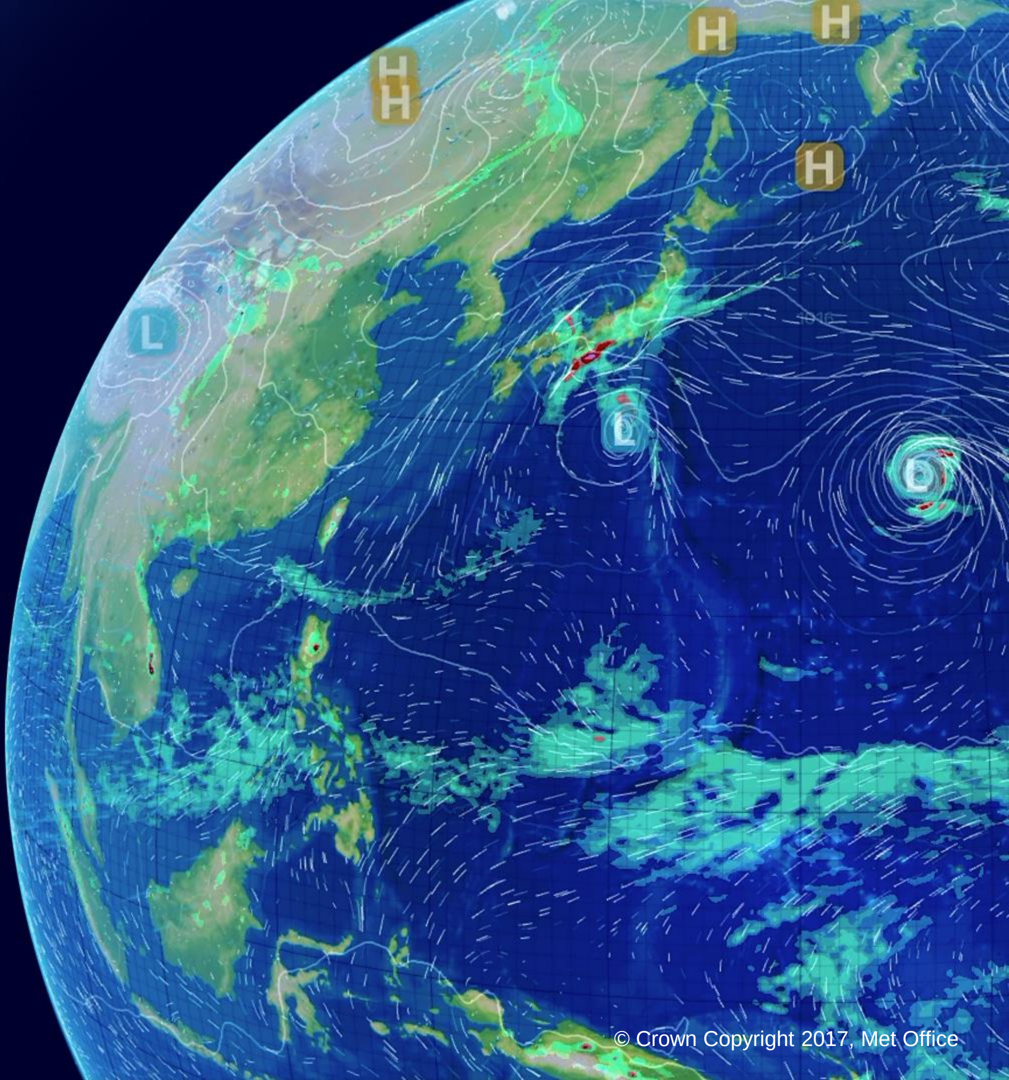
EnKF methods to initialize ensembles

Neill Bowler

ECMWF annual seminar



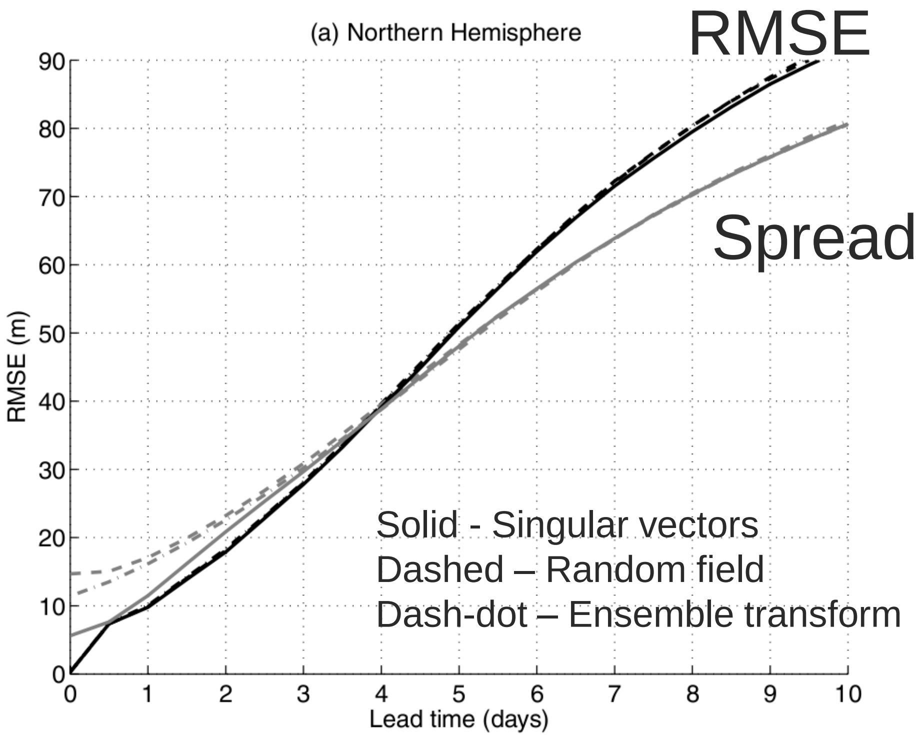
Contents

… before we begin

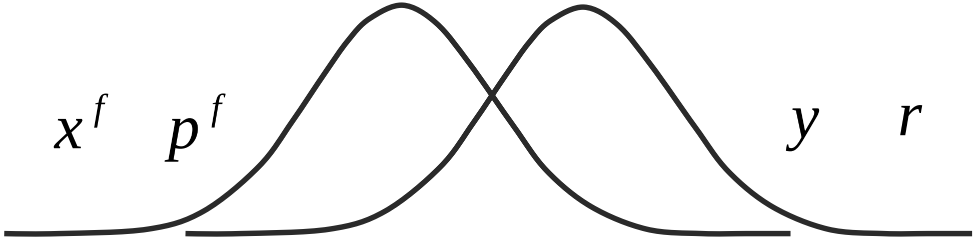
So what?

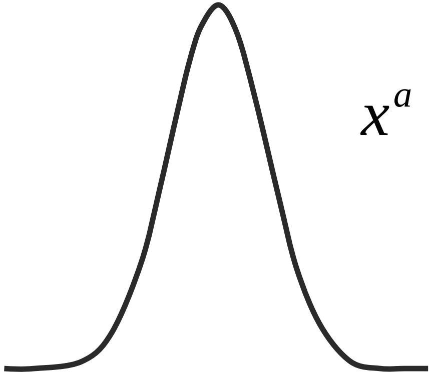
Magnusson et al (2009) showed that initialisation method matters little for medium-range ensemble forecasting

Increased focus on data assimilation as a way to measure performance



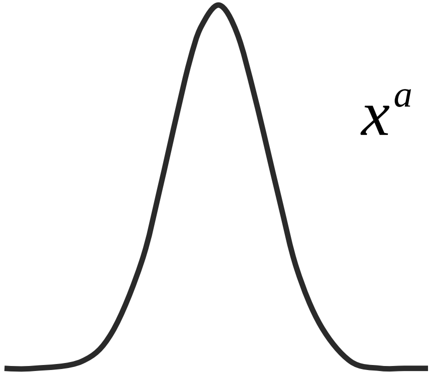
How to make an optimal estimate





$$x^a \propto \frac{x^f}{p^f} + \frac{y}{r}$$

How to make an optimal estimate



$$x^a \propto \frac{x^f}{p^f} + \frac{y}{r}$$

$$x^a = \frac{rp^f}{r + p^f}\left(\frac{x^f}{p^f} + \frac{y}{r}\right) = x^f + \frac{p^f}{r + p^f}\left(y - x^f\right)$$

$$p^a = \left(1 - \frac{p^f}{r + p^f}\right)p^f$$

Kalman filter equations

Update

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

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

$$\mathbf{P}_n^a = \mathbf{P}_n^f - \mathbf{K}_n \mathbf{H} \mathbf{P}_n^f$$

Forecast

$$\mathbf{x}_n^f = \mathbf{M} \mathbf{x}_{n-1}^a \qquad \mathbf{P}_n^f = \mathbf{M} \mathbf{P}_{n-1}^a \mathbf{M}^{\mathrm{T}} + \mathbf{Q}_n$$

NWP approximations

Update

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

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

$$\mathbf{P}_n^a = \mathbf{P}_n^f - \mathbf{K}_n \mathbf{H} \mathbf{P}_n^f$$

Forecast

$$\mathbf{x}_n^f = \mathbf{M} \mathbf{x}_{n-1}^a \qquad \mathbf{P}_n^f = \mathbf{M} \mathbf{P}_{n-1}^a \mathbf{M}^{\mathrm{T}} + \mathbf{Q}_n$$

The model-size problem

Operational model has 2.7×10^9 variables

$\mathbf{P}^f$ has 7×10^{18} entries (thousands of peta-bytes)

Use an ensemble to sample from this

$$\mathbf{X}_n^f = \frac{1}{\sqrt{N-1}} \left(\mathbf{x}_n^{f,1} - \overline{\mathbf{x}_n^f} \quad \mathbf{x}_n^{f,2} - \overline{\mathbf{x}_n^f} \quad \ldots \quad \mathbf{x}_n^{f,N} - \overline{\mathbf{x}_n^f} \right)$$

Ensemble Kalman filter equations

Update

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

$$\mathbf{x}_n^{a,i} = \mathbf{x}_n^{f,i} + \mathbf{K}_n \left(\mathbf{y}_n + \gamma_n^i - H\left(\mathbf{x}_n^{f,i} \right) \right)$$

$$\gamma_n^i \sim N\left(0, \mathbf{R}_n\right)$$

Forecast

$$\mathbf{x}_n^{f,i} = M\left(\mathbf{x}_{n-1}^{a,i} \right) + \eta_n^i \qquad \mathbf{P}_n^f = \mathbf{L} \circ \mathbf{X}_n^f \mathbf{X}_n^{f,\mathrm{T}}$$

Danger!

In making the switch to using ensembles and nonlinear models we have introduced many potential problems, mostly related to sampling error

- Localisation
- Inflation
- Perturbed observations
- Inbreeding

Perturbations or analyses?

Update state

$$\mathbf{x}_n^{a,i} = \mathbf{x}_n^{f,i} + \mathbf{K}_n\left(\mathbf{y}_n + \gamma_n^i - H\left(\mathbf{x}_n^{f,i}\right)\right)$$

Like an EDA

Update mean

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

Update perturbations

$$\hat{\mathbf{x}}_n^{a,i} = \hat{\mathbf{x}}_n^{f,i} + \mathbf{K}_n\left(\gamma_n^i - \mathbf{H}\hat{\mathbf{x}}_n^{f,i}\right)$$

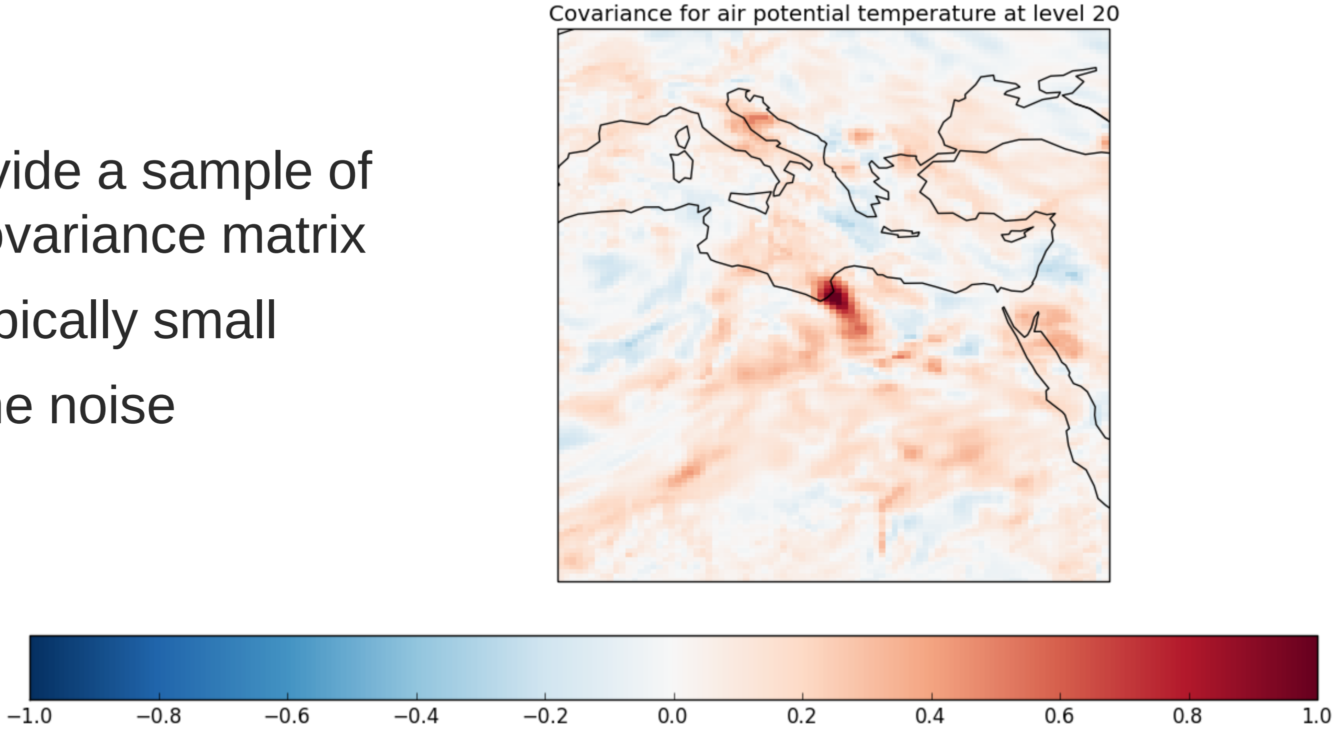
Localisation

Localisation

An ensemble can provide a sample of a background-error covariance matrix

These samples are typically small

We need to remove the noise



Localisation

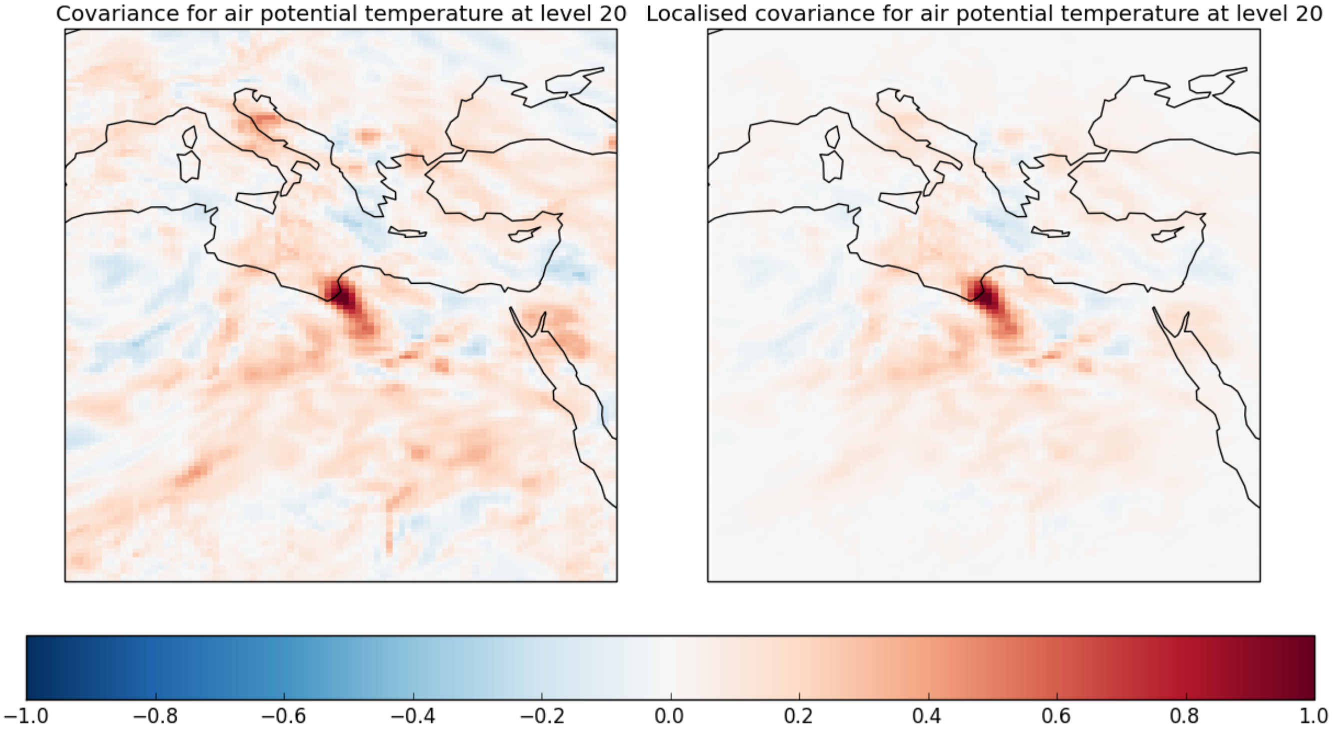
Covariance matrices have certain properties

- Positive semi-definite
- Symmetric, etc

Hadamard (Schur, elementwise) product of two covariance matrices is a covariance matrix

Assume that distant points are uncorrelated, and define a localising covariance matrix which enforces this

Localisation



Met Office

Perturbed observations

Perturbed observations and square-root filters

$$\mathbf{x}_n^{a,i} = \mathbf{x}_n^{f,i} + \mathbf{K}_n\left(\mathbf{y}_n + \gamma_n^i - H\left(\mathbf{x}_n^{f,i}\right)\right)$$

$$\gamma_n^i \sim N\left(0, \mathbf{R}_n\right)$$

Perturbed observations -> extra sampling error

Avoid this using square-root filters

$$\mathbf{P}_n^a = \left(\mathbf{I} - \mathbf{K}_n\mathbf{H}\right)\mathbf{P}_n^f$$

$$\mathbf{X}_n^a = \left(\mathbf{I} - \mathbf{K}_n\mathbf{H}\right)^{1/2}\mathbf{X}_n^f$$

EnSRF

Ensemble square-root filter (Whitaker & Hamill 2002)

Treat observations one at a time

$$\hat{\mathbf{x}}_n^{a,i} = \left(\mathbf{I} - \alpha \mathbf{K}_n \mathbf{H}\right)\hat{\mathbf{x}}_n^{f,i}$$

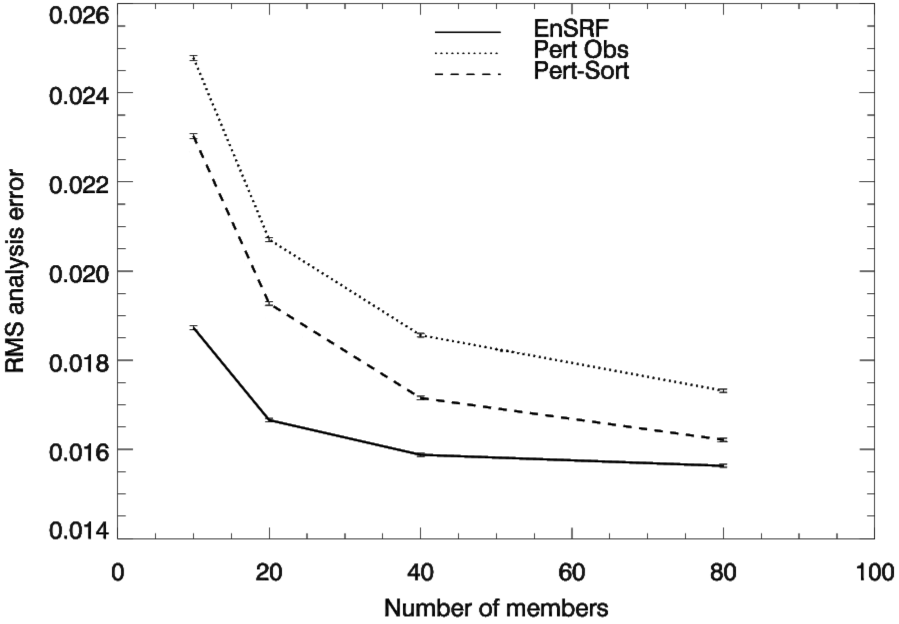
Gain reduction factor

$$\alpha = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R}}}\right)^{-1}$$

Perturbed observations and square-root filters

Comparison within Lorenz (1996) model (40 variables, mild nonlinearity)

Observations every grid-point (Bowler & Flowerdew, 2013)



Perturbed observations may be good

Square-root filters use simplified analysis-error covariance

$$\mathbf{P}_n^a = \left(\mathbf{I} - \mathbf{K}_n\mathbf{H}\right)\mathbf{P}_n^f$$

Perturbed-observations actually samples from

$$\mathbf{P}_n^a = \left(\mathbf{I} - \mathbf{K}_n\mathbf{H}\right)\mathbf{P}_n^f\left(\mathbf{I} - \mathbf{K}_n\mathbf{H}\right)^{\mathrm{T}} + \mathbf{K}_n\mathbf{R}\mathbf{K}_n^{\mathrm{T}}$$

In a nonlinear system the perturbations can become substantially non-Gaussian. Perturbed observations help maintain Gaussianity (Lawson & Hansen, 2004)

Met Office

Inflation

The need for inflation

Tight localisation -> Imbalance in perturbations, slow growth

Broad localisation -> Over-estimation of observation impact, small spread

There is no single correct answer

Typically inflation is needed to increase spread

Model error?

Met Office

Multiplicative inflation

- Simplest method to counter lack of spread in the ensemble
- Multiply perturbations by inflation factor

$$\mathbf{x}^{a,i} \rightarrow \overline{\mathbf{x}^{a}} + \beta\left(\mathbf{x}^{a,i} - \overline{\mathbf{x}^{a}}\right)$$

- Tuning required

Adaptive inflation

Wang and Bishop (2003) proposed a simple adaptive scheme

$$\beta_n = \beta_{n-1}\sqrt{\frac{\left(\mathbf{d}_f^o\right)^{\mathrm{T}}\mathbf{d}_f^o - Tr(\mathbf{R})}{Tr\left(\mathbf{HP}^f\mathbf{H}^{\mathrm{T}}\right)}}$$

This can be estimated for different regions (Bowler et al (2009), Flowerdew & Bowler (2013))

An alternative adaptive inflation scheme was developed by Anderson (2008)

Inflation oscillations

The observing network varies (more sondes 0, 12 UTC)

Wang & Bishop method based on what the inflation factor should have been

$$\beta_n = \beta_{n-1}\sqrt{\frac{\left(\mathbf{d}_f^o\right)^{\mathrm{T}}\mathbf{d}_f^o - Tr\left(\mathbf{R}\right)}{Tr\left(\mathbf{HP}^f\mathbf{H}^{\mathrm{T}}\right)}}$$

Larger inflation factor needed at 0, 12 UTC, but applied at 6, 18 UTC

Adaptive inflation

Ying & Zhang (2015) proposed a different method

$$\beta_n = \sqrt{\frac{\left(\mathbf{d}_a^f\right)^{\mathrm{T}} \mathbf{d}_o^a}{Tr\left(\mathbf{HP}_a\mathbf{H}^{\mathrm{T}}\right)}}$$

$$\mathbf{d}_o^a = \mathbf{y} - H\left(\mathbf{x}_a\right)$$

$$\mathbf{d}_a^f = H\left(\mathbf{x}_a\right) - H\left(\mathbf{x}_f\right)$$

Ratio of measured analysis spread to actual analysis spread

Should avoid oscillation issues, since dealing with analysis spread at current time

Met Office

Inflation in the Météo-France system

A global factor to counter under-spread in the ensemble system (Raynaud et al, 2012)

Uses ratio of cost-function minimum to optimal minimum

$$\beta_n = \sqrt{\frac{\mathbf{V}_s}{\sigma_f^2} \frac{J_b(\mathbf{x}_a)}{J_b^{theo}(\mathbf{x}_a)}}$$

$\mathbf{V}_s$ Specified variance from a climatological ensemble

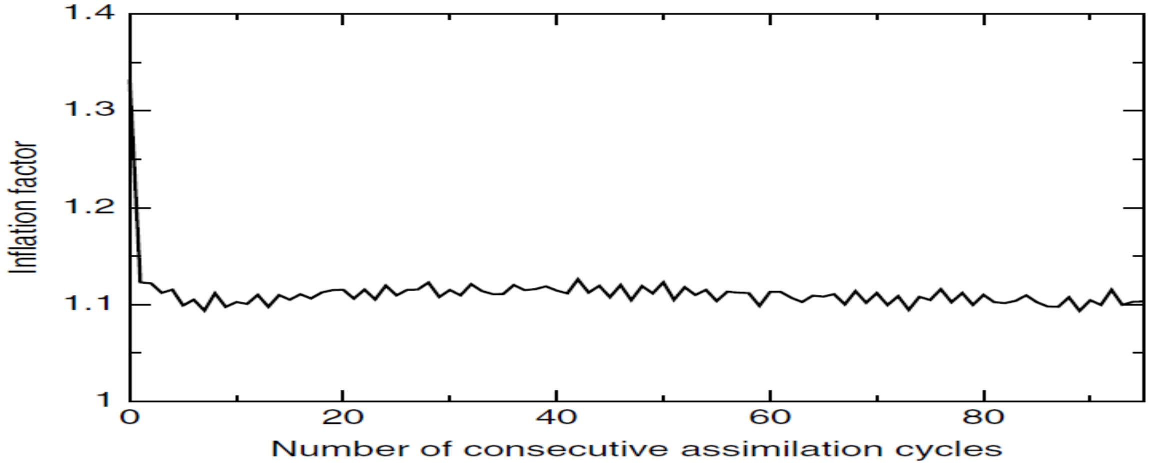
$$J_b^{theo}(\mathbf{x}_a) = Tr(\mathbf{HK})$$

Theoretical cost-function minimum, calculated from the EDA (Desroziers et al., 2009)

Inflation in the Météo-France system

Compensation for model error neglect in En-DA

Relatively stable



Relaxation methods

Multiplicative inflation can lead to over-spread in poorly observed regions

Relax perturbations back towards the forecast perturbations / spread

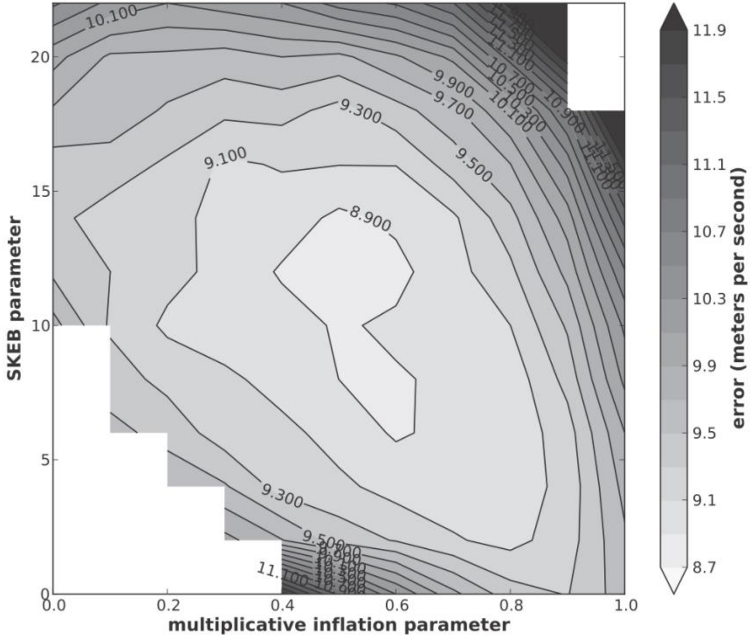
RTPP / RTPS

Popular

Inflation and model error

Whitaker and Hamill (2012) looked at combining model error representation with inflation methods

Showed that multiplicative inflation and representing model error are complementary – both are needed



Other issues

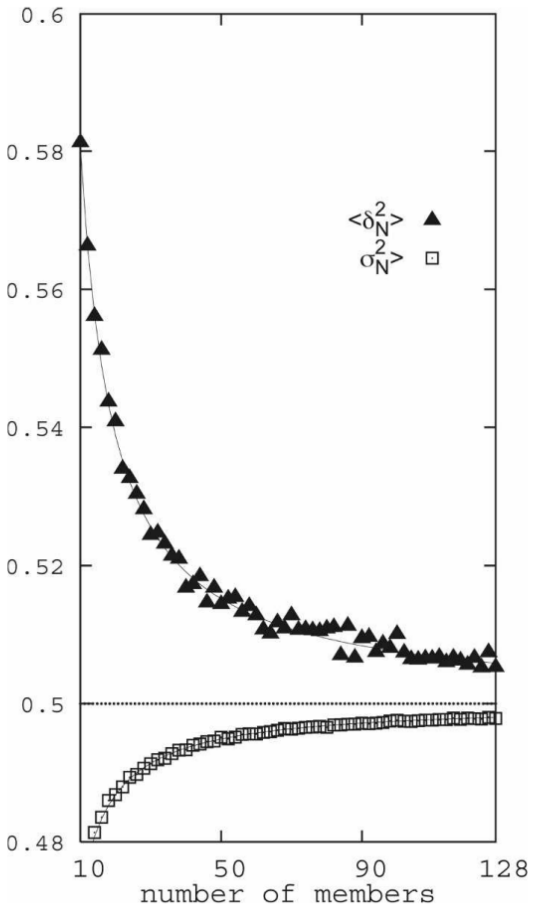
Inbreeding

The standard EnKF has a bias – with a finite ensemble the error is increased, but the spread decreased (Sacher & Bartello, 2009)

Inbreeding – using each ensemble perturbation in the covariance used to update that ensemble

Solution – split the ensemble into *M* sub-ensembles; use the *M-1* other sub-ensembles when updating

Introduces positive bias into ensemble spread



Met Office

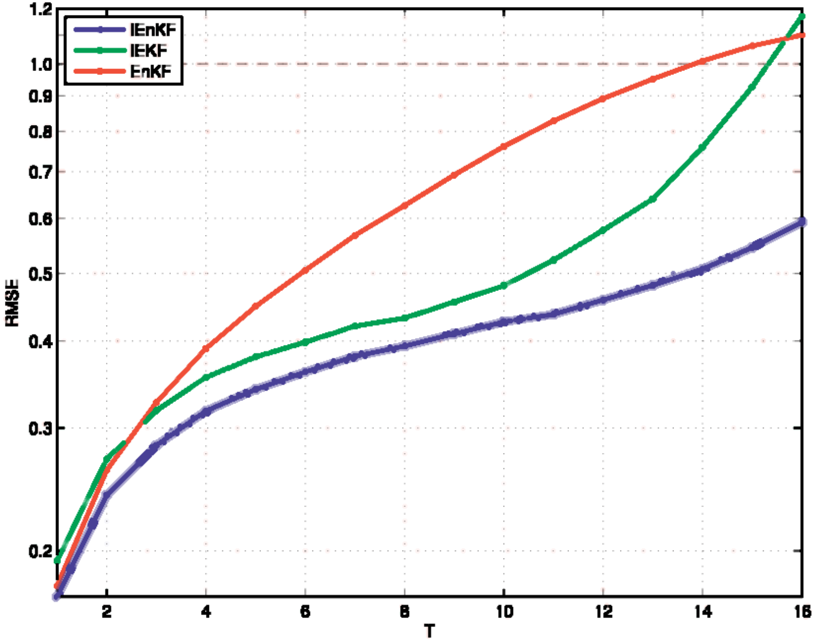
Iterative EnKF

EnKF struggles in non-linear systems

Variational DA can use outer-loops

EnKF can be iterated by re-running the ensemble member forecasts with updated information (Sakov et al, 2012)

Costly



Comparison with other methods

Increased focus on short-range

Compared with error-breeding or singular vectors

- Slower growth of perturbations
- Useful in data assimilation

Comparison with EDA

Essentially the same method

- Easier to set up
  - Better for coupled modelling
- Update algorithm cheaper, and very scalable

- Can’t use hybrid covariances
- Outer loop (iterative EnKF) expensive
- Either batches of observations, or observation-space localisation
  - Extracts less benefit from satellite observations

References

Anderson JL, 2008: *Tellus A*, **61**: 72-83, DOI: 10.1111/j.1600-0870.2008.00361.x

Bowler NE, Arribas A, Beare SE, Mylne KR, Shutts G, 2009: *Q. J. R. Meteorol. Soc.*, **135**: 767-776

Bowler NE, Flowerdew J, 2013: *Q. J. R. Meteorol. Soc.*, **139**: 1505-1519, DOI:10.1002/qj.2055

Flowerdew J, Bowler NE, 2013: *Q. J. R. Meteorol. Soc.*, **139**: 1863-1874, DOI:10.1002/qj.2072

Lawson WG, Hansen JA, 2004: *Mon. Wea. Rev.*, **132**, 1966–1981, DOI: 10.1175/1520-0493(2004)132<1966:IOSADF>2.0.CO;2

Lorenz EN, 1996: *Proceedings of the seminar on predictability*, ECMWF: pp. 1–18

Magnusson L, Nycander J, Källén E, 2009: *Tellus A,* **61**: 194-209, DOI: 10.1111/j.1600-0870.2008.00385.x

Raynaud L, Berre L, Desroziers G, 2012: *Q. J. R. Meteorol. Soc.*, **138**: 249-262, DOI: 10.1002/qj.906

Sacher W, Bartello P, 2008: *Mon. Wea. Rev.,* **136**: 3035-3049, DOI: 10.1175/2007MWR2323.1

References

Sakov P, Oliver DS, Bertino L, 2012: *Mon. Wea. Rev,* **140**: 1988-2004, DOI: 10.1175/MWR-D-11-00176.1

Wang X, Bishop CH, 2003: *J. Atmos. Sci.,* **60**: 1140-1158, DOI: 10.1175/1520-0469(2003)060<1140:ACOBAE>2.0.CO;2

Whitaker JS, Hamill TM: 2002, *Mon. Wea. Rev.* **130**: 1913-1924, DOI: 10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2

Whitaker JS, Hamill TM: 2012, *Mon. Wea. Rev.* **140**: 3078-3089, DOI: 10.1175/MWR-D-11-00276.1

Ying Y, Zhang F: 2015, *Q. J. R. Meteorol. Soc.* **141**: 2898–2906, DOI:10.1002/qj.2576

Zhang F, Snyder C, Sun J: 2004, *Mon. Wea. Rev.* **132**: 1238-1253, DOI: 10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2

Met Office

Other developments

Proposal for hybrid EnKF

Successive covariance localisation

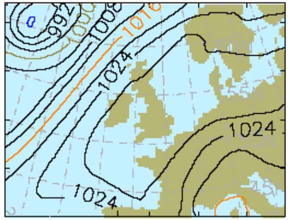
Relaxation to prior perturbations (RTPP)

Multiplicative inflation can lead to $\sigma_a^2 > \sigma_f^2$

Therefore apply a relaxation rather than inflation (Zhang et al, 2004)

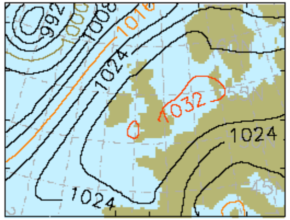
$$\mathbf{x}^{a,i} \rightarrow \overline{\mathbf{x}^{a}} + (1-\beta)\left(\mathbf{x}^{a,i} - \overline{\mathbf{x}^{a}}\right) + \beta\left(\mathbf{x}^{f,i} - \overline{\mathbf{x}^{f}}\right)$$

$(1-\beta)$



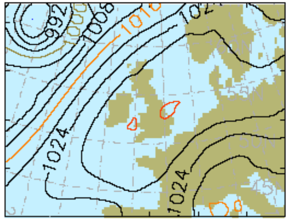
Analysis perturbation

$+ \ \beta$



Forecast perturbation

$=$



Relaxed perturbation

Relaxation to prior spread (RTPS)

RTPP mixes analysis and forecast perturbations

Forecast perturbations are larger-scale, more balanced

Therefore relax the spread, not the perturbations (Whitaker & Hamill, 2012)

$$\mathbf{x}^{a,i} \rightarrow \overline{\mathbf{x}^{a}} + \beta\left(\mathbf{x}^{a,i} - \overline{\mathbf{x}^{a}}\right)$$

$$\beta = \frac{\varphi\sigma_{f} + (1-\varphi)\sigma_{a}}{\sigma_{a}}$$