Differently observed scales and ensemble covariance localization

Anna Shlyaeva\textsuperscript{1,2}, Mark Buehner\textsuperscript{3} and Chris Snyder\textsuperscript{4}

\textsuperscript{1}NOAA/Earth System Research Lab, Boulder, CO, USA
\textsuperscript{2}Colorado University/CIRES, Boulder, CO, USA
\textsuperscript{3}Environment and Climate Change Canada, Montreal, QC, Canada
\textsuperscript{4}NCAR, Boulder, CO, USA
Motivation

• Background space:
  • In global high-resolution systems background may have errors on scales from
global to convective
  • Simple one-scale spatial localization approaches might not be optimal
  • Research on spectral or scale-dependent types of localizations (talks by A.
    Lorenc and J-F Caron yesterday)

• Observation space:
  • Some observations may have correlated errors
  • Background, observation error covariances and observation operator shape
    influence how different scales in the analysis are resolved
  • Taking into account correct observation error statistics can benefit the
    analysis
Simple 1D problem used in this talk

1D periodic domain, 50 points, H=I: fully observed network

True $B$: Gaussian correlation with lengthscale = 3

- $B$: covariance with the 1st gridpoint
- $B$: diagonal elements of spectral transform of $B$

Note: for isotropic homogeneous covariance as in this example, spectral transform is diagonal
Simple 1D problem used in this talk

1D periodic domain, 50 points, $H=I$: fully observed network
True $R$: Gaussian correlation with known lengthscale

$R$: covariance with the 1$^{\text{st}}$ gridpoint

$R$: diagonal elements of spectral transform of $R$

Note: for isotropic homogeneous covariance as in this example, spectral transform is diagonal
Background, **observation** and **analysis** error spectral variances

$$R = \text{identity}$$  
$$R \text{ has correlations (ls}=4)$$

**Diagonal of Kalman gain** (observation weight) in spectral space
R = identity \hspace{1cm} R \text{ has correlations (ls=4)}

\textbf{Background, observation and analysis} error spectral variances

\textbf{Diagonal of Kalman gain (observation weight)} in spectral space
## Summary on differently observed scales

<table>
<thead>
<tr>
<th>Examples of possible observations</th>
<th>How scales are observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>R has positive correlations</td>
<td>Small scales are observed better than large scales</td>
</tr>
<tr>
<td>- Atmospheric Motion Vectors</td>
<td></td>
</tr>
<tr>
<td>- Doppler radial wind observations</td>
<td></td>
</tr>
<tr>
<td>H is a footprint (averaging kernel) operator</td>
<td>Large scales are observed better than small scales</td>
</tr>
<tr>
<td>- Satellite radiances (in vertical)</td>
<td></td>
</tr>
<tr>
<td>- Satellite observations measuring area average (sea ice retrievals from PM)</td>
<td></td>
</tr>
</tbody>
</table>
Using ensembles for B with diagonal R

Spectral space error variances of
background mean,
analysis mean (using infinite ensemble),
analysis mean (40-mem ens, no loc)

H is diagonal, R is diagonal
Results averaged over 1000 computations
Using ensembles for B when R has positive correlations - I

Spectral space error variances of
- background mean,
- analysis mean (using infinite ensemble),
- analysis mean (40-mem ens, no loc)

Why is analysis error so high when using correlated R and non-localized B?

H is diagonal, R has correlations (ls=4)
Results averaged over 1000 computations
Background error covariance in spectral space

True B (infinite ensemble)  Ensemble B (no localization)

Finite ensemble B has spurious *cross-scale* covariances; spurious covariance between largest and smallest scale can be larger than smallest scale variance
Spurious *cross-scale* covariances can significantly degrade analysis quality at large scales if small scales are observed much better and information from the small scales is incorrectly propagated into large scale analysis increment.
• Spurious *cross-scale* covariances can significantly degrade analysis quality at large scales if small scales are observed much better and information from the small scales is incorrectly propagated into large scale analysis increment.

• Spectral localization can reduce the cross-scale correlations.
Using ensembles for B when R has positive correlations - II

Spectral space error variances of
background mean,
analysis mean (using infinite ensemble),
analysis mean (40-mem ens, no loc),
analysis mean (40-mem ens, spectral loc)

If spectral localization is applied, the negative effect on the large scales is reduced; the stronger the spectral localization, the better the analysis at large scales.

Note: for this case spatial localization gives worse results than spectral: since B is diagonal in spectral space, even strongest spectral localization incurs no loss of signal, whereas any amount of spatial localization will introduce a bias.

H is diagonal, R has correlations (ls=4)
Results averaged over 1000 computations
Using ensembles for B, varying R lengthscale

RMSE of background mean, analysis mean (using infinite ensemble), analysis mean (40-mem ens, no loc) analysis mean (40-mem ens, spectral loc) analysis mean (40-mem ens, spatial loc)

- No localization: analysis errors can be big when R has significant correlations.
- Spectral localization: the effect is mostly gone.
- Spatial localization (ls=7) is slightly worse than spectral localization.
Conclusions

• If observation errors are strongly correlated, large and small scales are observed with very different accuracy.

• For ensemble DA spurious cross-scale correlations in background error covariance can incorrectly propagate information between scales.

• As a result, if the observations are more accurate than the background at small scales, large spurious analysis increments can be generated at large scales.

• Spectral localization can be effective in reducing analysis error by damping spurious background error cross-scale covariances.
Simple 1D problem used in this talk

1D periodic domain, 50 points

True H: fully observed network, averaging kernel operator (normalized Gaussian correlation)

H for observation in the middle of the domain
Background, observation and analysis error spectral variances

- **H = identity;**
  - **R = identity**
- **H = identity;**
  - **R has correlations (ls=4)**
- **H = footprint op (ls=4)**
  - **R = identity**

Diagonal of I-KH (background weight) in spectral space
H = identity; 
R = identity

H = identity; 
R has correlations (ls=4)

H = footprint op (ls=4) 
R = identity

Background, observation and analysis error spectral variances in observation space ($H_B H_T$, $R$, $H_P a H_T$)

Diagonal of I-KH (background weight) in spectral space
Using ensembles for B when H is a footprint operator

Spectral space error variances of
background mean,
analysis mean (using infinite ensemble),
analysis mean (40-mem ens, no loc)

Analysis errors are not as bad as in the previous case.

Even though analysis at small and medium-scale is corrupted by sampling errors it doesn't have the extreme effect as in the case with correlated R:

- easier to see the effect for medium scales than for small scales, since small-scale errors are very small to begin with,
- observations at large scales are not exceptionally better than the background.

R is diagonal, H is a footprint operator (ls=4)
Results averaged over 1000 computations
Using ensembles for B, varying R and H lengthscales

RMSE of analysis mean

- No localization: analysis errors can be big when R has significant correlations; having footprint H counteracts it to some extent.
- Spectral localization: the effect is mostly gone.
- Spatial localization (ls=7) is slightly worse than spectral localization.