[1] [2]

An Adjoint-Based Adaptive Approach to Mitigate Background Limitations in EnKFs

Hajoon Song[1], Ibrahim Hoteit[2], Bruce Cornuelle[1],
Xiaodong Luo[2], and Aneesh Subramanian[1]

SCRIPPS INSTITUTION OF OCEANOGRAPHY

KAUST King Abdullah University of Science and Technology

Cefalu, Oct 12, 2010

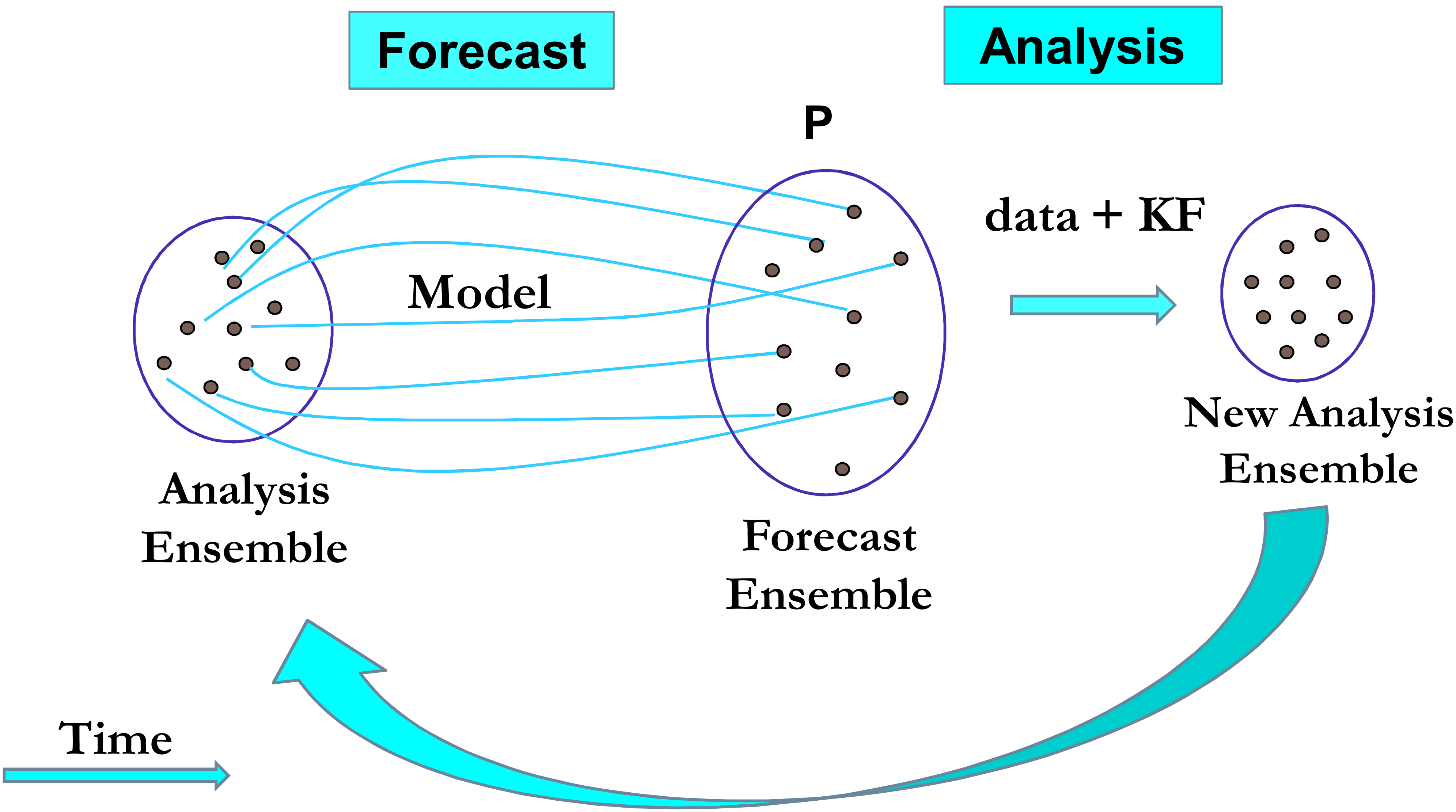
Context

- EnKF background is limited by small ensembles and poorly known model errors
- Hybrid, inflation and localization methods are used as some kind of estimates of the background "null space", but do not improve ensemble "diversity"
- The idea here is try to improve the EnKF background by incorporating new ensemble members estimated from the null space, using 3D and 4D-VAR!

Outline

- Ensemble Kalman Filtering (EnKF)
  - Background limitations
- Methods to mitigate background limitations
  - Inflation, Hybrid, localization …
- Adaptive EnKF (vs. Hybrid, and Inflation)
- Numerical Results

EnKF Algorithm



EnKF

- Monte-Carlo Approach: Represents uncertainties by an ensemble of vectors

$$P = \frac{1}{N-1}\sum_i \left(\overline{x} - x^i\right)\left(\overline{x} - x^i\right)^T$$

- Update the ensemble instead of **P** :
  - Solves storage problem
  - Significantly reduces number of model integrations
  - Suitable for nonlinear systems

- *Accurate description of* ***P*** *is critical (Lorenc 2003)!*

EnKF Background Limitation

- The accuracy of the background covariance matrix is mainly limited by:

  "*Small ensembles*" & "*Model deficiencies*"

  - Rank deficiency: not enough degrees of freedom to fit data
  - Spurious correlations: unreliable statistics from small sample
  - Underestimated background covariance (weak spread)

Inflation, Localization and Hybrid …

- The EnKF background is only an approximation of the ‘true’ uncertainties

$$\mathbf{P}^t = \mathbf{P} + \mathbf{B}$$

- Inflation, localization and hybrid (LIH) are all used to somehow represent estimates of $\mathbf{B}$ *can be simultaneously used!*

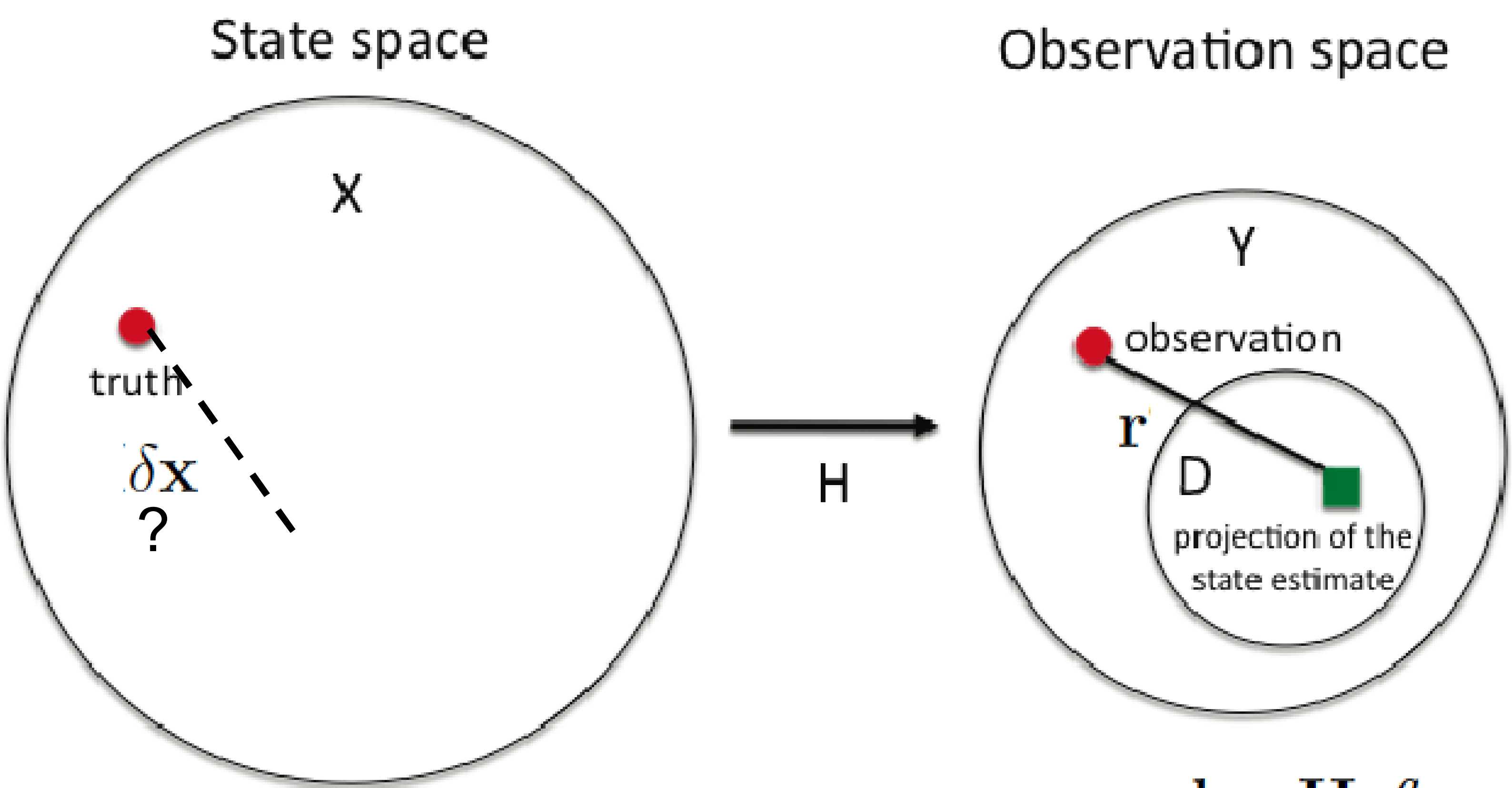
- <u>Hybrid:</u> “Relax OI-/3D-VAR background to flow-dependent EnKF background

$$\tilde{\mathbf{P}}^t = \alpha \mathbf{P} + (1-\alpha)\mathbf{B}$$

- <u>Additive Inflation</u>: Add some perturbations to the members

Background Limitation – Geometric Interp.

- EnKF solution is a linear combination of the ensemble members



$$\mathbf{r} = \mathbf{d} - \mathbf{H}\mathbf{x}^a$$

Adaptive EnKF (AEnKF)

- If **P**is not accurately estimated, the residuals are the result of missing directions in the ensemble

- Residuals carry information about EnKF background deficiency

- *The idea is then to back-project the residuals from the observation space to the state space and use that as new member in the ensemble!*

AEnKF

- Enrich the EnKF ensemble with new members estimated in the ensemble null space!

- To estimate the new members:

$$\mathbf{d} - \mathbf{H}\mathbf{x}^f = \mathbf{H}(\mathbf{x}^a - \mathbf{x}^f) + \mathbf{H}\delta\mathbf{x} + \mathbf{r}^e$$

which is equivalent to $\mathbf{d} - \mathbf{H}\mathbf{x}^a = \mathbf{H}\delta\mathbf{x} + \mathbf{r}^e$

- Consider it as an inverse problem with prior **B**and cov. obs**R**

$$J(\delta\mathbf{x}^e) = \frac{1}{2}\delta\mathbf{x}^{eT}\mathbf{B}^{-1}\delta\mathbf{x}^e + \frac{1}{2}(\mathbf{r} - \mathbf{H}\delta\mathbf{x}^e)^T\mathbf{R}^{-1}(\mathbf{r} - \mathbf{H}\delta\mathbf{x}^e)$$

- A new member is then $\mathbf{x}^{a,e} = \mathbf{x}^a + \beta\delta\mathbf{x}^e$

AEnKF

- Another way to interpret it is to split the Kalman Gain into two parts:

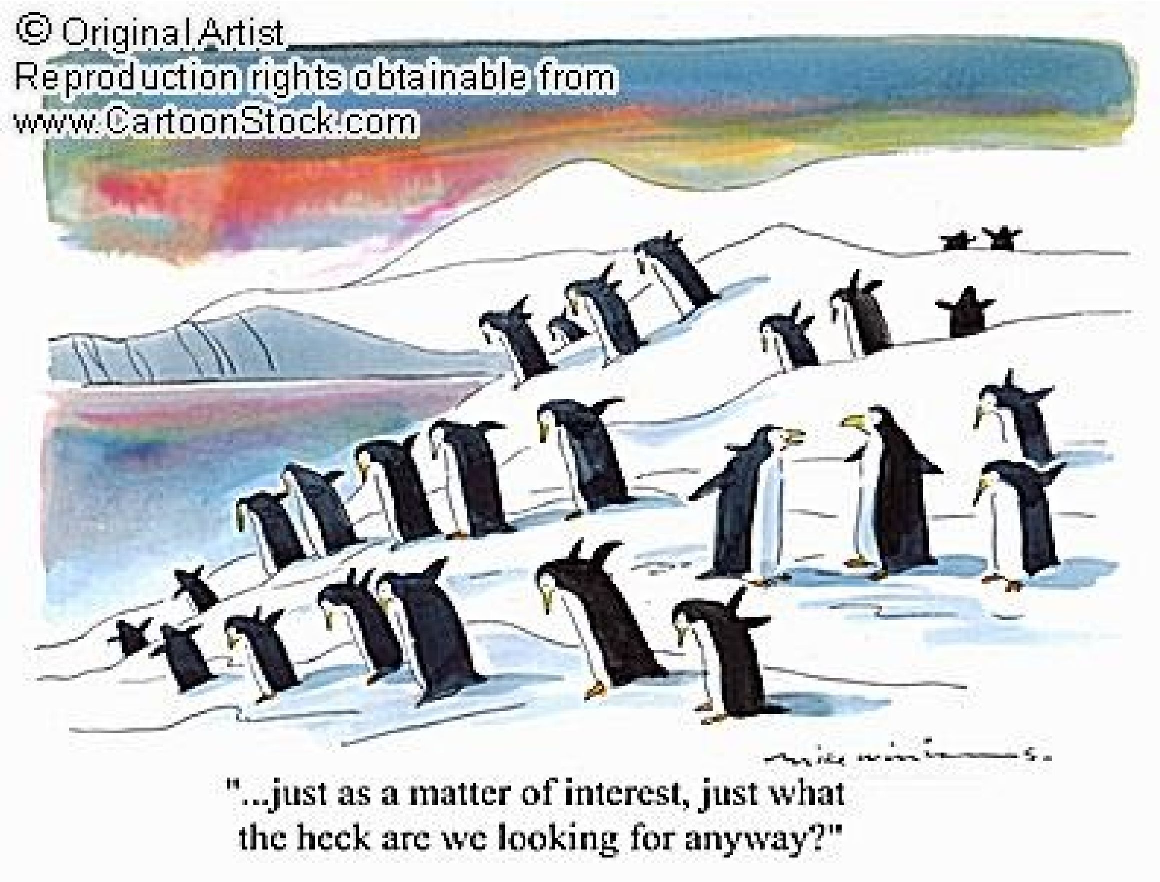
$$\mathbf{K} = \mathbf{P}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}^e\right)^{-1}$$

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

We use $\mathbf{K}$ *ıpdate the ensemble as in the regular EnKF, and we use t*$\mathbf{K}^r$ *stimate a new member.*

- We could use $\mathbf{K}^r$ for each member, as in LIH methods, so that same increments are added to all members. This would however increase correlations between members and does not improve “diversity”.

AEnKF – Real Life Interpretation



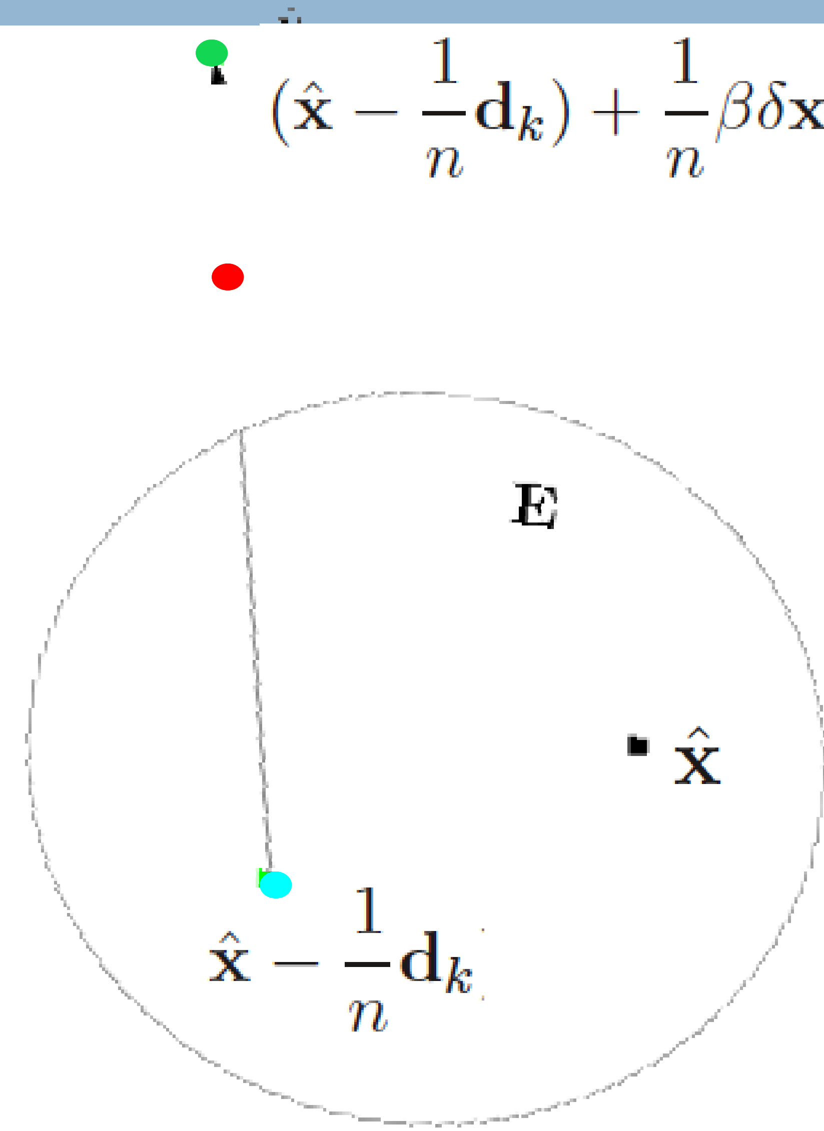
What are we looking for?

A "smart penguin" would lead the way …

Hajoon

New member – Geometric Interp.

- *For practical reasons, we need to remove members from the ensemble*
- We remove the members that distort the least the background distribution
- As for the weighting factor, here we do it by trial and error, but we can chose it according to some optimum criterion



Adaptive *vs.* Hybrid

- Adaptive limits growth of the ensemble to directions indicated by the residuals and not represented in the ensemble
- Adaptive easier to implement: the “two assimilation” systems are applied independently
- Technically speaking, adaptive does not increase the background rank

4D - AEnKF

- Use 4D-VAR formulation to reduce dependency on **B**while including more information from dynamics and data

$$J_{4D}(\delta\mathbf{x}_{i-n}) = \frac{1}{2}\delta\mathbf{x}_{i-n}{}^T\mathbf{B}^{-1}\delta\mathbf{x}_{i-n} + \frac{1}{2}\sum_{j=i-n}^{i}\alpha_j\left(\mathbf{r}_j - \mathbf{G}_j\delta\mathbf{x}_{i-n}\right)^T\mathbf{R}_j^{-1}\left(\mathbf{r}_j - \mathbf{G}_j\delta\mathbf{x}_{i-n}\right)$$

$$\boxed{\mathbf{G}_j = \mathbf{H}_j\mathbf{M}_{j,i-n},}$$

  - Amounts to integrate residuals backward in time with the adjoint!
  - A new member at time t_{i-n} is then

$$\mathbf{x}_{i-n}^{a,e} = \mathbf{x}_{i-n}^{a} + \beta\delta\mathbf{x}_{i-n}^{e}$$

which is next integrated forward to provide new member at t_i.

Generating More Members

- Random sampling from the estimated distribution of the new member
- Descent directions from optimizing the cost function

Some Numerical Experiments

- Lorenz 96 model as a testbed for ensemble methods

$$\frac{dx_j}{dt} = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F \qquad \begin{cases} n = 40 \\ F = 8 \\ \Delta t = 0.05 \sim 6h \end{cases}$$

  - First, long free-run
  - Initial ensemble: mean of free-run + N(0,1)
  - **B**covariance of free-run (Hamill and Snyder, 2000)
  - Assimilation period: “reference” states from 1 year

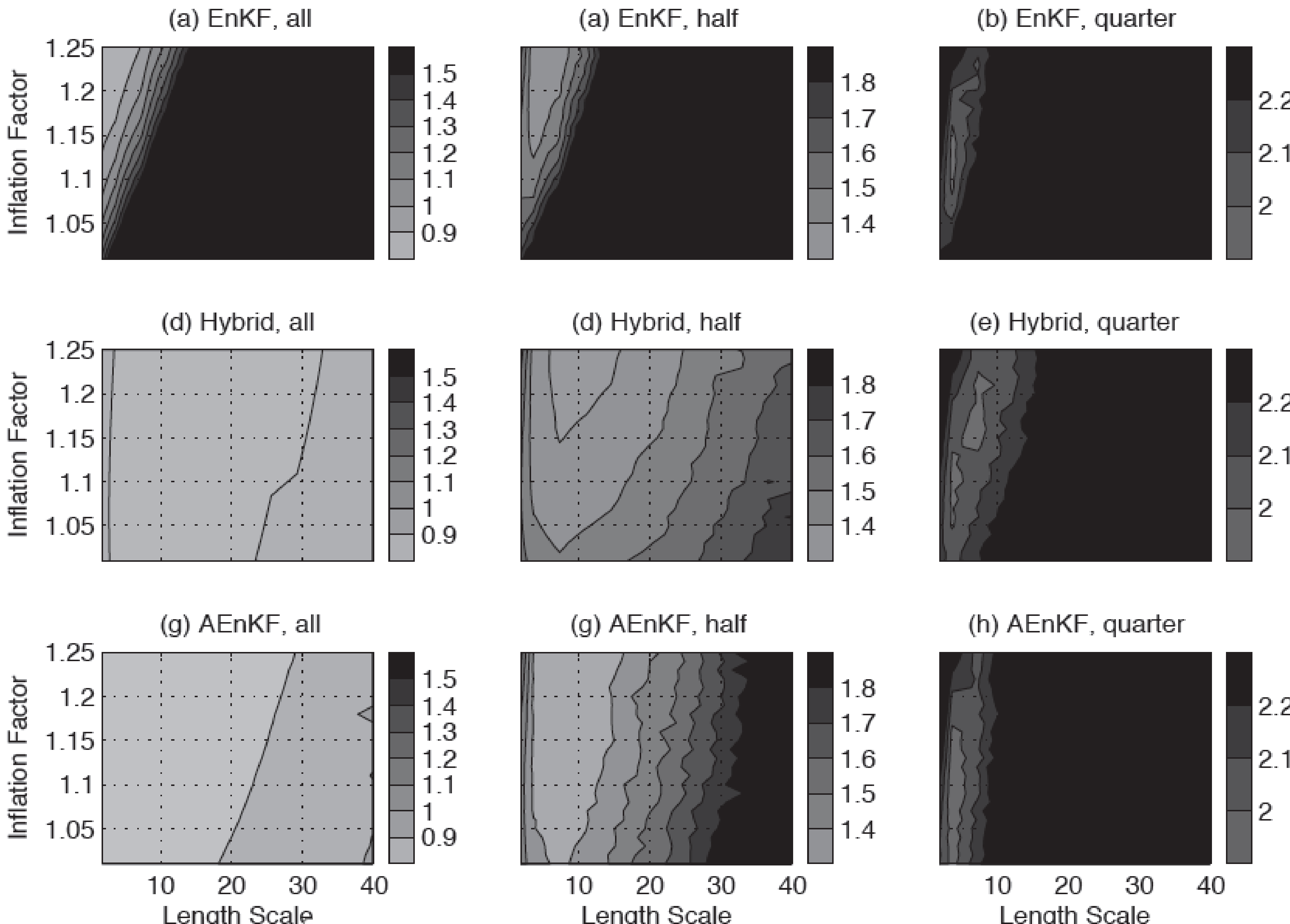
- Observations from reference states every day (All, Half, Quarter)
- Model error: F = 8 in reference model and F = 6 in forecast
- Sampling errors: only 10 members were used
- We compared EnKF, 3D-VAR hybrid, AEnKF, and 4D-AEnKF

Averaged RMS

EnKF

Hybrid

AEnKF

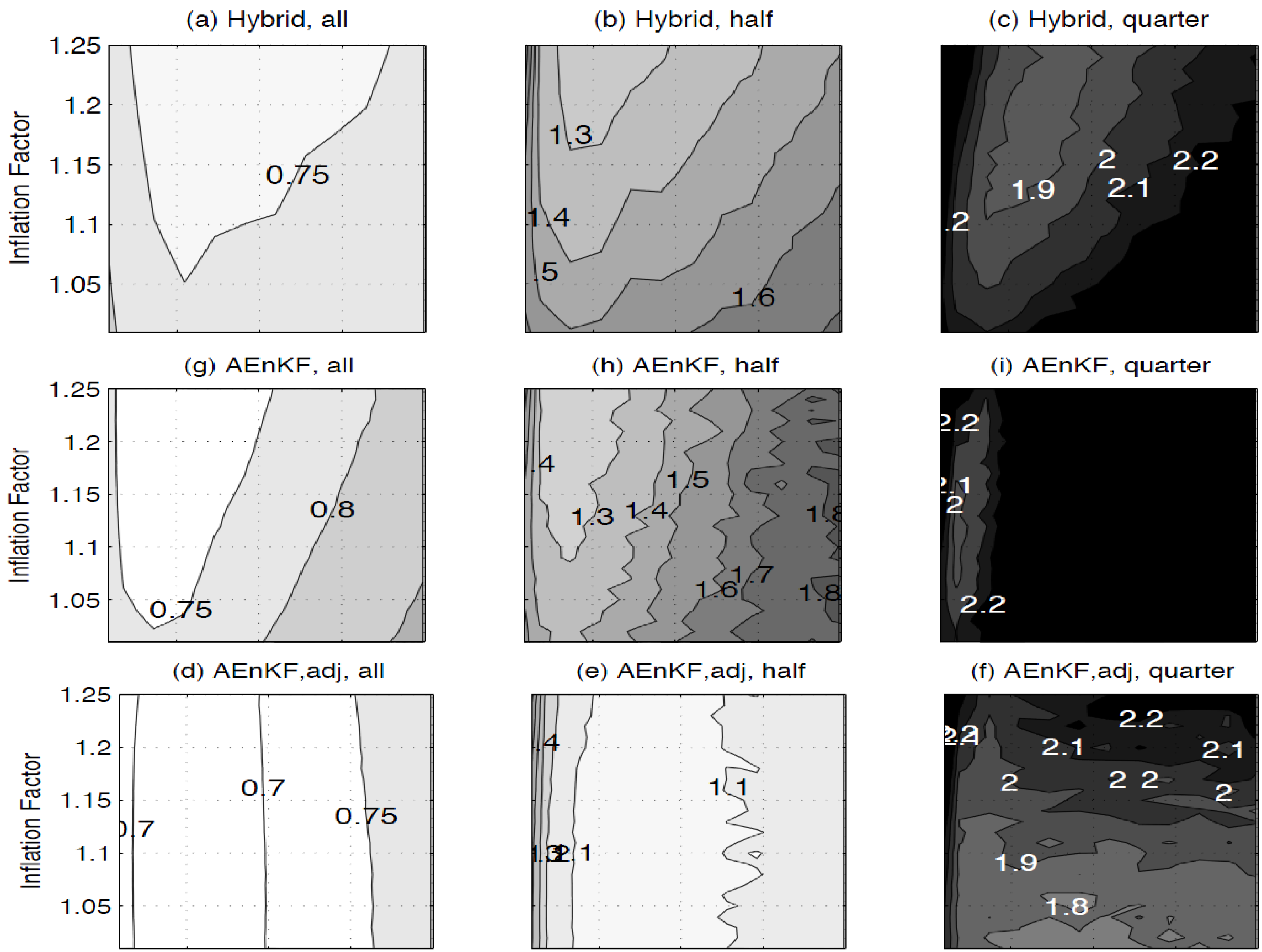


Averaged RMS (B = Identity)

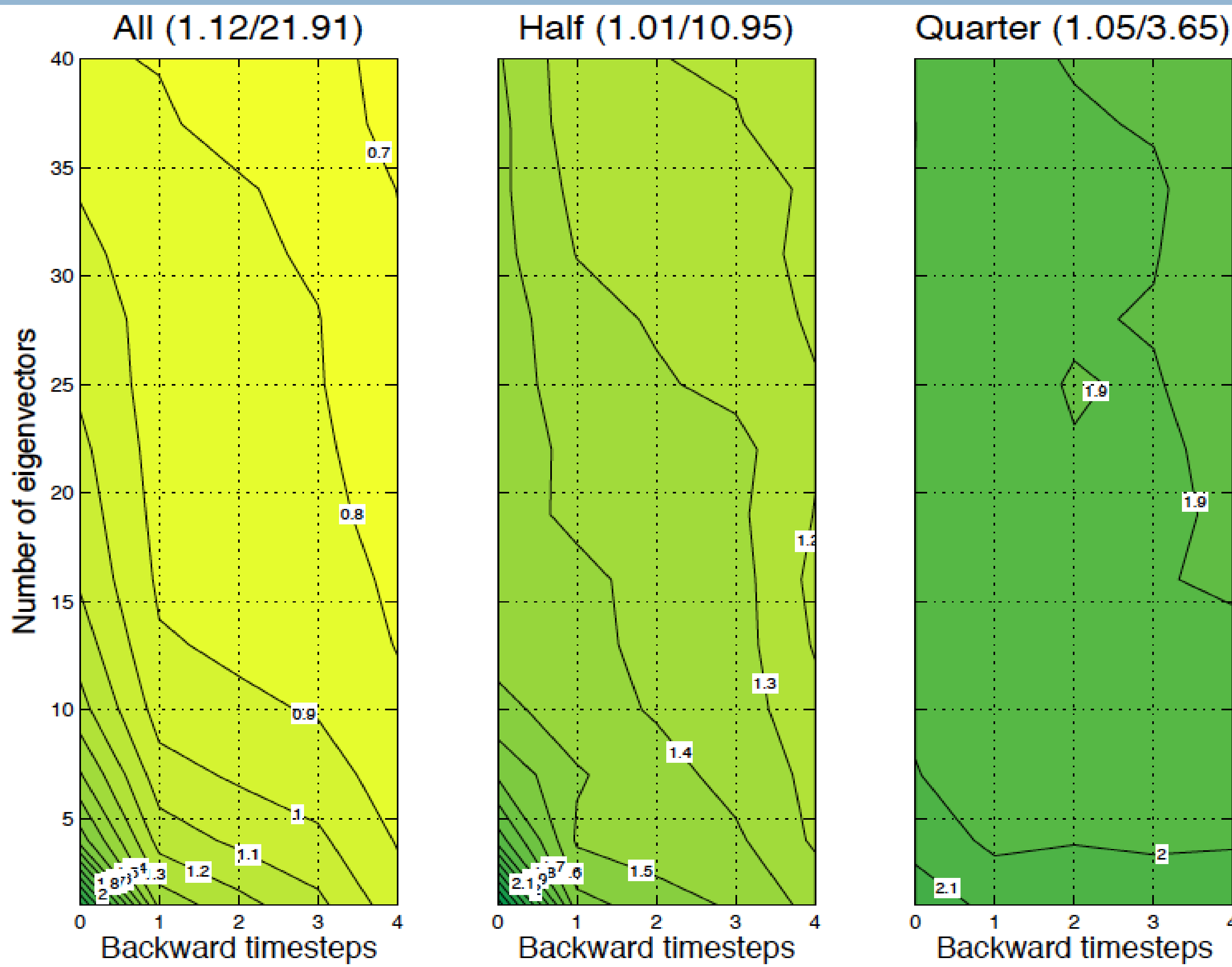
Hybrid

AEnKF

4D-AEnKF



4D-AEnKF % number of backward steps



Discussion

- Combining the good features of EnKF and 4D-VAR could enhance performances
- This however requires implementing the two methods; quite demanding!
- *Hybrid methods use EnKF to improve VAR background covariances (4D En-VAR). Here we use VAR to enrich EnKFsensembles (4D VAR-En) …*

Reference

- H. Song, I. Hoteit, B. Cornuelle, and A. Subramanian: *An adaptive approach to mitigate background covariance limitations in the ensemble Kalman filter*. Monthly Weather Review, 138, 2825-2845, 2010.

THANK YOU