Environment Canada Environnement Canada

Canada

Scale-dependent covariance localization for EnVar data assimilation

Mark Buehner and Anna Shlyaeva

Data Assimilation and Satellite Meteorology Research

23-26 February 2015

4th International Symposium on Data Assimilation, Kobe, Japan

Outline

- Overview of 4DEnVar operational implementation
- Scale-dependent covariance localization:
  - Motivation
  - Estimation of ensemble covariances in idealized 1D domain
  - Idealized assimilation test with 2D sea ice concentration ensemble

4D Ensemble-Variational assimilation: 4DEnVar

- 4DEnVar implemented in operational regional and global deterministic weather prediction systems on 18 November 2014, replacing 4DVar - in combination with numerous other changes (4D-IAU, more IR channels, improved used of RAOBS, improved radiance BC …)
- 4DEnVar uses a **variational assimilation approach** in combination with the already available **4D ensemble covariances** from the EnKF
- By making use of the 4D ensembles, 4DEnVar performs a 4D analysis without the need of the tangent-linear and adjoint of forecast model
- Consequently, it is **more computationally efficient** and **easier to maintain/adapt** than 4DVar
- Future improvements to EnKF will benefit both ensemble and deterministic prediction systems
- Increased incentive to improve EnKF and **improve how ensemble members used within 4DEnVar**

EnVar formulation

- In 4DVar the 3D analysis increment is evolved in time using the TL/AD forecast model (here included in $\mathbf{H}_{4D}$):

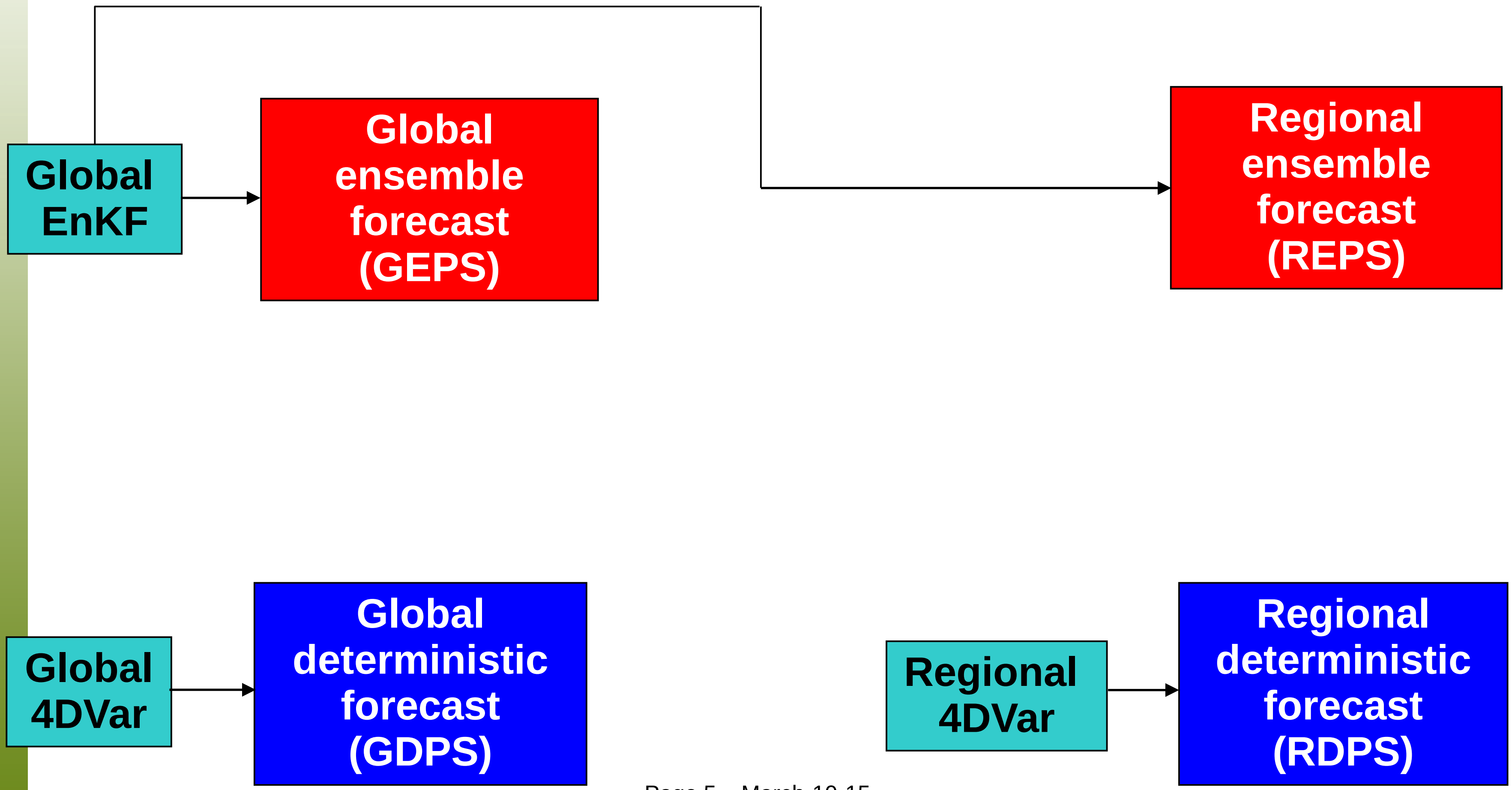
$$J(\Delta\mathbf{x}) = \frac{1}{2}(H_{4D}[\mathbf{x}_b] + \mathbf{H}_{4D}\Delta\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1}(H_{4D}[\mathbf{x}_b] + \mathbf{H}_{4D}\Delta\mathbf{x} - \mathbf{y}) + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{B}^{-1}\Delta\mathbf{x}$$

- In EnVar the background-error covariances and analysed state are explicitly 4-dimensional, resulting in cost function:

$$J(\Delta\mathbf{x}_{4D}) = \frac{1}{2}(H_{4D}[\mathbf{x}_b] + \mathbf{H}\Delta\mathbf{x}_{4D} - \mathbf{y})^T \mathbf{R}^{-1}(H_{4D}[\mathbf{x}_b] + \mathbf{H}\Delta\mathbf{x}_{4D} - \mathbf{y}) + \frac{1}{2}\Delta\mathbf{x}_{4D}{}^T\mathbf{B}_{4D}^{-1}\Delta\mathbf{x}_{4D}$$

2013-2017: Toward a Reorganization of the NWP Suites at Environment Canada

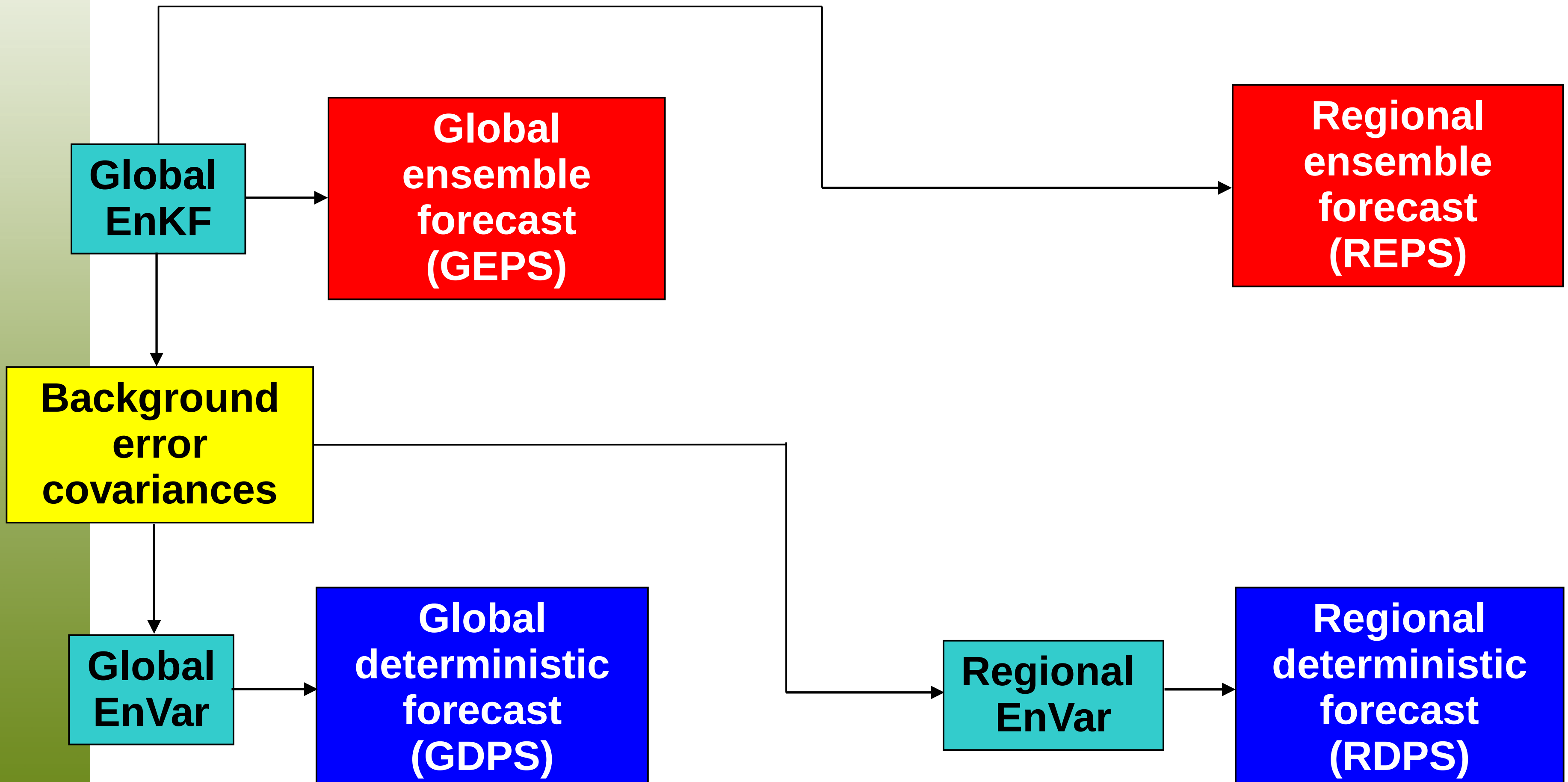
Previous operational systems



global system ***regional system***

2013-2017: Toward a Reorganization of the NWP Suites at Environment Canada

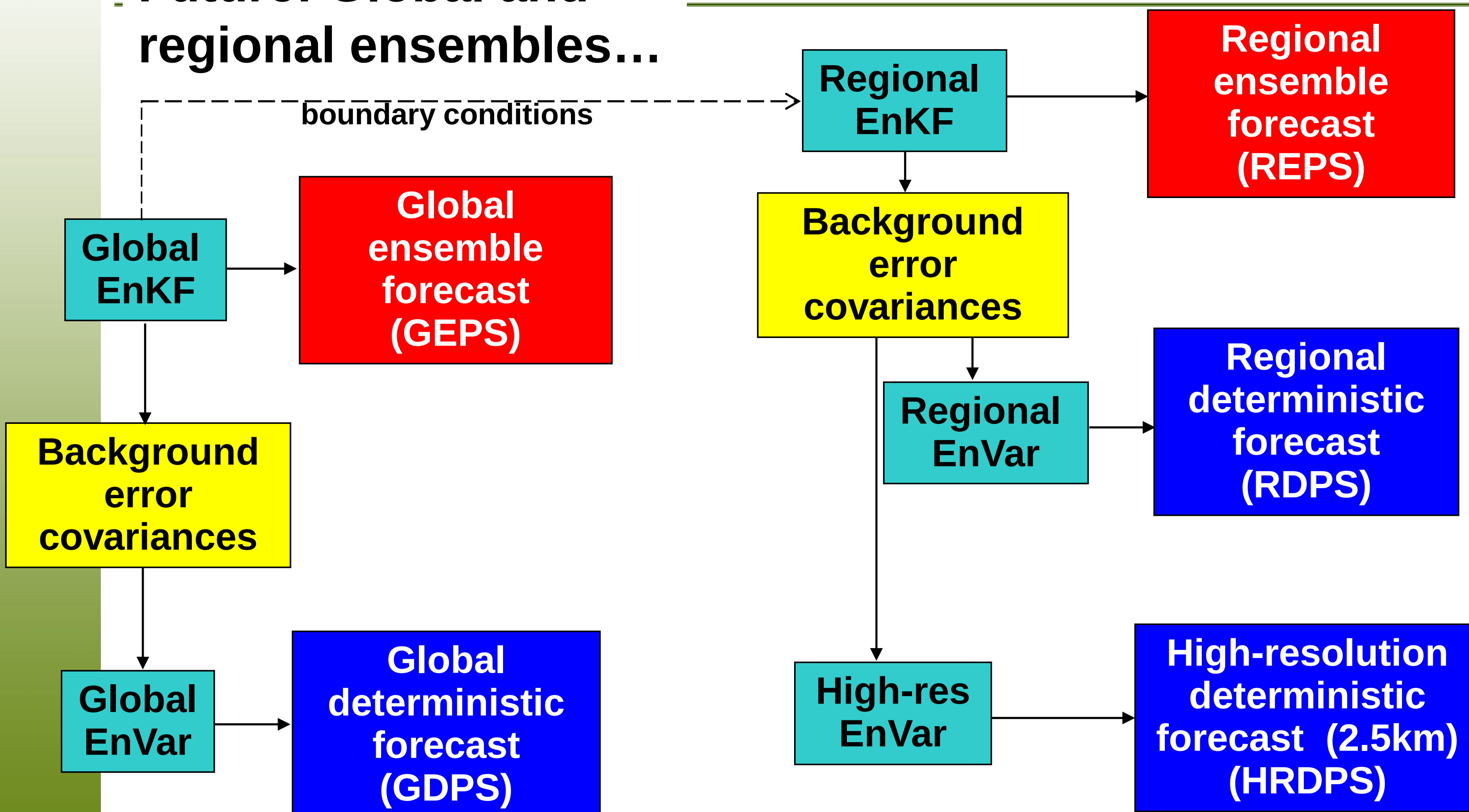
Nov. 2014 implementation: **Increasing role of global ensembles**



global system *regional system*

2013-2017: Toward a Reorganization of the NWP Suites at Environment Canada

Future: Global and regional ensembles…

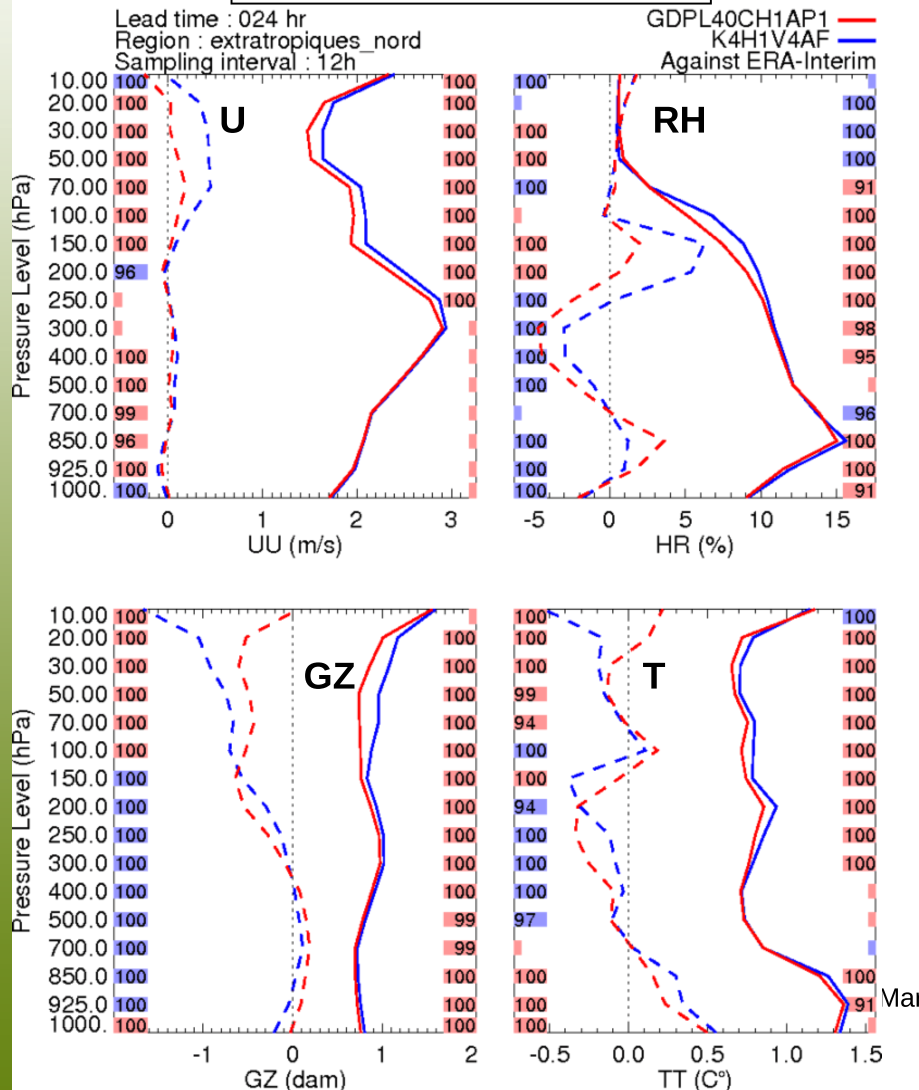


global system *regional system*

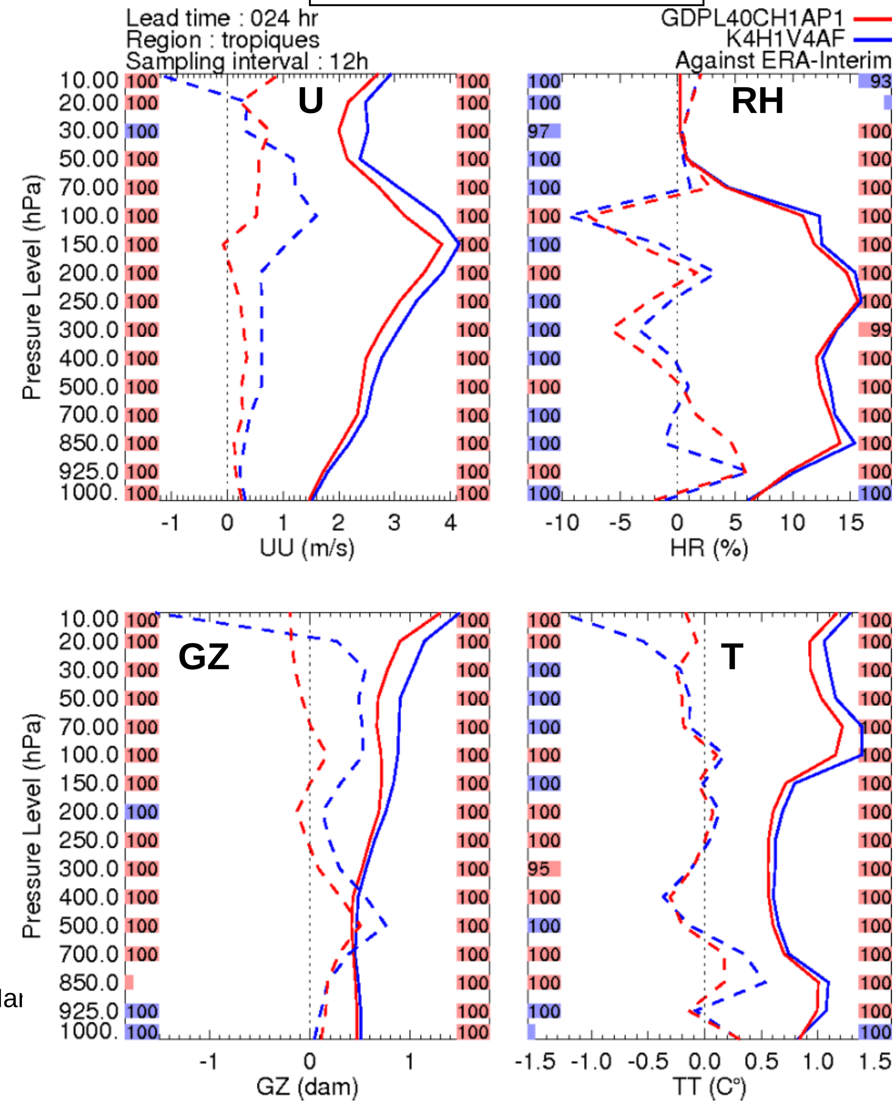
Forecast Results: GDPS 4 (EnVar) vs GDPS 3 (4DVar)

Verification vs. ERA-Interim analyses – 24h, Feb-March, 2011

Northern extratropics



Tropics



Scale-dependent covariance localization

Motivation

- Currently, EnVar uses simple horizontal and vertical localization of ensemble covariances, very similar to EnKF
- Comparing various studies, seems it is best to use different amount of localization depending on application:
  - convective-scale assimilation: ~10km
  - mesoscale assimilation: ~100km
  - global-scale assimilation: ~1000km – 3000km
- In the future, global systems may resolve convective scales
- Therefore, need a general approach for applying appropriate localization to wide range of scales in a single analysis procedure: **Scale-dependent localization**
- Possible in EnVar, since localization acts directly on model-space covariances (not $\mathbf{BH}^T$ and $\mathbf{HBH}^T$ or $\mathbf{R}$)

Scale-dependent covariance localization
General Approach

- Ensemble perturbations decomposed with respect to a series of overlapping spectral wavebands (filter coefficients sum to 1 for each wavenumber)
- Apply scale-dependent spatial localization to the scale-decomposed ensemble covariances, both **within-scale** and **between-scale** covariances
- Keeping the between-scale covariances is necessary to maintain heterogeneity of ensemble covariances (Buehner and Charron, 2007; Buehner, 2012)
- Motivation different than spectral localization where the between-scale covariances are set to zero

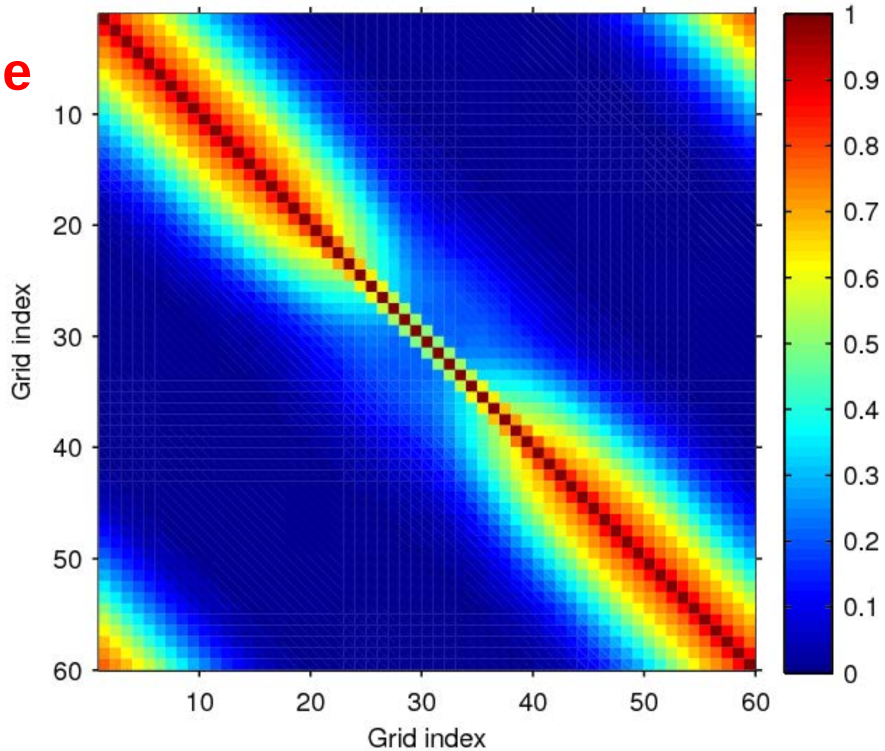
Scale-dependent covariance localization

1D Idealized System

- Idealized system on 1D periodic domain
- Assume a simple “true” heterogeneous covariance function that is a spatially varying weighted average of 2 Gaussian functions with different length scales:

  total = **small scale** + **large scale**

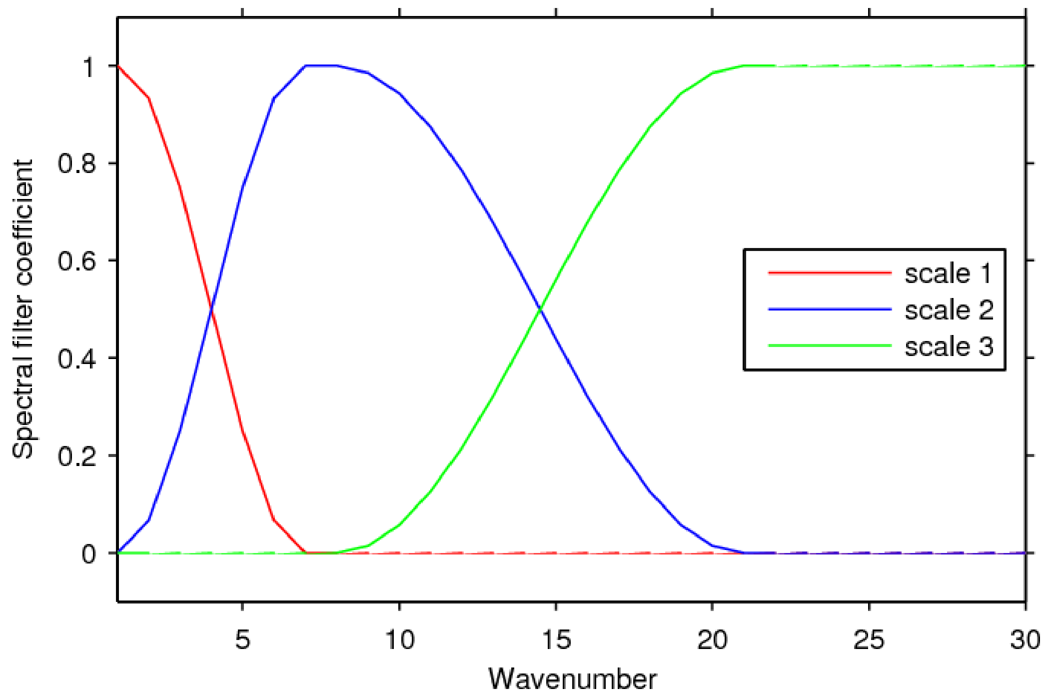
- Length scales of Gaussian functions:
  - large scale: 7 grid points
  - small scale: 0.7 grid points
- Middle of domain dominated by small scale errors, both ends dominated by large scales



Scale-dependent covariance localization

1D Idealized System

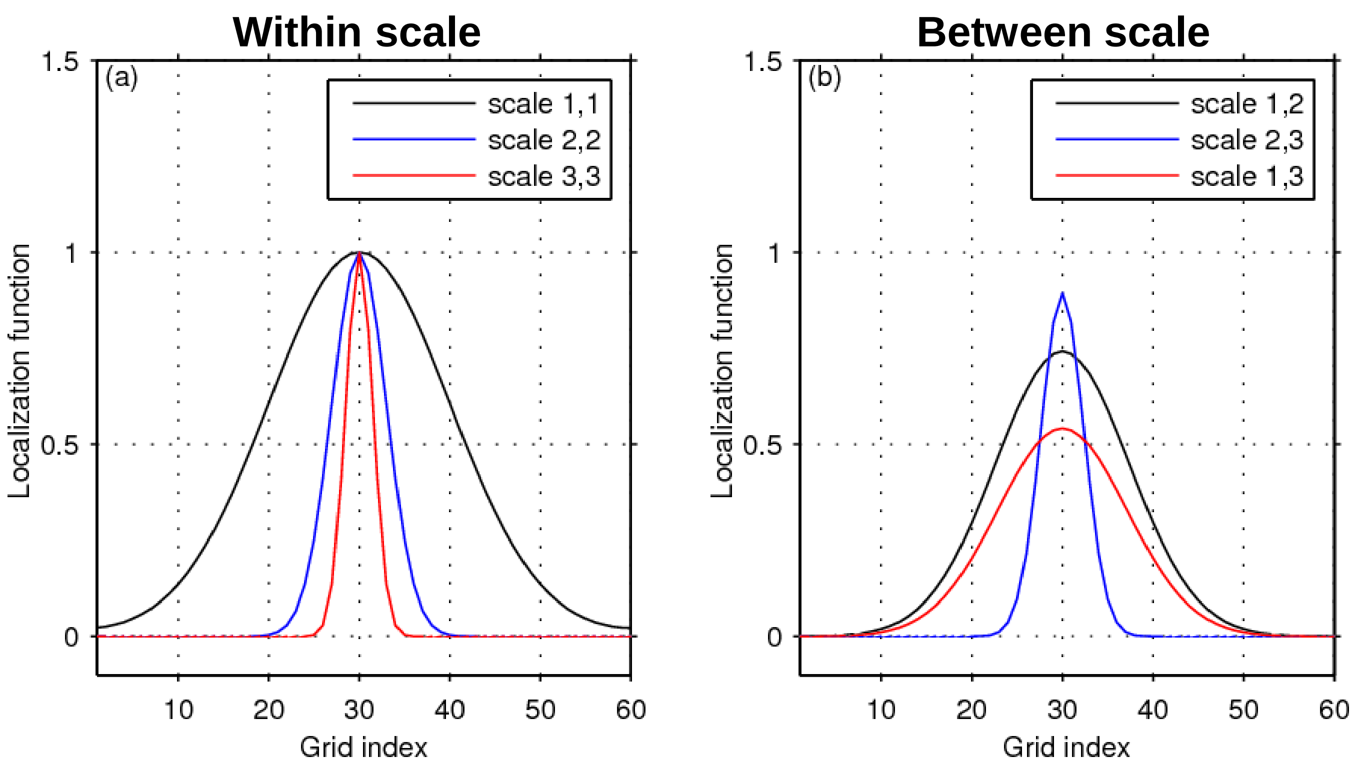
- Ensemble perturbations decomposed with respect to 3 overlapping spectral wavebands:



Scale-dependent covariance localization

1D Idealized System

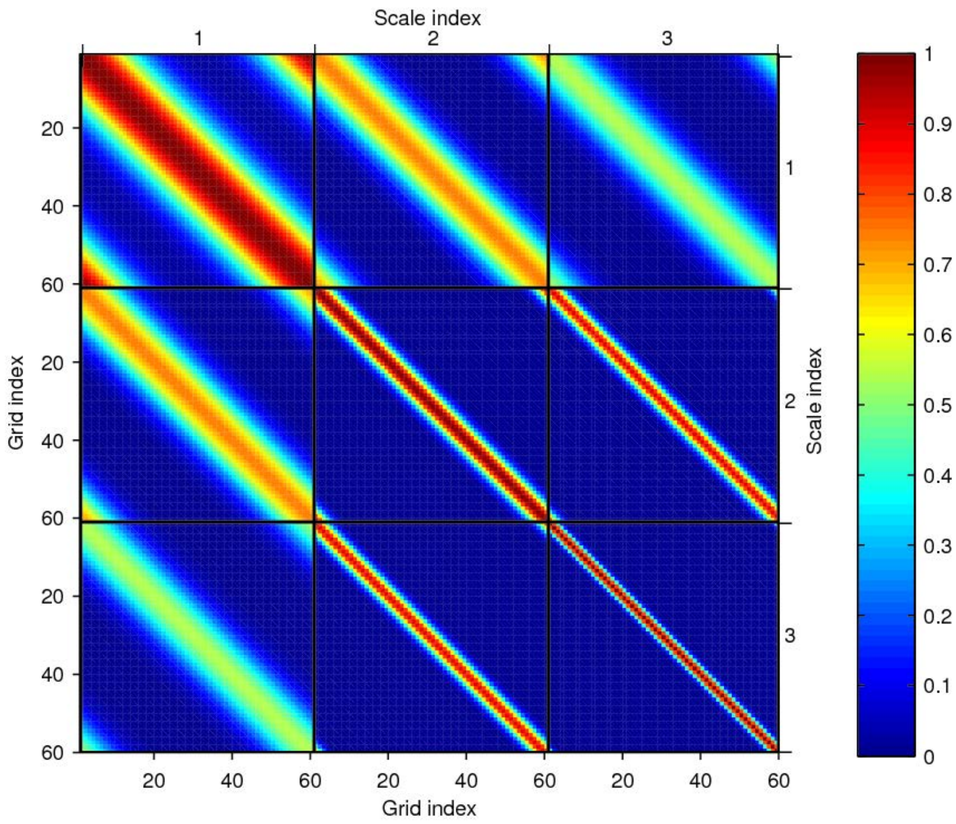
- Scale-dependent homogeneous spatial localization functions (Gaussian) are specified with length scales: **10**, **3**, and **1.5** grid points
- Localization of between-scale covariances constructed to ensure full matrix is positive-semidefinite: $\mathbf{L}_{i,j} = (\mathbf{L}_{i,i})^{1/2}(\mathbf{L}_{j,j})^{T/2}$ → btwn scales i & j



Scale-dependent covariance localization

1D Idealized System

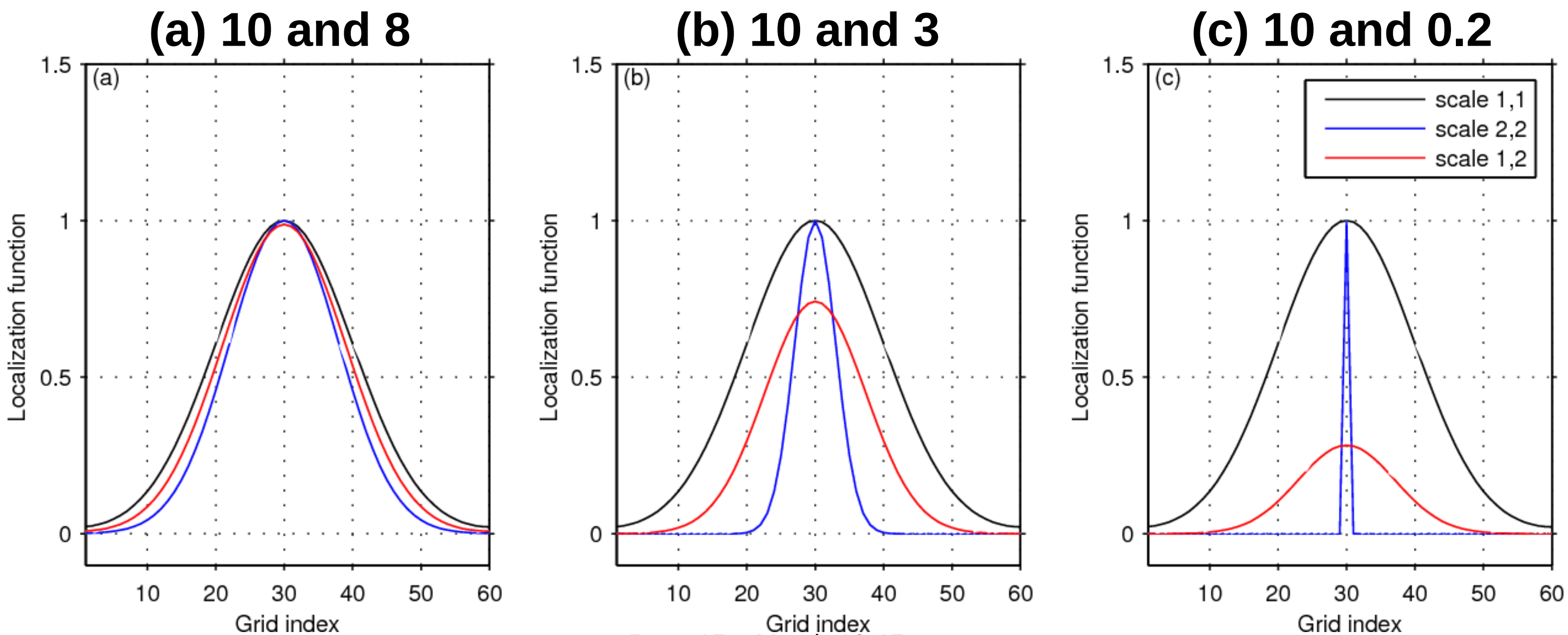
- Within-scale and between-scale localization matrices combined into a single “multi-scale” localization matrix
- Note that between-scale blocks have diagonal values less than 1!
- This is a necessary consequence of requiring the matrix to be positive-semidefinite (we have no choice!)



Scale-dependent covariance localization

1D Idealized System

- The greater the change in localization length scale, the more reduced the between-scale localization function
- This "spectral localization" corresponds with a local spatial averaging of the covariance function, i.e. loss of heterogeneity
- Examples with 2 scales: reduction in between-scale localization

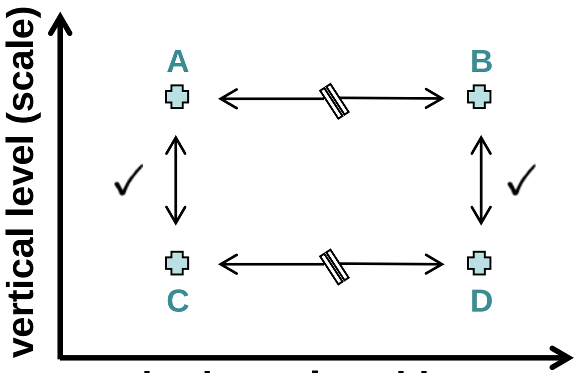


Scale-dependent covariance localization

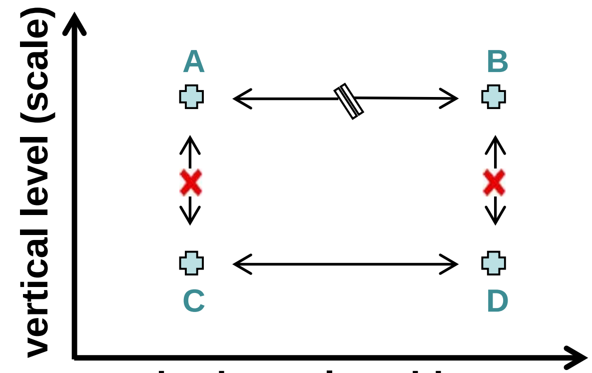
1D Idealized System

- Positive-semidefiniteness required for physically realizable correlations, without it, localized matrix no longer guaranteed to be a covariance matrix
- Reduction in the between-scale localization function from changes in horizontal localization also occurs for between-vertical level localization, which is easier to interpret physically:

Same severe horizontal localization for each level,
Vertical correlations can be maintained:
A ≠ B and C ≠ D, so A = C and B = D possible



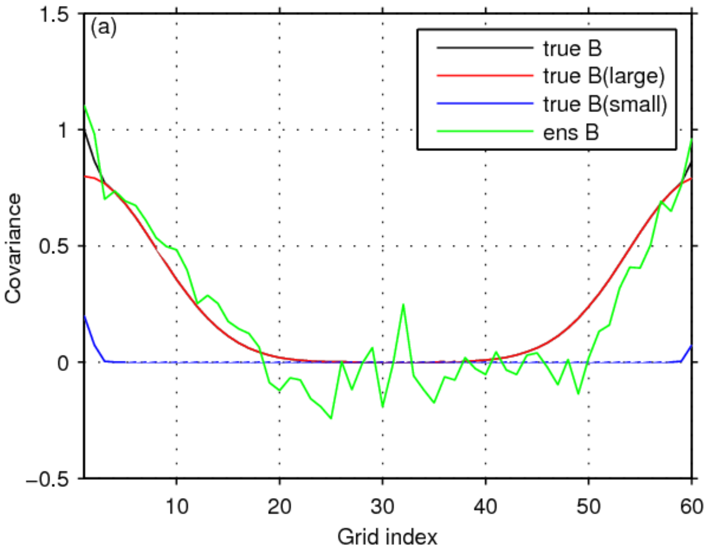
Very different horizontal localization at 2 levels,
Impossible to maintain vertical correlations:
A ≠ B and C = D, so A = C and B = D not possible

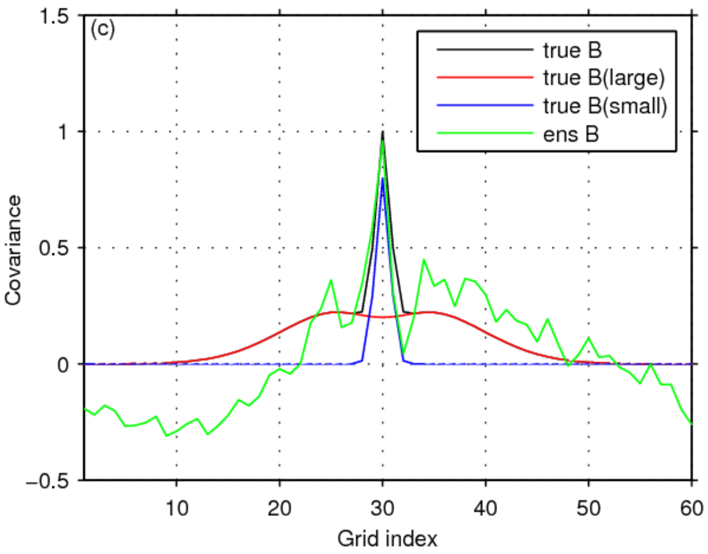


Scale-dependent covariance localization

1D Idealized System

- Mostly large scale at both ends of the domain (top panel) and mostly small scale at the middle (bottom panel)
- Generate a random sample of 50 ensemble members and compute raw sample **ensemble covariance**

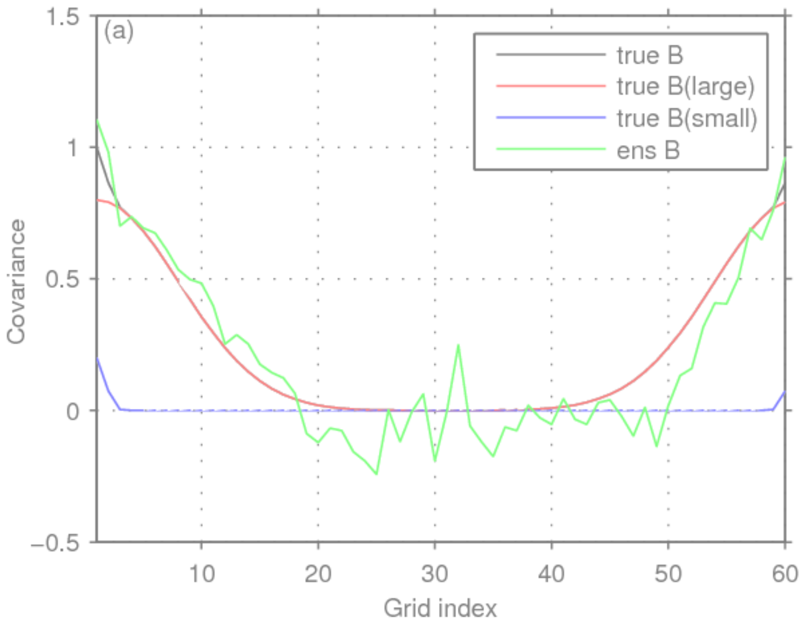


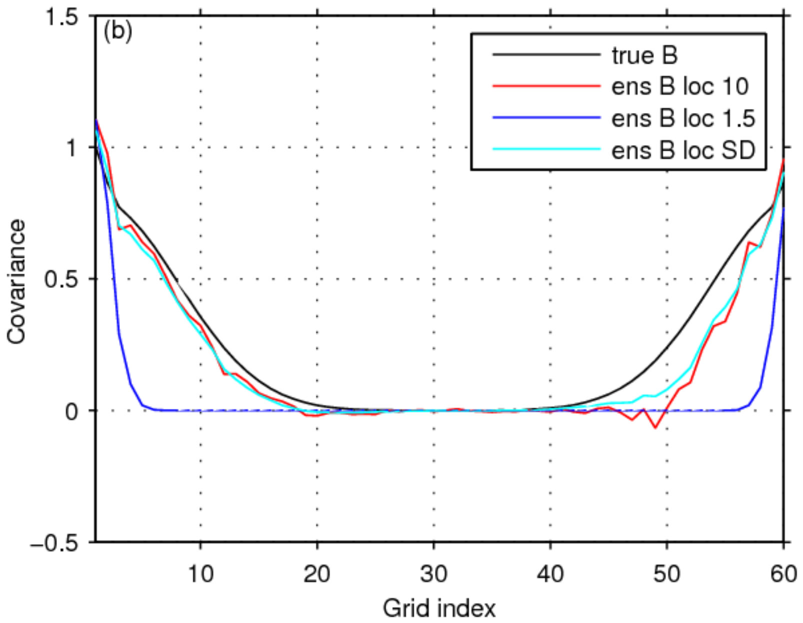


Scale-dependent covariance localization

1D Idealized System

- Apply various localization functions (right panel) and compare with true covariances:
  - **single, large-scale localization: 10 grid points**
  - **single, small-scale localization: 1.5 grid points**
  - **scale-dependent localization: 10, 3, 1.5 grid points**
- For location where true covariances are dominated by large scale component: the small-scale localization does a very bad job, large-scale localization is ok, scale-dependent similar, but less noisy

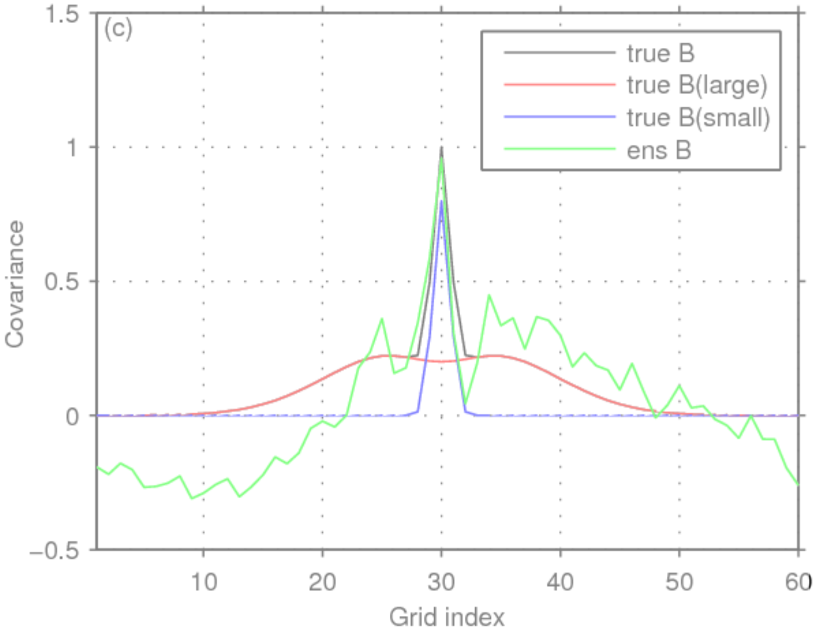


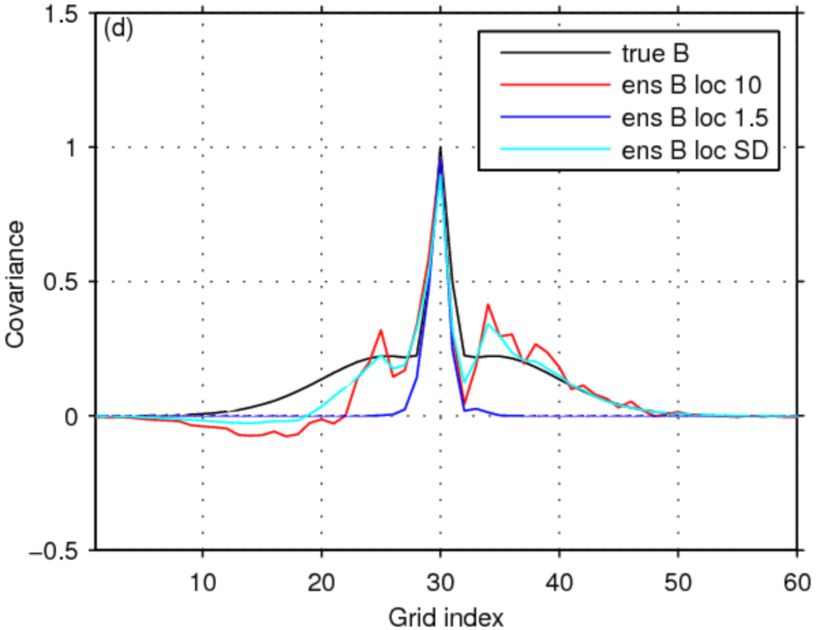


Scale-dependent covariance localization

1D Idealized System

- Apply various localization functions (right panel) and compare with true covariances:
  - **single, large-scale localization: 10 grid points**
  - **single, small-scale localization: 1.5 grid points**
  - **scale-dependent localization with 3 scales: 10, 3, 1.5 grid points**
- Where true covariances are dominated by small scale component: small-scale loc. not quite as bad, but removes large-scale component, scale-dependent smoother and looks better than large-scale loc.

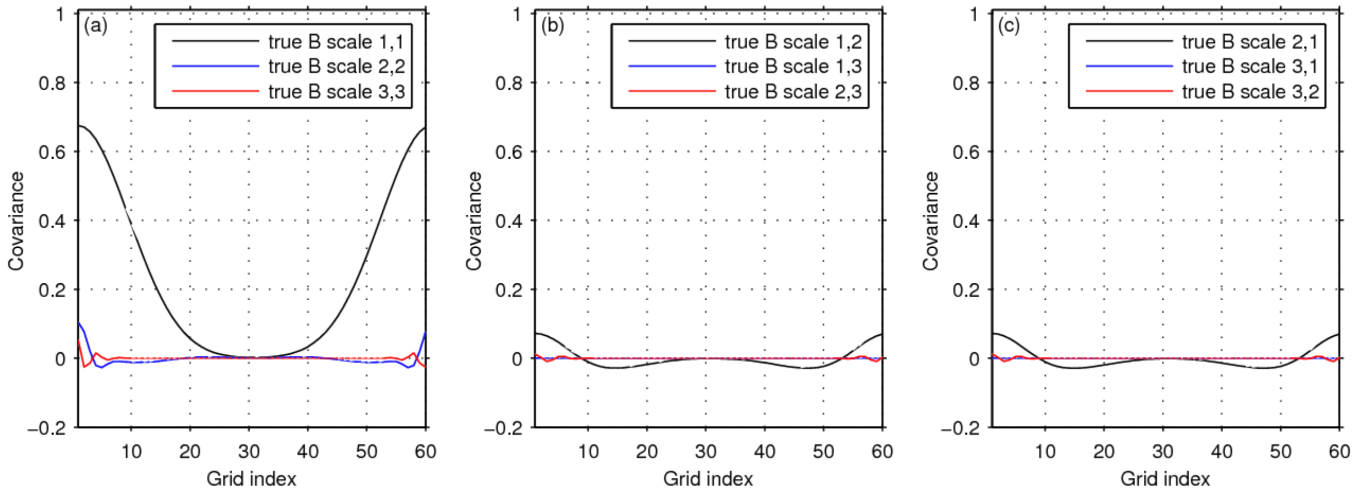




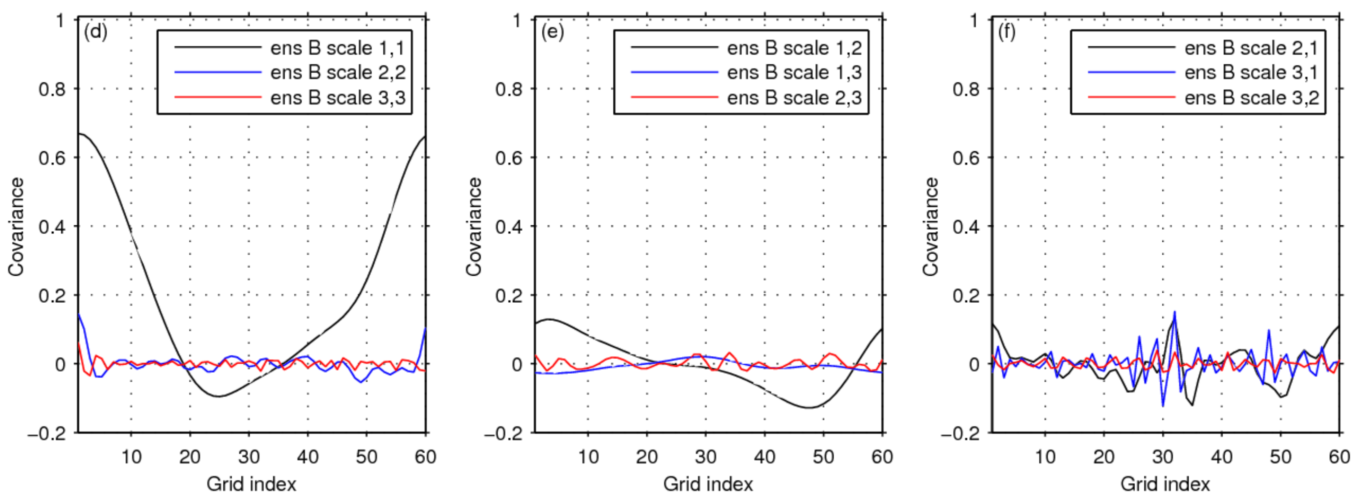
Scale-dependent covariance localization

1D Idealized System

- Look in more detail by dividing true covariance into 3 scales:



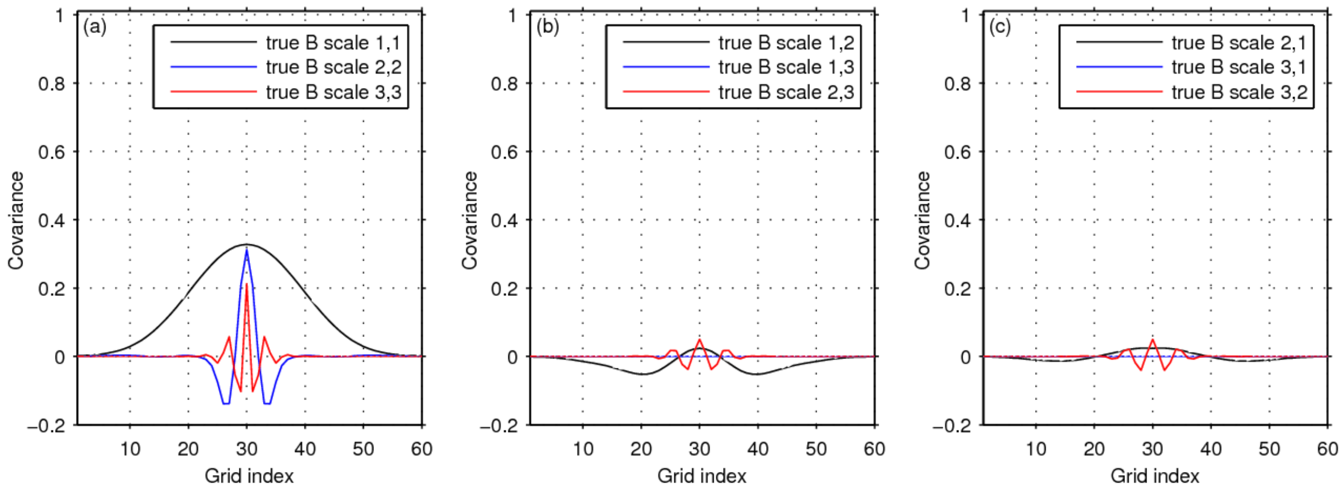
- Same for the ensemble covariance (between-scale cov. not symmetric!):



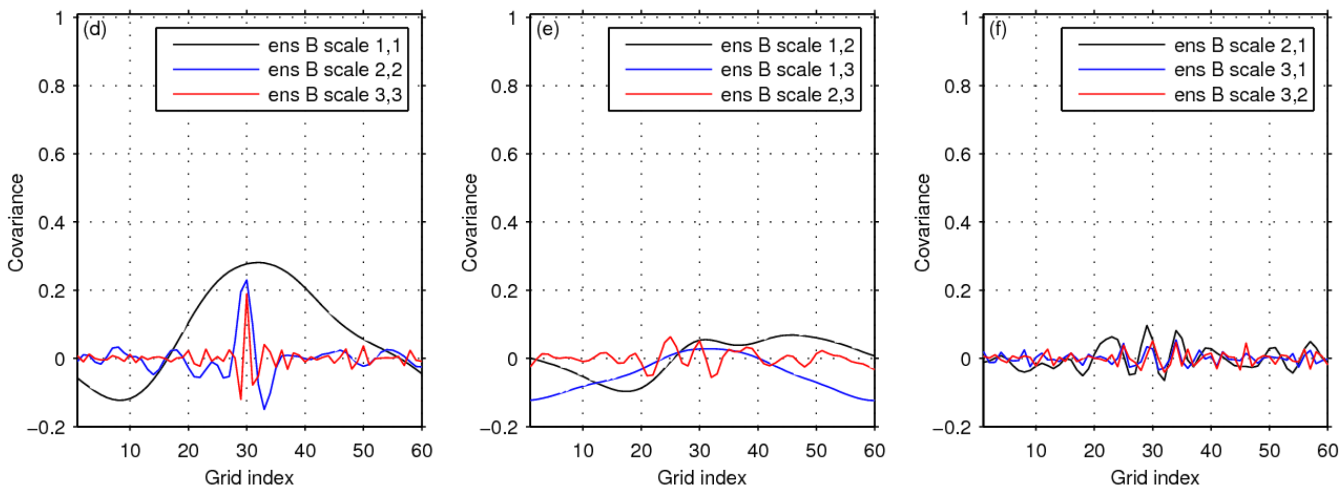
Scale-dependent covariance localization

1D Idealized System

- Look in more detail by dividing true covariance into 3 scales:



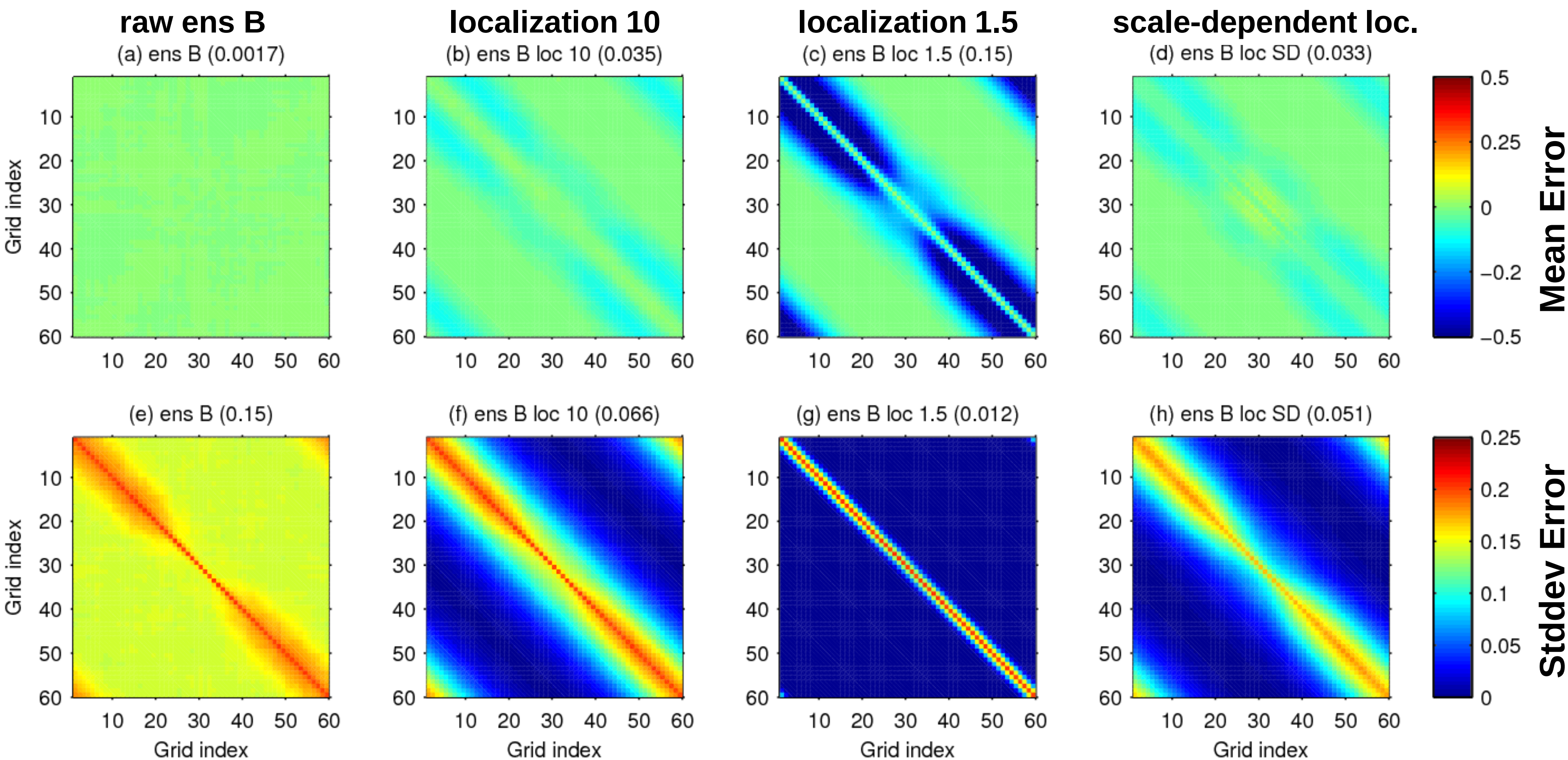
- Same for the ensemble covariance (between-scale cov. not symmetric!):



Canada

Scale-dependent covariance localization
1D Idealized System

- From 5000 random realizations, compute mean and stddev of the error of 50-member ensemble covariances with each type of localization:

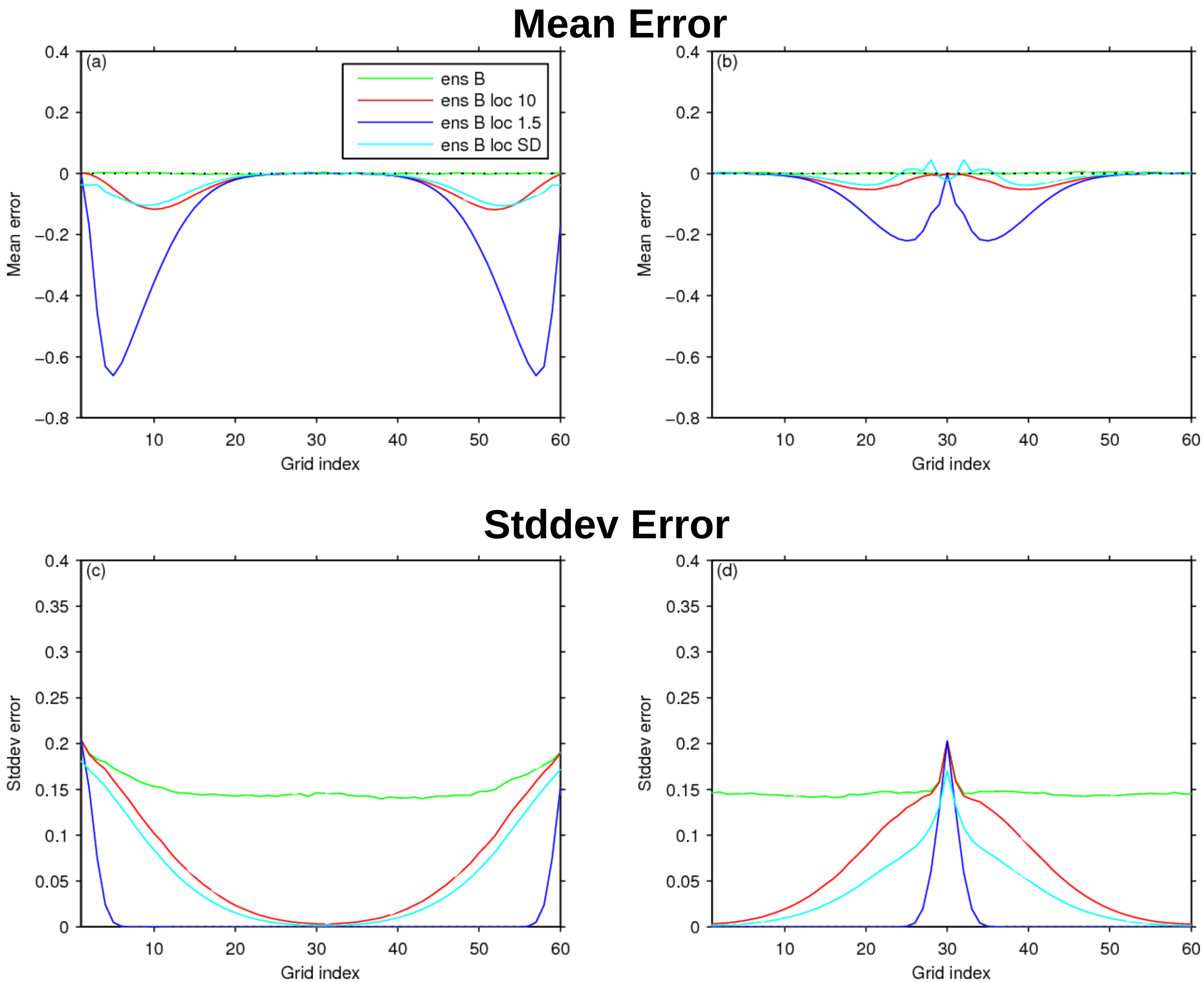


Scale-dependent covariance localization

1D Idealized System

- Same as previous slide, but shown separately for 2 locations only
- Note that unavoidable spectral localization with S-D spatial localization improves variance estimate

- **no localization**
- **large-scale loc.: 10 grid points**
- **small-scale loc.: 1.5 grid points**
- **scale-dependent loc.: 10, 3, 1.5 grid points**



Scale-dependent covariance localization

1D Idealized system: Conclusions

- When decomposed with respect to spatial scale:
  - small-scale component of true covariances are local
  - large-scale component of true covariances are not local
  - small-scale component of raw ensemble has spurious long-range covariances → benefit from more severe localization of **only** small-scale component of ensemble covariances
- Variation in the amount of localization as a function of scale:
  - reduces the between-scale covariances → spectral localization
  - this reduction corresponds with reducing the spatial heterogeneity
  - not possible to keep all heterogeneity and severely increase localization of small scales
- Using scale-dependent spatial localization results in:
  - similar mean error of covariance vs. only large-scale localization (both are much better than only small-scale localization)
  - better stddev error of covariance vs. only large-scale localization, especially in areas where true covariances dominated by small-scales

Scale-dependent covariance localization

Implementation in EnVar

- Analysis increment computed from control vector (**B**$^{1/2}$ preconditioning) using:

$$\Delta \mathbf{x} = \sum_k \sum_j \mathbf{e}_{k,j} \circ \left(\mathbf{L}_j^{1/2} \boldsymbol{\xi}_k \right)$$

where $\mathbf{e}_{k,j}$ is scale *j* of normalized member *k* perturbation

- Varying amounts of smoothing applied to same set of amplitudes for a given member

- Compare with wave-band localization (Buehner, 2012) – independent amplitudes for each member **and** scale (**square** of filter coefficients sum to 1):

$$\Delta \mathbf{x} = \sum_k \sum_j \mathbf{e}_{k,j} \circ \left(\mathbf{L}_j^{1/2} \boldsymbol{\xi}_{k,j} \right)$$

- And compare with standard EnVar approach – one set of amplitudes per member:

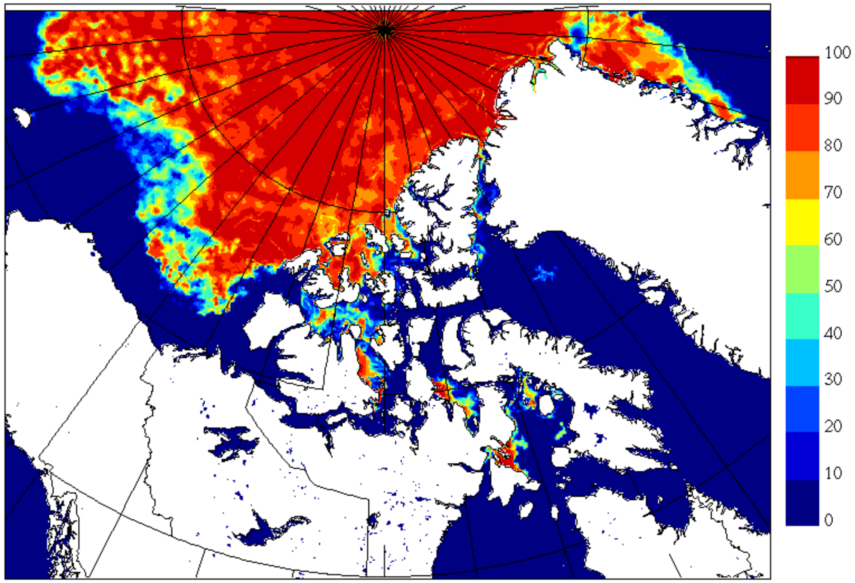
$$\Delta \mathbf{x} = \sum_k \mathbf{e}_k \circ \left(\mathbf{L}^{1/2} \boldsymbol{\xi}_k \right)$$

Scale-dependent covariance localization

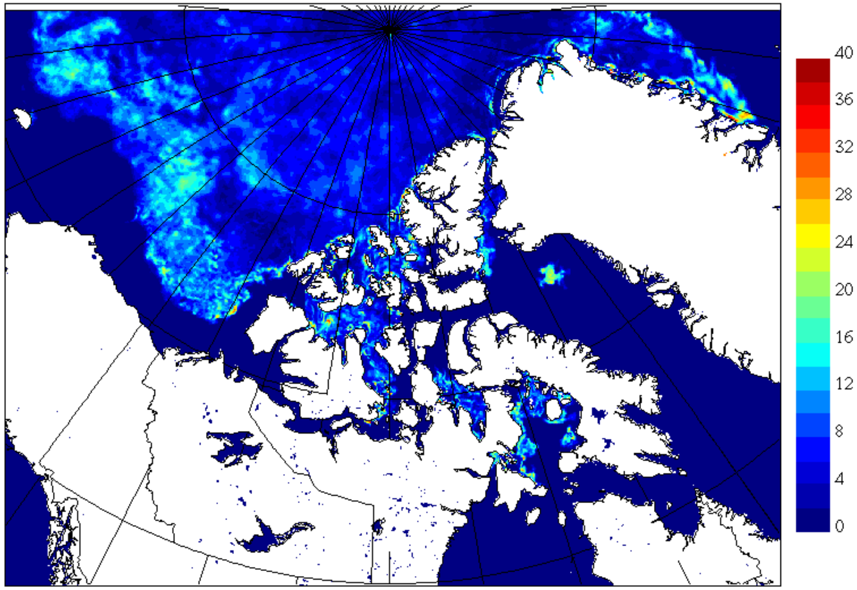
2D Sea Ice Ensemble

- Ensemble of sea ice concentration background fields (60 members, time-lagged ensemble) from the Canadian Regional Ice Prediction System ensemble of 3DVar analyses experiment
- Note the higher spread in the marginal ice zone (MIZ), compared with open water and pack ice areas

Ensemble mean ice concentration



Ensemble spread



Scale separation of ensemble perturbations with diffusion operator

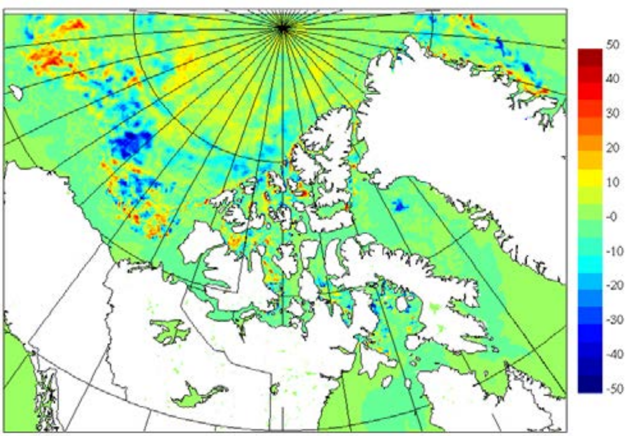
- Apply diffusion with increasing length scales to the original ensemble perturbations
- Decompose into different scales by taking differences between perturbations before and after each level of diffusion
- Example: **e** – original ensemble perturbation; D_n – diffusion with lengthscale *n*

  $\mathbf{e}_1 = D_{10km}(\mathbf{e})$ $\quad$ $\mathbf{e}_2 = D_{30km}(\mathbf{e}_1)$ $\quad$ $\mathbf{e}_3 = D_{100km}(\mathbf{e}_2)$

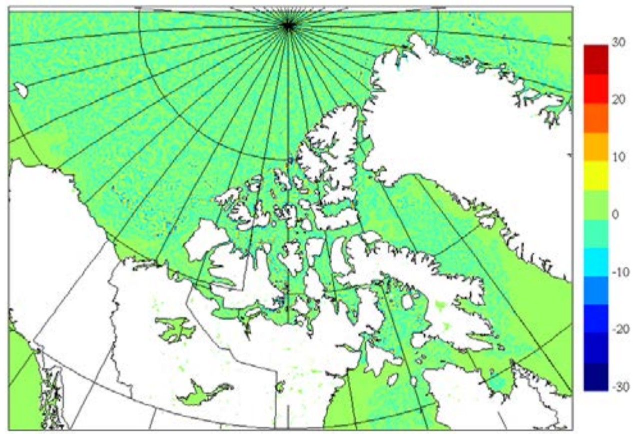
  – Scale 4 (smallest): $\mathbf{e} - \mathbf{e}_1$
  – Scale 3: $\mathbf{e}_1 - \mathbf{e}_2$
  – Scale 2: $\mathbf{e}_2 - \mathbf{e}_3$
  – Scale 1 (largest): $\mathbf{e}_3$
- Scale-decomposed perturbations sum up to the original **e**

Scale separation with diffusion operator: Example of one ensemble perturbation

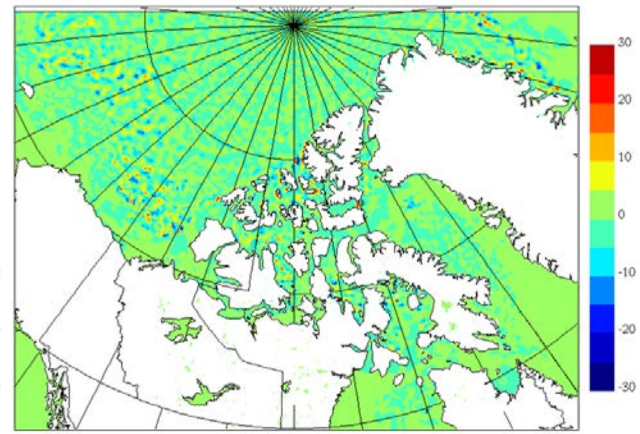
Original perturbation



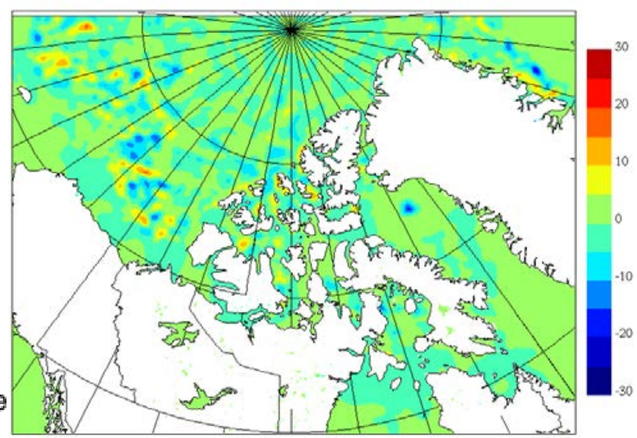
Scale 4 (smallest)



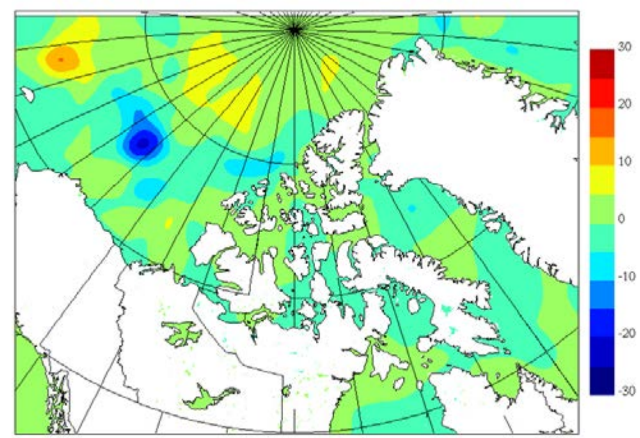
Scale 3



Scale 2

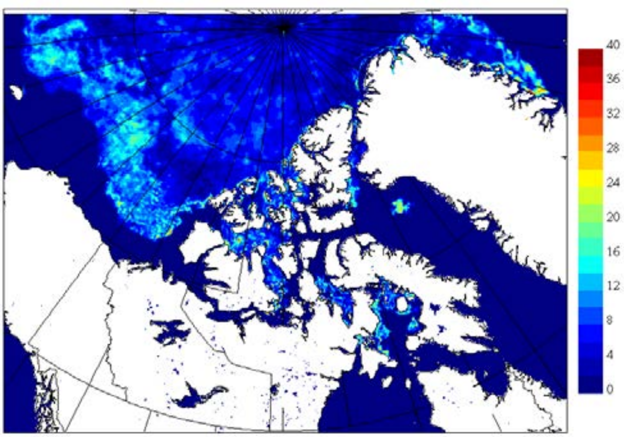


Scale 1 (largest)

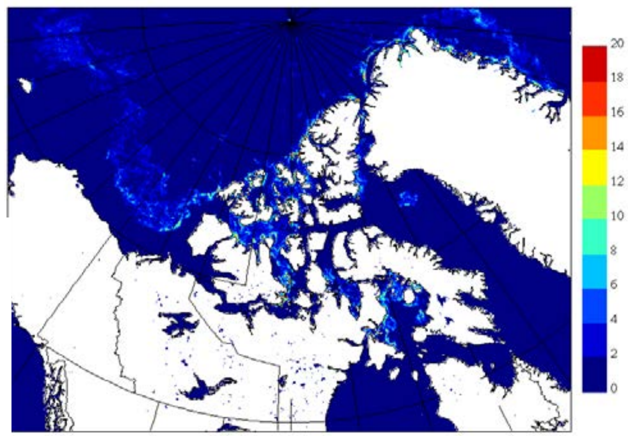


Scale separation with diffusion operator: Ensemble spread for each scale

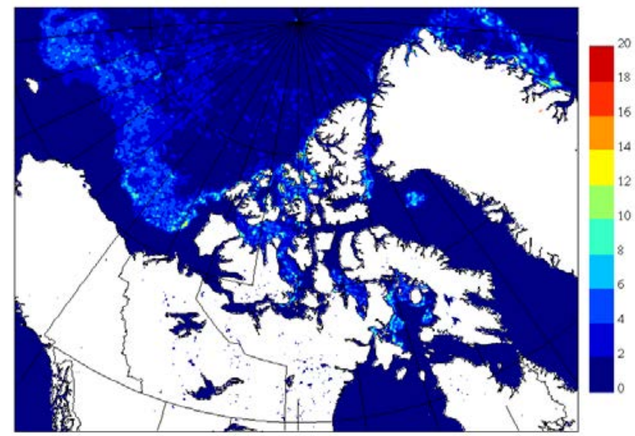
Full ensemble spread



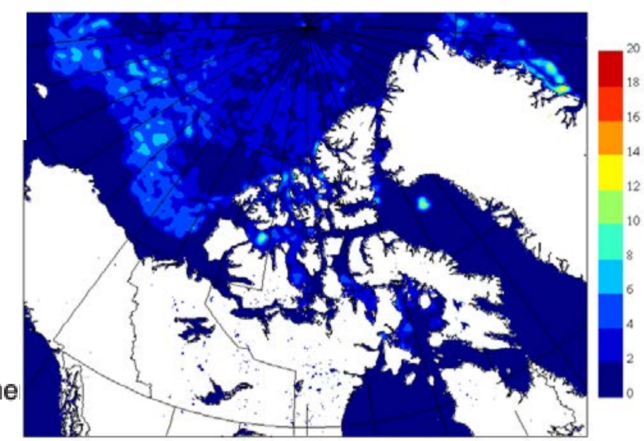
Scale 4 (smallest)



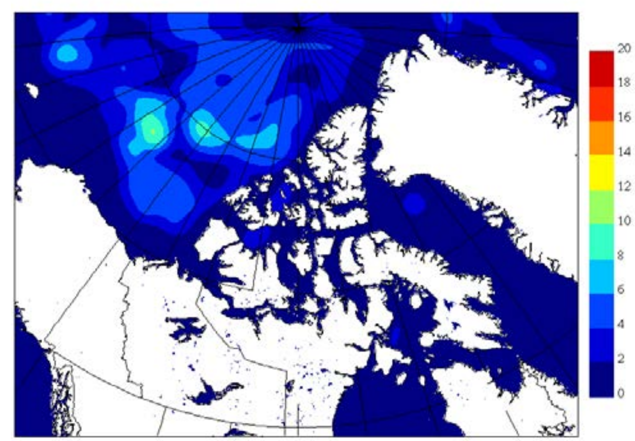
Scale 3



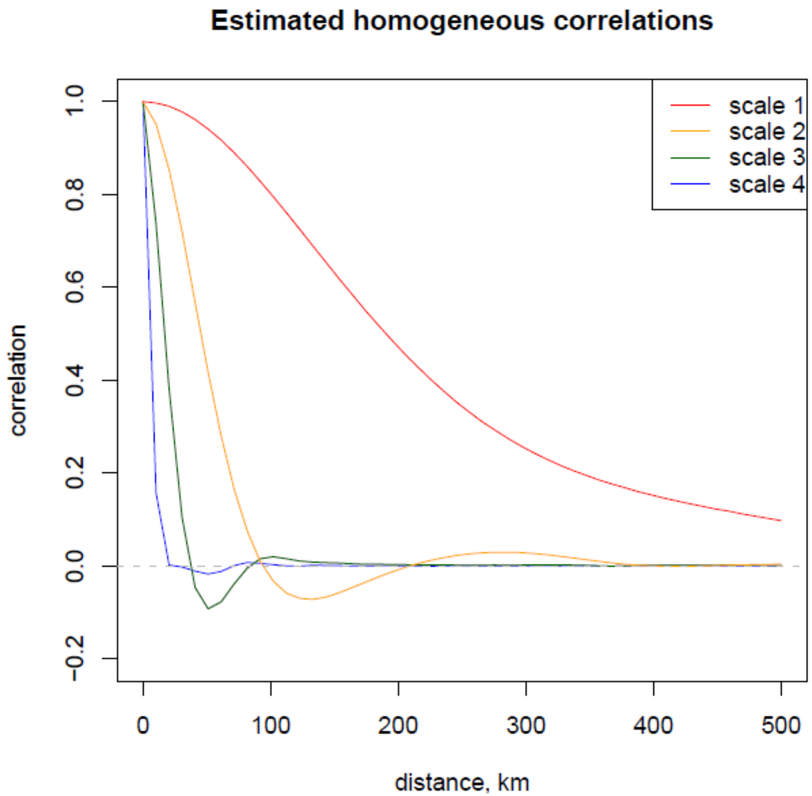
Scale 2

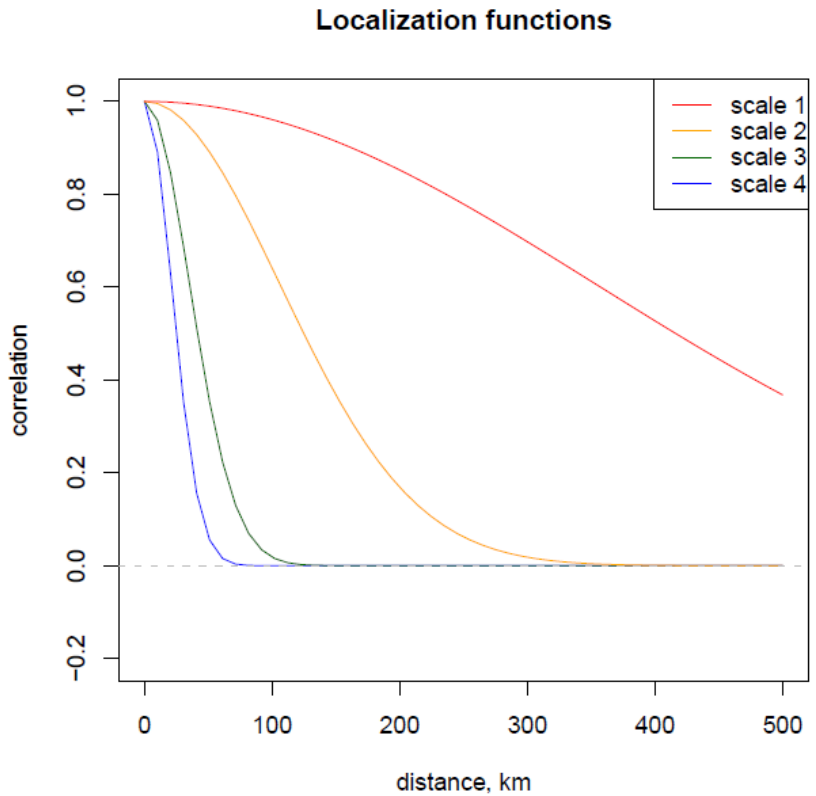


Scale 1 (largest)



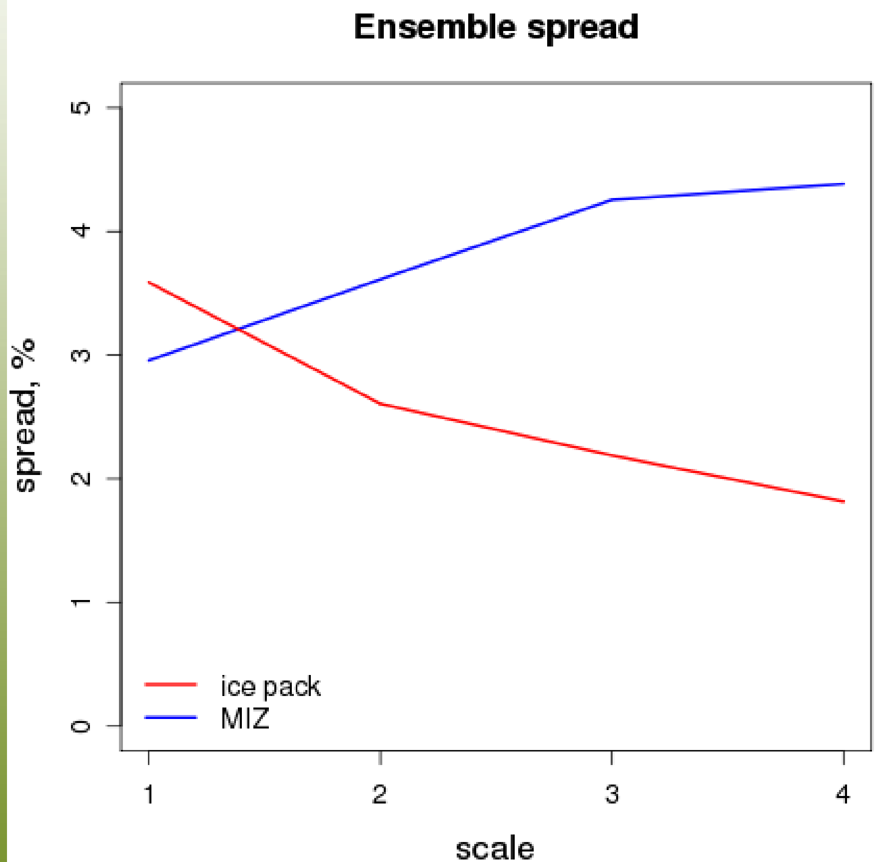
Homogeneous correlation functions and chosen localization functions for each scales





Localization length scales:
500km, 150km, 50km, 30km
(Gaussian-like functions)

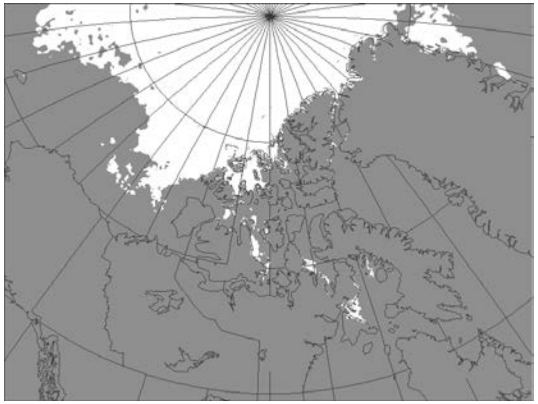
Average standard deviation for different scales in the ice pack and MIZ



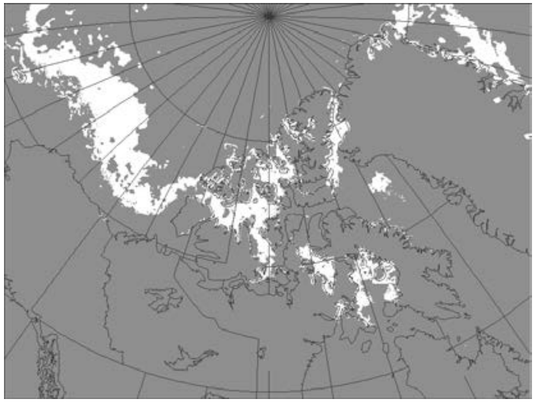
Largest scale

Smallest scale

Ice pack: large scales dominate



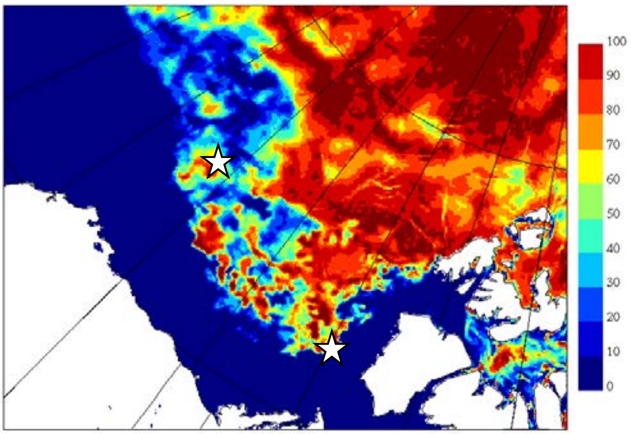
MIZ: small scales dominate



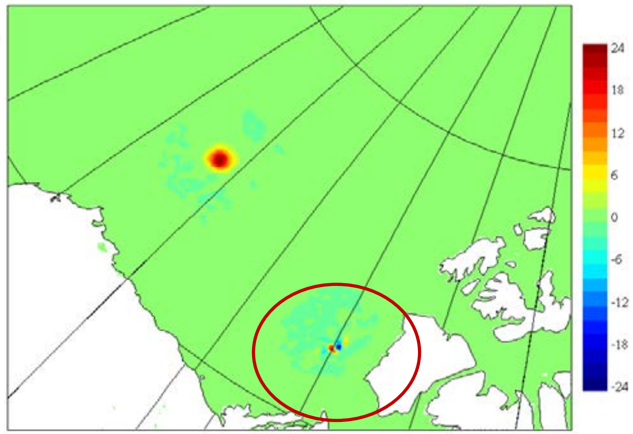
Assimilation of 2 observations

One obs in area dominated by large-scale error, other in area of small-scale error

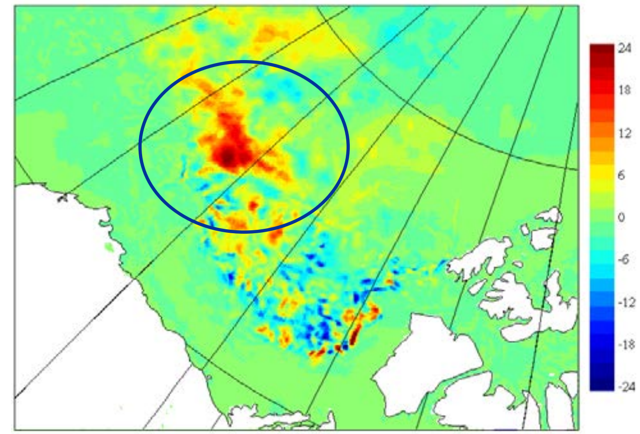
Background field and obs



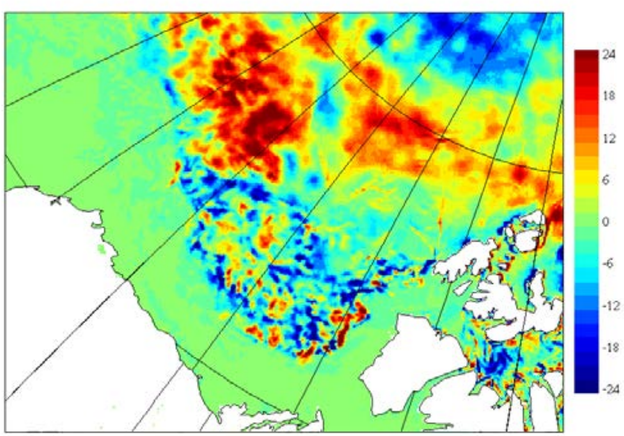
30km localization



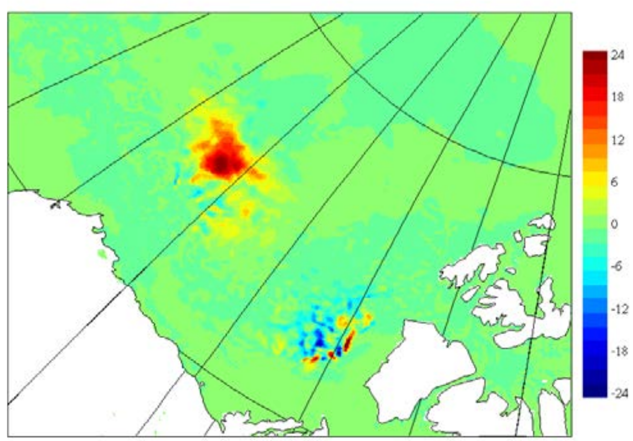
500km localization



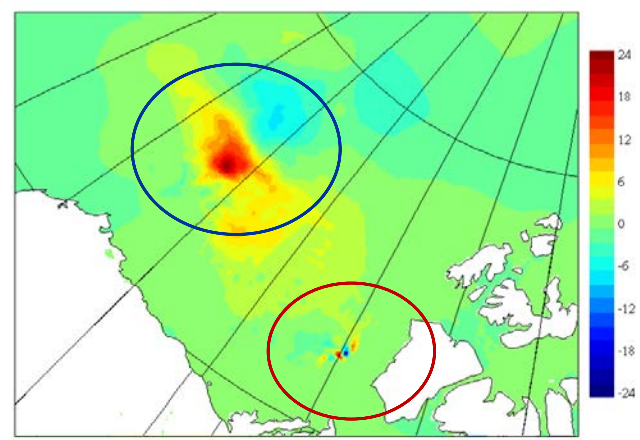
No localization



150km localization



Scale-dep. localization

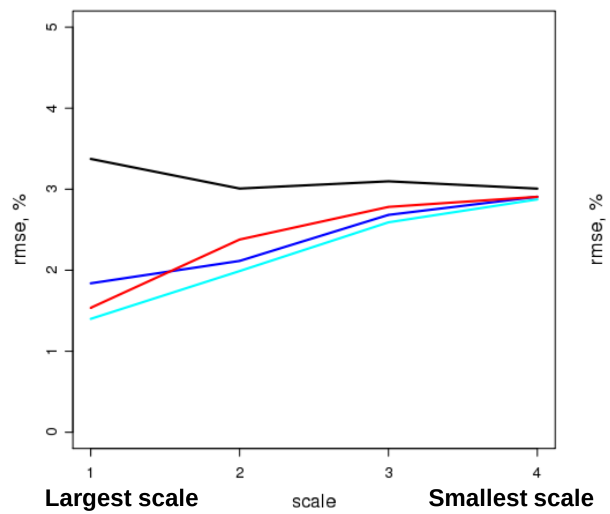


Data assimilation experiment setup

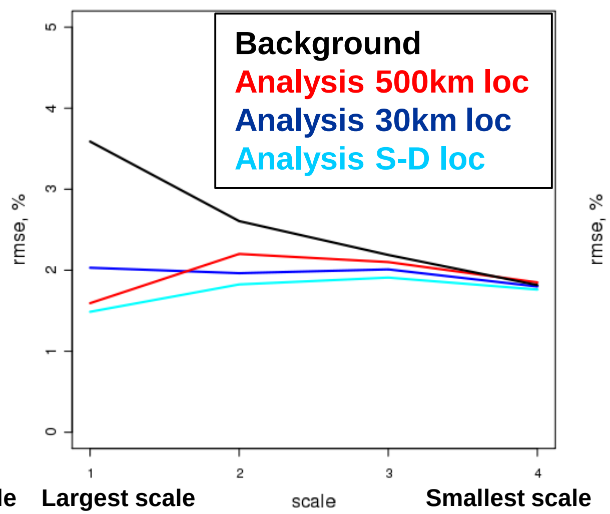
- Assume ensemble represents true error covariance: $\mathbf{B}^t$
- **True state**: $\mathbf{x}^t = \mathbf{x}^i$ (i th member) → mean($\mathbf{x}$) = $\mathbf{x}^t - \mathbf{e}^i$, $\mathbf{e}^i \sim N(0,\mathbf{B}^t)$
- **Background**: $\mathbf{x}^b = \mathbf{x}^t + \mathbf{e}^j$ → $\mathbf{e}^j = \mathbf{x}^j$ – mean($\mathbf{x}$), $\mathbf{e}^j \sim N(0,\mathbf{B}^t)$
- **Observations**: simulated by perturbing true state using same **R** as for assimilation - observation network is every 4th grid point, with random gaps to simulate "clouds“
- $\mathbf{e}^j$ (real background error) not used in ensemble **B** for assimilation
- **Verification**: compute analysis error for each scale (by decomposing $\mathbf{x}^a – \mathbf{x}^t$ by scale); averaged over 60 experiments (each ensemble member used as the true state)

Results of the experiment

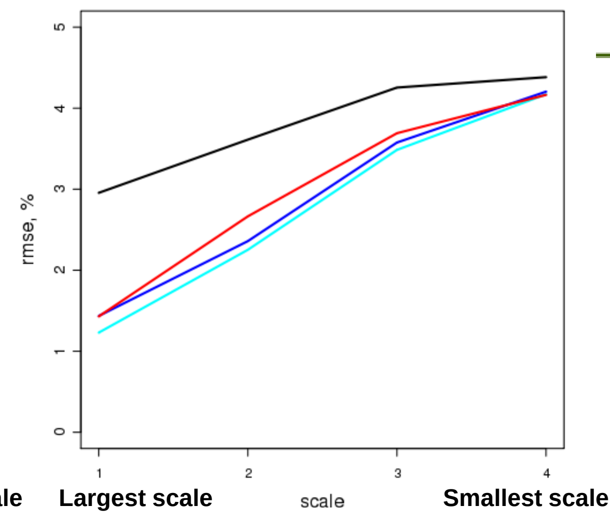
Error



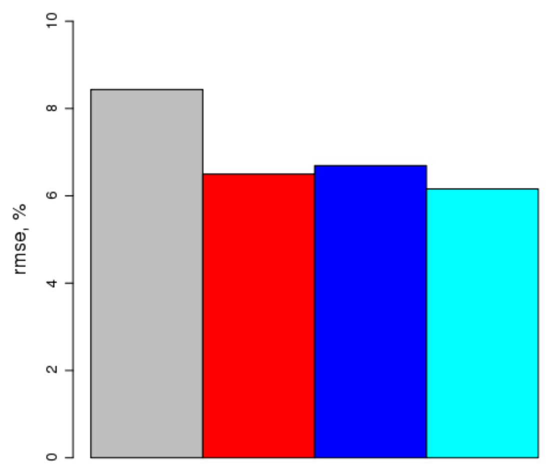
Error (Ice Pack)



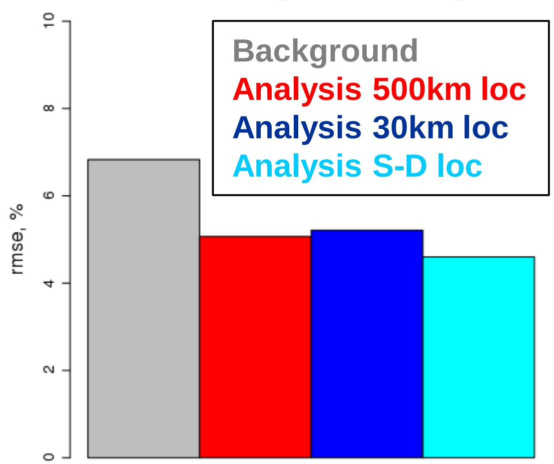
Error (MIZ)



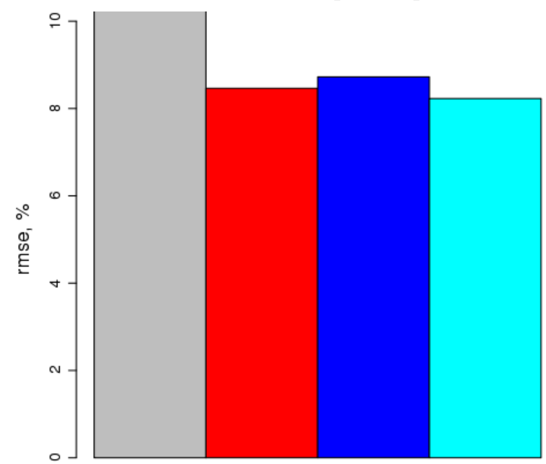
Error



Error (Ice Pack)



Error (MIZ)



Scale-dependent covariance localization

2D Sea Ice Ensemble: Conclusions

- Scale separation can be performed using a diffusion operator (convenient for variational systems that use diffusion operator or recursive filter instead of spectral transform for modelling **B**)
- Strong spatial variation in partition of error wrt scale leads to strong spatial variation in strength of effective localization when using scale-dependent localization (similarity with adaptive localization approaches)
- Scale-dependent localization results in lowest analysis RMSE for all scales in regions dominated by either small-scale or large-scale error

Scale-dependent covariance localization

General Conclusions

- Scale-dependent localization is feasible, but more expensive than single-scale localization: more spectral transforms or applications of diffusion operator per iteration
- But may provide net benefit by appropriately resolving error over wide range of scales with relatively small ensemble when assimilating all obs simultaneously
- Alternative is to use broad localization (to avoid messing up large scales) with huge ensemble (to reduce sampling error for small scales)
- Need to examine impact from unavoidable spectral localization on reducing estimation error vs. loss of heterogeneity