A new Hierarchical Bayes approach to ensemble-variational data assimilation

M Tsyrulnikov and A Rakitko

HydroMetCenter of Russia

Eretria, 8 Sep 2014

Methodological problems in the existing data assimilation approaches we intend to alleviate with the new technique

1. All existing Var, EnKF, and EnVar analysis equations assume that the effective background-error covariance matrix B is exact. But this is never the case.
2. EnVar takes a *linear combination* of static and ensemble covariances to specify B. This is ad hoc.
3. EnKF and EnVar use an ad-hoc localization. This is not theoretically optimal.
4. In the Var, EnKF, and EnVar analysis equations, there is no intrinsic feedback from observations to background-error statistics. This requires external adaptation or manual tuning.

The proposed paradigm

In words: Acknowledge that B is uncertain and random and update it along with the state.

Observations for B: both the ensemble and the ordinary observations contain info on B.

Level	Gnrlzd bckg	Prior	Gnrlzd obs	Update
2	Static $\mathbf{B}_0$	$p(\mathbf{B}\mid\mathbf{B}_0)$	Ensm $\mathbf{X^e}$, $\mathbf{x^{obs}}$	$\mathbf{B}_0 \Rightarrow \boxed{\mathbf{B}} \Leftarrow \mathbf{X^e}, \mathbf{x^{obs}}$
1	$\mathbf{x^b} = \mathbf{x^f}$ or $\overline{\mathbf{x^e}}$	$\mu\mathbf{B}_0 + (1-\mu)\mathbf{X^e}\mathbf{X^{eT}} \circ \mathbf{C}$	$\mathbf{x^{obs}}$	$\mathbf{x^b} \Rightarrow \boxed{\mathbf{x}} \Leftarrow \mathbf{x^{obs}}$

Level 2: extension by the new approach.

Level 1: the existing EnVar technique.

Hierarchical Bayes EnVar (HB-EnVar): principle

$$p(\mathbf{x}, \mathbf{m}, \mathbf{B} | \mathbf{x^f}, \mathbf{X^e}, \mathbf{x^{obs}}) \propto p(\mathbf{m}|\mathbf{x^f}) p(\mathbf{B}|\mathbf{B}_0) p(\mathbf{x}|\mathbf{m}, \mathbf{B}) p(\mathbf{X^e}|\mathbf{m}, \mathbf{B}) p(\mathbf{x^{obs}}|\mathbf{x})$$

The goal is the *posterior distribution* of $\mathbf{x}$ and its mean $\hat{\mathbf{x}}$ in particular.

Ensemble likelihood

$$p(\mathbf{X^e}|\mathbf{m},\mathbf{B}) \propto |\mathbf{B}|^{-N/2} \mathrm{e}^{-\frac{1}{2}\sum_{\mathrm{k=1}}^{\mathrm{N}} (\mathbf{x_k^e}-\mathbf{m})^\top \mathbf{B^{-1}} (\mathbf{x_k^e}-\mathbf{m})},$$

– no need and no room for approximations.

The prior pdf for $\mathbf{B}$: square-root Gaussian

$$\mathbf{B} = \mathbf{W}\mathbf{W}^\top$$

$\mathbf{W}$ is a Gaussian random matrix.

1) Its pdf:

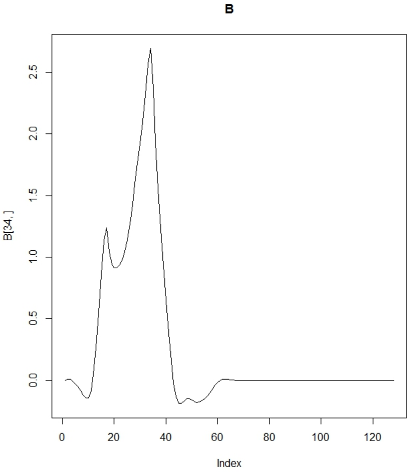
$$p(\mathbf{W}) \propto \mathrm{e}^{-\frac{1}{2}\,\mathrm{tr}[(\mathbf{W}-\mathbf{W_0})\mathbf{U^{-1}}(\mathbf{W}-\mathbf{W_0})^\top\mathbf{U^{-1}}]}$$

2) Sampling:

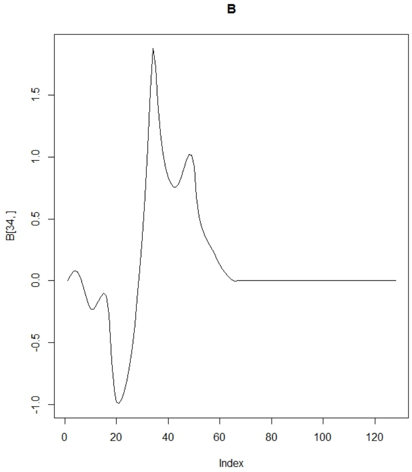
$$\mathbf{W} = \mathbf{W}_0 + \mathbf{\Phi}\mathbf{Y}\mathbf{\Phi}^\top,$$

where $\mathbf{Y}$ is the pure-noise matrix, with $\mathcal{N}(0,1)$ independent entries.

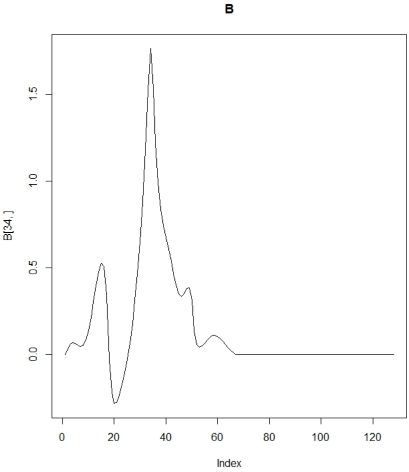
Random samples from $p(\mathbf{B})$



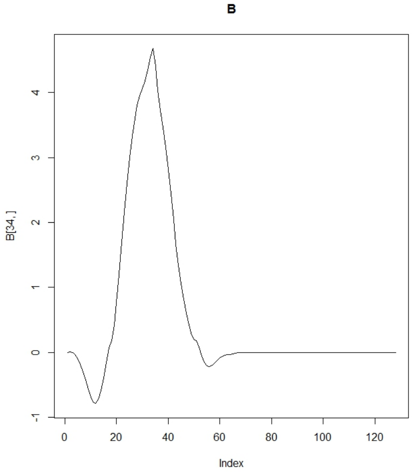
Random samples from $p(\mathbf{B})$



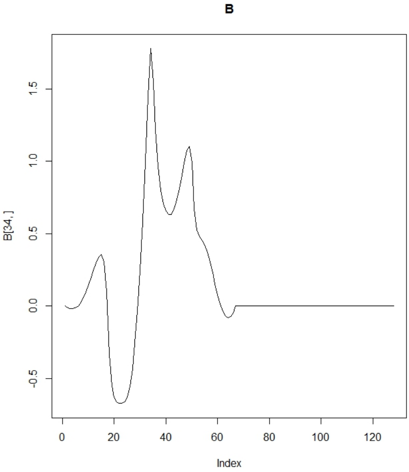
Random samples from $p(\mathbf{B})$



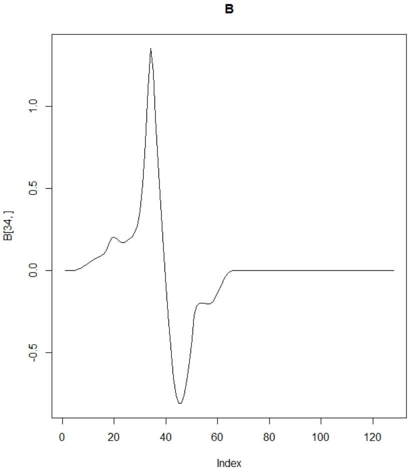
Random samples from $p(\mathbf{B})$



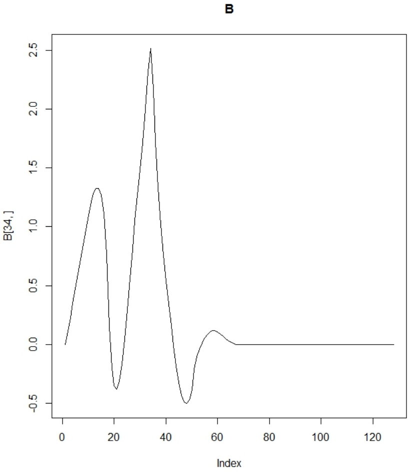
Random samples from $p(\mathbf{B})$



Random samples from $p(\mathbf{B})$



Random samples from $p(\mathbf{B})$



The posterior pdf

$$p_{post}(\mathbf{x}) = \int p_{post}(\mathbf{W}) \cdot p_{post}(\mathbf{x}|\mathbf{W})\, \mathrm{d}\mathbf{W}$$

$$p_{post}(\mathbf{x}|\mathbf{W}) \propto \mathcal{N}(\mathbf{x}^{\mathbf{a}}(\mathbf{m}, \mathbf{W}), \mathbf{B}^{\mathbf{a}}(\mathbf{W})),$$

$$p_{post}(\mathbf{W}) \propto p(\mathbf{W})\phi(\mathbf{W}),$$

$$\phi(\mathbf{W}) := |\mathbf{W}|^{-N} \cdot \mathrm{e}^{-\frac{1}{2}\sum \mathbf{f}_{\mathrm{k}}^{\top}\mathbf{f}_{\mathrm{k}}} \cdot \frac{1}{|\mathbf{G}|^{1/2}} \cdot \mathrm{e}^{\frac{1}{2}\mathbf{v}^{\top}\mathbf{R}^{-1}\mathbf{H}\mathbf{W}\mathbf{G}^{-1}\mathbf{W}^{\top}\mathbf{H}^{\top}\mathbf{R}^{-1}\mathbf{v}}$$

$$\mathbf{W} \cdot \mathbf{f}_k = \mathbf{x}_k^e - \mathbf{m}$$

$$\mathbf{G} := \mathbf{I} + \mathbf{W}^{\top}\mathbf{H}^{\top}\mathbf{R}^{-1}\mathbf{H}\mathbf{W}$$

Here **v** is the innovation vector, hence the adaptive effect.

Importance sampling

$$\hat{\mathbf{x}} = \int p_{post}(\mathbf{W}) \cdot \mathbf{x}^{\mathbf{a}}(\mathbf{m}, \mathbf{W})\, \mathrm{d}\mathbf{W}$$

$$\hat{\mathbf{x}} = \mathrm{E}\, \mathbf{x}^{\mathbf{a}}(\mathbf{m}, \mathbf{W}) = \mathrm{E}_{\,q(\mathbf{W})} \frac{p_{post}(\mathbf{W})}{q(\mathbf{W})} \mathbf{x}^{\mathbf{a}}(\mathbf{m}, \mathbf{W})$$

$$\hat{\mathbf{x}} \approx \overline{\mathbf{x}^{\mathbf{a}}} := \sum_{m=1}^{M} w_m \cdot \mathbf{x}^{\mathbf{a}}(\mathbf{m}, \mathbf{W}_m^{+})$$

=Selection of q: Gaussian pdf centered at the EnVar $\mathbf{W}_{EV}$.

=Localization is achieved by introducing *sparsity* within the *proposal distribution* q for $\mathbf{W}$.

=The ordinary analysis step is included in the importance sampling analysis.

=EnVar can be reproduced within HB-EnVar by nullifying the prior uncertainty in $\mathbf{B}$.

Analysis: Monte-Carlo sampling from the posterior

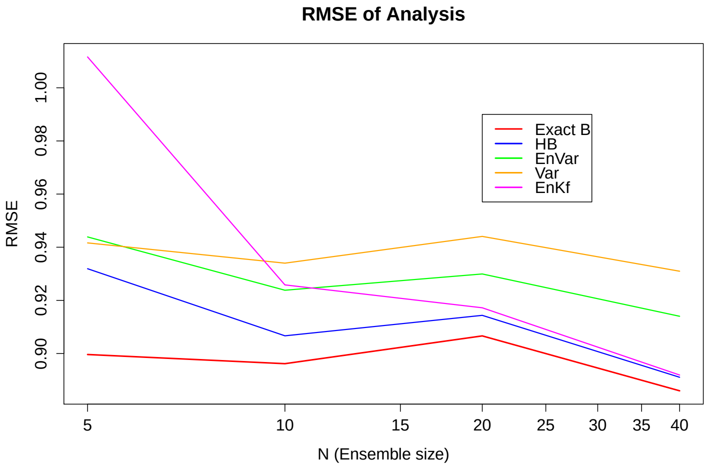
1. Draw M samples $\mathbf{W}_m^+$ from the proposal density of $\mathbf{W}$.
2. Compute their non-normalized importance weights:

$$\phi'(\mathbf{W}_m^+) := \frac{p_{post}(\mathbf{W}_m^+)}{q(\mathbf{W}_m^+)}$$

3. Normalize the importance weights, so that they sum up to one—getting w_m.
4. Perform m ordinary optimal analyses $\mathbf{x}_m^a$ with background-error covariance matrices $\mathbf{B}_m = \mathbf{W}_m^+ \cdot (\mathbf{W}_m^+)^\top$.
5. Average $\mathbf{x}_m^a$ with weights w_m:

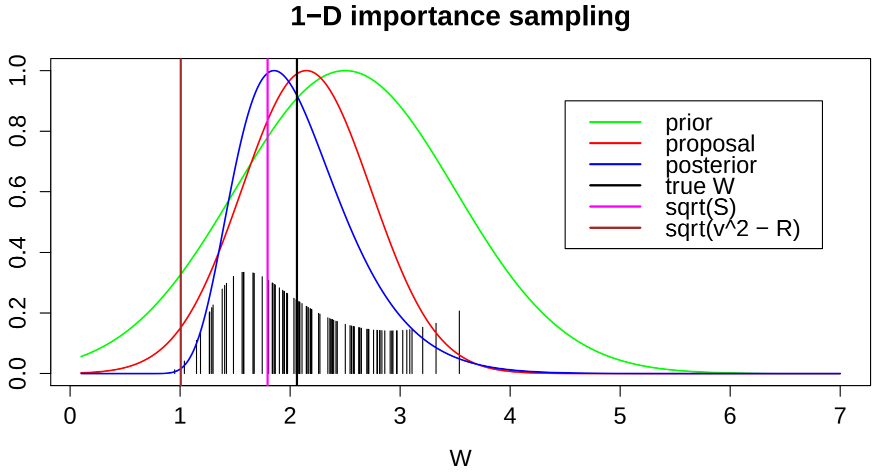
$$\overline{\mathbf{x}^\mathbf{a}} = \sum_{m=1}^{M} w_m \cdot \mathbf{x}^\mathbf{a}(\mathbf{B}_m)$$

1-D illustrative example: dependence on the ensemble size



In the toy problem, the deterministic HB-EnVar analysis outperforms Var, EnKF, and EnVar.

PDFs: prior, posterior, proposal, and importance weights



Conclusions

Main aspects HB-EnVar

- Background-error covariance matrix **B** is treated as a sparse *random matrix* and updated in the optimal scheme along with the state.
- The key element of the proposed technique is the *prior distribution* of **B**.
- Ensemble members are treated as *observations* on the background-error covariance matrix and *assimilated* along with ordinary observations.
- The technique is computationally expensive.

Potential benefits of HB-EnVar

- Optimized hybridization of static and ensemble covariances.
- Optimized combination of deterministic background and ensemble mean.
- Optimized localization.
- Optimized feedback from observations to background-error covariances.
- Uncertainty in **B** is explicitly accounted for in generation of the analysis ensemble, resulting in the *increased spread*.