# Ensemble $J_b$ Modelling

Meteorological Research Division
Environment Canada

Mark Buehner
11 June 2007

# Background

- Canadian NWP centre currently has both a global 4D-Var (for deterministic forecasts) and EnKF (for probabilistic forecasts)
- Provides good opportunity to evaluate use of flow-dependent ensemble background error covariances in a variational system
- Goal of presentation:

1. describe earlier experiments of using EnKF-estimated flow-dependent error covariances in 3D-Var
2. discuss complementary effects of spatial and spectral localization applied to ensemble-estimated error covariances
3. describe plans for incorporating flow-dependent error covariances in a global 4D-Var system (part of comparison of EnKF and 4D-Var)

# Ensemble-based error covariances in 3D-Var

- Approach [described in Buehner (2005), QJRMS]:
  - no localization: elements of control vector determine **global** contribution of each ensemble member to the analysis increment:

    $\Delta x = \sum (e^k - \langle e \rangle)\ \xi^k$ ($\xi^k$ is a scalar)

  - spatial localization: elements of control vector determine **local** contribution of each ensemble member to the analysis increment:

    $\Delta x = \sum (e^k - \langle e \rangle) \circ (C^{1/2}\ \xi^k)$ ($\xi^k$ is a vector)

  - can also combine with standard **B** matrix:

    $\Delta x = \beta^{1/2} \sum (e^k - \langle e \rangle) \circ (C^{1/2}\ \xi^k) + (1-\beta)^{1/2}\ B^{1/2}\ \xi_{HI}$

  - in each case, Jb is Euclidean inner product:

    $J_b = 1/2\ \xi^T\ \xi$

# Sampling error in ensemble-based error covariances

- Test ability of ensemble-based error covariances to reproduce “true” covariances as function of ensemble size and amount of localization

- Use operational **B** matrix as truth (with homog/isotr. correlations for main analysis variables), generate ensemble members:

$$e^k = B^{1/2} \varepsilon^k \qquad \text{where } \varepsilon^k = N(0,\mathbf{I})$$

- Final value of $J_o$ (including all operational data) is used as a simple measure of accuracy of ensemble-based covariance estimate: ability to fit to observations

- Assume resulting covariances are affected by sampling error in a similar way as with an EnKF-generated ensemble

# Sampling error in ensemble-based error covariances

Final value of Jo (normalized by value from using “true” **B**) as a function of localization radii and ensemble size:

| Localization radii | | Ensemble size | | | |
|---|---|---|---|---|---|
| **Horizontal (km)** | **Vertical (ln(P))** | **32** | **128** | **512** | ∞ |
| ∞ | ∞ | 3.15 | 3.10 | 2.98 | 1.00 |
| **10 000** | ∞ | 2.72 | 2.30 | 1.77 | 0.96 |
| **5 000** | ∞ | 2.42 | 1.83 | 1.35 | 0.91 |
| **2 800** | ∞ | 2.09 | 1.46 | 1.12 | 0.84 |
| **1 000** | ∞ | 1.53 | 1.04 | | 0.65 |
| **10 000** | **2** | 2.23 | 1.73 | 1.31 | 0.94 |
| **5 000** | **2** | 1.82 | 1.35 | 1.08 | 0.89 |
| **2 800** | **2** | 1.47 | 1.11 | 0.97 | 0.82 |
| **1 000** | **2** | 1.04 | 0.88 | | 0.63 |

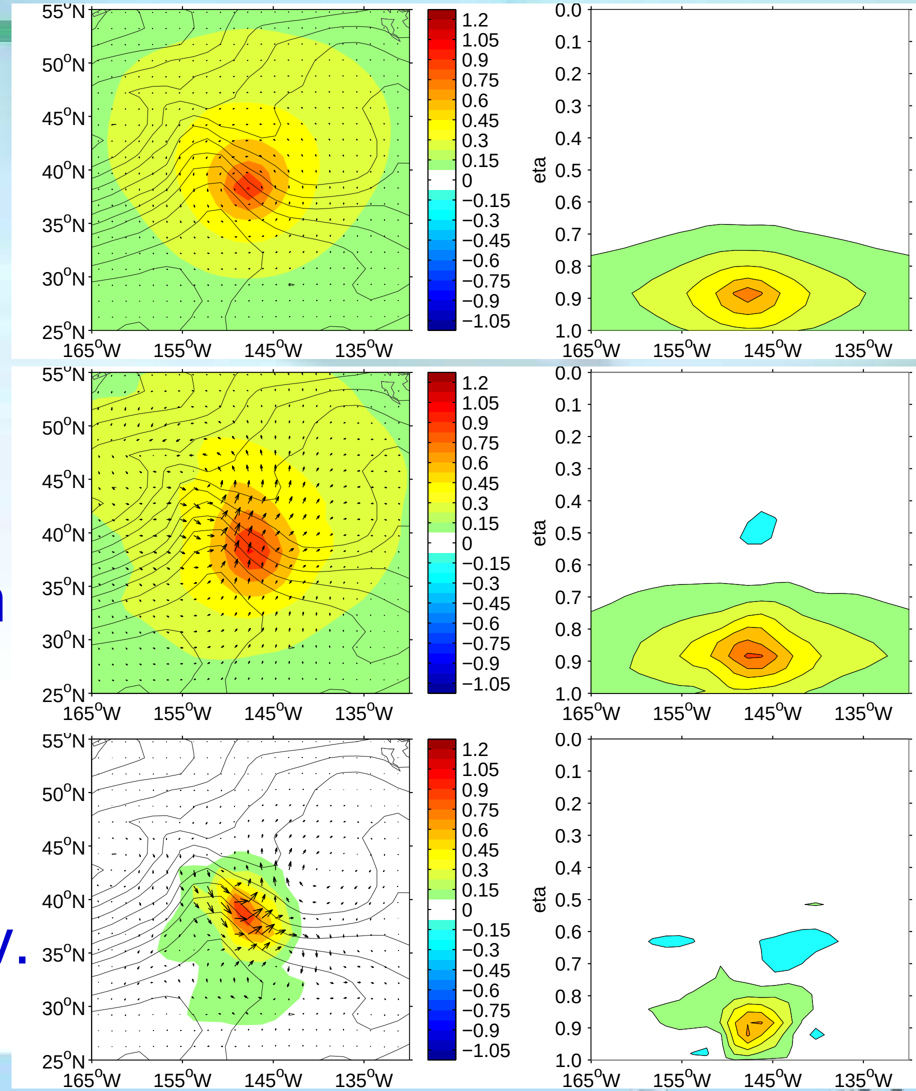
# Earlier tests with EnKF error covariances

- Corrections to T and UV in response to a single obs of T near the surface
- Black contours show background T
- EnKF error covariances from 128 ensemble members and horizontal and vertical localization

3D-Var

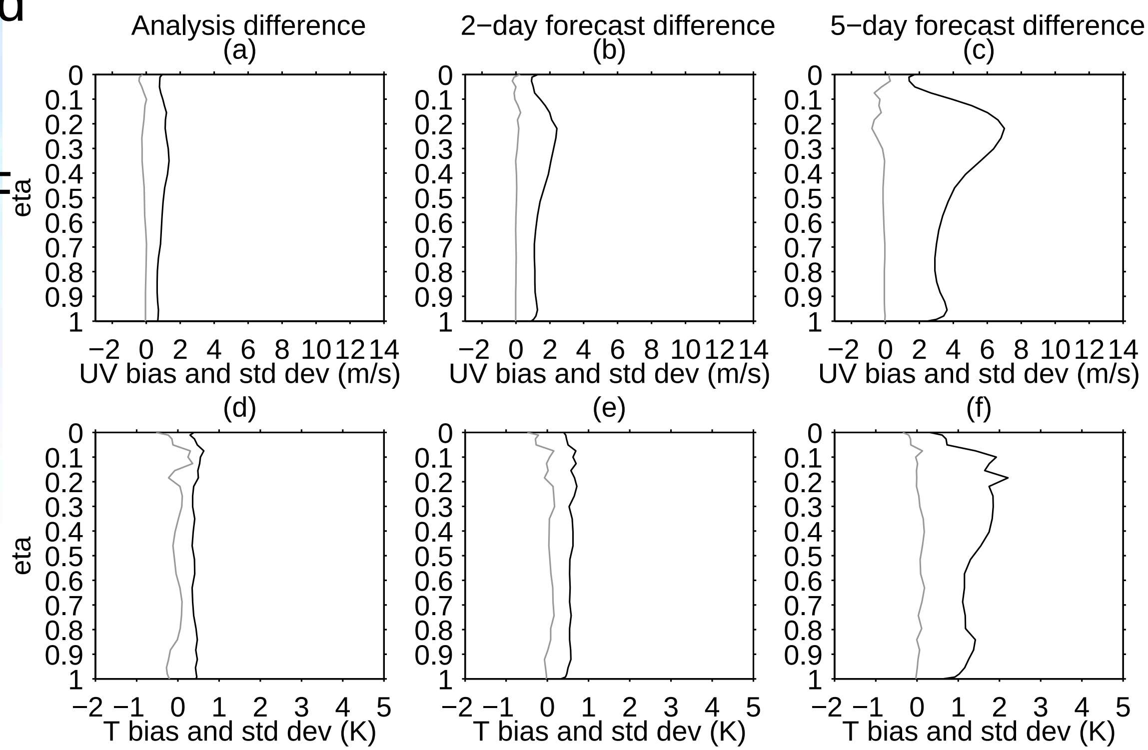
4D-Var (obs at end of 6h window)

EnKF error cov.



# Earlier tests with EnKF error covariances

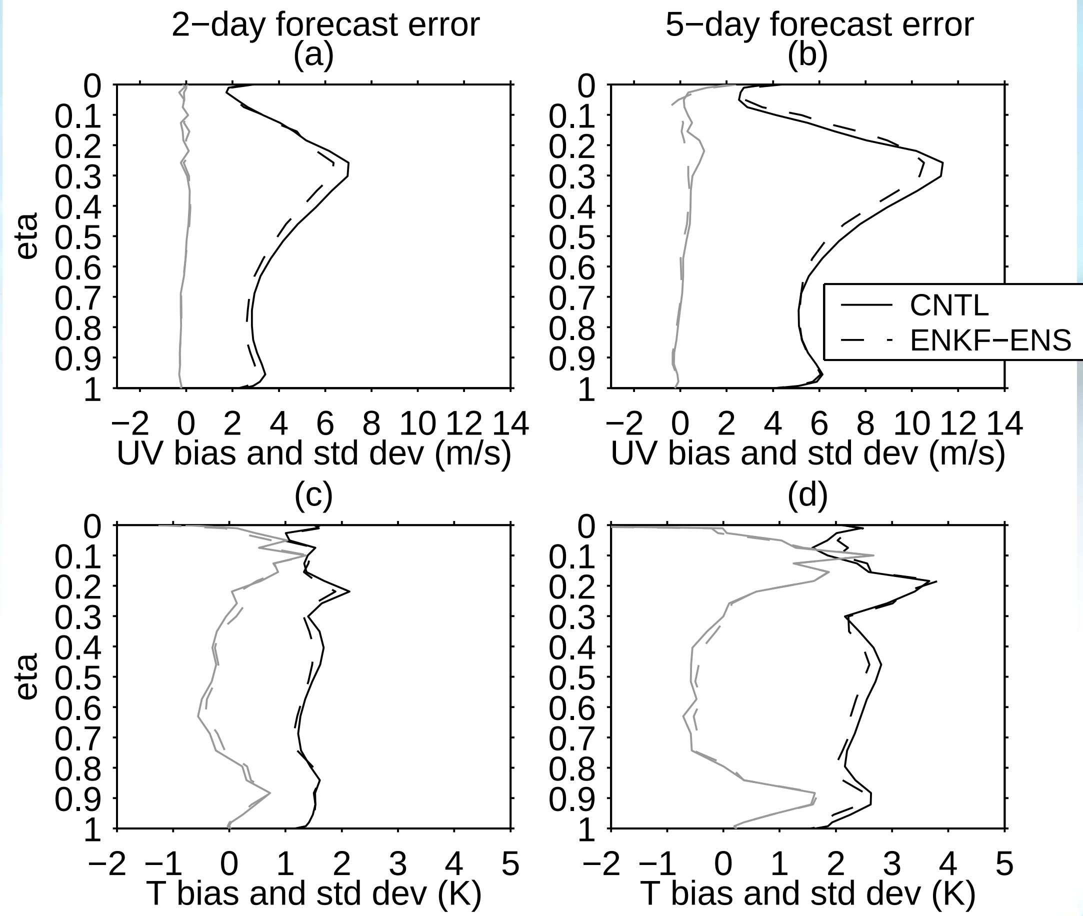
- Impact of EnKF vs. standard 3D-Var error covariances
- Horizontal and vertical localization applied to EnKF covariances
- Single case of rapidly developing system over Pacific (12 UTC, 27 May 2002)
- Bias (grey curves) and std dev (black curves) of the analysis and forecast differences



# Earlier tests with EnKF error covariances

- Forecast error measured vs. analyses from CNTL assimilation experiment
- General improvement from using EnKF error covariances

- Small improvement also seen in scores averaged over 2 week forecast-analysis experiments
- Should revisit, now 4D-Var and EnKF has also been improved

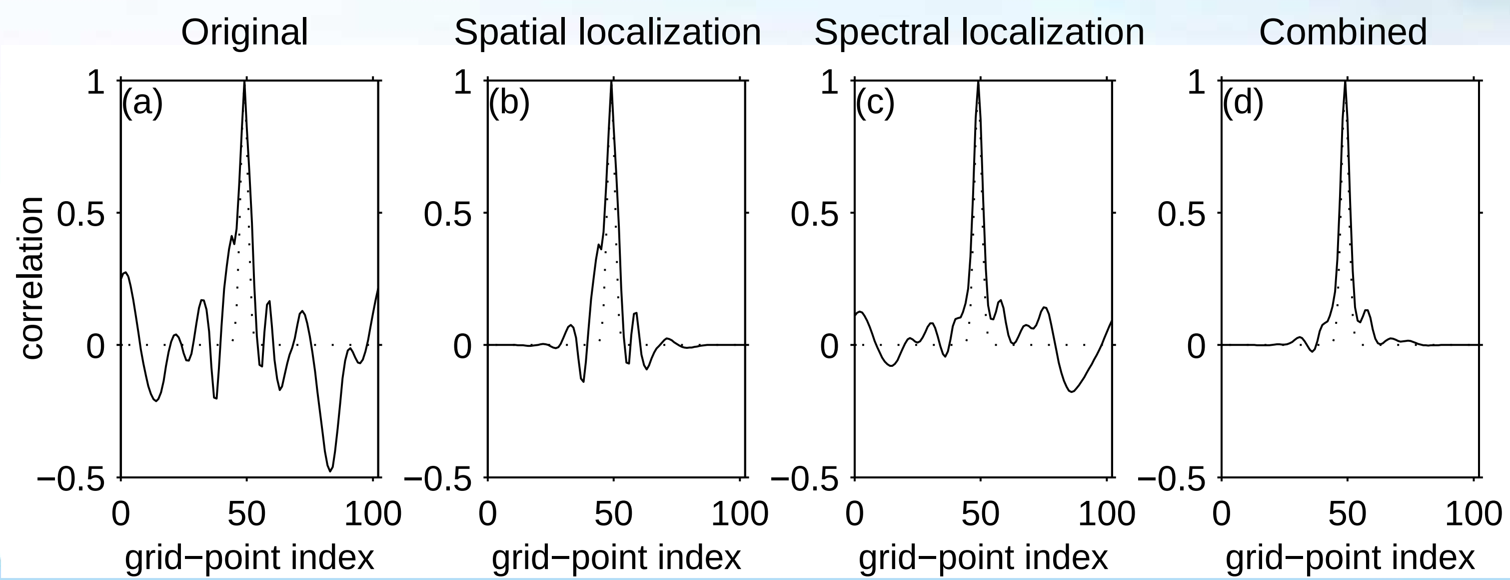


# Spatial and spectral correlation localization

- Approach for modelling correlations very different in operational variational system vs. EnKF:
  - homogeneous correlations (Var)
  - independently estimated at each grid-point (EnKF)
- When correlations estimated from a finite sample size, neither is likely to be optimal
- Averaging of correlations over a **local** region should be better:
  - reduce sampling error through averaging
  - maintain most of spatial/flow dependence of correlations
- Spatial averaging of correlations (convolution) is equivalent to localization of correlations in spectral space (multiplication) [from Buehner and Charron (2007) QJRMS]

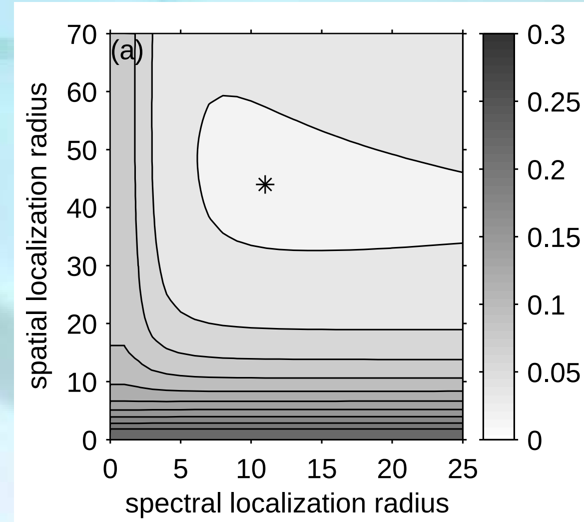
# Spatial and spectral correlation localization

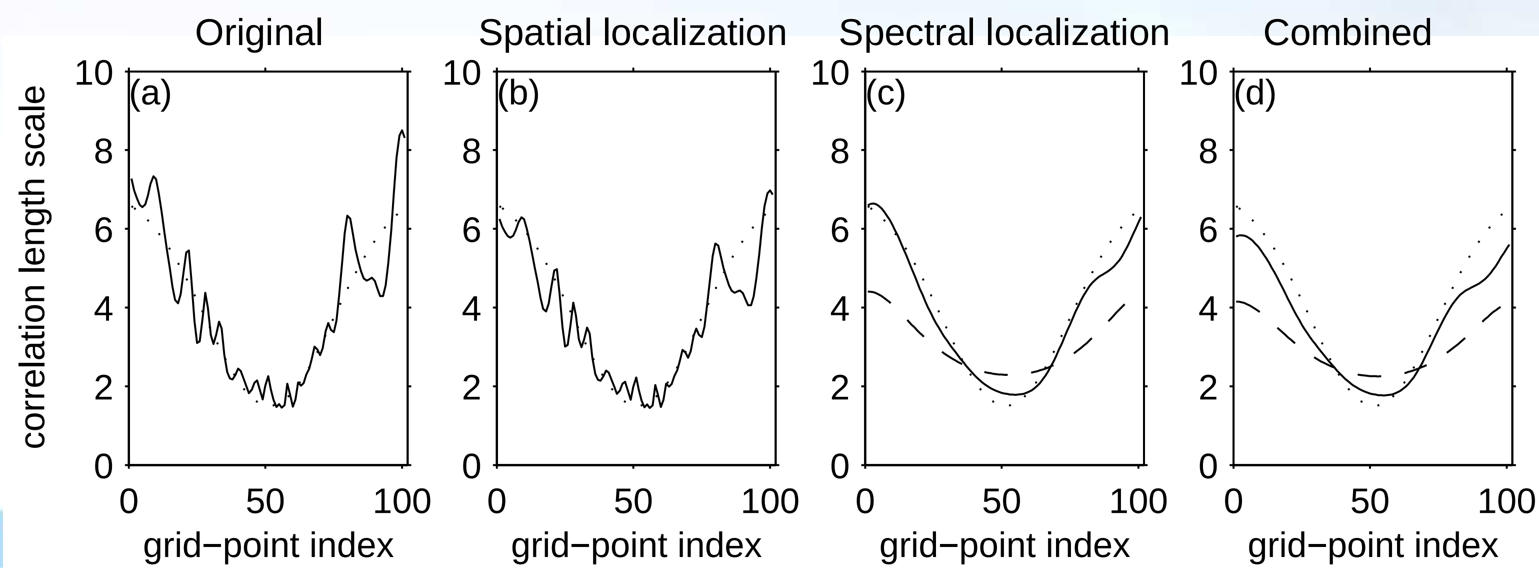
- Idealized 1-D example using prescribed "true" heterogeneous correlations and estimated correlations from 30 realizations
- Spatial localization cannot improve short-range correlations
- Spectral localization cannot remove long-range spurious correlations
- Combination seems to give best result



# Spatial and spectral correlation localization

- For this example, a unique optimal combination of spatial and spectral localization exists (minimum rms error of correlations)
- Spectral localization dramatically improves local estimate of correlation length scale: $(-d^2C/dx^2)^{-1/2}$
- With too much spectral localization, start to loose heterogeneity (dashed)



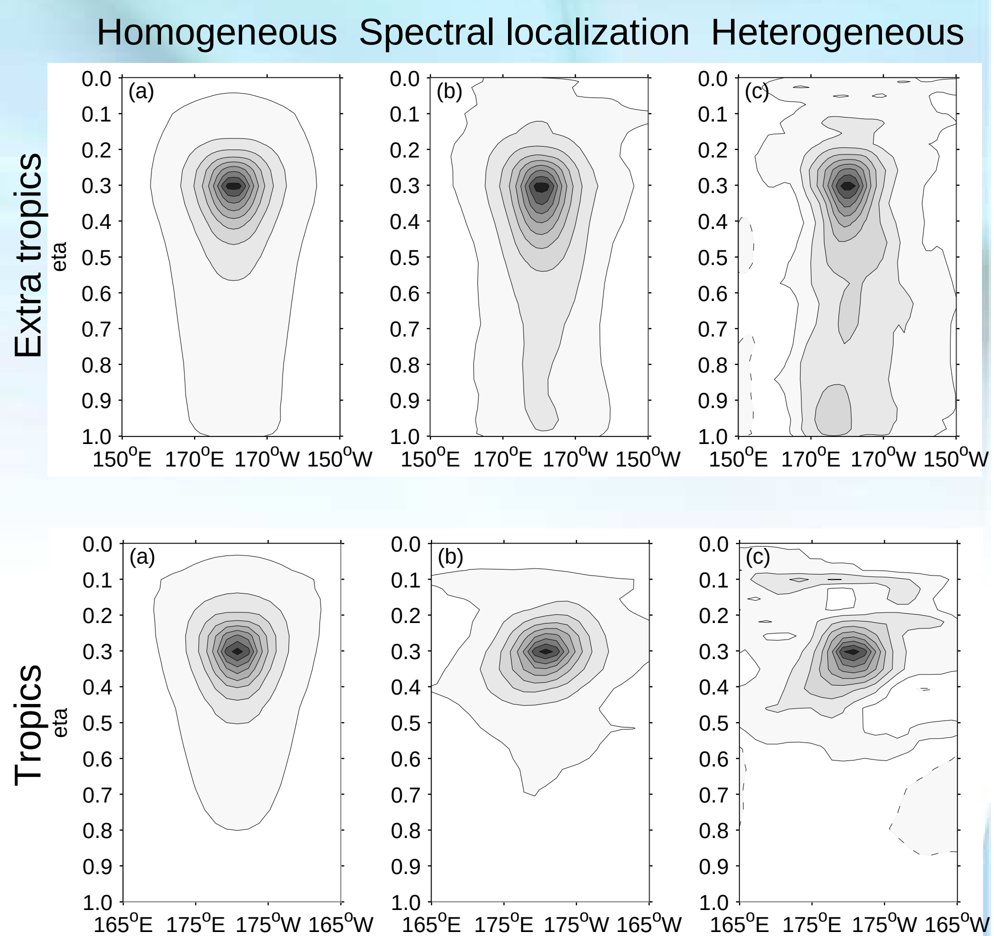



# Spatial and spectral correlation localization

- Apply in variational system, similar technique as spatial localization
- Elements of control vector determine local contribution (in spectral space) of members to analysis increment:

$$\Delta x = S^{-1} \sum (S(e^k - \langle e \rangle)) \circ (C_{sp}^{1/2} \xi^k)$$

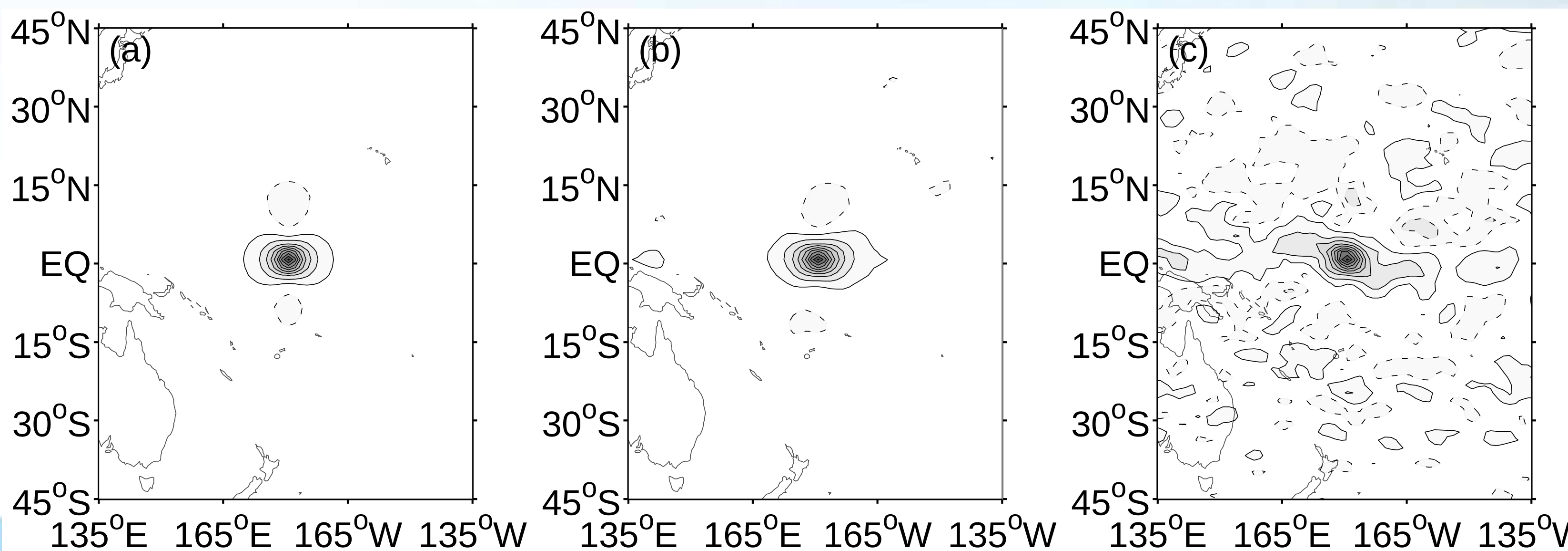
- Use ensemble of 266 members generated from an ensemble of 3D-Var forecast-analysis experiments
- Spectral correlations forced to zero beyond total wavenumber difference of 10



# Spatial and spectral correlation localization

- Like with homogeneous and isotropic correlations, still need to apply spatial localization to damp long-range spurious correlations
- However, with current approach combining spectral and spatial localization would result in very large control vector
- Wavelet-diagonal approach has similar spectral-spatial localization

Homogeneous | Spectral localization | Heterogeneous



# Plan for testing EnKF covariances in 4D-Var

- Prompted by workshop planned for November 2008 in Argentina
- Currently, EnKF and 4D-Var are too different to allow useful comparison: horizontal resolution, deterministic vs. probabilistic, etc.
- Design experiments to isolate specific differences:

1) **standard EnKF**: use ensemble mean for verification (low-res)
2) **"deterministic" EnKF**: additional member with no perturbations to simulate obs or model error (low-res)
3) **incremental "deterministic" EnKF**: additional deterministic member at higher horizontal resolution than EnKF ensemble
4) **incremental 4D-Var with ensemble-based B**: ensemble-based error covariances at beginning of assimilation window with same localization as EnKF
5) **incremental 4D-Var with static B**: same as operational deterministic analysis system

# Plan for testing EnKF covariances in 4D-Var

Specific differences whose impact could be evaluated:

- **smoothing of ensemble mean relative to deterministic forecast**:

  1) standard EnKF vs. 2) “deterministic” EnKF at same resolution

- **different analysis approach with equal covariances at beginning of assimilation window**:

  3) incremental “deterministic” EnKF vs. 4) incremental 4D-Var with ensemble-based B

- ***impact of flow-dependent ensemble-based covariances in 4D-Var**:

  4) ensemble-based error covariances vs. 5) static covariances