Localization and Inflation techniques in ensemble data assimilation

Takemasa Miyoshi
RIKEN Advanced Institute for Computational Science
Takemasa.Miyoshi@riken.jp

With many thanks to
UMD Weather-Chaos group,
Data Assimilation Research Team,
Data Assimilation (DA)

Data assimilation best combines observations and a model, and brings synergy.
Data assimilation in NWP

Numerical Weather Prediction

Model Simulation

ANL

FCST

OBS

ANL

FCST

OBS

True state (unknown)

FCST-ANALYSIS Cycle accumulates past observations (4-D data assimilation)
Direct application to high-dimensional systems is prohibitive.

\[ P_{t1}^f = M_{x_t^a} P_{t0}^a M_{x_t^a}^T \]
Ensemble Kalman Filter (EnKF)

Analysis ensemble mean

Obs.

R

T\approx t_1

An approximation to KF with ensemble representations

P_{t_1}^f \approx \frac{\delta X_{t_1}^f (\delta X_{t_1}^f)^T}{m-1}

Analysis w/ errors

FCST ensemble mean

T=t_0

T=t_1

T=t_2
**LETKF** (Local Ensemble Transform Kalman Filter)

*Analysis is given by a linear combination of forecast ensemble:*

\[
X^a = \bar{x}^f + \delta X^f T
\]

Ensemble Transform Matrix
(ETKF, Bishop et al. 2001; LETKF, Hunt et al. 2007)

\[
T = \tilde{P}^a (\delta Y)^T R^{-1} (y^o - H(x^f)) + [(m - 1)\tilde{P}^a]^{1/2}
\]

ensemble mean update uncertainty update

\[
\tilde{P}^a = [(m - 1)I / \rho + (\delta Y)^T R^{-1} \delta Y]^{-1}
\]

Analysis error covariance in the ensemble subspace
LETKF algorithm (Hunt et al. 2007)

Local Ensemble Transform Kalman Filter

Each grid point is treated independently.

→ essentially perfectly parallel

Multiple observations are treated simultaneously.

T matrix is computed at each grid point independently.
Covariance localization
Covariance localization (e.g., Houtekamer and Mitchell 1998)

Empirical treatment for...
- reducing sampling noise
- increasing the rank

$P^f$

Localized $\rho$

$\rho \circ P^f$
Localization in LETKF

Analysis of the $i$-th variable:

$$x^a_i = \bar{x}^f_i \mathbf{1}_{1 \times m} + \delta x^f_i \mathbf{T}_i (\delta Y^f_i, R_i, d_i)$$

Two steps of localization:

1. Selecting a subset of global obs for the $i$-th variable

   $\delta Y^f_i, R_i, d_i$ are composed of only selected local obs.

2. Obs error std. is weighted by the localization factor

   $$R_i \leftarrow \tilde{\rho}_i^{-1} \circ R_i$$

   so that far-away obs have large error.

   R-localization, Hunt et al. (2007)
Difficulties of localization

Difficulties include...

- dependence on (x, y, z, t)
- loss of flow-dependence

\[ \sigma = 0.95 \]

\[ \sigma = 0.51 \]

\[ \sigma = 0.08 \]
Adaptive localization approaches

• Hierarchical filter by Anderson (2007)
  – Cross-validation by groups of ensembles

\[
\frac{\delta x \delta y^T}{m-1} \quad \frac{\delta x \delta y'}{m-1} \quad \frac{\delta x \delta y'}{m-1} \quad \frac{\delta x \delta y'}{m-1}
\]

Cross-validation \rightarrow Localization weights

• ECO-RAP by Bishop and Hodyss (2009)
  – Smooth the sample correlations raised to a power

  • High sample correlation = more reliable = more weight
  • Spatial smoothing to reduce noisiness of the sample correlation
A challenge: multi-scale localization

Localization plays an essential role in an EnKF to cope with limited ensemble size.

Higher resolution requires more localization, limiting the use of observations.

We look for better use of observations by separating the scales.

Analysis increment from a single profile observation (20 members)
Scale-separated analysis increments

We will construct analysis increments at high \((h)\) and low \((l)\) resolutions separately.

\[ \delta x = \delta x_h + \delta x_l \]
Longer-range covariance

Motivated by Buehner (2012), we apply spatial smoothing to the ensemble perturbations to reduce noise in longer-range covariance.
Larger-scale localization

Applying a 1000-km (larger scale) localization.

Full-range (T30) analysis increment

Analysis increment from reduced-resolution (T21) ensemble perturbations

Noisier in distance
Smaller-scale structure

Applying a 500-km (smaller scale) localization.

More structure in short range
Merging the two scales

Original covariance with 500-km (smaller scale) localization

\[ \delta x_h \]

Preserve the smaller-scale structure in short range

Large-scale covariance with 1000-km (larger scale) localization

\[ \delta x_l \]

Removing the short-range structure

\[ \delta x = \delta x_h + \delta x_l \]
Merged analysis increment

\[ \delta x = \delta x_h + \delta x_l \]
Review: the algorithm

1. Compute the analysis increment regularly
   (with smaller-scale localization)

2. Compute the analysis increment with smoothed ensemble perturbations
   (with larger-scale localization)

3. Compute the analysis increment with smoothed ensemble perturbations
   (with smaller-scale localization)

4. Take the difference between 2 and 3

5. Add 1 and 4
Results are promising.

Experiments with the T30L7 SPEEDY model (Molteni, 2003)

**Global-average RMSE**

- **Mid-level U**
  - CTRL
  - DLOC

- **Low-level T**
  - CTRL
  - DLOC

- **Near-surface Q**
  - CTRL
  - DLOC

- **Surface pressure**
  - CTRL
  - DLOC

- Regular localization (700 km)
- Dual localization (600-900 km)
Improved almost everywhere
ENSEMBLE-BASED OBS IMPACT

Ota, Kalnay, Miyoshi and Derber (2013, *Tellus*)
With FSO approaches, observation impacts can be estimated without performing expensive data denial experiments (or OSEs).

Kunii, Miyoshi, Kalnay (2012)
Forecast sensitivity to observations

**Observation impact** can be calculated using an adjoint model (Langland and Baker 2004)

\[ J = e_{t|0} - e_{t|-6} \]

This difference comes from obs at 00hr

\[ e_{t|0} = x_{t|0}^f - x_t^a \]

The error reduction (or increase) due to obs at 00 (i.e., *obs impact*):

\[
J = (e_{t|0}^T C e_{t|0} - e_{t|-6}^T C e_{t|-6}) = (e_{t|0} - e_{t|-6})^T C (e_{t|0} + e_{t|-6})
\]

\[
x_{t|0}^f - x_{t|-6}^f \approx M (x_0^a - x_0^f)
\]

**analysis increment!**

\[
J \approx \delta y^T K^T M^T C (e_{t|0} + e_{t|-6})
\]

\[
K(y_0 - H x_0^f)
\]
Forecast sensitivity to observations

**Observation impact** can be calculated without an adjoint model

\[(Liu \text{ and Kalnay } 2008; \text{ Li et al. } 2009; \text{ Kalnay et al. } 2012)\]

\[J \approx \delta y^T K^T M^T C\left(e_{t|0} + e_{t|-6}\right) \quad (Langland \text{ and Baker } 2004)\]

In the ensemble Kalman filter (EnKF),

\[K = \frac{X_0 a X_0 a^T}{N_{ens} - 1} H^T R^{-1} = \frac{X_0 a Y_0 a^T R^{-1}}{N_{ens} - 1}\]

\[J \approx \delta y^T K^T M^T C\left(e_{t|0} + e_{t|-6}\right) = \frac{\delta y^T R^{-1} Y_0 a X_0 a^T M^T}{N_{ens} - 1} C\left(e_{t|0} + e_{t|-6}\right)\]

\[J \approx \frac{1}{N_{ens} - 1} \delta y^T R^{-1} Y_0 a X_{t|0} f^T C\left(e_{t|0} + e_{t|-6}\right)\]

We just need an ensemble of forecasts. \hspace{1cm} \textit{Kalnay et al. (2012)}
Denying negative impact data improves forecast!

Estimated observation impact

Typhoon track forecast is actually improved!!

Improved forecast

36-h forecasts

Observed track

TY Sinlaku

Kunii, Miyoshi, Kalnay (2012)
Impact of WC-130J dropsondes

Kunii, Miyoshi, Kalnay (2012)
An issue on localization

\[ J \approx \frac{1}{N_{ens} - 1} \delta y^T R^{-1} Y_0^a X_{t|0}^f T C(e_{t|0} + e_{t|-6}) \]

The ensemble-based covariance needs localization.

We need to consider a “mobile” localization function.

cf. Bishop and Hodyss (2009)
Ideas for “mobile” localization

a) Nonlinear incremental evolution

b) Constant advection

Kalnay et al. (2012)
Impact of mobile localization

Results from idealized experiments with the Lorenz-96 model.

Kalnay et al. (2012)
Variance inflation
Covariance inflation (e.g., Houtekamer and Mitchell 1998)

Empirical treatment for...
- variance underestimation

Error variance is underestimated due to various sources of imperfections:
- limited ensemble size
- nonlinearity
- model errors

Lorenz96-LETKF(10MEM) OBER=1.0

No inflation

5% inflation
Variance underestimation

Forecast ensemble tends to be under-dispersive.
Covariance inflation inflates the underestimated variance.

\[ P \rightarrow (1+\alpha)P \]
Previous inflation methods

1. Multiplicative inflation: $\delta x^f \leftarrow \alpha \cdot \delta x^f$

~50 hPa T ensemble spread

Problematic in real applications

Dense obs $\rightarrow$ under-dispersive
Sparse obs $\rightarrow$ over-dispersive

Tuned constant
Previous inflation methods

1. **Multiplicative inflation:** $\delta x^f \leftarrow \alpha \cdot \delta x^f$

   - $\sim 50 \text{ hPa}$ T ensemble spread
   - Tuned constant

   - Dense obs $\rightarrow$ under-dispersive
   - Sparse obs $\rightarrow$ over-dispersive

   Problematic in real applications

2. **Additive inflation:** $\delta x^a \leftarrow \delta x^a + \delta x^{rnd}$

   - $\sim 50 \text{ hPa}$ T ensemble spread

   - This brings new directions to span,
     but it is not trivial to have proper random fields.

   - Generally better spread, but an unfavorable high-frequency pattern appears.
Previous inflation methods

1. **Multiplicative inflation**: $\delta x^f \leftarrow \alpha \cdot \delta x^f$

   $\sim 50$ hPa T ensemble spread

   Tuned constant

   Dense obs $\rightarrow$ under-dispersive

   Sparse obs $\rightarrow$ over-dispersive

   Problematic in real applications

2. **Additive inflation**: $\delta x^a \leftarrow \delta x^a + \delta x^{rnd}$

   $\sim 50$ hPa T ensemble spread

   This brings new directions to span, but it is not trivial to have proper random fields.

   Generally better spread, but an unfavorable high-frequency pattern appears.

3. **Relaxation to prior**: $\delta x^a \leftarrow (1 - \beta) \cdot \delta x^a + \beta \cdot \delta x^f$ $\beta \sim 0.7$

   Zhang et al. (2004) showed this worked well.
Adaptive inflation (Anderson’s Bayesian approach)

Anderson (2007; 2009) applied the Bayesian estimation theory to adaptive inflation.

\[
p(\alpha_i^a) = p(y_i | \alpha_i) p(y_{i-1} | \alpha_i) \prod p(y_{i-p+1} | \alpha_i) p(\alpha_i^b) / \text{norm.}
\]

Obs PDF is Gaussian w.r.t. obs \(y\):

\[
p(y_i | \alpha_i) = \frac{1}{\sqrt{2\pi(\alpha_i H_i P_i H_i^T + R_i)}} \exp\left(-\frac{(y_i - H_i x_i)^2}{2(\alpha_i H_i P_i H_i^T + R_i)}\right)
\]

This is not Gaussian w.r.t. \(\alpha_i\)

Gaussian prior PDF is assumed:

\[
p(\alpha_i^b) = N(\bar{\alpha}_i^b, \nu_i^b)
\]

This is a tuning parameter!!
Adaptive inflation

*Anderson (2007; 2009)* applied the Bayesian estimation theory to estimate the inflation parameter \( \alpha \) adaptively.

\[
p(\alpha_i^a) = p(y_i | \alpha_i) p(y_{i-1} | \alpha_i) \prod p(y_{i-p+1} | \alpha_i) p(\alpha_i^b) / \text{norm.}
\]

*Li et al. (2009)* applied the Gaussian assumption.

\[
p(\alpha_i^a) = N(\overline{\alpha}_i^o, v_i^o) p(\alpha_i^b) / \text{norm.}
\]

The Gaussian approach is adopted, with additional enhancements of \( v_i^o \) and localization (next slide).

*Miyoshi (2011)*

The non-Gaussianity is very weak.
Localization of inflation estimates

Apply the maximum likelihood estimate at each grid point independently.

\[ \alpha = \alpha(x, y, z, t) \]

Miyoshi (2011)
First step to test the new idea

1. **Toy models** (e.g., Lorenz model)
2. **Intermediate AGCM** (SPEEDY model, Molteni 2003)
3. **Real systems** (e.g., operational models)

Easy to implement, fast to run, accumulating experiences by trial and error.
Results with the Lorenz model

Adaptive inflation reduces the RMS errors significantly.

Fast computations allow many kinds of sensitivity tests.

Adaptive inflation with various prior variance

Miyoshi (2011)

Adaptive inflation reduces the RMS errors significantly.
Results with the Lorenz model

Fast computations allow many kinds of sensitivity tests.

Adaptive inflation improves the ensemble spread.

Adaptive inflation reduces the RMS errors significantly.

Miyoshi (2011)
Step 2: more realistic testing

1. Toy models (e.g., Lorenz model)
2. Intermediate AGCM (SPEEDY model, Molteni 2003)
3. Real systems (e.g., operational models)

Testing applicability and feasibility with a single PC!
Spatial pattern of inflation

Generally large inflation over densely observed areas

Miyoshi (2011)
Improvements due to adaptive inflation

~20% improvements of analysis RMSE

Miyoshi (2011)
Improvements due to adaptive inflation

Miyoshi (2011)

Adaptive inflation improves the ensemble spread.

sparse obs, large spread

Adaptive inflation

dense obs, small spread

~20% improvements of analysis RMSE

Adaptive inflation

Miyoshi (2011)
Step 3: real applications

1. Toy models (e.g., Lorenz model)
2. Intermediate AGCM (SPEEDY model, Molteni 2003)
3. Real systems (e.g., operational models)
Adaptive inflation accounts for imperfections such as model errors and limited ensemble size.
Large adaptive inflation > 100% (2.0) appears occasionally and is appropriate in limited regions.

Miyoshi and Kunii (2011)
Ensemble spread (T500)

Adaptive inflation improves the ensemble spread.

under-dispersive with dense obs
over-dispersive with sparse obs

Miyoshi and Kunii (2011)
Adaptive inflation reduces the RMSE and BIAS consistently.

Miyoshi and Kunii (2011)
Benefits to many applications

1. Toy models (e.g., Lorenz model)
2. Intermediate AGCM (SPEEDY model, Molteni 2003)
3. Real systems (e.g., operational models)

Benefits to diverse research, world-wide users, and even operations

The LETKF code is available at http://code.google.com/p/miyoshi/
Application to CO$_2$ data assimilation

Near-surface CO$_2$ concentration error

Adaptive inflation improved the CO$_2$ analysis

RMSE=0.91

RMSE=0.59

 Courtesy of J.-S. Kang
Results of Mars GCM

Zonal mean temperature RMSE

Global average RMSE

Clear advantage of adaptive inflation

Courtesy of S. Greybush
Adaptive inflation improved the global 9-day forecasts significantly.

Courtesy of Y. Ota (JMA)
Experience from Brazilian CPTEC

**Goal:** To produce operational analysis and forecast using LETKF - MCGA/CPTEC

**Implementation steps:**

- **LETKF ERIC**
  - Interfaces: model MCGA/CPTEC$\rightarrow$ LETKF (It was not easy as we were thinking before to start);
  - Coding resolution changes to be friendly (original T062L28);
  - Coding for use of NCEP PREPBUF observations;

- **LETKF TAKEMASA** (parallel version, T062L28)

- Actual effort: tuning experiments.

---

Solange S. de Souza; José A. Aravêquia; Paulo Kubota
Feedbacks inspire future studies

1. Toy models (e.g., Lorenz model)
2. Intermediate AGCM (SPEEDY model, Molteni 2003)
3. Real systems (e.g., operational models)

Feedbacks are essential!

Benefits to diverse research, world-wide users, and even operations

- Inspiring new ideas
- Demands from operations
- Technical improvements
A challenge: better use of satellite data

### CTRL
Conventional (NCEP PREPBUFR)

### AIRS: Atmospheric Infrared Sounder
Conv. + AIRS retrievals (AIRX2RET - T, q)

Larger inflation is estimated due to the AIRS data.

- Adaptive inflation method was newly developed (Miyoshi 2011).
Improved adaptive inflation

Four separate inflation fields are estimated, considering diurnal variation of the AIRS tracks. (similar to Kang 2011)
AIRS impact on TC forecasts

- TC track forecasts for Typhoon Sinlaku (2008) were significantly better, particularly in longer leads.

Too deep to resolve by 60-km WRF

- TC track forecasts for Typhoon Sinlaku (2008) were significantly better, particularly in longer leads.