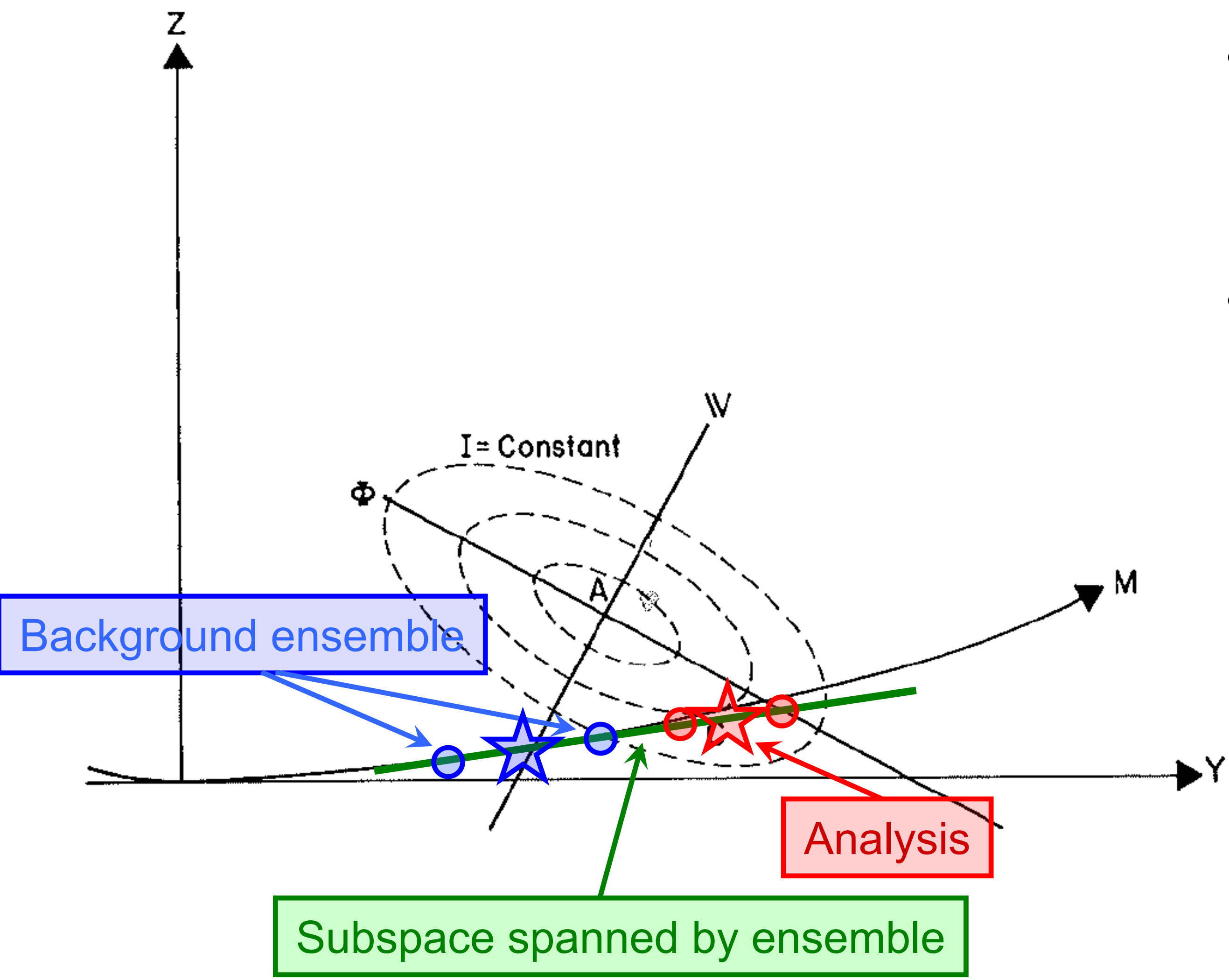
Balance in the EnKF

Jeffrey D. Kepert
Centre for Australian Weather and Climate Research
A partnership between the Australian Bureau of Meteorology and CSIRO

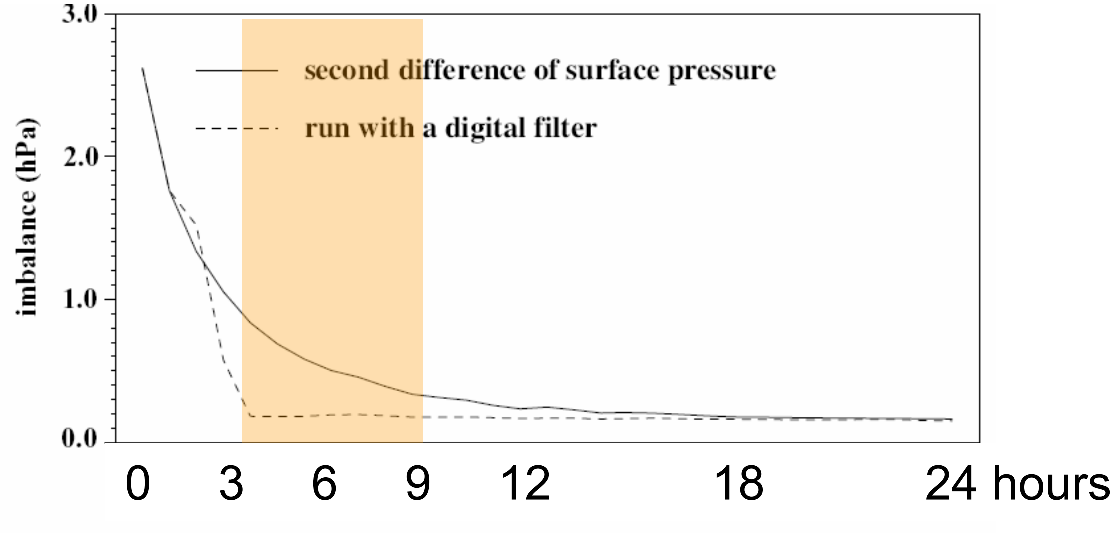
WWRP/Thorpex Workshop on 4D-Var and EnKF Intercomparisons
10-13 November, 2008

EnKF and Slow Manifold



- Analysis is from the subspace spanned by the background.
- If the background is on the manifold, and the manifold curvature is small (relative to the analysis increment), then the analysis will be close to balanced.

Canadian experience

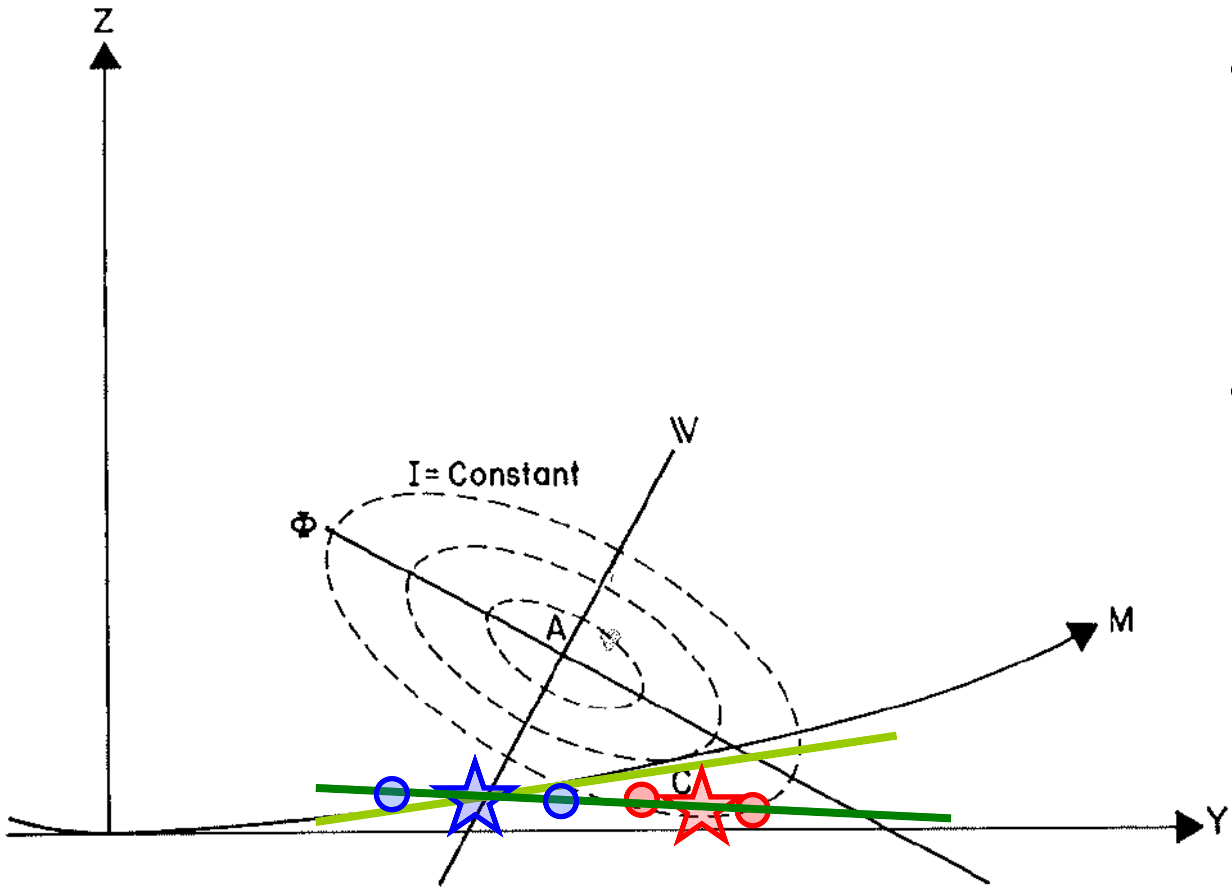


- An initialisation step is necessary as gravity waves in 3 to 9-hr forecast affect the ensemble covariances.
- Even though their model strongly damps gravity waves. (Houtekamer and Mitchell 2005)

Causes of imbalance in an EnKF

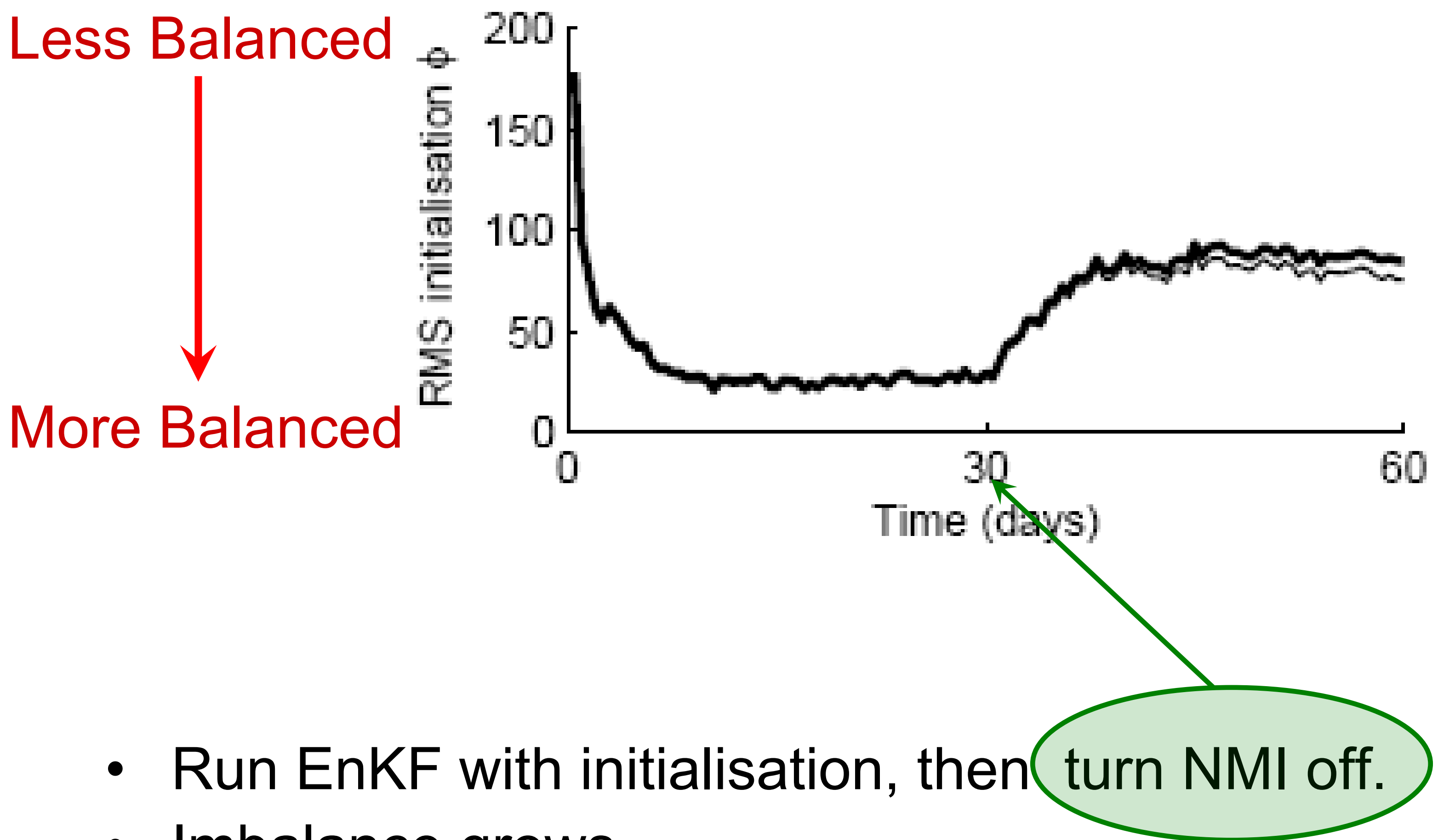
- Linear operations when balances are nonlinear
  - Analysis is in subspace spanned by ensemble
  - Covariance inflation (small)
- Background ensemble
- Approximations
  - Sampling error
  - Covariance localisation
- Mixed-mode (i.e. real) observations
- Introduced imbalance
  - Model error term
  - Blend with 3d-var covariances ("hybrid schemes")
  - Initialisation balance ≠ model balance (hopefully small)
- Can be hard to diagnose cause in a full system (Houtekamer and Mitchell 2005)

EnKF and Slow Manifold



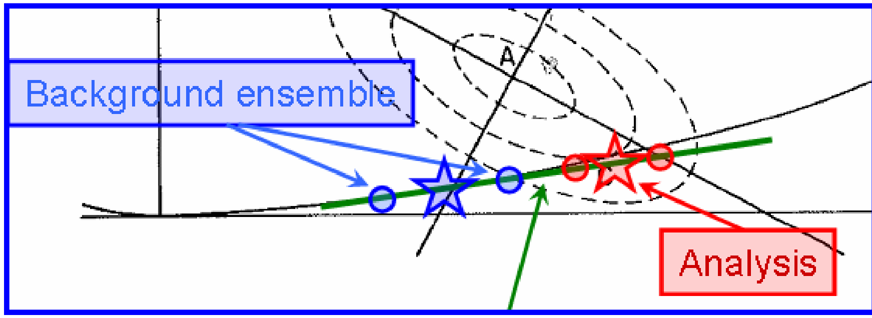
- Analysis is from the subspace spanned by the background.
- If the background is unbalanced, then the analysis probably will be also.

Positive feedback on imbalance



- Run EnKF with initialisation, then turn NMI off.
- Imbalance grows.
- Positive feedback: gravity waves describe increasingly more of the background variance (Lorenc 2003).

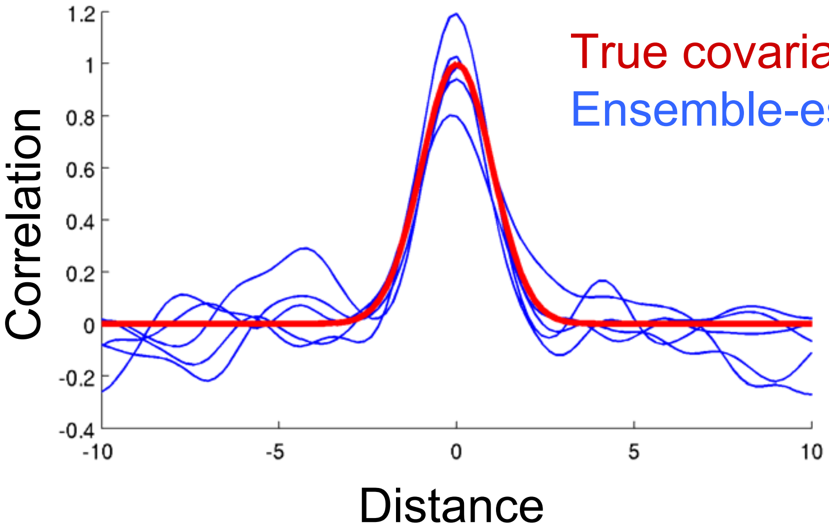
The problem with:



- The model has dimension ~10^6.
- The manifold has dimension ~10^4.
- The ensemble has dimension ~10^2.

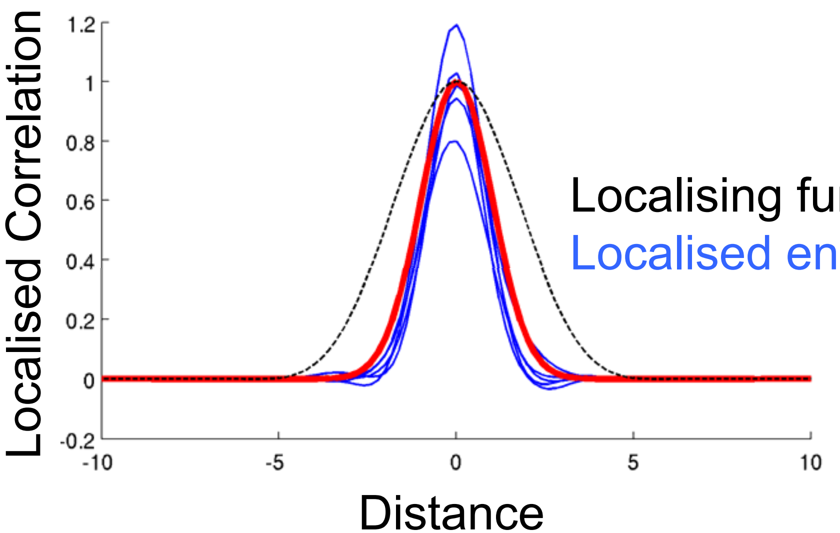
- Hence the closest point to the observations on the ensemble-subspace will almost certainly be a long way from the closest point on the manifold.
  - i.e. The analysis will fit the observations poorly (e.g. Lorenc 2003).
- Thus localisation is necessary.

Covariance Localisation



True covariance (red)
Ensemble-estimated covariance (blue, 100 members)

- Localisation eliminates the effect of sampling error at large distance (but not small).
- See Tom Hamill's talk.

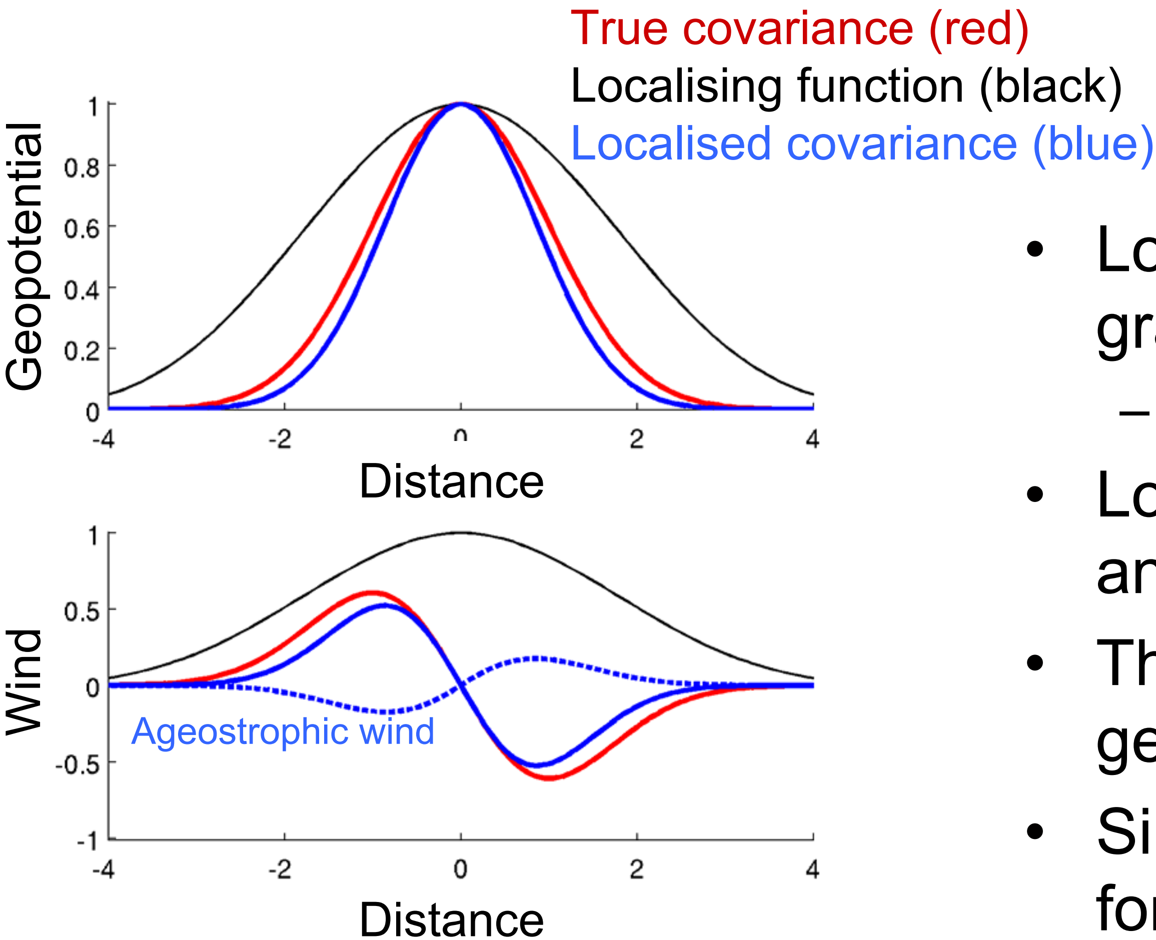


Localising function (black, Gaspari and Cohn 1999)
Localised ensemble covariances (blue)

Localisation and balance

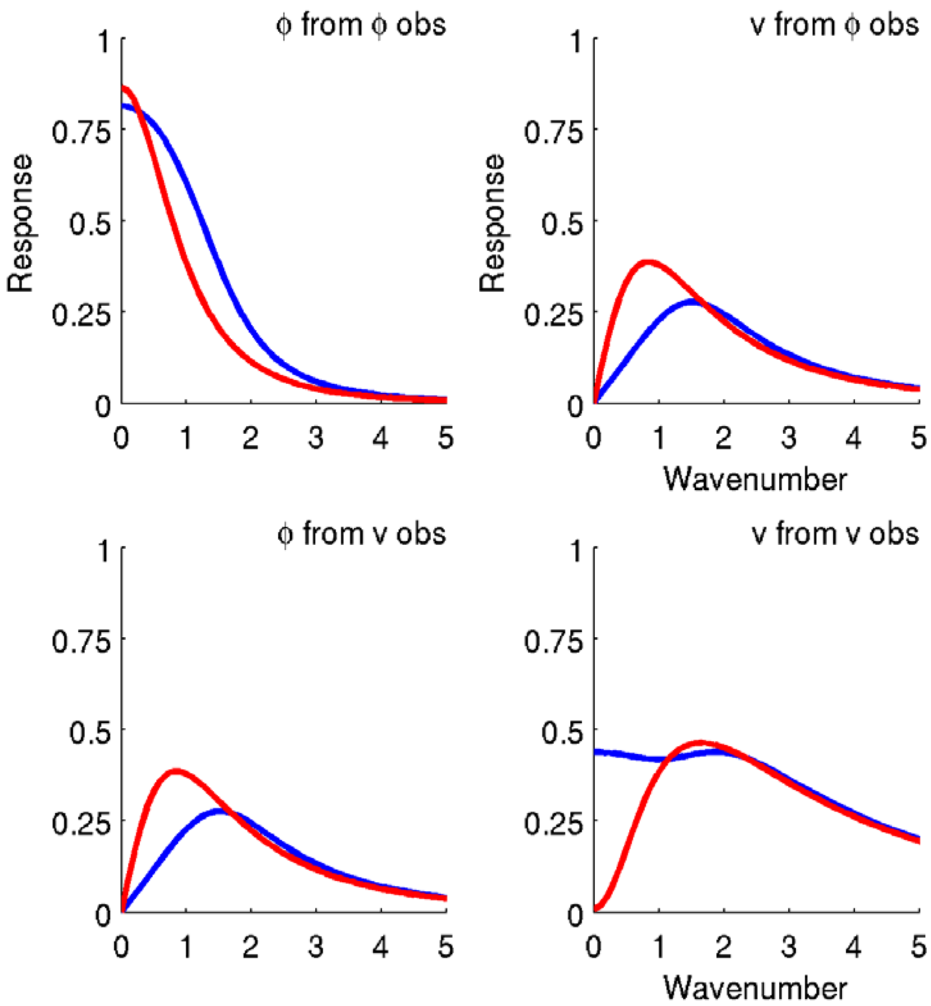
- Balance defined by null space of $\mathbf{P}^f$ (Lorenc 2003)
- Localisation increases the rank of $\mathbf{P}^f$
  - necessary to fit observations
- Does localisation reduce the null space in a way that is consistent with balance?

Effect of localisation on geostrophic balance



- Localisation increases gradient of geopotential
  - Increases geostrophic wind
- Localisation reduces analysed wind speed
- Thus localisation weakens geostrophic balance
- Similar arguments apply for other balance relations

Dense observing system



- 1-dimensional infinite domain, observe wind and geopotential
- Geostrophic covariances
- Red: Kalman gain in spectral space, no localisation
- Blue: with localisation

Better localisation

- Covariances involving streamfunction ψ and velocity potential χ are more isotropic than those involving (u, v)
  - Long assimilation experience
- Balance equations relate grad(φ) to grad(ψ)
  - e.g. geostrophy, nonlinear balance equation
- Localising in (ψ, χ) rather than (u, v) space will be less severe on balance, because grad(φ) and grad(ψ) will increase by similar amounts and because covariances in (ψ, χ) are more isotropic than in (u, v).

Localisation in (φ, ψ, χ)-space

Write the velocity components in terms of streamfunction ψ and velocity potential χ:

$$
\begin{aligned}
u &= -\frac{\partial \psi}{\partial y} + \frac{\partial \chi}{\partial x} \\
v &= \frac{\partial \psi}{\partial x} + \frac{\partial \chi}{\partial y}
\end{aligned} \tag{1}
$$

Write the two-point covariance

$$
\langle u_1, u_2 \rangle = \langle u(x_1, y_1), u(x_2, y_2) \rangle \tag{2}
$$

in terms of ψ and χ:

$$
\begin{aligned}
\langle u_1, u_2 \rangle = & \frac{\partial}{\partial y_1}\frac{\partial}{\partial y_2}\langle \psi_1, \psi_2 \rangle - \frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\langle \psi_1, \chi_2 \rangle \\
& - \frac{\partial}{\partial x_1}\frac{\partial}{\partial y_2}\langle \chi_1, \psi_2 \rangle + \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}\langle \chi_1, \chi_2 \rangle
\end{aligned} \tag{3}
$$

(φ, ψ, χ)-localisation (cont'd)

Let $C = C(x_1, y_1, x_2, y_2)$ be a localisation function, and apply the localisation be replacing $\langle\psi_1, \psi_2\rangle \leftarrow C\langle\psi_1, \psi_2\rangle$ in (3) and similarly for $\langle\psi_1, \chi_2\rangle$, $\langle\chi_1, \psi_2\rangle$ and $\langle\chi_1, \chi_2\rangle$. Then the *localised* two-point covariance $\langle u_1, u_2\rangle_L$ becomes

$$\begin{aligned}\langle u_1, u_2\rangle_L = {} & C\langle u_1, u_2\rangle \\ & + \frac{\partial C}{\partial x_1}\langle\chi_1, u_2\rangle - \frac{\partial C}{\partial y_1}\langle\psi_1, u_2\rangle + \frac{\partial C}{\partial x_2}\langle u_1, \chi_2\rangle - \frac{\partial C}{\partial y_2}\langle u_1, \psi_2\rangle \\ & + \frac{\partial^2 C}{\partial x_1 \partial x_2}\langle\chi_1, \chi_2\rangle - \frac{\partial^2 C}{\partial x_1 \partial y_2}\langle\chi_1, \psi_2\rangle - \frac{\partial^2 C}{\partial y_1 \partial x_2}\langle\psi_1, \chi_2\rangle + \frac{\partial^2 C}{\partial y_1 \partial y_2}\langle\psi_1, \psi_2\rangle\end{aligned} \tag{4}$$

The covariances in ψ, χ, u and v can all be calculated from the ensemble in the usual way. (The EnKF will need to read in ψ and χ in addition to u and v.)

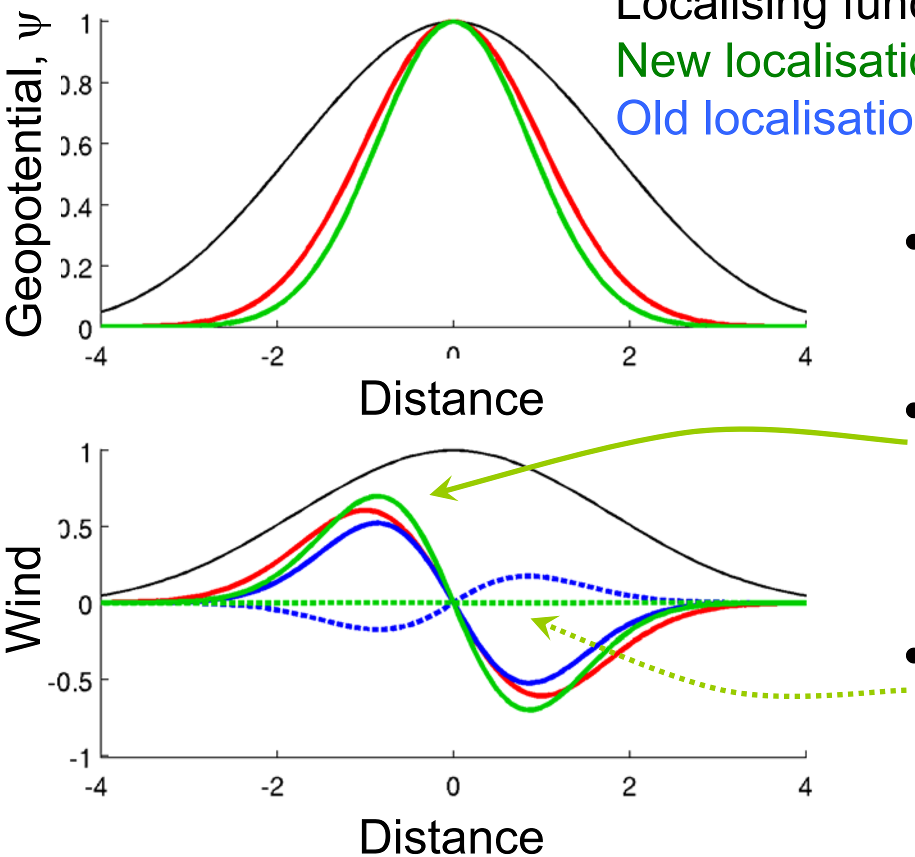
Other covariances are calculated similarly (Kepert, 2008).

Optional *intervariable localisation* with $\langle\psi_1, \chi_2\rangle = \langle\phi_1, \chi_2\rangle = 0$, etc.

Effect of localisation in (φ, ψ, χ)-space

Single height observation

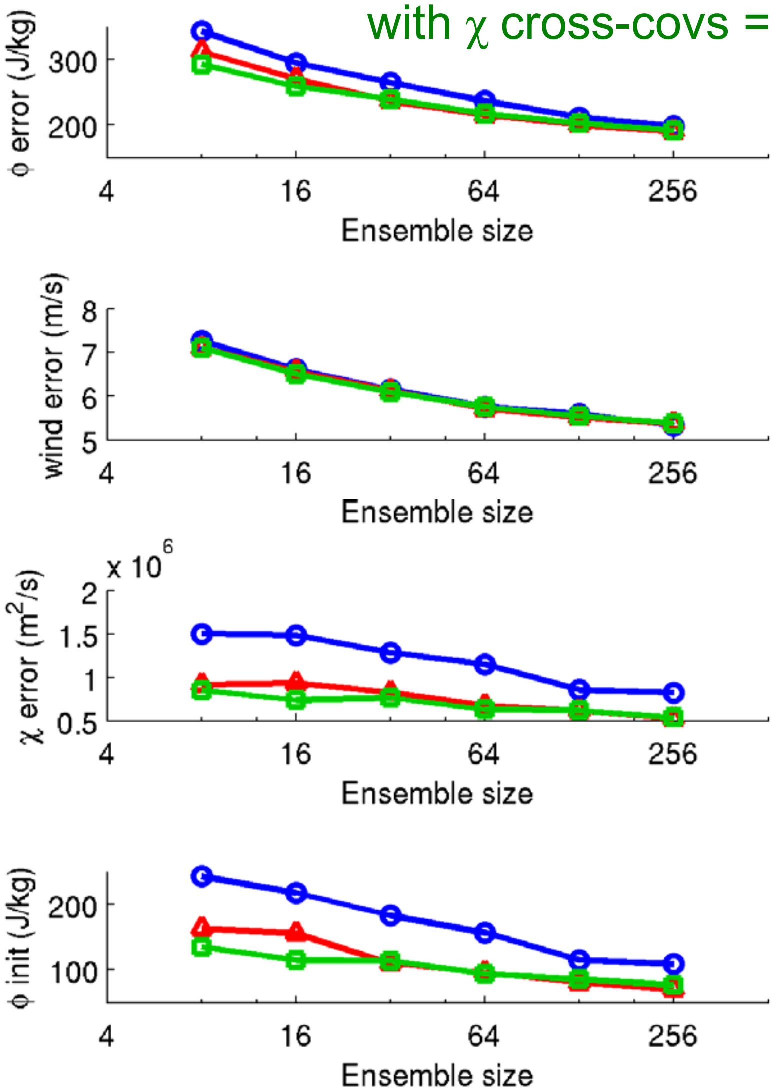
True covariance (red)
Localising function (black)
New localisation (green)
Old localisation (blue)



- Geopotential is same in both localisations.
- Wind response increases (because the ψ gradient increases)
- Ageostrophic wind in analysis is now zero.

Performance in test system

old localisation (blue)
new localisation (red)
with χ cross-covs = 0 (green)



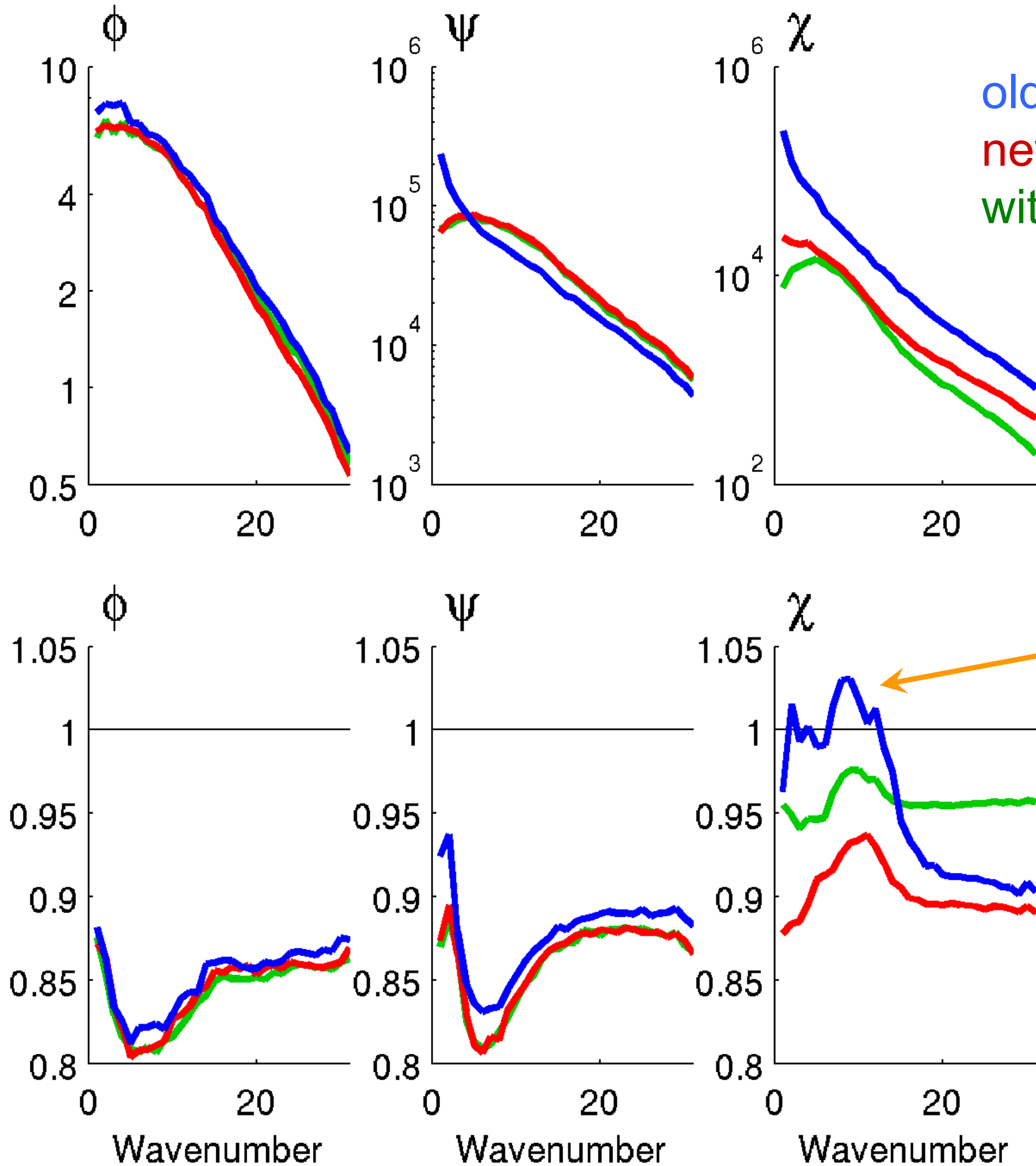
- Identical twin experiments, global spectral shallow-water model, various localisations.

- New localisations are substantially more accurate
- and better balanced.
- But analyses still contain some imbalance, and performance is improved when normal-mode initialisation is included (in this low-dissipation system).

Analysis increment spectra

increment in ensemble mean

relative increment in ensemble variance



old localisation (blue)
new localisation (red)
with χ cross-covs = 0 (green)

Analysis increases ens. variance ➔ suboptimal gain (old localisation).

Less variance reduction when χ cross-covs = 0 in new localisations.

Observing mixed variables

- Most real observations are a mix of vortical and gravity modes
  - e.g. wind, geopotential, temperature, …
  - assimilation has to correctly unscramble the mix
- Time-scale of vortical vs. gravity is variable
  - e.g. mid-latitudes vs. tropics, mesoscale
- Background variance of vortical vs. gravity modes is variable
  - e.g. troposphere vs. mesosphere
- Assimilation system needs to be robust in all these circumstances.

Idealised experiments with mixed-variable observations (Neef et al.)

- Accurate recovery of both modes requires
  - accurate fast-slow correlations
  - accurate error variances
- EnKF with enough members can do it
  - but tends to diverge in fast mode
  - and spurious projection onto fast mode can cause both to diverge
- Model error in gravity waves can be detrimental to EnKF performance
  - issue with semi-implicit time stepping schemes?
- Neef et al. (2006, 2008)

Observing mixed variables

Two-variable system, background covariance matrix

$$\mathbf{P}^b = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Observe the sum $\mathbf{H} = \begin{bmatrix} 11 \end{bmatrix}$ with error $\mathbf{R} = \begin{bmatrix} \sigma_o^2 \end{bmatrix}$. The analysis covariance is

$$\mathbf{P}^a = \frac{1}{\sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2 + \sigma_o^2} \begin{bmatrix} \sigma_1^2(\sigma_2^2(1-\rho^2) + \sigma_o^2) & \sigma_1\sigma_2(\sigma_1\sigma_2(1-\rho^2) + \rho\sigma_o^2 \\ \sigma_1\sigma_2(\sigma_1\sigma_2(1-\rho^2) + \rho\sigma_o^2 & \sigma_2^2(\sigma_1^2(1-\rho^2) + \sigma_o^2) \end{bmatrix}$$

so the variance ratio is

$$\frac{\sigma_1^2(\sigma_2^2(1-\rho^2) + \sigma_o^2)}{\sigma_2^2(\sigma_1^2(1-\rho^2) + \sigma_o^2)} = \frac{\frac{\sigma_1^2\sigma_2^2}{\sigma_o^2}(1-\rho^2) + \sigma_1^2}{\frac{\sigma_1^2\sigma_2^2}{\sigma_o^2}(1-\rho^2) + \sigma_2^2}$$

which is closer to 1 than σ_1^2/σ_2^2.

If ρ^2 is too large due to sampling error, then the variable with least variance (e.g. gravity waves) will become underspread and may diverge.

Improving balance in an EnKF

- Keep everything as linear as possible
  - small analysis increments, small model error term, small covariance inflation, etc
  - eliminate or account for model bias (in all modes)
  - *requires an accurate system*
- Localise more carefully
- (Balanced) model error term
  - Better introduced at beginning of forecast?
- Initialise
  - Only way to stop the positive feedback?
- Generally less options than in VAR.

Bugbears

- Balance often not reported in literature
  - operational Canadian EnKF is an honourable exception.
- Relevant model settings (e.g. diffusion) sometimes not reported.
- Data voids are a good way of testing that covariances are ok (but not always done).
- Understanding EnKF balance properties may require different systems to practical applications (e.g. well-chosen toy models, very low dissipation, no initialisation).

Summary

- Balance is a significant issue in atmospheric EnKFs
- Multiple causes of imbalance, hard to diagnose in a full system. (Houtekamer and Mitchell 2005)
- Inherent to algorithm
  - because balances are nonlinear
- Various approximations also contribute
  - sampling error, localisation, model error, etc
- Some unbalanced flows are important and need to be captured
  - mesoscale, tropics, mesosphere
  - a global atmospheric system has to manage all these situations at once.