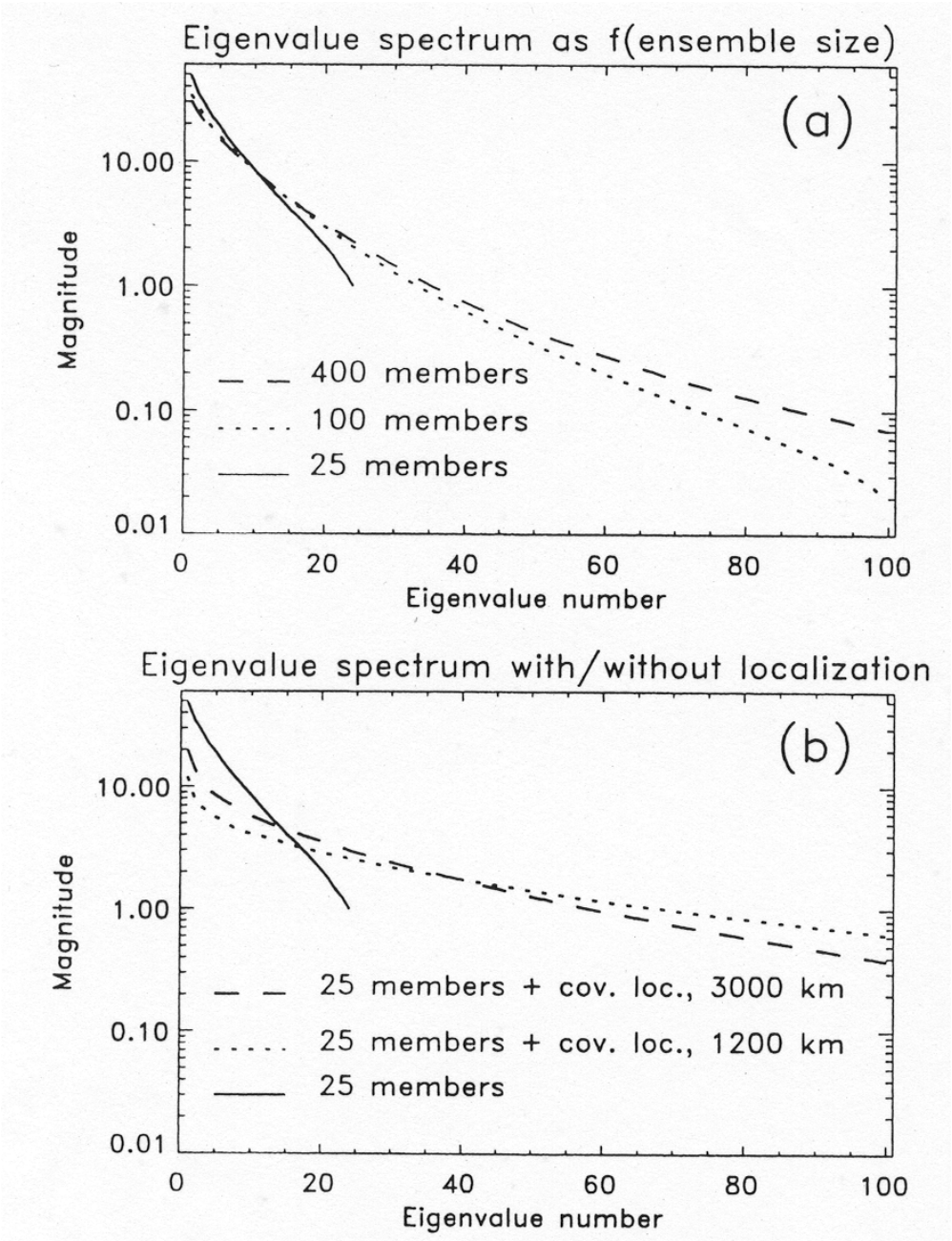


FIG. 4. Analysis error as a function of r_l (km), the distance beyond which ρ is zero, for various ensemble sizes. (left) The rms spread in the ensemble, and (right) the rms error of the ensemble mean.

Smaller ensembles achieve lowest error and comparable spread/error with a tighter localization function

(from Houtekamer and Mitchell, MWR, Jan 2001)

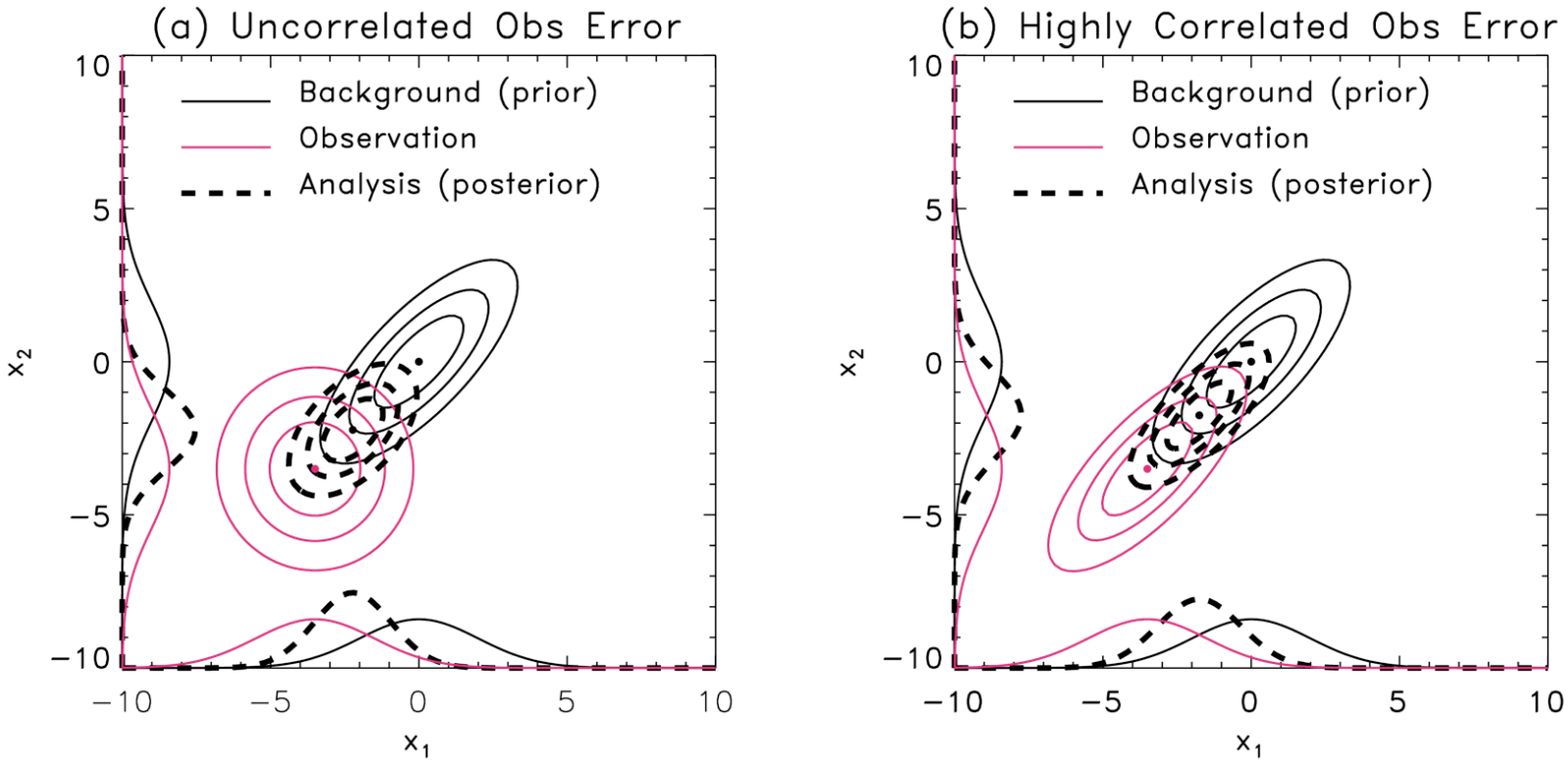
How does covariance localization make up for larger ensemble?



(a) eigenvalue spectrum from small ensemble too steep, not enough variance in trailing directions. (b) Covariance localization adds directions, flattens eigenvalue spectrum)

source: Hamill et al, MWR, Nov 2001

Observation-error covariance effect on posterior & spread growth



Common assumption of independence of observation errors may inappropriately whiten the posterior, creating small-scale noise that contributes to a lack of spread growth (speculation).