# Some comparisons

Javier Amezcua
Data Assimilation Research Centre
University of Reading
March 2018

# Bayes' theorem in DA

**Likelihood**. Pdf of the observations given a value of the state variable.

**Prior pdf**. Pdf of the state variables coming from the model

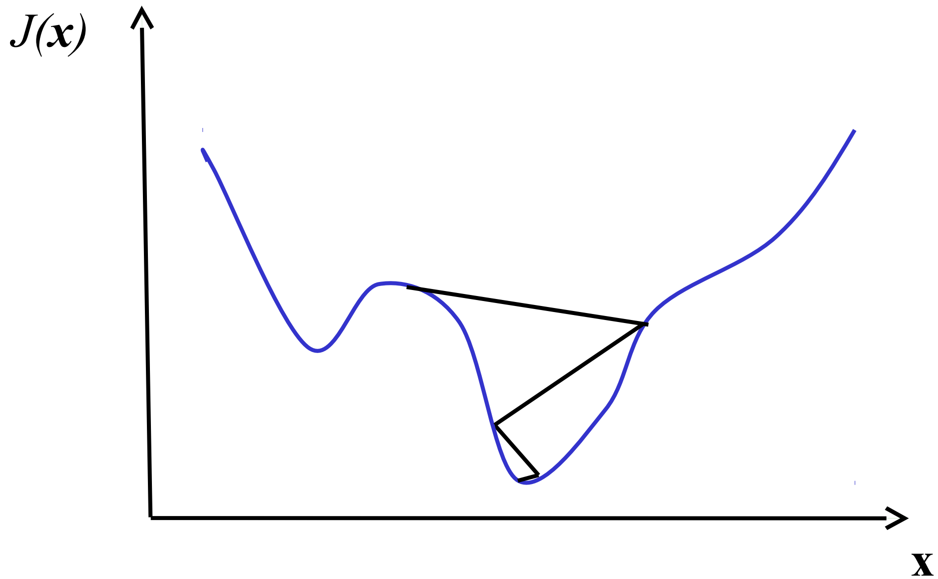
$$pdf(\mathbf{x}|\mathbf{y}) = \frac{pdf(\mathbf{y}|\mathbf{x})pdf(\mathbf{x})}{p(\mathbf{y})}$$

**Posterior pdf.** Pdf of the state variables given the observations.

**Marginal pdf of the observations**. It is often the case we do not need to compute this, since it acts as a normalisation constant.

# 1. Variational methods

$$J(\mathbf{x}) = (\mathbf{x}-\mathbf{x}_b)^T\mathbf{P}^{-1}(\mathbf{x}-\mathbf{x}_b) + (\mathbf{y}-H(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x}))$$

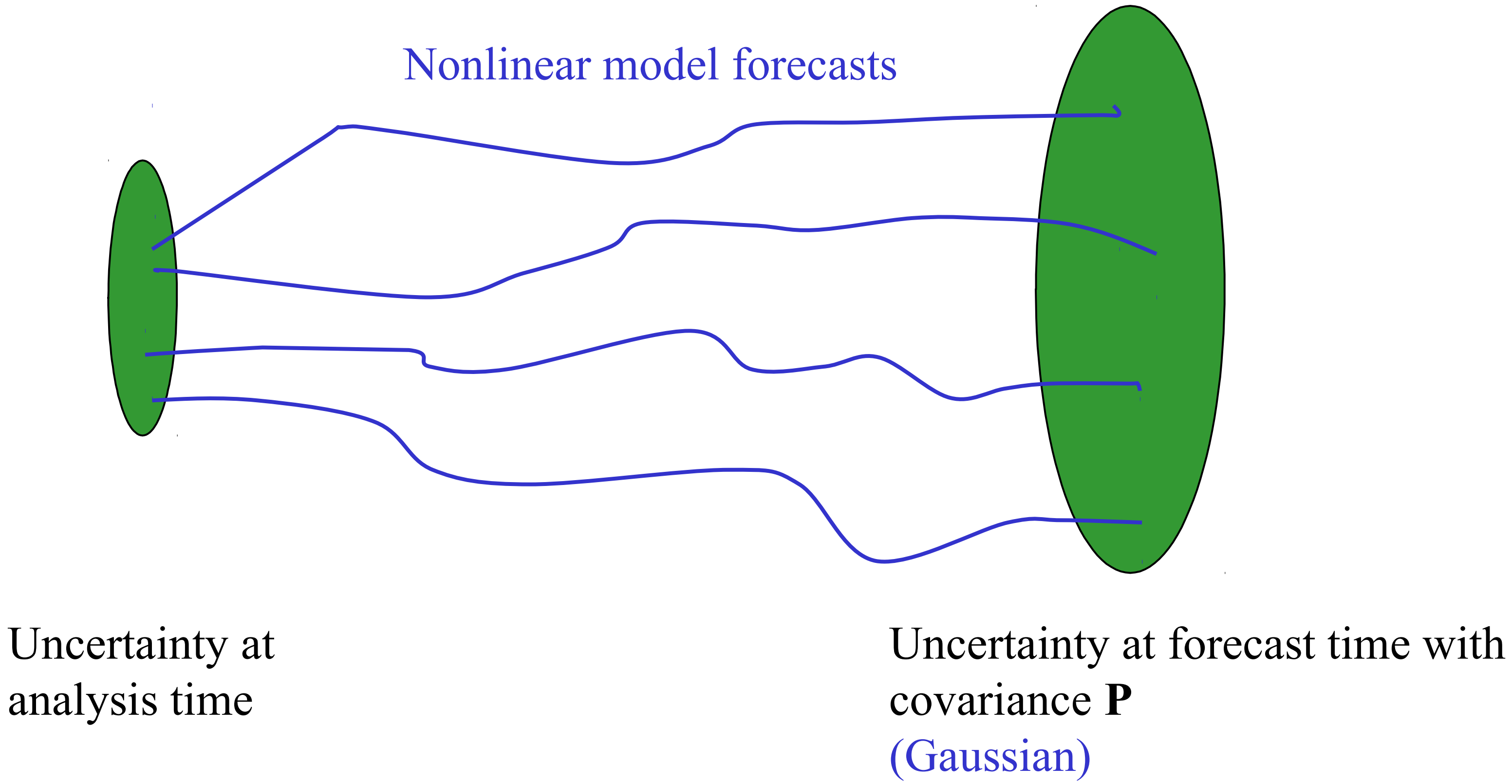



Find the minimum of the cost function via (iterative) optimisation techniques. One needs the gradient of the cost function.

The background error covariance is static.

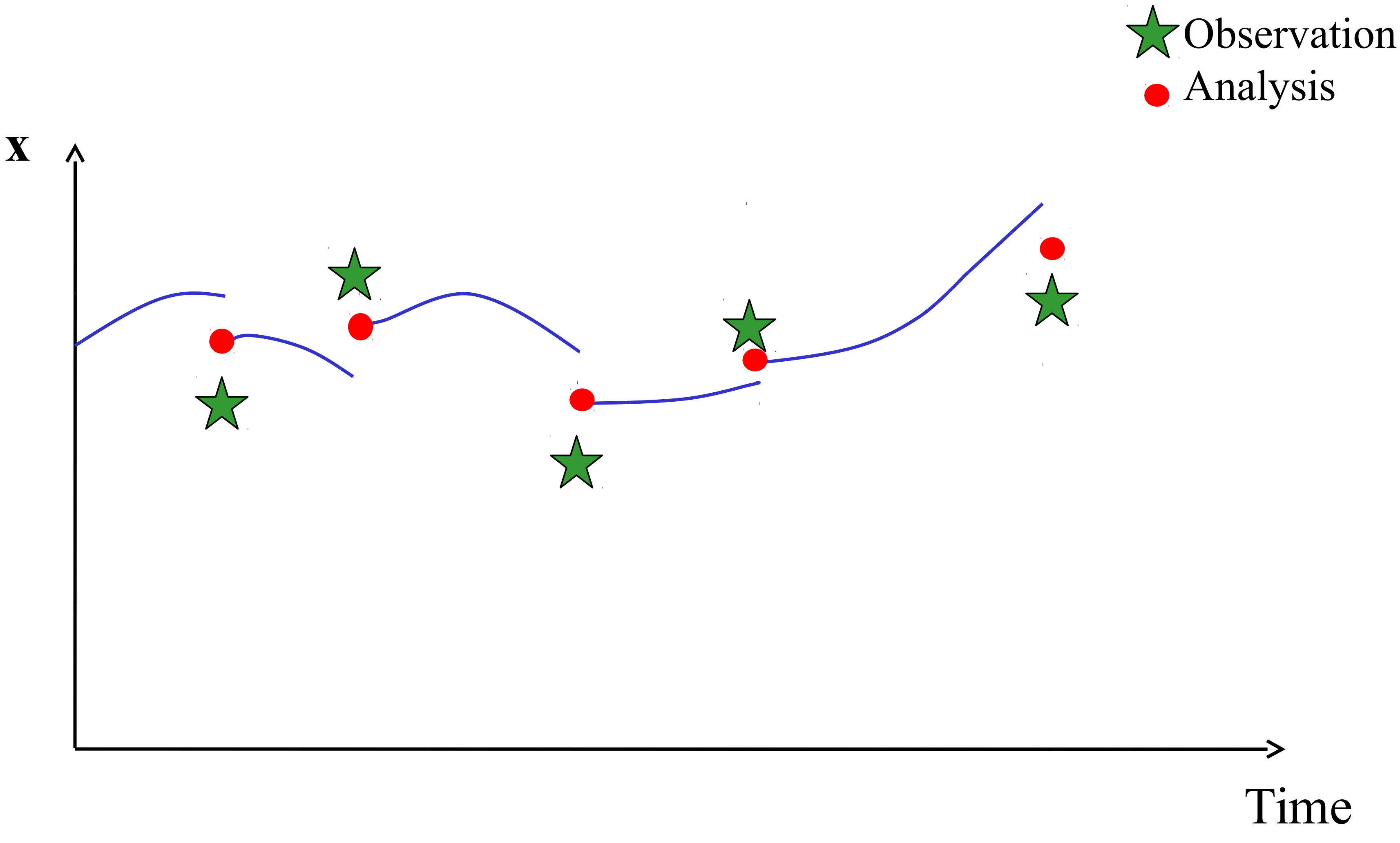
# 2. Ensemble Kalman filter

Use sample estimators with the KF equations.



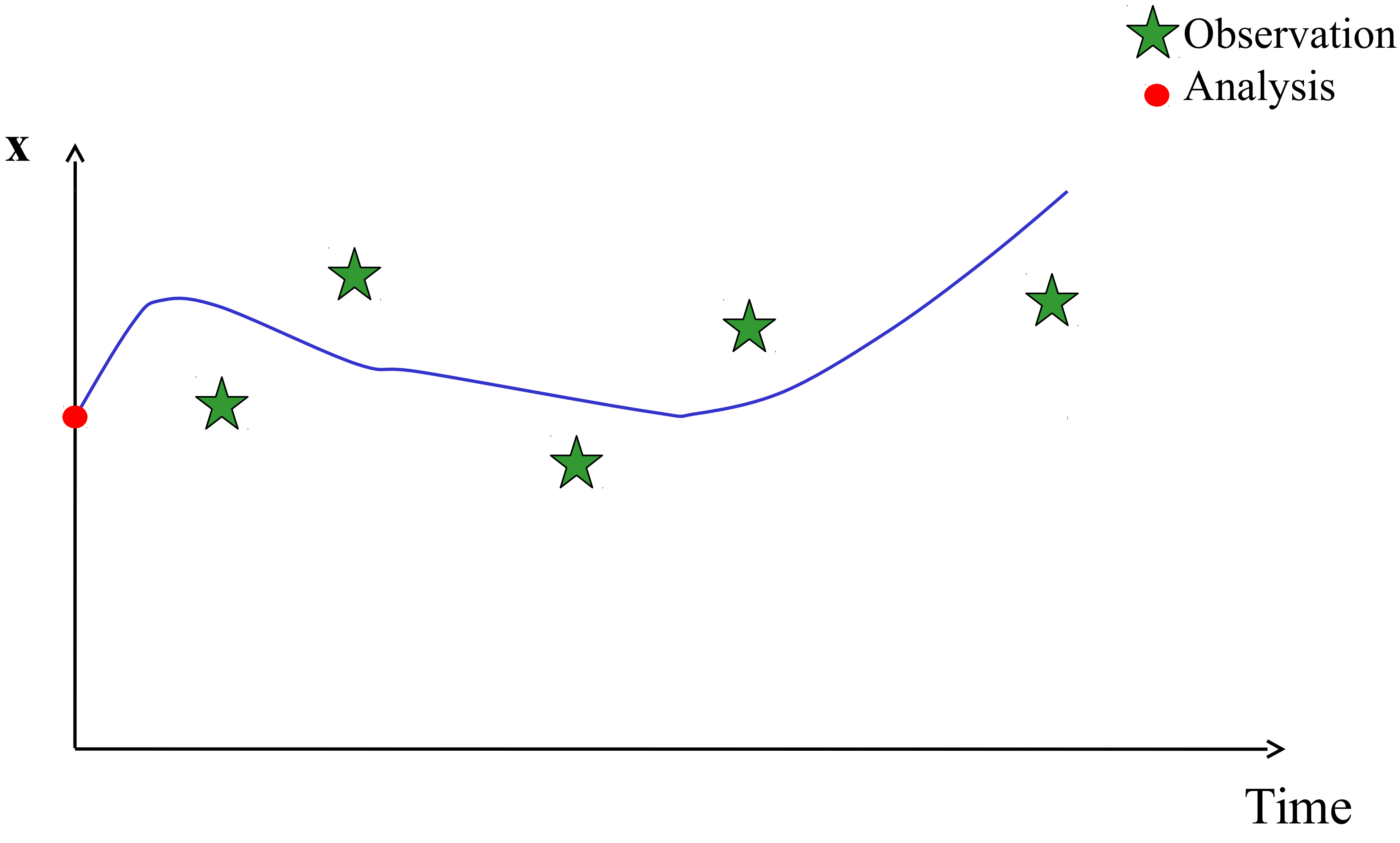
# Filters

Assimilate every time observations are available.



# Smoothers

Assimilate observations over a time window.



# Characteristics of traditional DA methods

| | Method | | Observations | | Covariance | |
|---|---|---|---|---|---|---|
| | Variational | Kalman | Sequential | Smoother | Static | Dynamic |
| 3DVar | ✓ | | ✓ | | ✓ | |
| 4DVar | ✓ | | | ✓ | (✓) | ✓ |
| Optimal Interpolation | | ✓ | ✓ | | ✓ | |
| Kalman Filters | | ✓ | ✓ | | | ✓ |
| Kalman Smoother | | ✓ | | ✓ | | ✓ |

Solution is got using (iterative) **minimisation** techniques.

Solution is got using explicit **linear algebra**.

Estimation is done for an **instant**.

Estimation is done within a **time window**.

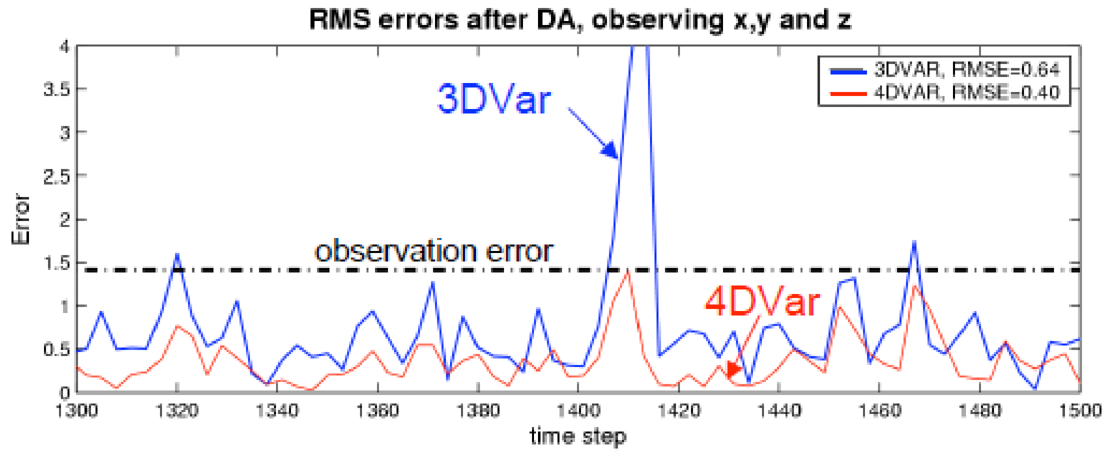
**Uncertainty** is considered **fixed** in time.

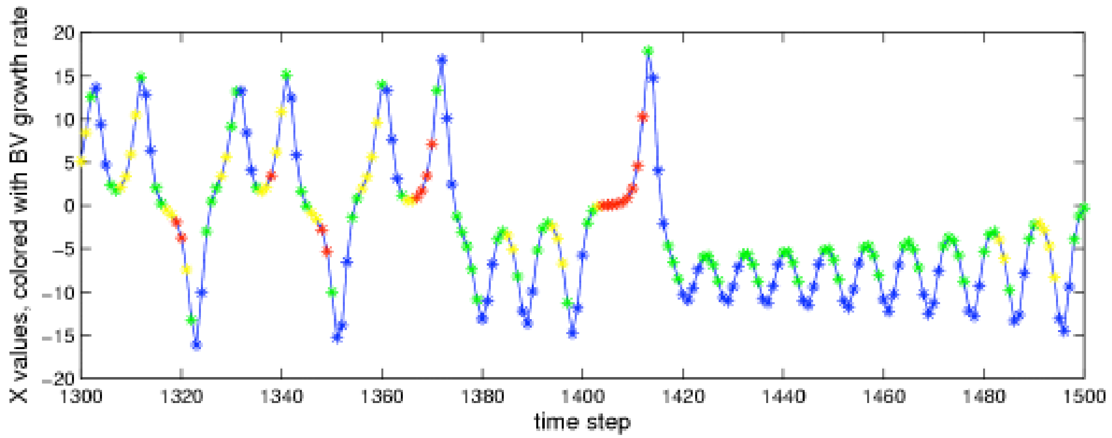
**Uncertainty evolves** in time.

# 3D vs 4DVar

4DVar has important information from the future (after all, it is a smoother), 3DVar does not.

The figure shows a comparison of the performance of the two methods. Taken from Evans et al, 2005.



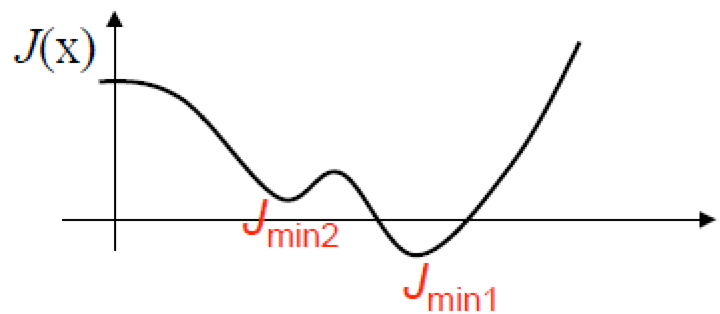

DA cycle and observations: $8\Delta t$, $\mathbf{R}=2*\mathbf{I}$
4D-Var assimilation window: $24\Delta t$

# How long should the assimilation window be?

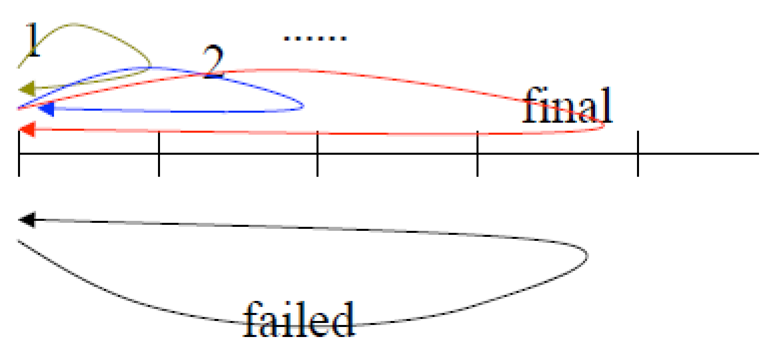
The longer the 4D assimilation window the more observations we'll have… but also the more nonlinear the forecast will be. The best should be somewhere in the middle.

| | Win=8 | 16 | 24 | 32 | 40 | **48** | 56 | 64 | 72 |
|---|---|---|---|---|---|---|---|---|---|
| Fixed window | 0.59 | 0.59 | 0.47 | 0.43 | 0.62 | **0.95** | 0.96 | 0.91 | 0.98 |
| Start with short window | 0.59 | 0.51 | 0.47 | 0.43 | 0.42 | **0.39** | 0.44 | 0.38 | 0.43 |

Performance of 4DVar using the Lorenz 1963 and different lengths of assimilation window (Kalnay *et al.*, 2007).



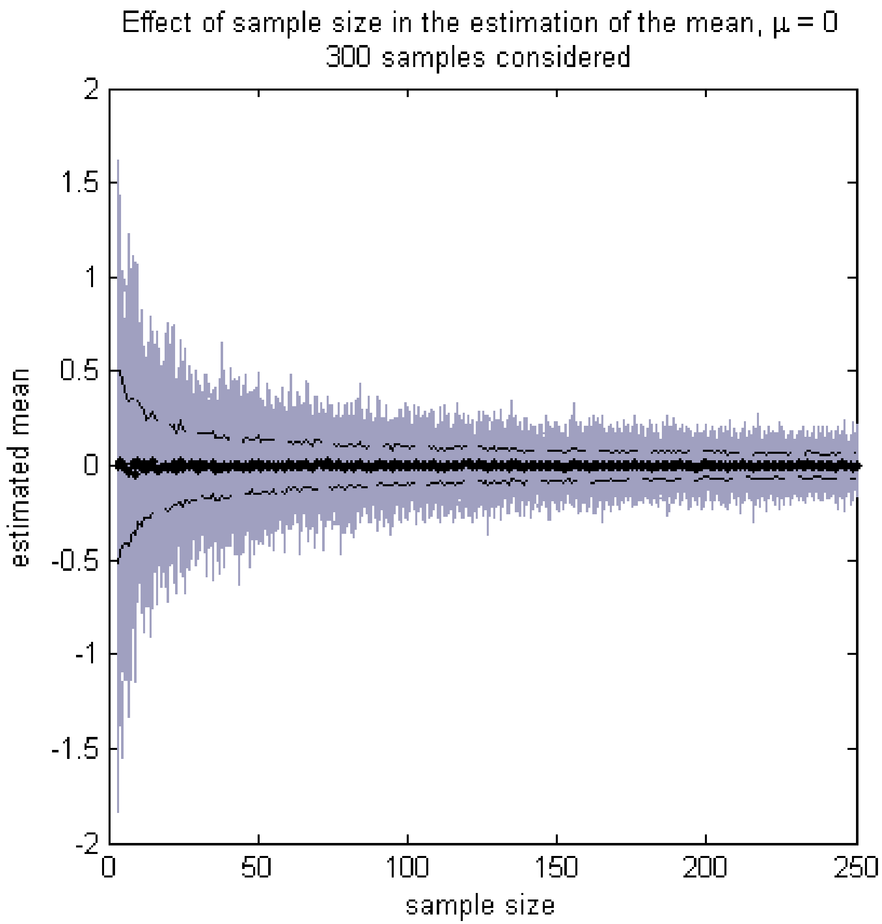

It is recommendable to do the minimization progressively while increasing the assimilation window (Pires *et al.*, 1996).

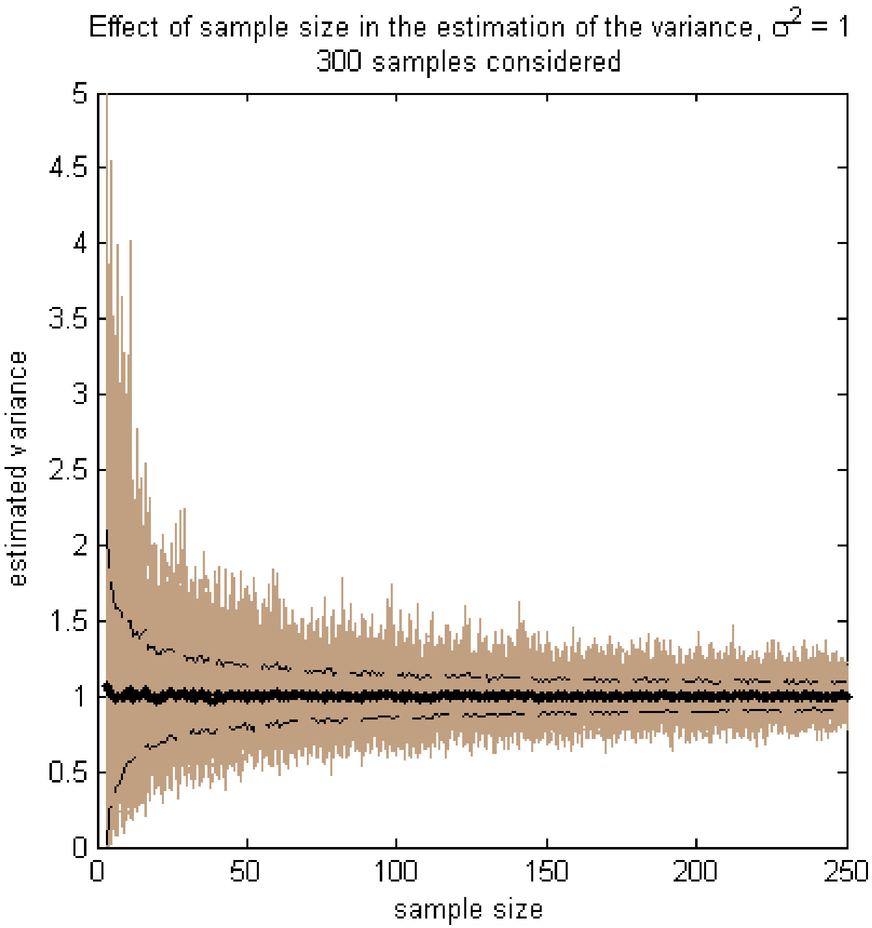


# Sampling

There is always sampling noise in the estimators, this reduces as the ensemble size increases.

Example with a univariate Gaussian distribution.



# Sampling

Two effects of **finite sample size**:

- **Underestimation** of sample **covariance**.
- **Spurious** long-range **correlations**.

Fixes:

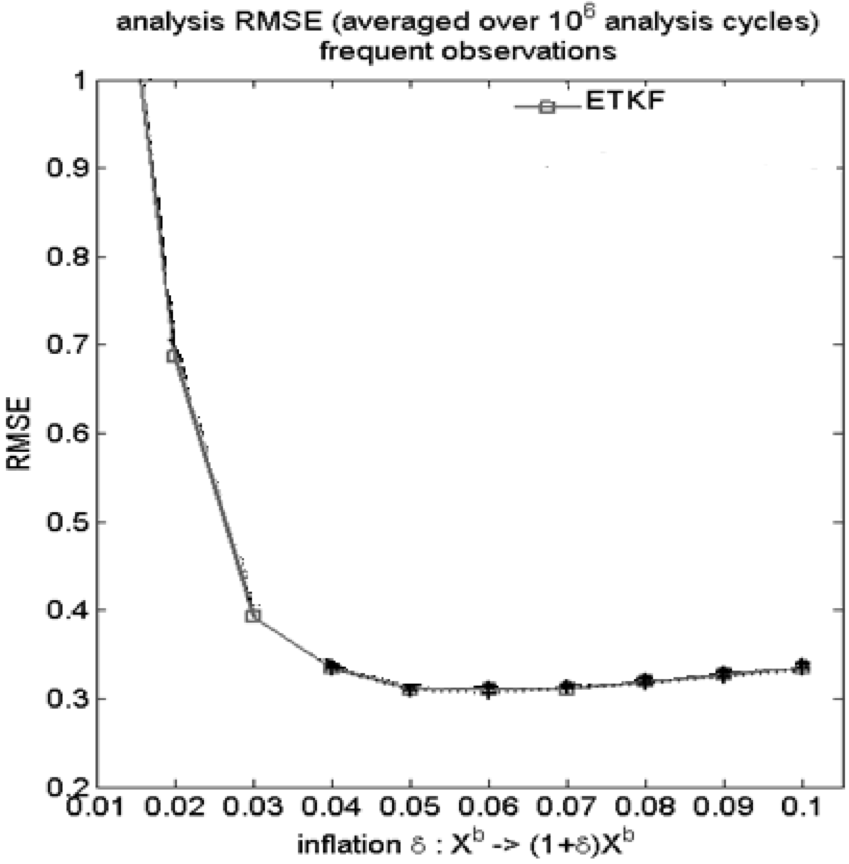
- Covariance inflation
- Covariance localization

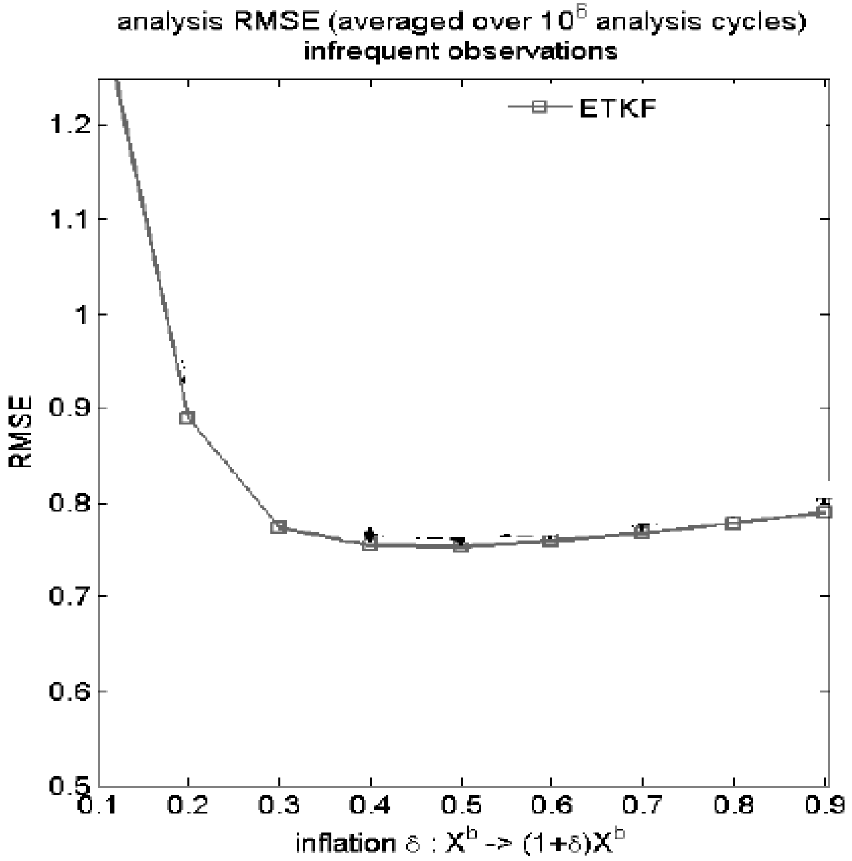
Also, the sample covariance matrix is singular for $N>M$…

**How many members** would we need? At least as many as the **number of unstable directions of error growth**?

# Covariance inflation and performance.

Lorenz 1963 $\mathbf{H}=\mathbf{I},\mathbf{R}=2\mathbf{I},M=3$




Amezcua et al, 2012.

# Covariance localization

- When forecast error covariance is mispecified (e.g., due to neglecting model error, or when $M << N$), it may include spurious correlations between very distant grid points.

- A common solution is to multiply each $\mathbf{P}^b$ element by an appropriate weight that reduces long-distance correlations.

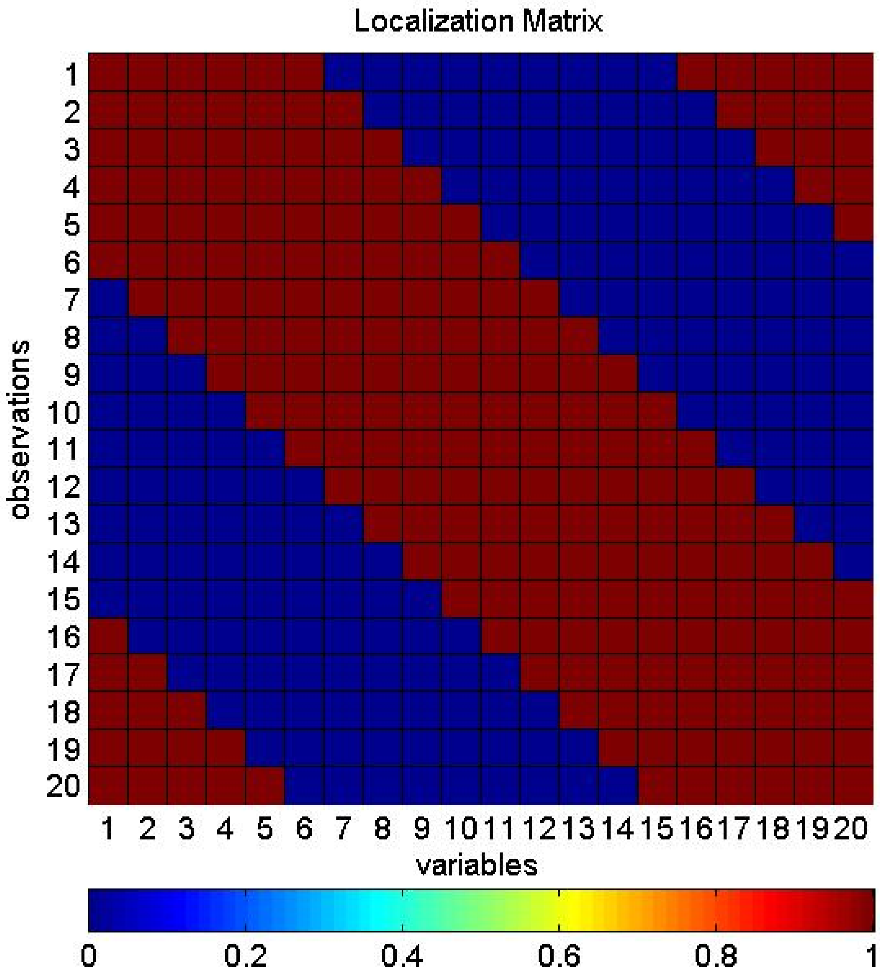
- This ensures that only the components of $\mathbf{P}^b$ believed to represent the corresponding components of $\mathbf{P}^b$ accurately are retained.

# Localization

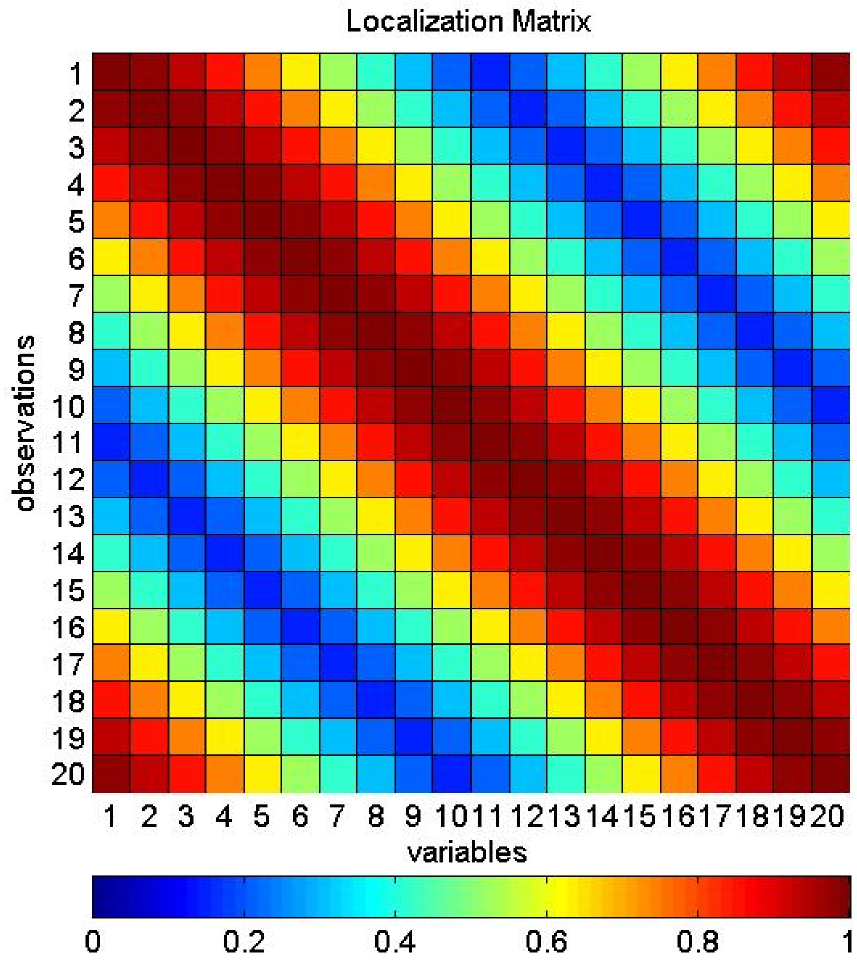
Example using Lorenz 1996

$$\mathbf{C} \circ \mathbf{P}^b$$

Cut-off



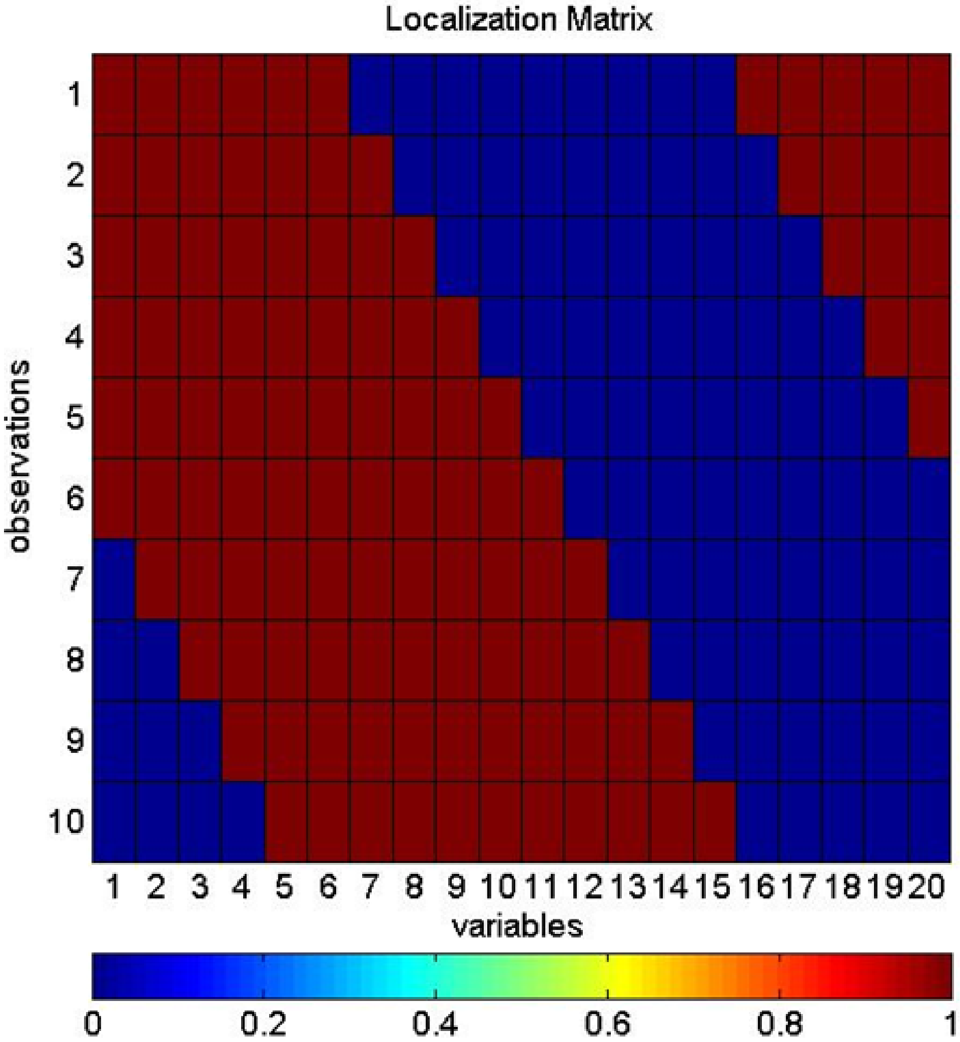

Gaspari-Cohn



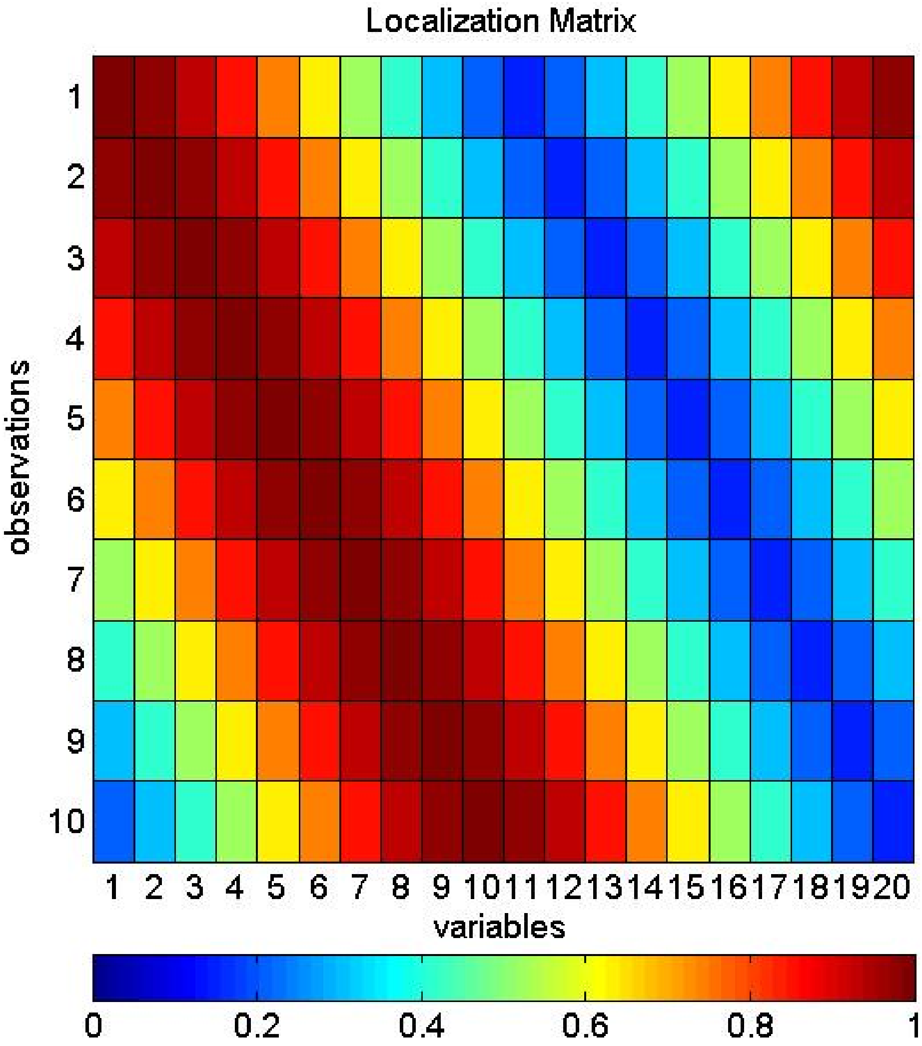
# Localization $\mathbf{C} \circ \left(\mathbf{P}^b \mathbf{H}^T\right)$

Example using Lorenz 1996, observing every other variable.

## Cut off

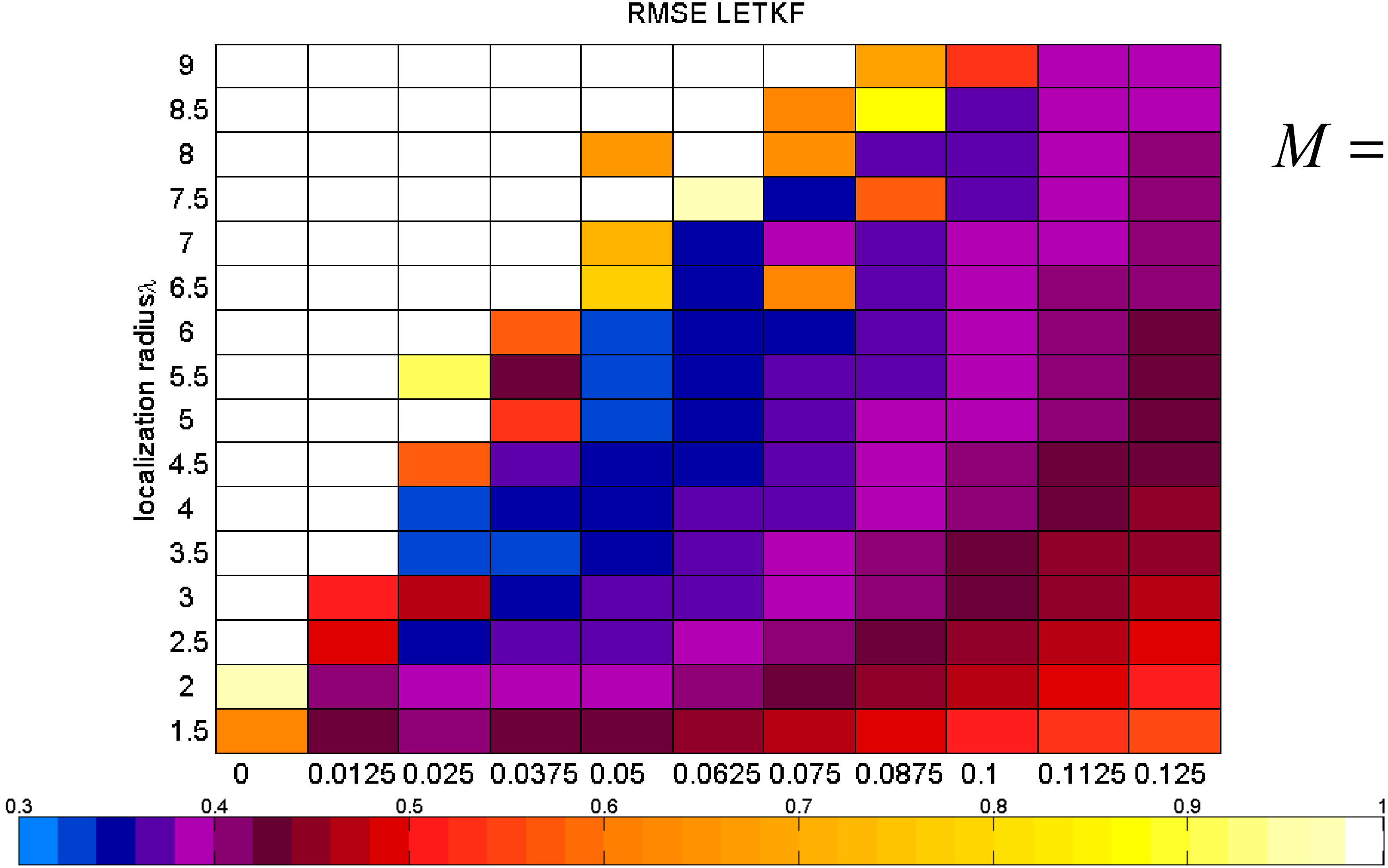



## Gaspari-Cohn



# Combined effects of inflation and localization

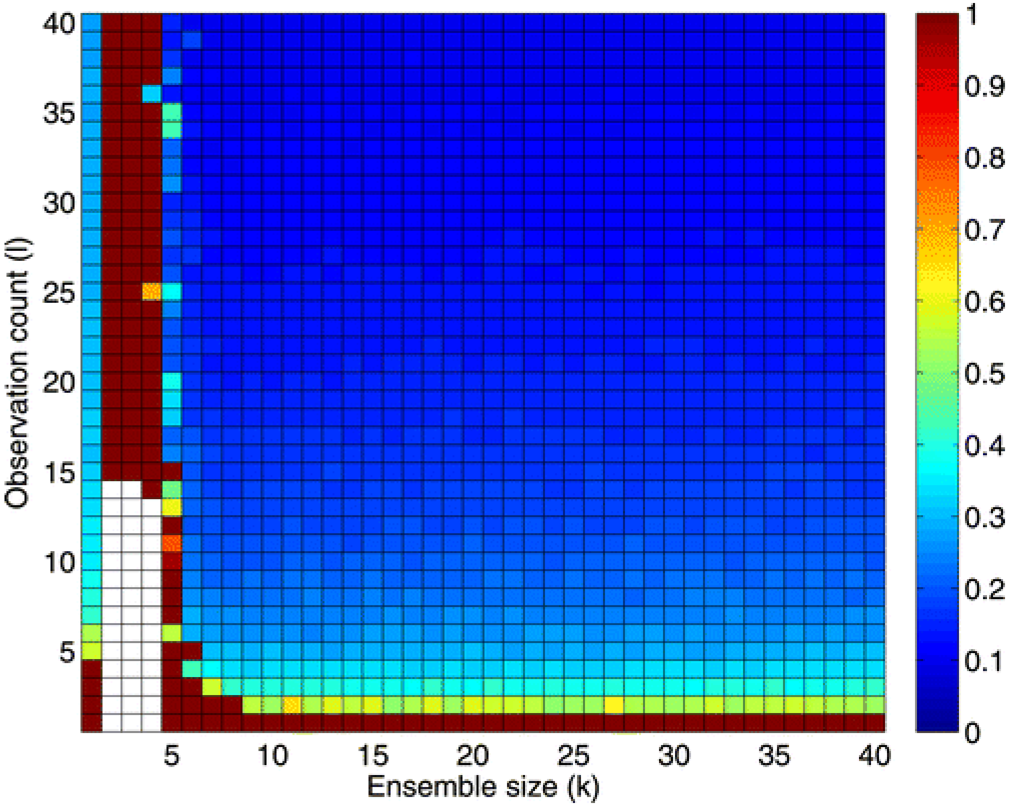
Experiments with Lorenz 1996 and 40 variables, observing every 2 time steps and every other variable.

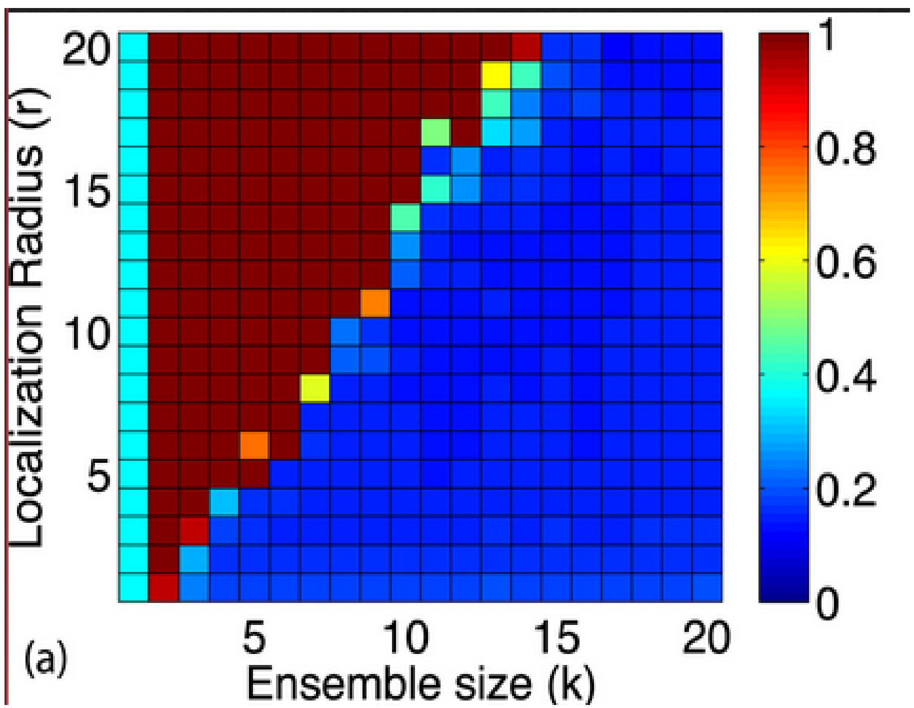



$$M = 10$$

Amezcua et al, 2012.

# Interactions of different parameters in the EnKF

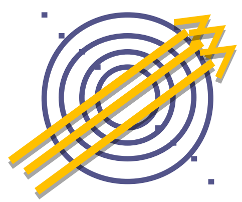



Penny, 2014

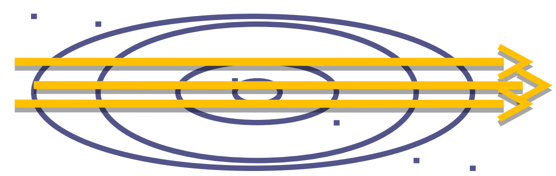
# Combining the best of 2 worlds?

A static covariance is full rank, it is invertible, it gives idea of the climatology.



Climatology

An ensemble covariance has information of the flow, but it can be singular and contains sampling errors.



Flow/State Dependence

$$\mathbf{B} = \alpha \mathbf{B}_{static} + (1-\alpha)\mathbf{B}_{ensemble}$$ → Compromise?

There are several ways to implement this.