EnKF Practical Considerations

Ross Bannister, r.n.bannister@reading.ac.uk

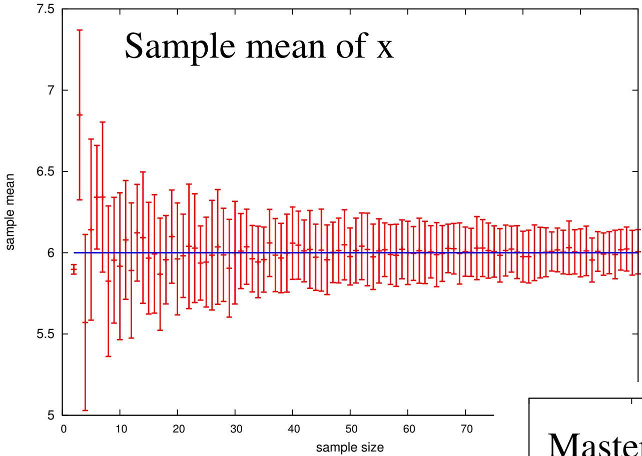
Variational DA vs. Ensemble KF

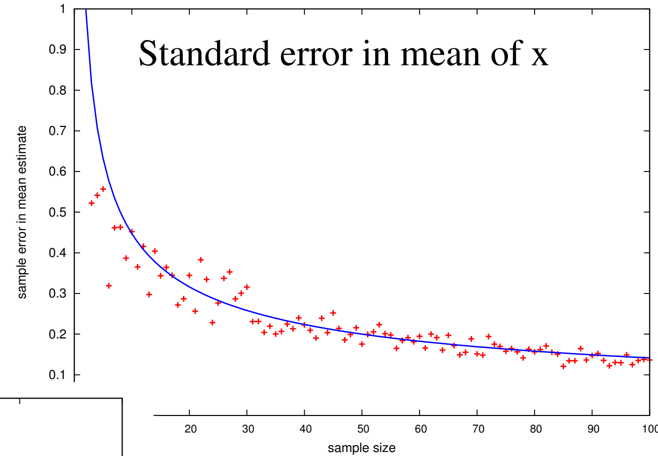
	Variational	**Bare ensemble KF**
Description	Minimize a cost function (maximum a-posteriori)	N ens. members of poss. background and analysis. Analyses derived from KF eqs.
Flavours	1D-Var, 3-D Var, 4-D Var (strong/weak const.)	Stochastic EnKF, Deterministic (square-root) forms, localized filters
Uncertainty	Respect obs and background uncertainty	Respect obs and background uncertainty
Stats	Gaussian	Gaussian
Operators	Allows weakly non-linear $\mathcal{M}$, $\mathcal{H}$ Allows direct and indirect observations Need $\mathbf{M}$, $\mathbf{H}$, and $\mathbf{M}^{\mathrm{T}}$, $\mathbf{H}^{\mathrm{T}}$	Allows weakly non-linear $\mathcal{M}$, $\mathcal{H}$ Allows direct and indirect observations Does not need $\mathbf{M}$, $\mathbf{H}$, and $\mathbf{M}^{\mathrm{T}}$, $\mathbf{H}^{\mathrm{T}}$
Obs types	Direct and indirect observations	Direct and indirect observations
A-priori error stats	$\mathbf{P}_{\mathrm{B}} \rightarrow \mathbf{B}$ (prescribed) $\mathbf{B}$ difficult to determine	$\mathbf{P}_{\mathrm{B}}$ adapts with flow (approx. from ens.) Initial ensemble difficult to determine Ens. tends to be under-spread (filter divergence)
Analysis	Smooth and balanced (according to $\mathbf{B}$) Analysis err. stats can be est. with extra procedures	Sampling noise in ens. leads to noisy analyses Appropriate balance properties Est. of analysis err stats from analysis ens.

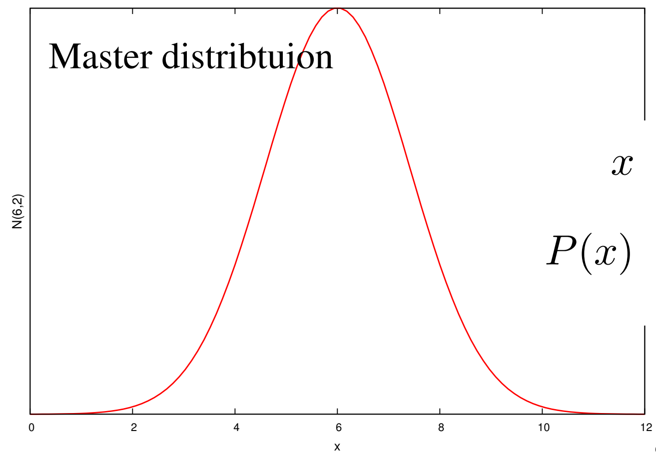
Lecture outline

- Sampling error for one variable
- Sampling error in the EnKF
- Ensemble spread
- Localization

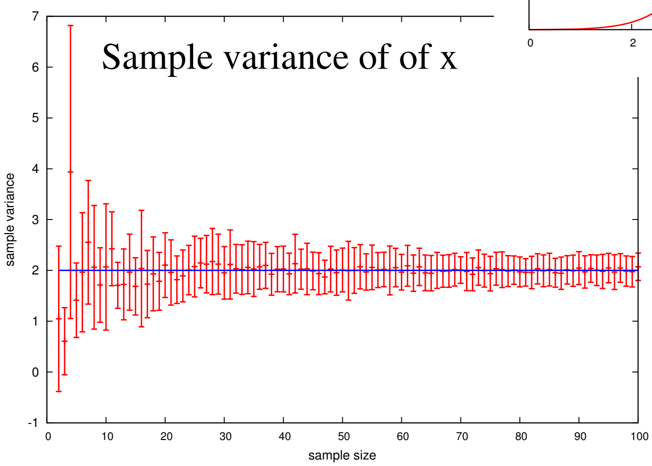
Sampling error for one variable

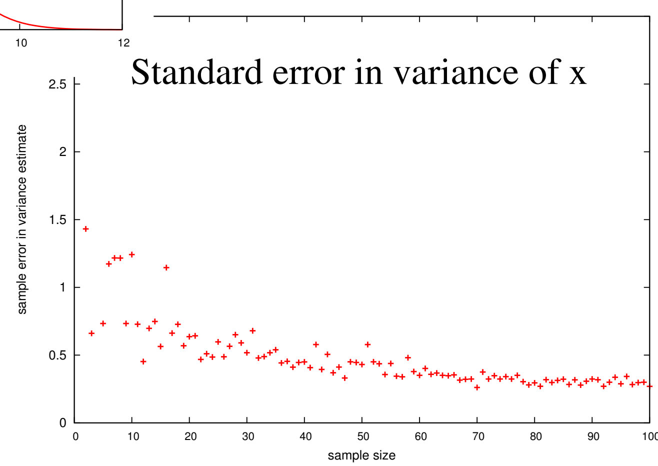






$$x \sim N(6,2),$$
$$P(x) = \frac{1}{4\pi} \exp - \frac{(x-6)^2}{4}.$$





Sampling error and covariances

Basic ensemble estimate of the forecast error covariance matrix:

1. Take ensemble analysis at $t = -T$ (N ensemble members stored in an $n \times N$ matrix):

$$\mathbf{X}_\mathrm{A}(-T) = \begin{pmatrix} \uparrow & & \uparrow \\ \mathbf{x}_\mathrm{A}^{(1)}(-T) & \cdots & \mathbf{x}_\mathrm{A}^{(N)}(-T) \\ \downarrow & & \downarrow \end{pmatrix}.$$

2. Propagate all members to $t = 0$ (with added noise to represent model error):

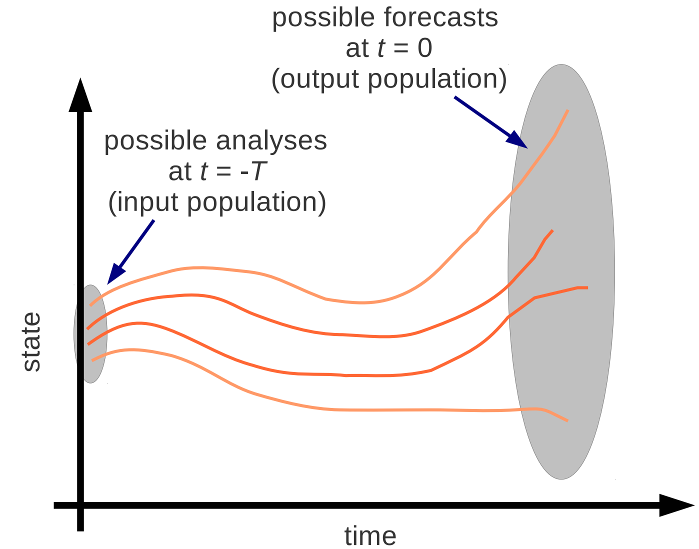
$$\mathbf{X}_\mathrm{B}(0) = \mathcal{M}\left[\mathbf{X}_\mathrm{A}(-T)\right] + \underline{\boldsymbol{\eta}} = \begin{pmatrix} \uparrow & & \uparrow \\ \mathbf{x}_\mathrm{B}^{(1)}(0) & \cdots & \mathbf{x}_\mathrm{B}^{(N)}(0) \\ \downarrow & & \downarrow \end{pmatrix}.$$

3. Calculate perturbations from the mean, $\mathbf{x}_\mathrm{B}^{\prime(i)} = \mathbf{x}_\mathrm{B}^{(i)} - \langle \mathbf{x}_\mathrm{B} \rangle$ (proxy for forecast errors):

$$\mathbf{X}'_\mathrm{B} = \mathbf{X}'_\mathrm{B}(0) = \begin{pmatrix} \uparrow & & \uparrow \\ \mathbf{x}_\mathrm{B}^{\prime(0)} & \cdots & \mathbf{x}_\mathrm{B}^{\prime(N)} \\ \downarrow & & \downarrow \end{pmatrix}.$$

4. Formula for the sample error covariance:

$$\mathbf{P}_\mathrm{B} \approx \mathbf{P}_\mathrm{B}^{(N)} = \frac{1}{N-1} \sum_{i=1}^{N} \mathbf{x}_\mathrm{B}^{\prime(i)} \mathbf{x}_\mathrm{B}^{\prime(i)\mathrm{T}}, \qquad = \frac{1}{N-1} \mathbf{X}'_\mathrm{B} \mathbf{X}_\mathrm{B}^{\prime\mathrm{T}}.$$



This matrix is not calculated explicitly for large systems, but we can use the formula to explore the consequences of $N \ll n$.

Sampling error and covariances (continued)

Reminder (example with stochastic formulation)

$$\begin{aligned}\text{Analysis increment formula for member } i: \mathbf{x}_{\mathrm{A}}^{(i)} - \mathbf{x}_{\mathrm{B}}^{(i)} &= \mathbf{P}_{\mathrm{B}}^{(N)}\mathbf{H}^{\mathrm{T}}\left(\mathbf{R}+\mathbf{H}\mathbf{P}_{\mathrm{B}}^{(N)}\mathbf{H}^{\mathrm{T}}\right)^{-1}\left(\mathbf{y}_{\mathrm{o}}^{(i)} - \mathcal{H}(\mathbf{x}_{\mathrm{B}}^{(i)})\right),\\ &= \mathbf{P}_{\mathrm{B}}^{(N)}\mathbf{v}^{(i)},\\ \text{Reminder: sample forecast error cov. matrix: } \mathbf{P}_{\mathrm{B}}^{(N)} &= \frac{1}{N-1}\sum_{i=1}^{N}\mathbf{x}_{\mathrm{B}}^{\prime(i)}\mathbf{x}_{\mathrm{B}}^{\prime(i)\mathrm{T}}.\end{aligned}$$

1. Analysis increments ($\mathbf{x}_{\mathrm{A}}^{(i)} - \mathbf{x}_{\mathrm{B}}^{(i)}$) lie in the subspace of the forecast error ensemble

Approximate $\mathbf{P}_{\mathrm{B}}$ with $\mathbf{P}_{\mathrm{B}}^{(N)}$ in the analysis increment formula:

$$\begin{aligned}\mathbf{x}_{\mathrm{A}}^{(i)} - \mathbf{x}_{\mathrm{B}}^{(i)} &\approx \frac{1}{N-1}\sum_{i=1}^{N}\mathbf{x}_{\mathrm{B}}^{\prime(i)}\mathbf{x}_{\mathrm{B}}^{\prime(i)\mathrm{T}}\mathbf{v}^{(i)},\\ &\approx \frac{1}{N-1}\sum_{i=1}^{N}\mathbf{x}_{\mathrm{B}}^{\prime(i)}\alpha^{(i)},\\ \text{where } \alpha^{(i)} &= \mathbf{x}_{\mathrm{B}}^{\prime(i)\mathrm{T}}\mathbf{v}^{(i)} = \mathbf{x}_{\mathrm{B}}^{\prime(i)}\cdot\mathbf{v}^{(i)}.\end{aligned}$$

Even if the observations indicate otherwise, the analysis increments are restricted to be a linear combination of the forecast error ensemble. This is a low-dimensional space ($N-1$ dimensions at most).

2. The forecast error covariance matrix is rank deficient

The rank of $\mathbf{P}_{\mathrm{f}}^{(N)}$ is an indication of the size of the state space spanned by the forecast error ensemble.

$$\mathrm{rank}\left(\mathbf{P}_{\mathrm{B}}^{(N)}\right) \leq N - 1.$$

This is a guide to the severity of the sampling problem in point 1.

3. The forecast ensemble spread will be subject to sampling error

- If the spread is too large then the analysis ens. will over-fit the obs. - too little trust in the fc. ens.
- If the spread is too small then the analysis ens. will under-fit the obs. - too much trust in the fc. ens.
  - Once in this regime, it is difficult to escape as the ens. will (effectively) ignore the obs..
  - This is called "filter divergence" (because we diverge from reality).

Filter divergence means that each ensemble member will (effectively) be free running.

4. The correlations will be subject to sampling error

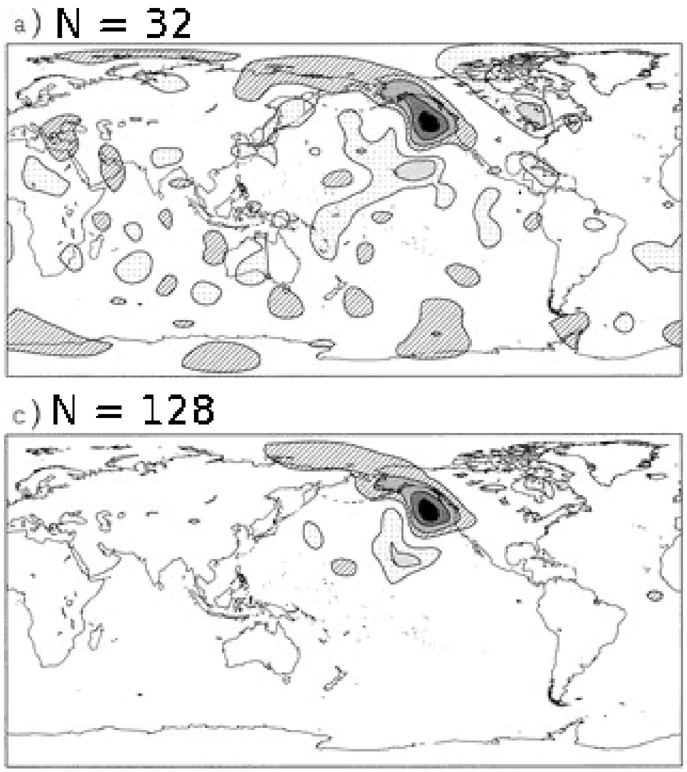
- The error in the sample correlation between errors at locations i and j has expectation:

$$[\mathcal{E}(\delta \mathbf{C}_{\mathrm{B}}^{(N)})]_{ij} \sim \frac{1}{\sqrt{N}} \left(1 - ([\mathbf{C}_{\mathrm{B}}]_{ij})^2\right),$$

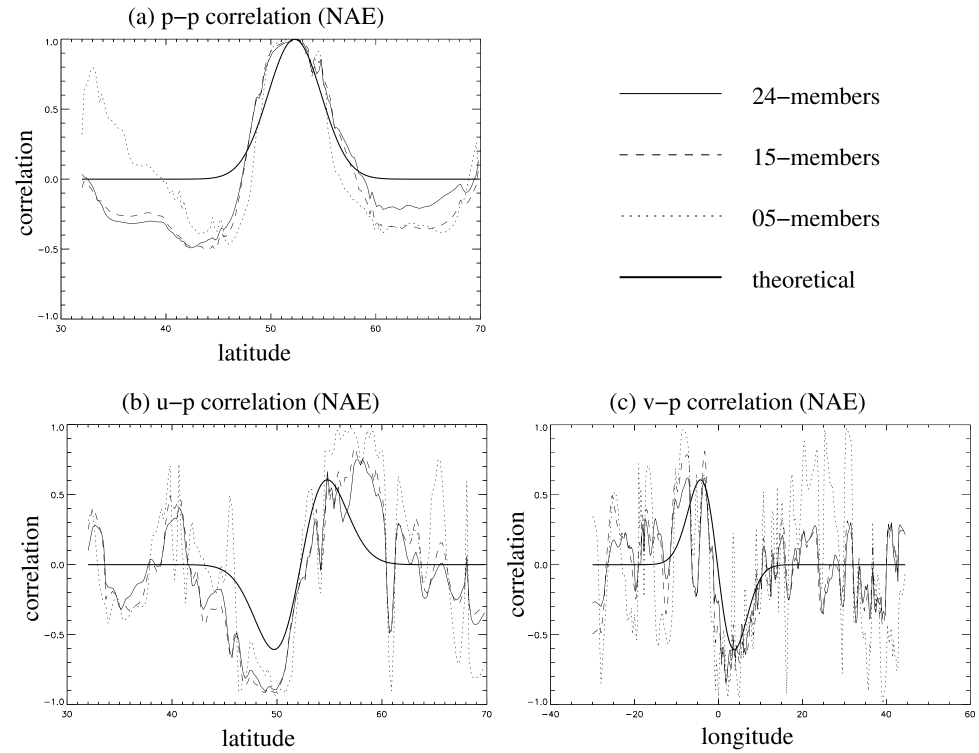
  (errors are expected to be large when N small and/or correlations are close to zero).

- Pairs of distant points would be expected to have correlations close to zero.

Sampling error means that we can't trust distant correlations. Left untreated this noise will destroy the benefits of DA (analysis increments will be influenced by distant observations).



From Houtekamer & Mitchell (1998)

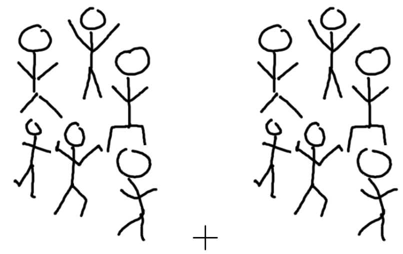


From Bannister, Migliorini & Dixon (2011)

Making progress

What can be done to reduce/mitigate this problem $N \ll n$?

- **Use more ensemble members**.
  - This is expensive.
  - How many is 'enough'?



- **Ensemble inflation**.
  - Artificially increase the size of each $\mathbf{x}_{\mathrm{B}}'^{(i)}$.
  - How do we know what the ensemble spread should be?

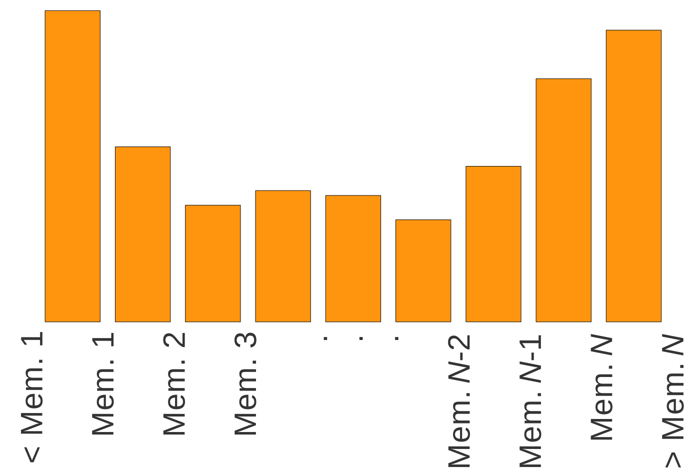
- **Localization**.
  - Eliminate far-field correlations.
  - How should this be done?
  - Does this have any other consequences?

- **Combine ensemble with variational approaches**.
  - Adopt a hybrid method.
  - How to do this? See next lecture.

Ensemble inflation: how do we know what the ens spread should be?

Method 1: Rank histograms (Talagrand diagrams)

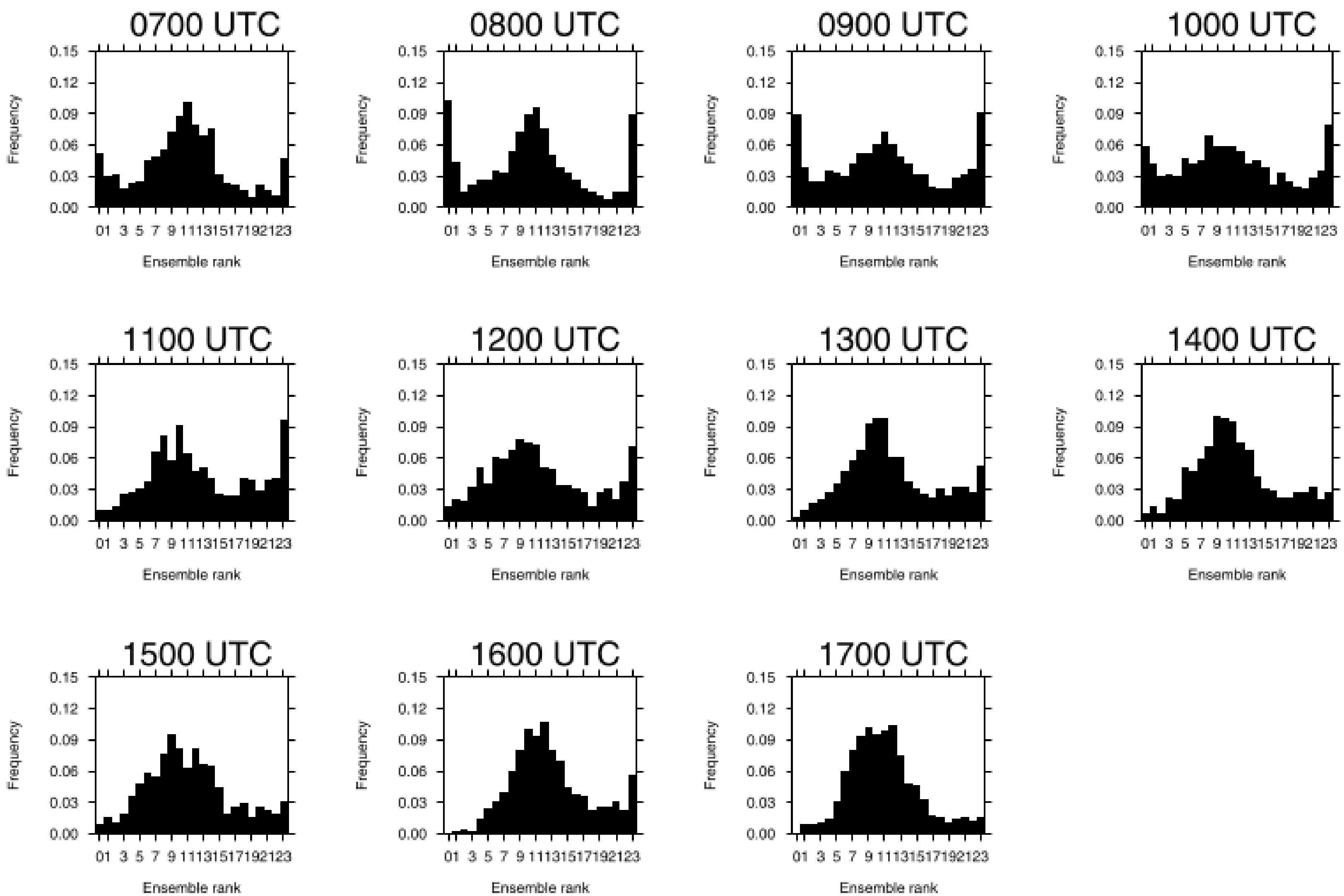
- Each ensemble member should be equally likely.
- Consider a point in space that has many observations:
  - Rank values of ensemble members at that point from lowest to highest ($N-1$ bins). Add an extra bin at each end to give $N+1$ bins.
  - Bin each observation to give a frequency histogram.
- Interpretation:
  - ∪-shaped: the ensemble is under-spread.
  - ∩-shaped: the ensemble is over-spread.
  - Flat: the ensemble is correctly spread.
  - Asymmetric: the ensemble is biased.



N.B. When observational error is significant, noise needs to first be added to each ensemble member (noise is sampled from the observation error distribution). Otherwise even a correctly distributed ensemble would look under-spread. See Hamill (2001).

Ensemble inflation: how do we know what the ens spread should be?

Example rank histograms



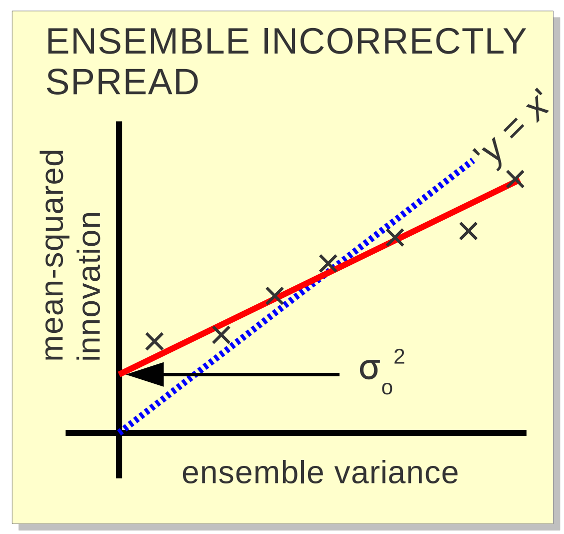
Rank histograms for surface precipitation rate rate. From Migliorini et al. (2011).

Ensemble inflation: how do we know what the ens spread should be?

Method 2: Spread/skill diagrams

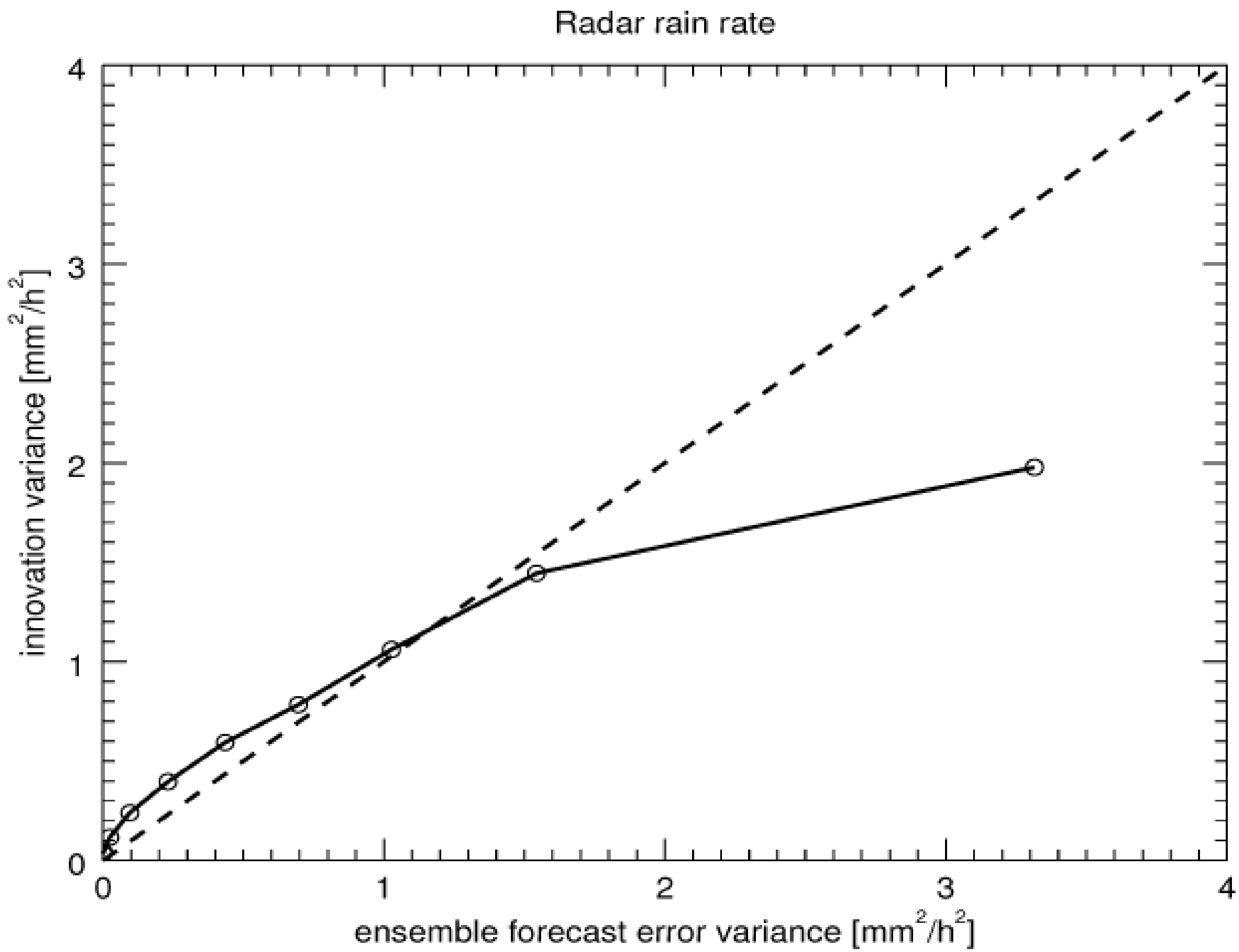
Suppose that we have an ensemble of N forecasts, $x_{(i)}$, and a large number of observations, y_j.

1. For each ob (ob index j):
   (a) Calculate ens mean model-ob:
   $$\frac{1}{N}\sum_{i'=1}^{N} H_j(x_{(i')}).$$
   (b) Calculate ens variance:
   $$\sigma^2_{\text{ens},j} = \frac{1}{N-1}\sum_{i=1}^{N}\left(H_j(x_{(i)}) - \frac{1}{N}\sum_{i'=1}^{N} H_j(x_{(i')})\right)^2.$$
   (c) Calculate the ens mean-square innovation:
   $$d_j^2 = \frac{1}{N}\sum_{i=1}^{N}\left(y_j - H_j(x_{(i)})\right)^2.$$
2. Chose a number of bins spanning the range of ens variance values. Let there be M obs per bin. Each ob (and the associated ens variance and innovation) is associated with a given bin.
3. For each bin (bin index k):
   (a) Calculate the mean ens variance averaged over the bin:
   $$\sigma^2_{\text{ens}}(k) = \frac{1}{M}\sum_{j\in\text{bin } k}\sigma^2_{\text{ens},j}.$$
   (b) Calculate the mean-square innovation averaged over the bin:
   $$\sigma^2_{\text{innov}}(k) = \frac{1}{M}\sum_{j\in\text{bin } k} d_j^2.$$
4. Assuming that N is sufficiently large, ensemble is unbiased, obs and ensemble errors are uncorrelated, all obs have same error stats (σ_o^2), and ensemble correctly spread:
   $$\sigma^2_{\text{innov}}(k) = \sigma_o^2 + \sigma^2_{\text{ens}}(k).$$
5. Plot $\sigma^2_{\text{innov}}(k)$ (the 'skill') against $\sigma^2_{\text{ens}}(k)$ (the 'spread').
6. Can use this information to derive an *inflation factor*. See Wang and Bishop (2003).



Ensemble inflation: how do we know what the ens spread should be?

Example spread/skill diagram



From Baker et al. (2014).

How to increase the spread of an ensemble

- Multiplicative inflation - multiply all ensemble members by a constant > 1. Hamill et al. (2000).
- Additive inflation - add noise to each member - like adding model error in the EnKF during the forecast stage.
  - Multi-physics - use a combination of different physics schemes. Berner et al. (2011).
  - Stochastic kinetic energy backscatter - add KE lost due to unresolved processes. Berner et al. (2009).
  - Stochastically perturbed physical tendencies - perturb physics increments. Bouttier et al. (2012).
  - Perturb model parameters - use different settings for model parameters. Bowler et al. (2008).
- Relax slightly to background - analysis ensemble has lower spread than background ensemble, so relaxing will increase spread. Zhang et al. (2004).

Degree of inflation may be prescribed or tuned to achieve correct spread.

Localization

Reminder

$$\text{Observation error cov. matrix:} \quad \mathbf{R}$$

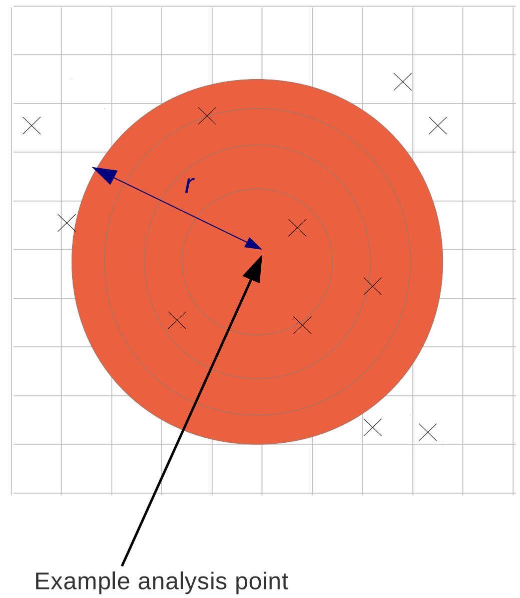
$$\text{Sample forecast error cov. matrix:} \quad \mathbf{P}_{\mathrm{B}}^{(N)} = \frac{1}{N-1}\sum_{i=1}^{N} \mathbf{x}_{\mathrm{B}}'^{(i)} \mathbf{x}_{\mathrm{B}}'^{(i)\mathrm{T}}.$$

Many ways of doing localization, e.g.:

- $\mathbf{R}$-localization
  - Restrict the observations that are allowed to influence each grid-point.
  - Used in the LETKF.
- $\mathbf{P}_{\mathrm{B}}$-localization
  - Modify $\mathbf{P}_{\mathrm{B}}^{(N)}$ with a localization/moderation function that decreases with separation.

R-localization:

- Perform a separate ens analysis at each grid point.
- Include obs inside a defined radius. Multiply obs error variance by a weight, $\rho > 1$ (increases with distance).
- Used in the LETKF.
- Difficult to use for non-local observations.



R-localization in the LETKF

The Localized Ensemble Transform Kalman Filter is a square-root formulation.

$$\begin{aligned}
\text{mean anal.: } \bar{\mathbf{x}}_\mathrm{A} &= \bar{\mathbf{x}}_\mathrm{B} + \mathbf{P}_\mathrm{A}^{(N)} \mathbf{H}^\mathrm{T} \mathbf{R}^{-1} \left(\mathbf{y} - \mathbf{H} \bar{\mathbf{x}}_\mathrm{B} \right) \in \mathbb{R}^n, \\
&= \bar{\mathbf{x}}_\mathrm{B} + \Delta \bar{\mathbf{x}}_\mathrm{A}, \\
\text{anal. err. cov.: } \mathbf{P}_\mathrm{A}^{(N)} &= \left(\mathbf{P}_\mathrm{B}^{(N)^{-1}} + \mathbf{H}^\mathrm{T} \mathbf{R}^{-1} \mathbf{H} \right)^{-1}, \\
&= \frac{1}{N-1} \mathbf{X}'_\mathrm{A} \mathbf{X}'^\mathrm{T}_\mathrm{A} \in \mathbb{R}^{n \times n}, \\
\text{bg. err. cov.: } \mathbf{P}_\mathrm{B}^{(N)} &= \frac{1}{N-1} \mathbf{X}'_\mathrm{B} \mathbf{X}'^\mathrm{T}_\mathrm{B} \in \mathbb{R}^{n \times n}, \\
\text{perts to mean anal.: } \mathbf{X}'_\mathrm{A} &= \sqrt{N-1}\, \mathbf{P}_\mathrm{A}^{(N)^{1/2}} \in \mathbb{R}^{n \times N}.
\end{aligned}$$

Describe all perts. as a linear combination of background perts:

$$\begin{aligned}
\text{in general: } \mathbf{X}' &= \mathbf{X}'_\mathrm{B} \mathbf{W} \qquad \mathbf{X}', \mathbf{X}'_\mathrm{B} \in \mathbb{R}^{n \times N},\ \mathbf{W} \in \mathbb{R}^{N \times N}, \\
\text{for bg. pert. ens.: } \mathbf{X}'_\mathrm{B} &= \mathbf{X}'_\mathrm{B} \mathbf{W}_\mathrm{B} \qquad \mathbf{W}_\mathrm{B} = \mathbf{I}_N, \\
\text{for anal. pert. ens.: } \mathbf{X}'_\mathrm{A} &= \mathbf{X}'_\mathrm{B} \mathbf{W}_\mathrm{A}.
\end{aligned}$$

Can transform the analysis equations to $\mathbf{W}$-space with the following:

$$\begin{aligned}
\mathbf{P}_\mathrm{A}^{(N)} &\rightarrow \widetilde{\mathbf{P}_\mathrm{A}^{(N)}} = \frac{1}{N-1} \mathbf{W}_\mathrm{A} \mathbf{W}_\mathrm{A}^\mathrm{T} \in \mathbb{R}^{N \times N}, \\
\mathbf{P}_\mathrm{B}^{(N)} &\rightarrow \widetilde{\mathbf{P}_\mathrm{B}^{(N)}} = \frac{1}{N-1} \mathbf{W}_\mathrm{B} \mathbf{W}_\mathrm{B}^\mathrm{T} = \frac{1}{N-1} \mathbf{I}_N \in \mathbb{R}^{N \times N}, \\
\mathbf{H} &\rightarrow \widetilde{\mathbf{H}} = \mathbf{H} \mathbf{X}'_\mathrm{B} \in \mathbb{R}^{p \times N}, \\
\mathbf{R} &\rightarrow \widetilde{\mathbf{R}} = \boldsymbol{\rho} \mathbf{R} \boldsymbol{\rho} \in \mathbb{R}^{p \times p}.
\end{aligned}$$

The relevant update equations in $\mathbf{W}$-space become:

$$\begin{aligned}
\text{mean anal. update: } \widetilde{\Delta\bar{\mathbf{x}}_{\mathrm{A}}} &= \widetilde{\mathbf{P}_{\mathrm{A}}^{(N)}}\widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{R}}^{-1}\left(\mathbf{y}-\mathbf{H}\bar{\mathbf{x}}_{\mathrm{B}}\right) \in \mathbb{R}^{N}, \\
\text{anal. err. cov.: } \widetilde{\mathbf{P}_{\mathrm{A}}^{(N)}} &= \left((N-1)\mathbf{I}_N + \widetilde{\mathbf{H}}^{\mathrm{T}}\widetilde{\mathbf{R}}^{-1}\widetilde{\mathbf{H}}\right)^{-1} \in \mathbb{R}^{N\times N}, \\
\text{perts to mean anal.: } \mathbf{W}_{\mathrm{A}} &= \sqrt{N-1}\widetilde{\mathbf{P}_{\mathrm{A}}^{(N)}}^{1/2} \in \mathbb{R}^{N\times N}.
\end{aligned}$$

The analysis states in model space are then:

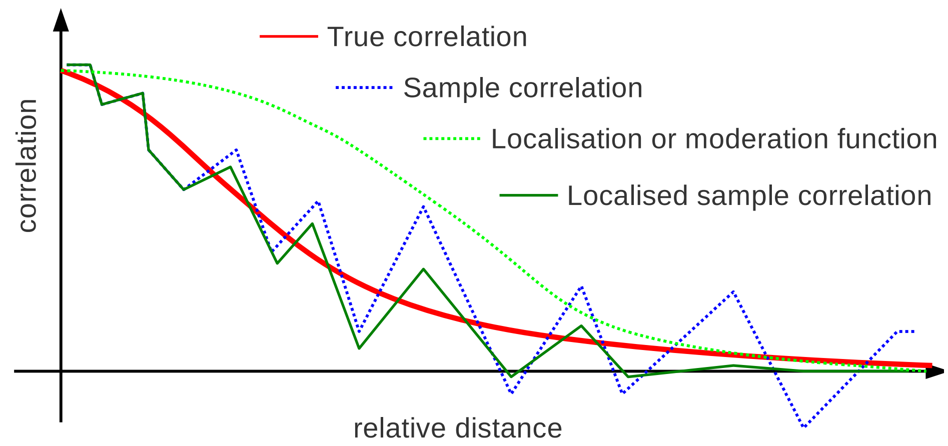
$$\begin{aligned}
\text{mean anal.: } \bar{\mathbf{x}}_{\mathrm{A}} &= \bar{\mathbf{x}}_{\mathrm{A}} + \mathbf{X}'_{\mathrm{B}}\widetilde{\Delta\bar{\mathbf{x}}_{\mathrm{A}}}, \\
\text{anal. perts: } \mathbf{X}'_{\mathrm{A}} &= \mathbf{X}'_{\mathrm{B}}\mathbf{W}_{\mathrm{A}}.
\end{aligned}$$

E&OE

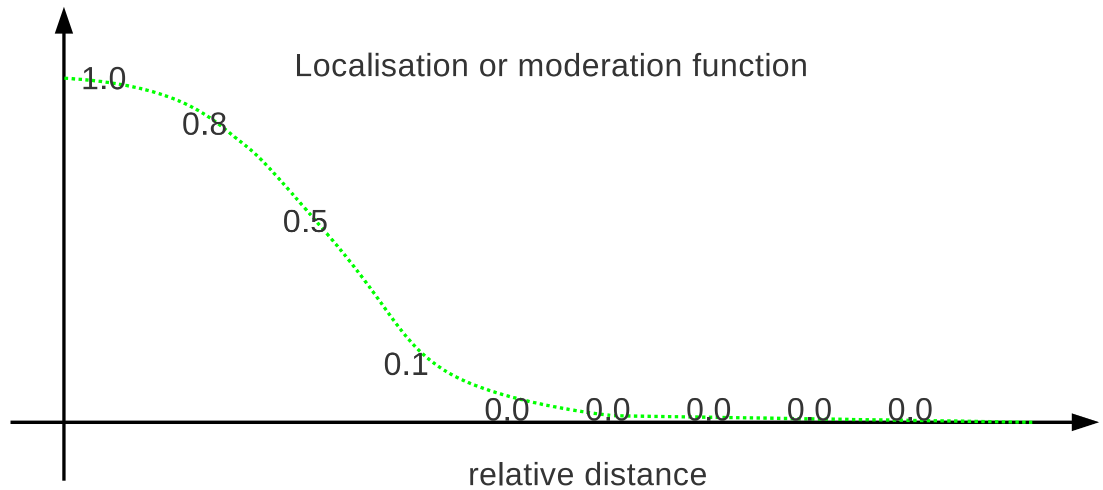
- The key aspect of the LETKF is that we may allow only the weights in $\mathbf{W}_{\mathrm{A}}$ and $\widetilde{\Delta\bar{\mathbf{x}}_{\mathrm{A}}}$ to update one point in space (or vertical column) that was used to determine the $\boldsymbol{\rho}$ function.
- A different set of weights can be determined for other points (with a different $\boldsymbol{\rho}$ function).
- For efficiency the observations used for a particular point can be restricted to a cut-off radius.
- See Hunt et al. (2007).

$\mathbf{P}_\mathrm{B}$-localization

- Modify $\mathbf{P}_\mathrm{B}^{(N)}$ with a localization/moderation function that decreases with separation.
- What length-scale? How to do multivariate aspects?
- Has side effects (e.g. affects length-scales, affects balance).



$\mathbf{P}_{\mathrm{B}}$-localization (univariate)



$$
\begin{aligned}
\mathbf{P}_{\mathrm{B}}^{\mathrm{Loc}} &= \mathbf{P}_{\mathrm{B}} \circ \boldsymbol{\Omega}, \\
&= \begin{pmatrix}
P_{\mathrm{B}11}^{(N)} & P_{\mathrm{B}12}^{(N)} & \cdots & \cdots & P_{\mathrm{B}15}^{(N)} & \cdots & \cdots & P_{\mathrm{B}18}^{(N)} & P_{\mathrm{B}19}^{(N)} \\
P_{\mathrm{B}21}^{(N)} & P_{\mathrm{B}22}^{(N)} & \cdots & \cdots & P_{\mathrm{B}25}^{(N)} & \cdots & \cdots & P_{\mathrm{B}28}^{(N)} & P_{\mathrm{B}29}^{(N)} \\
\vdots & \vdots & \ddots & \cdots & P_{\mathrm{B}35}^{(N)} & \cdots & \cdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & P_{\mathrm{B}45}^{(N)} & \vdots & \vdots & \vdots & \vdots \\
P_{\mathrm{B}51}^{(N)} & P_{\mathrm{B}52}^{(N)} & P_{\mathrm{B}53}^{(N)} & P_{\mathrm{B}54}^{(N)} & P_{\mathrm{B}55}^{(N)} & P_{\mathrm{B}56}^{(N)} & P_{\mathrm{B}57}^{(N)} & P_{\mathrm{B}58}^{(N)} & P_{\mathrm{B}59}^{(N)} \\
\vdots & \vdots & \vdots & \vdots & P_{\mathrm{B}65}^{(N)} & \ddots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \cdots & \cdots & P_{\mathrm{B}75}^{(N)} & \cdots & \ddots & \vdots & \vdots \\
P_{\mathrm{B}81}^{(N)} & P_{\mathrm{B}82}^{(N)} & \cdots & \cdots & P_{\mathrm{B}85}^{(N)} & \cdots & \cdots & P_{\mathrm{B}88}^{(N)} & P_{\mathrm{B}89}^{(N)} \\
P_{\mathrm{B}91}^{(N)} & P_{\mathrm{B}92}^{(N)} & \cdots & \cdots & P_{\mathrm{B}95}^{(N)} & \cdots & \cdots & P_{\mathrm{B}98}^{(N)} & P_{\mathrm{B}99}^{(N)}
\end{pmatrix}
\circ
\begin{pmatrix}
1.0 & 0.8 & 0.5 & 0.1 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.8 & 1.0 & 0.8 & 0.5 & 0.1 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.5 & 0.8 & 1.0 & 0.8 & 0.5 & 0.1 & 0.0 & 0.0 & 0.0 \\
0.1 & 0.5 & 0.5 & 1.0 & 0.8 & 0.5 & 0.1 & 0.0 & 0.0 \\
0.0 & 0.1 & 0.1 & 0.8 & 1.0 & 0.8 & 0.5 & 0.1 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.5 & 0.8 & 1.0 & 0.8 & 0.5 & 0.1 \\
0.0 & 0.0 & 0.0 & 0.1 & 0.5 & 0.8 & 1.0 & 0.8 & 0.5 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.1 & 0.5 & 0.8 & 1.0 & 0.8 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.1 & 0.5 & 0.8 & 1.0
\end{pmatrix}, \\
P_{\mathrm{B}ij}^{\mathrm{Loc}} &= P_{\mathrm{B}ij}\Omega_{ij}.
\end{aligned}
$$

Can be extended to multivariate localization. But - we rarely have access to explicit $\mathbf{P}_{\mathrm{B}}$ or $\boldsymbol{\Omega}$ matrices $(n \times n)$.

Localization without explicit $\mathbf{P}_\mathrm{B}$ and $\boldsymbol{\Omega}$ matrices

$$
\begin{aligned}
\text{Sample forecast error cov. matrix from } N \text{ members}: \mathbf{P}_\mathrm{B}^{(N)} &= \frac{1}{N-1}\sum_{l=1}^{N} \mathbf{x}_\mathrm{B}^{\prime(l)} \mathbf{x}_\mathrm{B}^{\prime(l)\mathrm{T}}, \\
\text{Sample localisation/moderation matrix from } K \text{ members}: \boldsymbol{\Omega}^{(K)} &= \frac{1}{K-1}\sum_{k=1}^{K} \boldsymbol{\omega}^{(k)} \boldsymbol{\omega}^{(k)\mathrm{T}}, \\
\text{One matrix element: } P_{\mathrm{B}ij}^{(N)} &= \frac{1}{N-1}\sum_{l=1}^{N} x_{\mathrm{B}i}^{\prime(l)} x_{\mathrm{B}j}^{\prime(l)}, \\
\text{One matrix element: } \Omega_{ij}^{(K)} &= \frac{1}{K-1}\sum_{k=1}^{K} \omega_i^{(k)} \omega_j^{(k)}.
\end{aligned}
$$

$$
\begin{aligned}
P_{\mathrm{B}ij}^{\mathrm{Loc}} &= P_{\mathrm{B}ij}^{(N)} \Omega_{ij}^{(K)}, \\
&= \left[\frac{1}{N-1}\sum_{l=1}^{N} x_{\mathrm{B}i}^{\prime(l)} x_{\mathrm{B}j}^{\prime(l)}\right]\left[\frac{1}{K-1}\sum_{k=1}^{K} \omega_i^{(k)} \omega_j^{(k)}\right], \\
&= \frac{1}{N-1}\frac{1}{K-1}\sum_{l=1}^{N}\sum_{k=1}^{K} \underbrace{x_{\mathrm{B}i}^{\prime(l)}\omega_i^{(k)}}_{\substack{\text{element } i \\ \text{of } \tilde{\mathbf{x}}^{\prime(m)}}} \quad \underbrace{x_{\mathrm{B}j}^{\prime(l)}\omega_j^{(k)}}_{\substack{\text{element } j \\ \text{of } \tilde{\mathbf{x}}^{\prime(m)}}}, \\
&= \frac{1}{N-1}\frac{1}{K-1}\sum_{m=1}^{M} \tilde{x}_i^{\prime(m)} \tilde{x}_j^{\prime(m)}, \qquad M = NK,
\end{aligned}
$$

$$
\tilde{\mathbf{x}}^{\prime(m)} = \mathbf{x}_\mathrm{B}^{\prime(l)} \circ \boldsymbol{\omega}^{(k)} = \begin{pmatrix} x_{\mathrm{B}1}^{\prime(l)} \\ \vdots \\ x_{\mathrm{B}n}^{\prime(l)} \end{pmatrix} \circ \begin{pmatrix} \omega_1^{(k)} \\ \vdots \\ \omega_n^{(k)} \end{pmatrix} = \begin{pmatrix} x_{\mathrm{B}1}^{\prime(l)}\omega_1^{(k)} \\ \vdots \\ x_{\mathrm{B}n}^{\prime(l)}\omega_n^{(k)} \end{pmatrix},
$$

$$
\begin{aligned}
m = 1 &\Rightarrow l = 1, k = 1 \\
m = 2 &\Rightarrow l = 1, k = 2 \\
\vdots & \qquad \vdots \\
m = K &\Rightarrow l = 1, k = K \\
m = K+1 &\Rightarrow l = 2, k = 1 \\
\vdots & \qquad \vdots \\
m = M &\Rightarrow l = N, k = K
\end{aligned}
$$

See Buehner (2005).

Summary

- Ensemble data assimilation schemes suffer from sampling error as $N \ll n$:
  - Analysis increments lie in subspace of ensemble.
  - Rank deficiency.
  - Filter divergence.
  - Anomalous far-field correlations.
- To make ensemble DA practical:
  - Ensemble inflation.
  - Localization.
  - Use with other schemes (hybrid - next lectures).

Ensemble KF vs variational data assimilation, Schur product localization

- Lorenc A.C., The potential of the ensemble Kalman filter for NWP - a comparison with 4d-Var, Q.J.R. Meteor. Soc. 129, 3183-3203 (2003).
- Ehrendorfer M., A review of issues in ensemble-based Kalman filtering, Meteorol. Z. 16, 795-818 (2007).

Impact of sampling error

- Houtekamer P.L., Mitchell H.L., Data assimilation using an ensemble Kalman Filter technique, Mon. Wea. Rev. 126, 796-811 (1998).
- Hamill T.M., Whitaker J.S., Synder C., Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filer, Mon. Wea. Rev. 129, 2776-2790 (2001).

Inflation and localization

- Bowler N.E., Arribas A., Mylne K.R., Robertson K.B., Beare S.E., The MOGREPS short-range ensemble prediction system, Q.J.R.Meteor.Soc. 134, 703-722 (2008).
- Berner J., Ha S.-Y., Hacker J.P., Fournier A., Snyder C., Model uncertainty in a mesoscale ensemble prediction system: stochastic versus multiphysics representations, Mon. Wea. Rev. 139, 1972-1995 (2011).
- Berner J., Schutts G.J., Leutbecher M., Palmer T.N., A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF Ensemble Prediction System, J. Atmos. Sci. 66, 602-629 (2009).
- Bouttier F., Vié B., Nuissier O., Raynaud L., Impact of stochastic physics in a convection-permitting ensemble, Mon. Wea. Rev. 140, 3706-3721 (2012).
- Zhang F., Snyder C., Sun J., Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter, Mon. Wea. Rev. 132, 1238-1253 (2004).
- Hunt B.R., Kostelich E.J., Szunyogh I., Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman Filter, Physica D 230, 112-126 (2007).
- Buehner M., Ensemble derived stationary and flow dependent background error covariances: Evaluation in a quasi-operational NWP setting, Q.J.R.Meteor.Soc. 131 pp.1013-1043 (2005).

Forecast verification

- www.cawcr.gov.au/projects/verification
- Hamill T.M., Interpretation of Rank Histograms for Verifying Ensemble Forecasts, Mon. Wea. Rev. 129., 550-560 (2001).
- Wang X., Bishop C.H., A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes, J. Atmos. Sci. 60, 1140-1158 (2003).
- Baker L.H., Rudd A.C., Migliorini S., Bannister R.N., Representation of model error in a convective-scale ensemble prediction system, Nonlin. Processes Geophys. 21, 19-39 (2014).
- Migliorini S., Dixon M., Bannister R.N., Ballard, S., Ensemble prediction for nowcasting with a convection-permitting model - I: description of the system and the impact of radar-derived surface precipitation rates, Tellus 63A, 468-496 (2011).
- Bannister R.N., Migliorini S., Dixon M. A. G., Ensemble prediction for nowcasting with a convection-permitting model - II: forecast error statistics, Tellus 63A, 497-512 (2011).