Is the Local Ensemble Transform Kalman Filter suitable for operational data assimilation?

M Tsyrulnikov

Offenbach, 7 Sept. 2009

Purpose of the review on LETKF

- Check the existing literature and experience.
- Critically examine the operational usability of the technique.

Advantages of LETKF

- Simplicity
  No explicit B matrix modeling is needed (no spatial covariances and no explicit balance operators).
- Computational efficiency
  - Working in low-dimensional ensemble space.
  - Locality: as in OI, LETKF analysis computations at all grid points are mutually independent.
- Flow-dependent covariances
- 4D-LETKF

LETKF. Analysis step: key features

1. At each analysis grid point, the analysis is performed *locally*: using only nearby observations from a box (cylinder, ellipsoid,. . .) surrounding the grid point

2. The analysis is performed in *ensemble space*: within each local box, the analysis increment belongs to the subspace spanned by ensemble deviations, $\mathbf{x}_i^f - \mathbf{x}^b$.

3. 'Observation localization': the $\mathbf{R}^{-1}$ entries are multiplied by a monotonically decaying function of the distance between the center of the box and the particular observation (within the box)

The ensemble-space approach

$$\mathbf{e}_i := \frac{1}{\sqrt{n_e - 1}}(\mathbf{x}_i^f - \mathbf{x}^b),$$

$$\mathbf{E} = (\mathbf{e}_1 \cdots \mathbf{e}_{n_e}).$$

$$\mathbf{z} = \sum_{i=1}^{n_e} \tilde{z}_i \mathbf{e}_i \equiv \mathbf{E}\tilde{\mathbf{z}}.$$

Observation operator in ensemble space

$$\mathbf{x}^o = \mathcal{H}(\mathbf{x}^b) + \mathbf{H}(\mathbf{x} - \mathbf{x}^b) + \eta + \eta_{lin},$$

$$\mathbf{x} - \mathbf{x}^b = \mathbf{E}\tilde{\mathbf{y}} + \eta_{et}$$

$$\tilde{\mathbf{H}} = \mathbf{H} \cdot \mathbf{E}$$

Analysis equations

$$\tilde{\mathbf{y}}^a = \tilde{\mathbf{K}} \cdot \mathbf{y}^o,$$

$$\tilde{\mathbf{K}} = (\mathbf{I} + \tilde{\mathbf{H}}^T \mathbf{R}^{-1} \tilde{\mathbf{H}})^{-1} \tilde{\mathbf{H}}^T \mathbf{R}^{-1},$$

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{E}_{loc} \cdot \tilde{\mathbf{y}}^a$$

1. A finite-difference approximation to $\mathbf{H}$

$$\tilde{\mathbf{H}}\tilde{\mathbf{y}} \equiv \mathbf{HE}\tilde{\mathbf{y}} \equiv \sum \tilde{y}_i \cdot \mathbf{He}_i$$

Device: approximate $\mathbf{He}_i$ by $\frac{1}{s}(\mathcal{H}(\mathbf{x}^b + s\mathbf{e}_i) - \mathcal{H}(\mathbf{x}^b))$, where $s := \sqrt{n_e - 1}$

Statement: the covariance matrix of the sum $\eta_{lin} + \eta'_{fd}$ is larger than that of η_{lin}

Conclusion: Not critical.

2. With non-local satellite observations, the effective box size becomes *large*

Global ensemble vectors are used to fit observations.

$$a_i \geq |\mathrm{supp}\mathcal{H}|_i$$

Radiances: supp=20-30 km in the vertical.
GPS: support up to 1000 km in the horizontal.

Conclusion: Important.

3. Within *large* effective boxes, affordable ensemble size implies poor analysis resolution and hence accuracy

Strict limitation of the ensemble size:

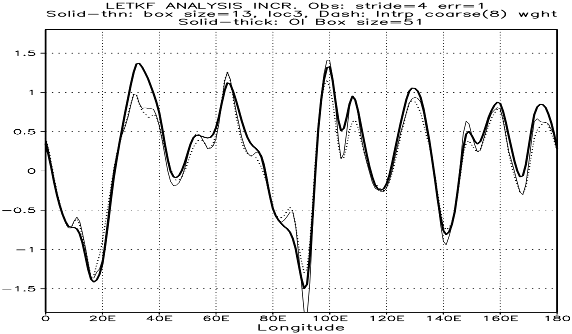
$$n_e \sim n_{odof}$$

Effective resolution:

$$h_i^{eff} \sim a_i / \sqrt[3]{n_e}.$$

Conclusion: Critical.

4. *Small* local boxes can led to small-scale noise in the analysis increments



Analysis increment, one realization: OI box diameter 50 (solid thick curve), LETKF box diameter 12, localization length 3 (solid thin), and LETKF on coarse grid with post-interpolated ensemble weights (dotted).
Conclusion: Important.

5. Observation-error correlations can destroy the LETKF computational efficiency

$$\tilde{\mathbf{K}} = (\mathbf{I} + \tilde{\mathbf{H}}^T \mathbf{R}^{-1} \tilde{\mathbf{H}})^{-1} \tilde{\mathbf{H}}^T \mathbf{R}^{-1},$$

An application of $\mathbf{R}^{-1}$ requires as large as $O(n_{obs}^3)$ flops (Golub and van Loan 1989). In this case, *the computational advantage of the LETKF algorithm disappears*

Conclusion: Critical if R is non-diagonal.

Conclusions

1. A finite-difference approximation to **H** increases the error but not critically.
2. With non-local satellite observations, the effective box size becomes *large*
3. Within *large* effective boxes, affordable ensemble size implies poor analysis resolution and hence accuracy
4. *Small* local boxes can led to small-scale noise in the analysis increments
5. Observation-error correlations can destroy the LETKF computational efficiency

B. Hunt:

- Many of the criticisms (e.g. Sections 4.1-4.3) seem to apply generally to ensemble square-root F. They are related to the reduced rank approximation of the background covariance, which necessitates some form of localization.
- Nonlocal observations undoubtedly negate some of the benefits of localization. We do what we can to overcome these issues.
- Small-scale noise due to localization doesn't strike me as hard to overcome; for LETKF this will be one benefit of the weight interpolation in our paper w/ Yang and Bowler, and if all else fails then maybe one has to filter the analysis increment.

B. Hunt:

Section 4.4: If the LETKF step that scales like the cube of the number of local observations becomes problematic, one may need to make an approximation, which could involve simply ignoring weak error correlations.
However, it's not clear to me that R^{-1} needs to be computed from scratch at each analysis time. Even though the observation data set changes, one may be able to efficiently compute a good approximation for R^{-1} using the R^{-1} (Also, weight interpolation will help by reducing the number of local analyses.)

B. Hunt:

Section 6, 1st paragraph: I disagree here. The LETKF analysis increment is NOT a linear combination of the ensemble perturbations, since it weights the perturbations differently at each grid point. Locally it is close to a linear combination of the ensemble perturbations, but I believe the same is true with other ensemble Kalman filters and other methods of localization.

B. Hunt:

Section 4.2, 2nd paragraph: I strongly disagree with the suggestion that the number of ensemble members must be comparable to the number of local observations in order to get an accurate analysis. If the number of local observations is larger than the number of degrees of freedom allowed by the local dynamics, then I think smoothing the observations is a good thing. If the ensemble represents the background covariance well, I don't see that the number of observations matters.

B. Hunt:

I think we will continue to disagree about the necessary size of the ensemble. I maintain that the relevant comparison is to the effective number of degrees of freedom of the (local) model dynamics, and that any additional "degrees of freedom"in the observations are largely an illusion. By effective number of degrees of freedom, I mean the number of dimensions in model space needed to capture most of the background uncertainty. I understand that this number may depend on model resolution, but I still believe it is much smaller than the number of available observations.

B. Hunt:

Perhaps we can agree that there's a trade-off here; ensemble Kalman filters may not resolve small-scale information as well as high-dimensional analyses, but they may quantify the large-scale "errors of the day"(temporal and spatial fluctuations in background uncertainty) better than is practical with a high-dimensional approach.

QJ referees:

I think it is generally very difficult to prove that an algorithm will never become operational. When several groups of clever people are working on an algorithm existing problems are likely to get resolved in one way or another.

...

I feel the question posed in the title is impossible to answer by a scientific study.

QJ referees:

Many, if not all, operational data assimilation systems assume that the covariance matrix for the observational error is diagonal.

...

It is possible (done in several operational systems) to perform the forward operator towards the GPS-RO observations by only considering a vertical column of model coordinates

...

The author is correct that localization in the ensemble-based schemes can generate small scale noise. But, the effect of such noise can be controlled by initialization.