An EnKF formulation that better respects atmospheric balance

www.cawcr.gov.au

Jeffrey D. Kepert

8th Adjoint Workshop, Pennsylvania,
May 18 – 22 2009.

Australian Government
Bureau of Meteorology

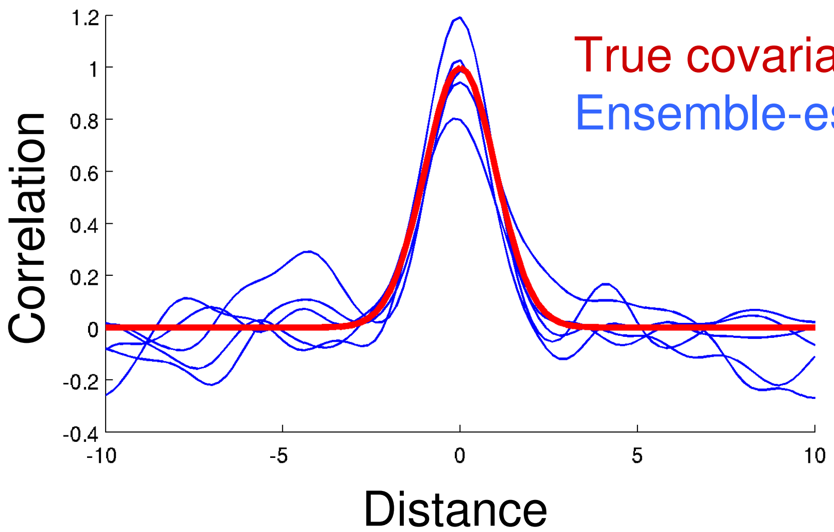
The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology

CSIRO

Sources 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
  - $\boxed{\text{Covariance localisation}}$
- 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)

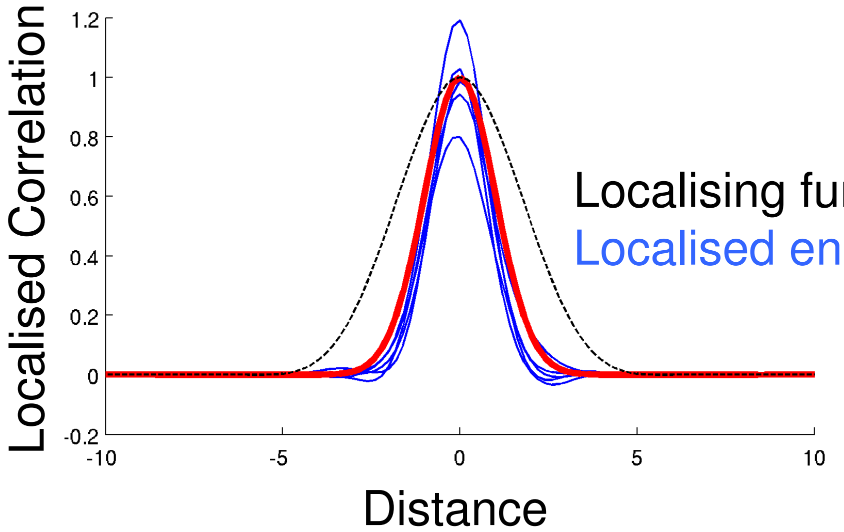
Covariance Localisation (A necessary evil)



True covariance (red)

Ensemble-estimated covariance (blue, 100 members)

- $\mathbf{P}^b{}_e\mathbf{H}^T$ and $\mathbf{HP}^b{}_e\mathbf{H}^T$ need localisation.
- Localisation eliminates the effect of sampling error at large distance (but not small).
- Localisation introduces imbalance.



Localising function (black, Gaspari and Cohn 1999)

Localised ensemble covariances (blue)

Better localisation

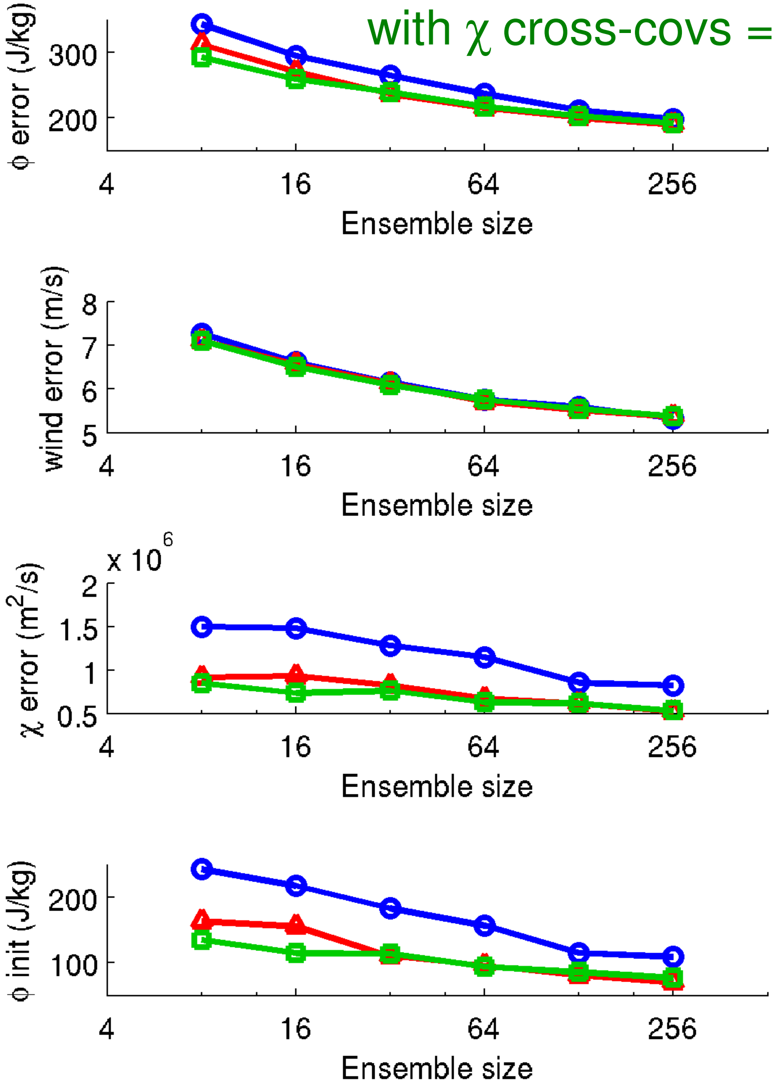
- Covariances involving streamfunction ψ and velocity potential χ are more isotropic than those involving (u, v)
  - Long assimilation experience
- Balance equations relate grad(φ) to grad(ψ)
  - e.g. geostrophy, nonlinear balance equation
- Localising in (ψ, χ) rather than (u, v) space will be less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

Performance in shallow-water test system

standard localisation in (φ,u,v)-space (blue)

improved localisation in (φ,ψ,χ)-space (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.

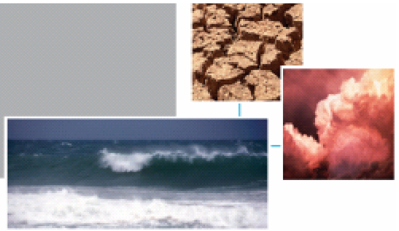
- Kepert (2009, QJRMS in press).

Questions

- Does using the full VAR-style variable transformations in the covariance localisation yield further gains?
  - i.e. using φ_{unbal} in place of φ.

- Is it better to analyse into $(\varphi,\ u,\ v)$-space, or into $(\varphi,\ \psi,\ \chi)$-space, or into $(\varphi_{\mathrm{unbal}},\ \psi,\ \chi_{\mathrm{unbal}})$-space?

The Basic Idea



- Write all variables as sum of balanced and residual parts,

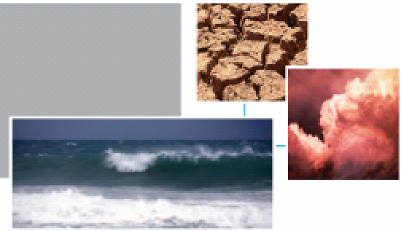
$$u = u_b + u_r\,,\ \phi = \phi_b + \phi_r\,,\ \text{etc.}$$

- Assume that forecast errors for balanced and unbalanced components are uncorrelated

$$\langle \cdot_b, \cdot_r \rangle = 0$$

  for any pair of variables u, v, ϕ, ψ, χ, where $\langle \cdot, \cdot \rangle$ is the two-point covariance.
- Assume (for simplicity) that $\psi = \psi_b$ and that all balanced variables are functions of ψ.

Choice of analysis variable



$$\mathbf{P}^b\mathbf{H}^T \approx \mathbf{P}_e^b\mathbf{H}^T \equiv \frac{1}{N-1}\left[\mathbf{X}^b - \overline{\mathbf{X}^b}\right]\left[\mathcal{H}(\mathbf{X}^b) - \overline{\mathcal{H}(\mathbf{X}^b)}\right]^T$$

The analysis-space part $\mathbf{X}^b$ can be represented in many ways: e.g. (ϕ, u, v), (ϕ, ψ, χ), (ϕ_r, ψ, χ_r), etc.

$$\text{e.g. } \frac{1}{N-1}\left[\mathbf{X}^b{}_{\phi_r,\psi,\chi_r} - \overline{\mathbf{X}^b{}_{\phi_r,\psi,\chi_r}}\right]\left[\mathcal{H}(\mathbf{X}^b{}_{\phi,u,v}) - \overline{\mathcal{H}(\mathbf{X}^b{}_{\phi,u,v})}\right]^T$$

would lead to an analysis increment in (ϕ_r, ψ, χ_r)-space from obs of (ϕ, u, v).

Linear Regression

$$\mathbf{X}^a - \mathbf{X}^b =$$

$$\boxed{\left[\mathbf{P}_e^b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_e^b\mathbf{H}^T\right)^{-1}\right]}\boxed{\left[\mathbf{H}\mathbf{P}_e^b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_e^b\mathbf{H}^T+\mathbf{R}\right)^{-1}\left(\mathbf{Y}-\mathcal{H}(\mathbf{X}^b)\right)\right]}$$

The EnKF can be regarded as

- a filtering step, which gives an analysis in observation space, followed by
- a linear regression, trained on the background ensemble, to take the analysis from observation space back into analysis space. (Anderson, 2003, MWR).

The effect of changing the analysis variable will depend on the quality of this linear regression.

Balance-aware Covariance Localisation

Recipe for covariance localisation

- Write everything as balanced (b) + unbalanced (r)
- Assume balanced-unbalanced covariances are zero.
- Only localise things that have quasi-isotropic covariances (ϕ, ψ, χ are ok, u, v are not).

For example,

$$\begin{aligned}
\langle \phi, \phi \rangle_L &= \langle \phi_b + \phi_r, \phi_b + \phi_r \rangle_L \\
&= \langle \phi_b, \phi_b \rangle_L + \langle \phi_b, \phi_r \rangle_L + \langle \phi_r, \phi_b \rangle_L + \langle \phi_r, \phi_r \rangle_L \\
&= \langle \phi_b, \phi_b \rangle_L + \langle \phi_r, \phi_r \rangle_L \\
&= C\langle \phi_b, \phi_b \rangle + C\langle \phi_r, \phi_r \rangle
\end{aligned}$$

- C is the localisation function, $\langle \cdot, \cdot \rangle$ is the two-point covariance, subscript L means localised.

Balance-aware Localisation (cont'd)

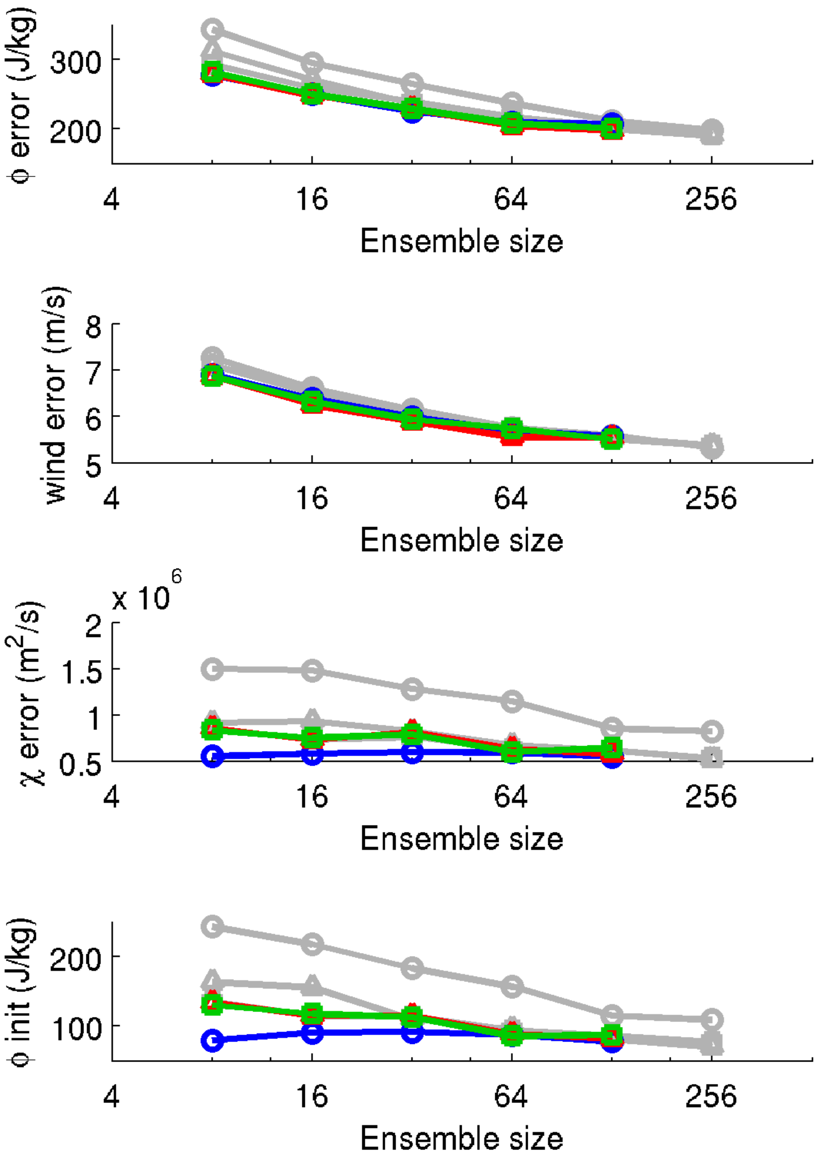
Another example ...

$$
\begin{aligned}
\langle \phi, u \rangle_L &= \langle \phi_b, u_b \rangle_L + \langle \phi_r, u_r \rangle_L \\
&= \left\langle \phi_b, -\frac{\partial \psi}{\partial y_2} \right\rangle_L + \left\langle \phi_r, \frac{\partial \chi}{\partial x_2} \right\rangle_L \\
&= -\frac{\partial \langle \phi_b, \psi \rangle_L}{\partial y_2} + \frac{\partial \langle \phi_r, \chi \rangle_L}{\partial x_2} \\
&= -\frac{\partial C \langle \phi_b, \psi \rangle}{\partial y_2} + \frac{\partial C \langle \phi_r, \chi \rangle}{\partial x_2} \\
&= C \left(-\left\langle \phi_b, \frac{\partial \psi}{\partial y_2} \right\rangle + \left\langle \phi_r, \frac{\partial \chi}{\partial x_2} \right\rangle \right) - \frac{\partial C}{\partial y_2} \langle \phi_b, \psi \rangle + \frac{\partial C}{\partial x_2} \langle \phi_r, \chi \rangle \\
&= C \left(\langle \phi_b, u_b \rangle + \langle \phi_r, u_r \rangle \right) - \frac{\partial C}{\partial y_2} \langle \phi_b, \psi \rangle + \frac{\partial C}{\partial x_2} \langle \phi_r, \chi \rangle
\end{aligned}
$$

Experimental Design

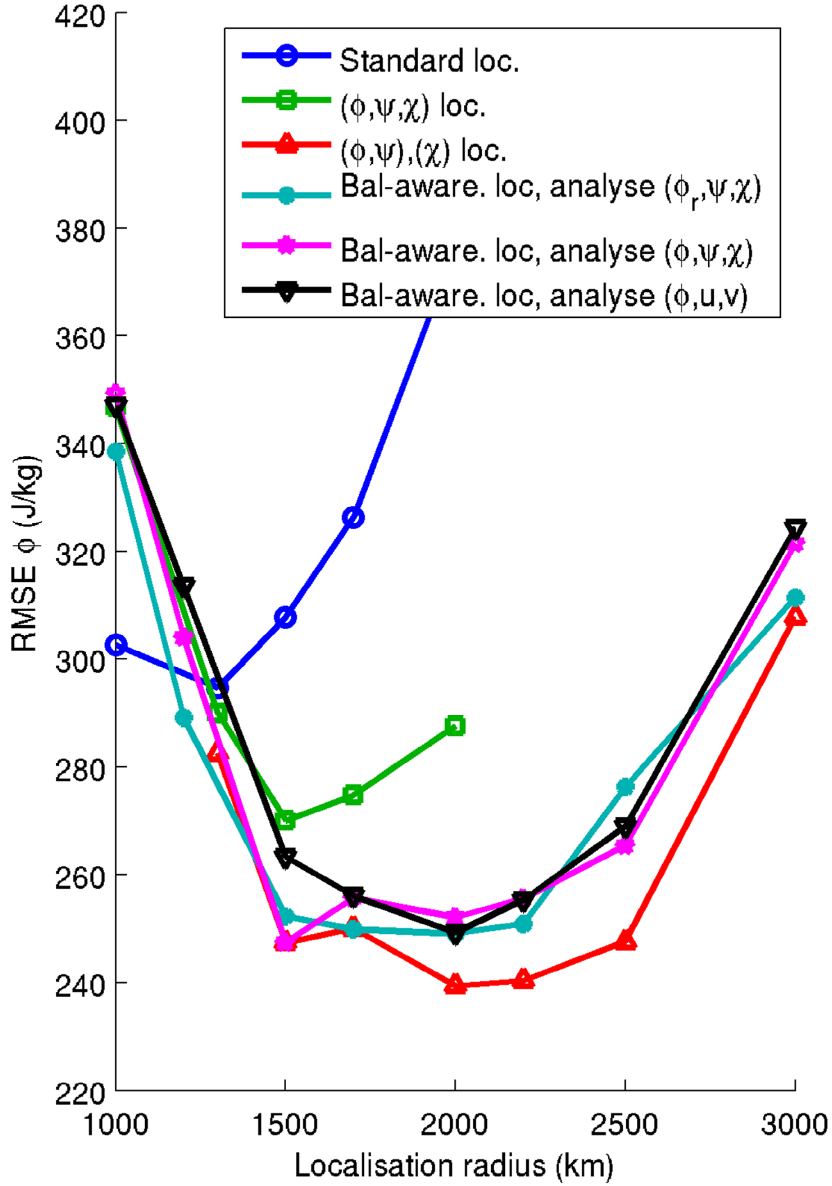
- Global spectral shallow water model of Bourke (1972) at T31.
  - Includes nonlinear balance equation solver.
  - Includes nonlinear normal modes initialisation.
  - Simplest model to have main atmospheric balances.
- Identical twin with 60-day truth run
  - Start from ERA-40 analysis for 500 hPa 18 Jan 1962.
  - Analyse last 30 days of assimilation cycle.
- Obs network based on 50% of global radiosondes.
  - Observe (φ, u, v) 12-hourly.
  - Height error = 100 J/kg, wind error = 5 m/s.
- Standard perturbed-obs EnKF with covariance inflation, various ensemble sizes.
  - Tune localisation length and covariance inflation to minimise errors.
  - Verify spread using rank histograms.
- Balance-aware localisation
- Analyse to (φ_r, ψ, χ), (φ, ψ, χ) or (φ, u, v)

Skill and balance scores



- Grey = previous localisations
  - circles = standard localisation
  - triangles, squares = localise in (ψ, χ)-space
- Colour = balance-aware localisation
  - blue = analyse to (φ_r, ψ, χ)
  - green = analyse to (φ, ψ, χ)
  - red = analyse to (φ, u, v)
- Balance-aware localisation outperforms previous effort.
- Analysing to (φ_r, ψ, χ) is slightly less accurate, but clearly better balanced.
- Little difference between analysis to (φ, ψ, χ) and (φ, u, v).

Ease of tuning



- 16 members
- covariance inflation = 2%
- various localisation lengths
- 6 different EnKFs

- Generally the new localisations are penalised less for mis-tuning than the old.
- The EnKFs with balance-aware localisation also have less variation in tuning parameters if tuned for tropics, northern hemisphere, or southern hemisphere instead of for global scores.

Impact of initialisation

- 32 members, comparing no-NNMI to with-NNMI assimilation.
- Largest difference for standard localisation.
- Generally smaller difference for balance-aware localisation.
- (φ_r, ψ, χ) is the best-balanced version, and its scores deteriorate when NNMI included.

	No NNMI			With NNMI	
	STDE φ	STDE (u,v)	Bal.	STDE φ	STDE (u,v)
Standard loc.	264	6.2	182	208	5.7
(φ, ψ, χ)-loc.	236	6.1	110	217	5.9
$(\varphi,\psi)(\chi)$-loc.	239	6.1	113	219	5.9
Bal-awr loc., (φ_r, ψ, χ)	225	6.0	91	230	6.1
Bal-awr loc., (φ, ψ, χ)	229	5.9	115	211	5.8
Bal-awr loc., (φ, u, v)	229	5.9	112	216	5.8

Conclusions (so far …)

- Balance-aware covariance localisation leads to more accurate and better balanced analyses.
  - Using the full VAR variable transformations in the localisation is a good idea.
- Analysis to unbalanced φ is slightly less accurate, but better balanced, than analysis to full φ.
  - Accuracy loss because φ analysis comes from …
    - φ obs impact ψ and φ_r via ensemble stats,
    - ψ then gives φ_b via NLB equation, $\varphi = \varphi_b + \varphi_r$.
  - Balance gain from explicit NLB.
- Balance-aware EnKFs seem to be less sensitive to tuning, observation density, etc than old.
- Incorporating NNMI makes (φ_r, ψ, χ)-analysis less accurate.
- (φ_r, ψ, χ)-analysis without balance-aware localisation is disastrous.