School on Data Assimilation Stockholm, 26 - 30 April 2011

Localisation in the EnKF

Pavel Sakov

Nansen Environmental and Remote Sensing Center, Norway

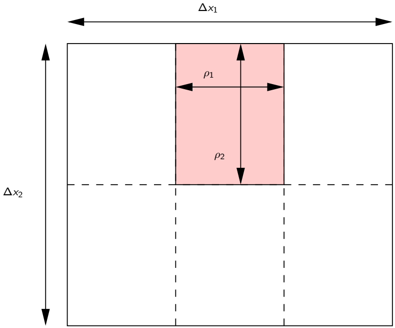
Outline

Reasoning

Two common reasons:

- "spurious covariances"
  $\sigma \sim \frac{1}{m^{1/2}}$
- ensemble rank versus model subspace dimension

$$m_{eff} \approx m \frac{\Delta \mathbf{x}_1}{\rho_1} \frac{\Delta \mathbf{x}_2}{\rho_2} \dots \frac{\Delta \mathbf{x}_N}{\rho_N}$$



Problems

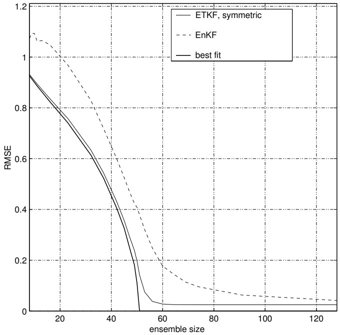
- Q: what is the "price" paid for localisation?
  - dynamical ballance
  - optimality
- Note two different problems:
  1. Mean update ↔ local curve fitting
  2. Covariance update

Reasoning

Two common reasons:

- "spurious covariances"
  $\sigma \sim \frac{1}{m^{1/2}}$
- ensemble rank versus model subspace dimension

$$m_{eff} \approx m \frac{\Delta \mathbf{x}_1}{\rho_1} \frac{\Delta \mathbf{x}_2}{\rho_2} \ldots \frac{\Delta \mathbf{x}_N}{\rho_N}$$



Problems

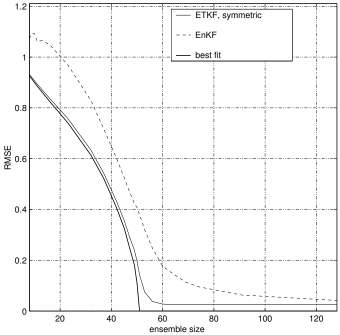
- Q: what is the "price" paid for localisation?
  - dynamical ballance
  - optimality
- Note two different problems:
  1. Mean update ↔ local curve fitting
  2. Covariance update

Reasoning

Two common reasons:

- "spurious covariances"
  $\sigma \sim \frac{1}{m^{1/2}}$
- ensemble rank versus model subspace dimension

$$m_{eff} \approx m \frac{\Delta \mathbf{x}_1}{\rho_1} \frac{\Delta \mathbf{x}_2}{\rho_2} \cdots \frac{\Delta \mathbf{x}_N}{\rho_N}$$



Problems

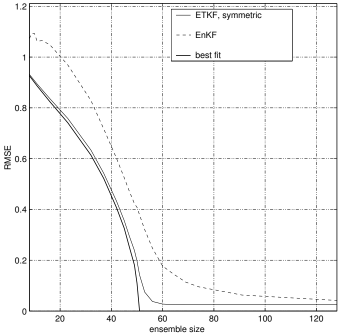
- Q: what is the "price" paid for localisation?
  - dynamical ballance
  - optimality
- Note two different problems:
  1. Mean update $\leftrightarrow$ local curve fitting
  2. Covariance update

Reasoning

Two common reasons:

- "spurious covariances"
  $\sigma \sim \frac{1}{m^{1/2}}$
- ensemble rank versus model subspace dimension

$$m_{eff} \approx m \frac{\Delta \mathbf{x}_1}{\rho_1} \frac{\Delta \mathbf{x}_2}{\rho_2} \dots \frac{\Delta \mathbf{x}_N}{\rho_N}$$



Problems

- Q: what is the "price" paid for localisation?
  - dynamical ballance
  - optimality
- Note two different problems:
  1. Mean update ↔ local curve fitting
  2. Covariance update

Historical overview

- Covariance localisation: Hamill and Whitaker (2001); Houtekamer and Mitchell (2001)
- Local analysis: Evensen (2003); Anderson (2003); Ott et al. (2004)
- Smooth tapering in parallel ESRFs: Hunt et al. (2007)
- Adaptive methods: Anderson (2007); Bishop and Hodyss (2007, 2009)

Historical overview

- Covariance localisation: Hamill and Whitaker (2001); Houtekamer and Mitchell (2001)
- Local analysis: Evensen (2003); Anderson (2003); Ott et al. (2004)
- Smooth tapering in parallel ESRFs: Hunt et al. (2007)
- Adaptive methods: Anderson (2007); Bishop and Hodyss (2007, 2009)

Historical overview

- Covariance localisation: Hamill and Whitaker (2001); Houtekamer and Mitchell (2001)
- Local analysis: Evensen (2003); Anderson (2003); Ott et al. (2004)
- Smooth tapering in parallel ESRFs: Hunt et al. (2007)
- Adaptive methods: Anderson (2007); Bishop and Hodyss (2007, 2009)

Historical overview

- Covariance localisation: Hamill and Whitaker (2001); Houtekamer and Mitchell (2001)
- Local analysis: Evensen (2003); Anderson (2003); Ott et al. (2004)
- Smooth tapering in parallel ESRFs: Hunt et al. (2007)
- Adaptive methods: Anderson (2007); Bishop and Hodyss (2007, 2009)

Some conventions

Indices

$(\mathbf{X})_i$ – ith column of $\mathbf{X}$
$(\mathbf{X})_{i,:}$ – ith row of $\mathbf{X}$
$(\mathbf{X})_{ij}$ – the element in ith row and jth column of $\mathbf{X}$
$(\mathbf{x})_i$ – ith element of $\mathbf{x}$

Matrix operations

$\mathbf{A} \circ \mathbf{B}$ – Schur (Hadamard, element-wise) product of $\mathbf{A}$ and $\mathbf{B}$

Localisation specific

$\overset{i}{\mathbf{A}}$ – ensemble tapered around ith element
$\overset{i}{\mathbf{f}}$ – taper function for the ith element
$\overset{i}{\mathbf{F}}$ – taper matrix, $\overset{i}{\mathbf{F}} = [\overset{i}{\mathbf{f}}, \overset{i}{\mathbf{f}} \dots \overset{i}{\mathbf{f}}]$, $\quad \overset{i}{\mathbf{A}} = \mathbf{A} \circ \overset{i}{\mathbf{F}}$
$\overset{\{o\}}{\mathbf{H}}$ – $\mathbf{H}$ with all rows but oth zeroed

Convenience

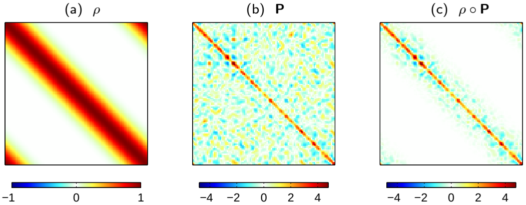
$\mathbf{A}$ – scaled ensemble anomalies: $\mathbf{A} \leftarrow \frac{1}{(m-1)^{1/2}} \mathbf{A}$

Non-adaptive localisation: the two methods

1. Covariance localisation (covariance filtering)

Houtekamer and Mitchell (2001); Whitaker and Hamill (2002)

$$\mathbf{P} \rightarrow \boldsymbol{\rho} \circ \mathbf{P}$$

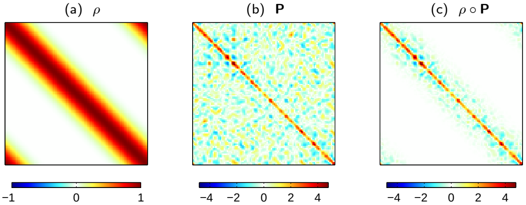


Non-adaptive localisation: the two methods

1. Covariance localisation (covariance filtering)

Houtekamer and Mitchell (2001); Whitaker and Hamill (2002)

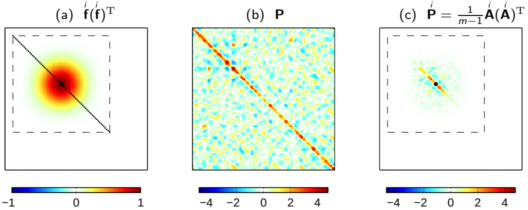
$$\mathbf{P} \rightarrow \boldsymbol{\rho} \circ \mathbf{P}$$



2. Local analysis

(Evensen, 2003; Anderson, 2003; Ott et al., 2004; Hunt et al., 2007)

$$i : \mathbf{A} \rightarrow \overset{i}{\mathbf{A}}$$



Covariance localisation: the technique

$$\boxed{\mathbf{P} \rightarrow \boldsymbol{\rho} \circ \mathbf{P}}$$

$$\mathbf{K} = (\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} \left[\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} + \mathbf{R}\right]^{-1}$$

$(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:

$$(\boldsymbol{\rho} \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx \sum_{o=1}^{p} \overset{i_o}{\mathbf{A}}{}^f\,(\overset{\{o\}}{\mathbf{HA}}{}^f)^{\mathrm{T}}$$

$\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:

$$\mathbf{H}(\boldsymbol{\rho} \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx (\tilde{\mathbf{H}}\,\boldsymbol{\rho}\,\tilde{\mathbf{H}}^{\mathrm{T}}) \circ \left[(\mathbf{HA}^f)(\mathbf{HA}^f)^{\mathrm{T}}\right]$$

CF and particular schemes:

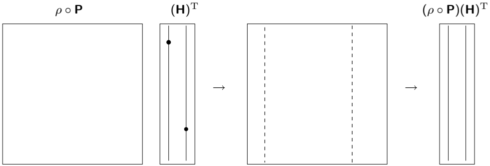
EnKF:	$\mathbf{A}^a = \mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{HA}^f)$	- OK
ESRF:	$\mathbf{A}^a = (\mathbf{I} - \mathbf{KH})^{1/2}\mathbf{A}^f$	- OK
ESRF:	$\mathbf{A}^a = (\mathbf{I} + \mathbf{PH}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1/2}\mathbf{A}^f$	- OK
DEnKF:	$\mathbf{A}^a = (\mathbf{I} - \frac{1}{2}\mathbf{KH})\mathbf{A}^f$	- OK
ETKF:	$\mathbf{A}^a = \mathbf{A}^f\left[\mathbf{I} + (\mathbf{HA^f})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{HA})\right]^{-1/2}$	- NOT POSSIBLE

Covariance localisation: the technique

$$\boxed{\mathbf{P} \longrightarrow \boldsymbol{\rho} \circ \mathbf{P}}$$

$$\mathbf{K} = (\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} \left[\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} + \mathbf{R} \right]^{-1}$$

$(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:



$$(\rho \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx \sum_{o=1}^{p} {}^{i_o}\mathbf{A}^f \,({}^{\{o\}}\mathbf{H}\mathbf{A}^f)^{\mathrm{T}}$$

$\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:

$$\mathbf{H}(\rho \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx (\tilde{\mathbf{H}}\,\rho\,\tilde{\mathbf{H}}^{\mathrm{T}}) \circ \left[(\mathbf{H}\mathbf{A}^f)(\mathbf{H}\mathbf{A}^f)^{\mathrm{T}} \right]$$

CF and particular schemes:

EnKF: $\mathbf{A}^a = \mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{H}\mathbf{A}^f)$ - OK

ESRF: $\mathbf{A}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})^{1/2}\mathbf{A}^f$ - OK

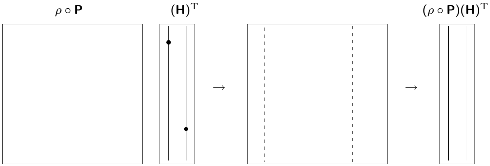
ESRF: $\mathbf{A}^a = (\mathbf{I} + \mathbf{P}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1/2}\mathbf{A}^f$ - OK

Covariance localisation: the technique

$$\boxed{\mathbf{P} \rightarrow \rho \circ \mathbf{P}}$$

$$\mathbf{K} = (\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} \left[\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} + \mathbf{R}\right]^{-1}$$

$(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:



$$(\rho \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx \sum_{o=1}^{p} \overset{i_o}{\mathbf{A}}{}^f \,(\overset{\{o\}}{\mathbf{HA}}{}^f)^{\mathrm{T}}$$

$\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:

$$\mathbf{H}(\rho \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx (\tilde{\mathbf{H}}\,\rho\,\tilde{\mathbf{H}}^{\mathrm{T}}) \circ \left[(\mathbf{HA}^f)(\mathbf{HA}^f)^{\mathrm{T}}\right]$$

CF and particular schemes:

EnKF: $\mathbf{A}^a = \mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{HA}^f)$ - OK

ESRF: $\mathbf{A}^a = (\mathbf{I} - \mathbf{KH})^{1/2}\mathbf{A}^f$ - OK

ESRF: $\mathbf{A}^a = (\mathbf{I} + \mathbf{PH}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1/2}\mathbf{A}^f$ - OK

Covariance localisation: the technique

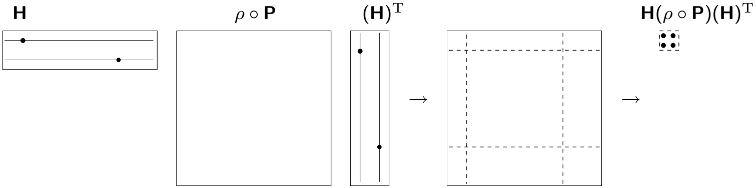
$$\boxed{\mathbf{P} \longrightarrow \boldsymbol{\rho} \circ \mathbf{P}}$$

$$\mathbf{K} = (\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} \left[\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} + \mathbf{R}\right]^{-1}$$

$(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:

$$(\boldsymbol{\rho} \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx \sum_{o=1}^{p} \overset{i_o}{\mathbf{A}}{}^f\,(\overset{\{o\}}{\mathbf{H}\mathbf{A}}{}^f)^{\mathrm{T}}$$

$\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:



$$\mathbf{H}(\rho \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx (\tilde{\mathbf{H}}\,\rho\,\tilde{\mathbf{H}}^{\mathrm{T}}) \circ \left[(\mathbf{H}\mathbf{A}^f)(\mathbf{H}\mathbf{A}^f)^{\mathrm{T}}\right]$$

CF and particular schemes:

Covariance localisation: the technique

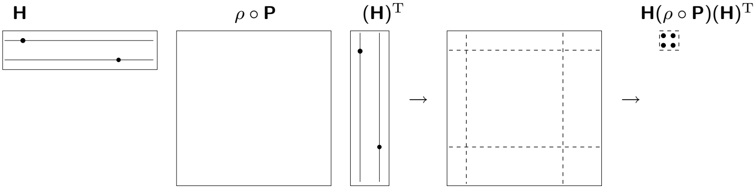
$$\boxed{\mathbf{P} \longrightarrow \boldsymbol{\rho} \circ \mathbf{P}}$$

$$\mathbf{K} = (\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} \left[\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} + \mathbf{R}\right]^{-1}$$

$(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:

$$(\boldsymbol{\rho} \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx \sum_{o=1}^{p} {}^{i_o}\mathbf{A}^f \; {}^{\{o\}}(\mathbf{HA}^f)^{\mathrm{T}}$$

$\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:



$$\mathbf{H}(\boldsymbol{\rho} \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx (\tilde{\mathbf{H}}\,\boldsymbol{\rho}\,\tilde{\mathbf{H}}^{\mathrm{T}}) \circ \left[(\mathbf{HA}^f)(\mathbf{HA}^f)^{\mathrm{T}}\right]$$

CF and particular schemes:

Covariance localisation: the technique

$$\boxed{\mathbf{P} \rightarrow \boldsymbol{\rho} \circ \mathbf{P}}$$

$$\mathbf{K} = (\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} \left[\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}} + \mathbf{R}\right]^{-1}$$

$(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:

$$(\boldsymbol{\rho} \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx \sum_{o=1}^{p} \overset{i_o}{\mathbf{A}}{}^f \, (\overset{\{o\}}{\mathbf{HA}}{}^f)^{\mathrm{T}}$$

$\mathbf{H}(\rho \circ \mathbf{P}^f)(\mathbf{H})^{\mathrm{T}}$:

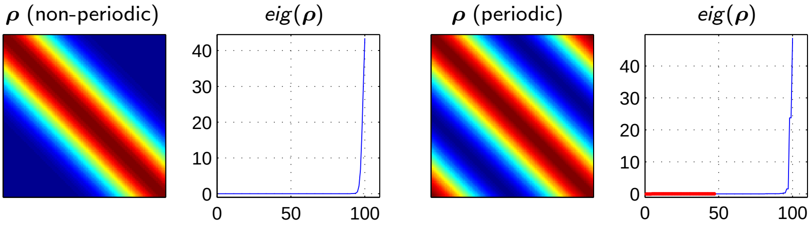
$$\mathbf{H}(\boldsymbol{\rho} \circ \mathbf{P}^f)\,(\mathbf{H})^{\mathrm{T}} \approx (\tilde{\mathbf{H}}\,\boldsymbol{\rho}\,\tilde{\mathbf{H}}^{\mathrm{T}}) \circ \left[(\mathbf{HA}^f)(\mathbf{HA}^f)^{\mathrm{T}}\right]$$

CF and particular schemes:

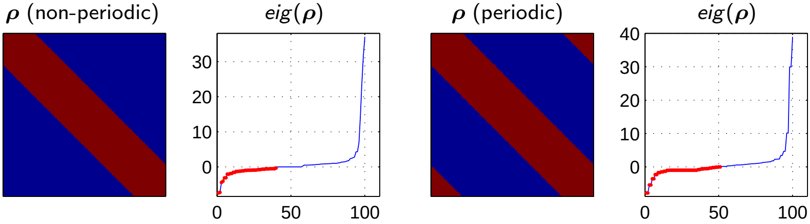
EnKF:	$\mathbf{A}^a = \mathbf{A}^f + \mathbf{K}(\mathbf{D} - \mathbf{HA}^f)$	- OK
ESRF:	$\mathbf{A}^a = (\mathbf{I} - \mathbf{KH})^{1/2}\mathbf{A}^f$	- OK
ESRF:	$\mathbf{A}^a = (\mathbf{I} + \mathbf{PH}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1/2}\mathbf{A}^f$	- OK
DEnKF:	$\mathbf{A}^a = (\mathbf{I} - \frac{1}{2}\mathbf{KH})\mathbf{A}^f$	- OK
ETKF:	$\mathbf{A}^a = \mathbf{A}^f \left[\mathbf{I} + (\mathbf{HA^f})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{HA})\right]^{-1/2}$	- NOT POSSIBLE

Possible inconsistency

1D case, $n = 100$, $r_{loc} = 20$, Gaspari and Cohn (1999) taper function

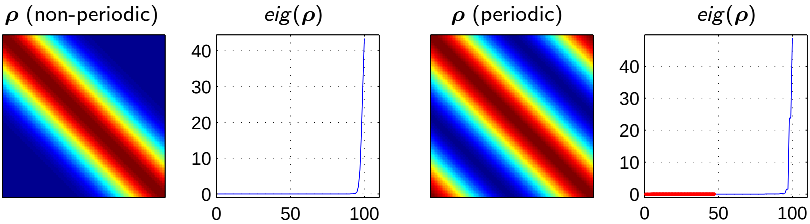


Step taper function

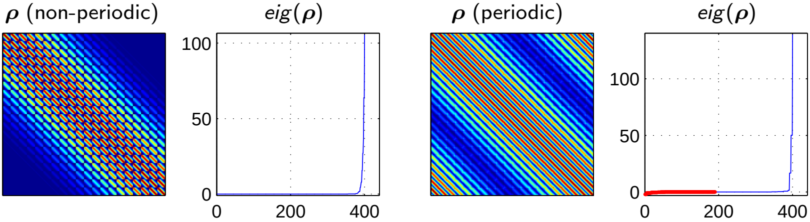


Possible inconsistency

1D case, $n = 100$, $r_{loc} = 20$, Gaspari and Cohn (1999) taper function

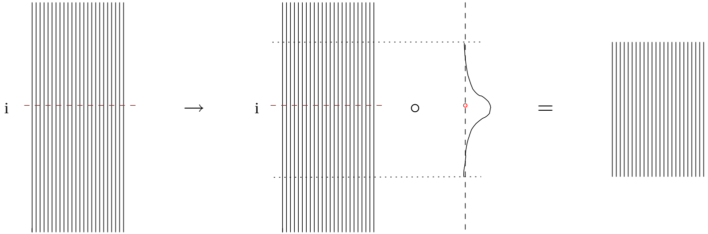


2D case, $n = 20 \times 20$, $r_{loc} = 5$, Gaspari & Cohn taper function



Local analysis

$$i: \quad \mathbf{A} \to \overset{i}{\mathbf{A}} \equiv \mathbf{A} \circ [\overset{i}{\mathbf{f}} \dots \overset{i}{\mathbf{f}}] \equiv \mathbf{A} \circ \overset{i}{\mathbf{F}}$$



$$\begin{aligned} \mathbf{x}^a &= \mathbf{x}^f + \mathbf{K}(\mathbf{d} - \mathbf{H}\mathbf{x}^f) &\to\quad& \mathbf{x}^a_i = \mathbf{x}^f_i + (\overset{i}{\mathbf{K}})_{i,:}(\overset{i}{\mathbf{d}} - \mathbf{H}\overset{i}{\mathbf{x}^f}) \\ \mathbf{A}^a &= \mathbf{A}^f\mathbf{T} &\to\quad& (\mathbf{A}^a)_{i,:} = (\mathbf{A}^f)_{i,:}\overset{i}{\mathbf{T}} \end{aligned}$$

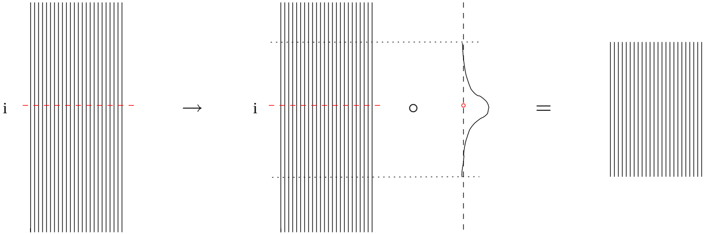
- Equivalence of ensemble tapering and observation scaling:

$$\begin{aligned} \overset{i}{\mathbf{K}} &= \mathbf{A}(\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}}\left[(\mathbf{H}\overset{i}{\mathbf{A}})(\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}} + \mathbf{R}\right]^{-1} \\ &= \mathbf{A}(\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}}\left[\mathbf{I} + (\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}})(\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}}\right]^{-1}\mathbf{R}^{-1/2} \end{aligned}$$

$$\overset{i}{\mathbf{T}} = \left[\mathbf{I} + (\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}}(\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}})\right]^{-1/2} \quad \blacksquare$$

Local analysis

$$i: \quad \mathbf{A} \to \overset{i}{\mathbf{A}} \equiv \mathbf{A} \circ [\overset{i}{\mathbf{f}} \dots \overset{i}{\mathbf{f}}] \equiv \mathbf{A} \circ \overset{i}{\mathbf{F}}$$



$$\begin{aligned} \mathbf{x}^a = \mathbf{x}^f + \mathbf{K}(\mathbf{d} - \mathbf{H}\mathbf{x}^f) \quad &\to \quad \mathbf{x}^a_i = \mathbf{x}^f_i + (\overset{i}{\mathbf{K}})_{i,:}(\overset{i}{\mathbf{d}} - \mathbf{H}\overset{i}{\mathbf{x}}^f) \\ \mathbf{A}^a = \mathbf{A}^f \mathbf{T} \quad &\to \quad (\mathbf{A}^a)_{i,:} = (\mathbf{A}^f)_{i,:} \overset{i}{\mathbf{T}} \end{aligned}$$

- Equivalence of ensemble tapering and observation scaling:

$$\begin{aligned} \overset{i}{\mathbf{K}} &= \mathbf{A}(\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}} \left[(\mathbf{H}\overset{i}{\mathbf{A}})(\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}} + \mathbf{R} \right]^{-1} \\ &= \mathbf{A}(\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}} \left[\mathbf{I} + (\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}})(\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}} \right]^{-1} \mathbf{R}^{-1/2} \\ \overset{i}{\mathbf{T}} &= \left[\mathbf{I} + (\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}})^{\mathrm{T}} (\mathbf{R}^{-1/2}\mathbf{H}\overset{i}{\mathbf{A}}) \right]^{-1/2} \quad \blacksquare \end{aligned}$$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

  experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

  experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Non-adaptive methods: conclusions

Covariance localisation:

- can only be applied to schemes formulated in state space
- can be inconsistent, depending on taper function
- can be formulated via either covariance filtering or tapering of ensemble anomalies
- computationally can be the most effective choice for intermediate systems

experiment

Local analysis:

- can be used with any scheme
- is locally consistent
- can be formulated via either tapering of ensemble anomalies or scaling of observation error variance
- computationally can be relatively expensive, but is scalable as $O(n)$

Adaptive methods

Anderson (2007):

- split ensemble into N subensembles
- calculate gain matrices for each sub-ensemble
- calculate “regression confidence factor” α
- use $\mathbf{K}_s \rightarrow \alpha \circ \mathbf{K}_s, \quad s = 1, \ldots, N$

Conclusions:

- does not require distance measure between observation and state vector element
- Consistent formulation
- Optimises for sub-ensemble

Bishop and Hodyss (2007, 2009):

Main idea:

$$\mathbf{P} \rightarrow \rho \circ \mathbf{P}, \qquad \rho = \rho(\mathbf{C})$$

Conclusions:

- does not require distance measure between observation and state vector element
- a lot of tuning to play with; rather expensive
- physically intuitive but *ad hoc*

Adaptive methods

Anderson (2007):

- split ensemble into N subensembles
- calculate gain matrices for each sub-ensemble
- calculate "regression confidence factor" α

For *each state vector element* and *each observation* find the optimal regression coefficient (Kalman gain element) assuming that it is a random draw from the same distribution as $\mathbf{K}_s$ for each sub-ensemble $s = 1, \ldots, N$, and find α to minimise

$$J = \sum_{i=1}^{N} \sum_{j \neq i}^{N} (\alpha \mathbf{K}_i - \mathbf{K}_j)^2$$

Solution:

$$\alpha(i, o) = \max \left\{ \frac{\frac{(\sum_s \mathbf{K}_s)^2}{\sum_s \mathbf{K}_s^2} - 1}{N - 1}, 0 \right\}$$

- use $\mathbf{K}_s \rightarrow \alpha \circ \mathbf{K}_s, \quad s = 1, \ldots, N$

Conclusions:

- does not require distance measure between observation and state vector element
- Consistent formulation
- Optimises for sub-ensemble

Bishop and Hodyss (2007, 2009):

Main idea:

Adaptive methods

Anderson (2007):

- split ensemble into N subensembles
- calculate gain matrices for each sub-ensemble
- calculate "regression confidence factor" α
- use $\mathbf{K}_s \to \alpha \circ \mathbf{K}_s, \quad s = 1, \ldots, N$

Conclusions:

- does not require distance measure between observation and state vector element
- Consistent formulation
- Optimises for sub-ensemble

Bishop and Hodyss (2007, 2009):

Main idea:

$$\mathbf{P} \to \rho \circ \mathbf{P}, \qquad \rho = \rho(\mathbf{C})$$

Conclusions:

- does not require distance measure between observation and state vector element
- a lot of tuning to play with; rather expensive
- physically intuitive but *ad hoc*

Adaptive methods

Anderson (2007):

- split ensemble into N subensembles
- calculate gain matrices for each sub-ensemble
- calculate “regression confidence factor” α
- use $\mathbf{K}_s \to \alpha \circ \mathbf{K}_s, \quad s = 1, \ldots, N$

Conclusions:

- does not require distance measure between observation and state vector element
- Consistent formulation
- Optimises for sub-ensemble

Bishop and Hodyss (2007, 2009):

Main idea:

$$\mathbf{P} \to \rho \circ \mathbf{P}, \qquad \rho = \rho(\mathbf{C})$$

Conclusions:

- does not require distance measure between observation and state vector element
- a lot of tuning to play with; rather expensive
- physically intuitive but *ad hoc*

Adaptive methods

Anderson (2007):

- split ensemble into N subensembles
- calculate gain matrices for each sub-ensemble
- calculate "regression confidence factor" α
- use $\mathbf{K}_s \to \alpha \circ \mathbf{K}_s, \quad s = 1, \ldots, N$

Conclusions:

- does not require distance measure between observation and state vector element
- Consistent formulation
- Optimises for sub-ensemble

Bishop and Hodyss (2007, 2009):

Main idea:

$$\mathbf{P} \to \rho \circ \mathbf{P}, \qquad \rho = \rho(\mathbf{C})$$

Conclusions:

- does not require distance measure between observation and state vector element
- a lot of tuning to play with; rather expensive
- physically intuitive but *ad hoc*

Adaptive methods

Anderson (2007):

- split ensemble into N subensembles
- calculate gain matrices for each sub-ensemble
- calculate "regression confidence factor" α
- use $\mathbf{K}_s \rightarrow \alpha \circ \mathbf{K}_s, \quad s = 1, \ldots, N$

Conclusions:

- does not require distance measure between observation and state vector element
- Consistent formulation
- Optimises for sub-ensemble

Bishop and Hodyss (2007, 2009):

Main idea:

$$\mathbf{P} \rightarrow \rho \circ \mathbf{P}, \qquad \rho = \rho(\mathbf{C})$$

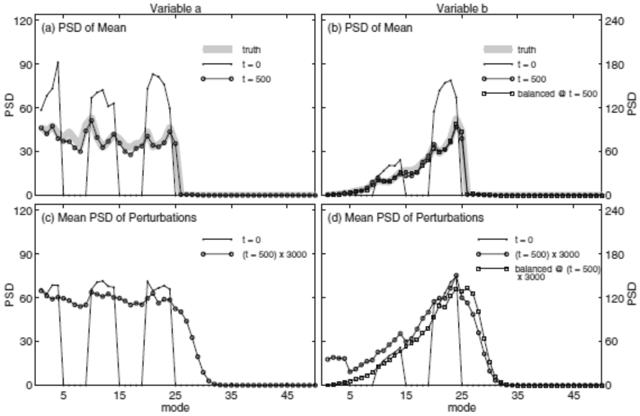
Conclusions:

- does not require distance measure between observation and state vector element
- a lot of tuning to play with; rather expensive
- physically intuitive but *ad hoc*

Diffusivity

Example: LA model (Oke et al., 2007)

- Linear advection (LA) model: Evensen (2004), http://enkf.nersc.no/Code/EnKF-Matlab
- $n = 1000,\ m = 20,\ r_{loc} = 100,\ 51\ \text{modes},\ L = 100$



Conclusions

- Localisation in the EnKF is an ad-hoc modification of the analysis scheme
- One must use localisation if the the ensemble size is smaller than the model subspace dimension
- Localisation makes the analysis schemes suboptimal and therefore inconsistent (sometimes - strongly) in regard to estimation of posterior covariance
- Localisation makes it possible to recover the modal structure from observations even with a rank-deficient ensemble experiment
- There are two main non-adaptive localisation schemes
- These schemes are mainly equivalent and should be chosen by algorithmic convenience

Conclusions

- Localisation in the EnKF is an ad-hoc modification of the analysis scheme
- One must use localisation if the the ensemble size is smaller than the model subspace dimension
- Localisation makes the analysis schemes suboptimal and therefore inconsistent (sometimes - strongly) in regard to estimation of posterior covariance
- Localisation makes it possible to recover the modal structure from observations even with a rank-deficient ensemble experiment
- There are two main non-adaptive localisation schemes
- These schemes are mainly equivalent and should be chosen by algorithmic convenience

Conclusions

- Localisation in the EnKF is an ad-hoc modification of the analysis scheme
- One must use localisation if the the ensemble size is smaller than the model subspace dimension
- Localisation makes the analysis schemes suboptimal and therefore inconsistent (sometimes - strongly) in regard to estimation of posterior covariance
- Localisation makes it possible to recover the modal structure from observations even with a rank-deficient ensemble experiment
- There are two main non-adaptive localisation schemes
- These schemes are mainly equivalent and should be chosen by algorithmic convenience

Conclusions

- Localisation in the EnKF is an ad-hoc modification of the analysis scheme
- One must use localisation if the the ensemble size is smaller than the model subspace dimension
- Localisation makes the analysis schemes suboptimal and therefore inconsistent (sometimes - strongly) in regard to estimation of posterior covariance
- Localisation makes it possible to recover the modal structure from observations even with a rank-deficient ensemble experiment
- There are two main non-adaptive localisation schemes
- These schemes are mainly equivalent and should be chosen by algorithmic convenience

Conclusions

- Localisation in the EnKF is an ad-hoc modification of the analysis scheme
- One must use localisation if the the ensemble size is smaller than the model subspace dimension
- Localisation makes the analysis schemes suboptimal and therefore inconsistent (sometimes - strongly) in regard to estimation of posterior covariance
- Localisation makes it possible to recover the modal structure from observations even with a rank-deficient ensemble experiment
- There are two main non-adaptive localisation schemes
- These schemes are mainly equivalent and should be chosen by algorithmic convenience

Conclusions

- Localisation in the EnKF is an ad-hoc modification of the analysis scheme
- One must use localisation if the the ensemble size is smaller than the model subspace dimension
- Localisation makes the analysis schemes suboptimal and therefore inconsistent (sometimes - strongly) in regard to estimation of posterior covariance
- Localisation makes it possible to recover the modal structure from observations even with a rank-deficient ensemble experiment
- There are two main non-adaptive localisation schemes
- These schemes are mainly equivalent and should be chosen by algorithmic convenience

References

Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642.

— 2007: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D*, **230**, 99–111.

Bishop, C. H. and D. Hodyss, 2007: Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Q. J. R. Meteorol. Soc.*, **133**, 2029–2044.

— 2009: Ensemble covariances adaptively localized with ECO-RAP. part 1: tests on simple error models. *Tellus*, **61A**, 84–96.

Evensen, G., 2003: The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343–367, doi:10.1007/sl10236-003-0036-9.

— 2004: Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics*, **54**, 539–560.

Gaspari, G. and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757.

Hamill, T. M. and J. S. Whitaker, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790.

Houtekamer, P. L. and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126.

Oke, P. R., P. Sakov, and S. P. Corney, 2007: Impacts of localisation in the EnKF and EnOI: experiments with a small model. *Ocean Dynamics*, **57**, 32–45, doi:10.1007/sl10236-006-0088-8.

Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428.

Sakov, P. and L. Bertino, 2010: Relation between two common localisation methods for the EnKF. *Comput. Geosci.*, in press, doi:10.1007/s10596-010-9202-6.

Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.