EFFICIENT ASSIMILATION OF ATMOSPHERIC DATA: A LOCAL ENSEMBLE TRANSFORM KALMAN FILTER

Brian Hunt

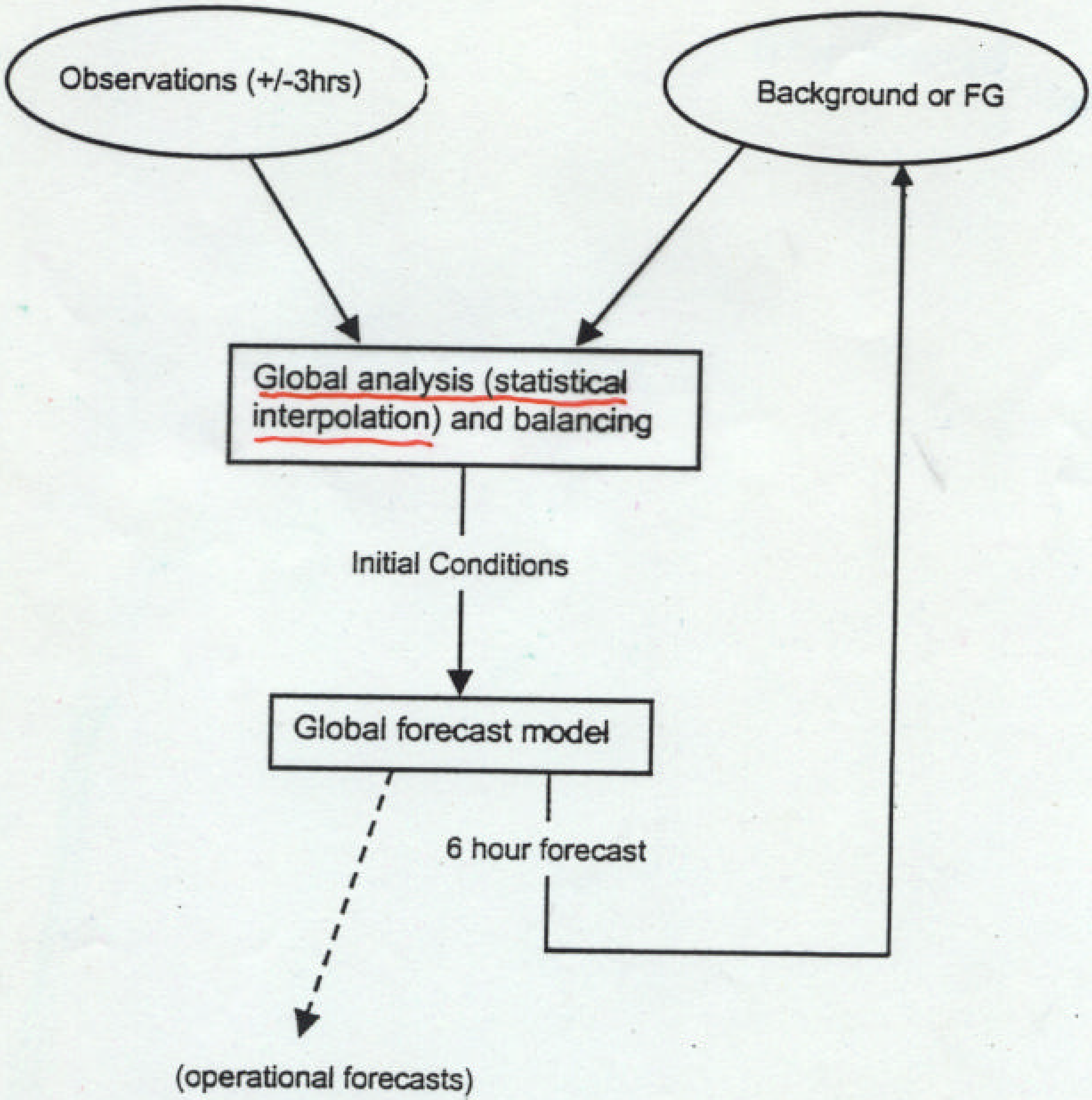
University of Maryland

Chaos/Weather Group

http://keck2.umd.edu/weather/

Thanks to: J.S. McDonnell Foundation
NSF
NOAA/THORPEX

Global 6-hour analysis cycle



ENSEMBLE DATA ASSIMILATION

- Inputs:
  - An ensemble $x_1^b, x_2^b, \ldots, x_k^b$ of forecast model states representing a background (prior) prob. distribution.
  - A vector y^o of observations.
  - A model relating model states to observations, e.g. $y = H(x) + \varepsilon$ where ε is Gaussian w/ mean 0, covariance R.
- Output:
  - An ensemble $x_1^a, x_2^a, \ldots, x_k^a$ representing the analysis (posterior) probability distribution.

ENSEMBLE KALMAN FILTERS

- Assume a Gaussian background distribution with mean $\bar{x}^b$ (= sample mean of bg ensemble) and covariance P^b (= sample covariance).

- Use Kalman filter equations to determine the analysis mean $\bar{x}^a$ and covariance P^a.
[Evensen 1994]

- Choose an analysis ensemble (not uniquely determined!) with the correct sample mean and covariance.

FLAVORS OF EnKF

- Double EnKF [Houtekamer & Mitchell 1998, 2001] – recently surpassed operational 3DVar at MSC.
- ETKF [Bishop, Etherton, Majumdar 2001]
- EAKF [Anderson 2001]
- EnSRF [Whitaker & Hamill 2002]
- Parallel EnKF [Keppenne & Rienecker 2002]
- LEKF [Ott et al. 2002, 2004]

GOALS FOR LETKF

- Practical for operational weather forecasting.
- Mathematical simplicity.
- As efficient as possible.
  - Run time linear in # of observations.
  - Embarassingly parallel.
- Extensible:
  - Localizes in a flexible manner.
  - 4D extension is almost trivial.
  - Nonlinear and/or nonlocal observation operator H.
  - Non-Gaussian distributions?

GAUSSIAN MODEL

- The background distribution of the forecast model state x is

$$p(x) \sim e^{-(x-\bar{x}^b)^T (P^b)^{-1} (x-\bar{x}^b)}.$$

- The conditional distribution of the observation vector y is

$$p(y|x) \sim e^{-(y-H(x))^T R^{-1} (y-H(x))}.$$

- The analysis distribution is

$$p(x|y^o) = \frac{p(y^o|x)\,p(x)}{p(y^o)} \sim e^{-J(x)}$$

where...

COST FUNCTION

$$J(x) = (x - \bar{x}^b)^T (P^b)^{-1} (x - \bar{x}^b) + (y^o - H(x))^T R^{-1} (y^o - H(x))$$

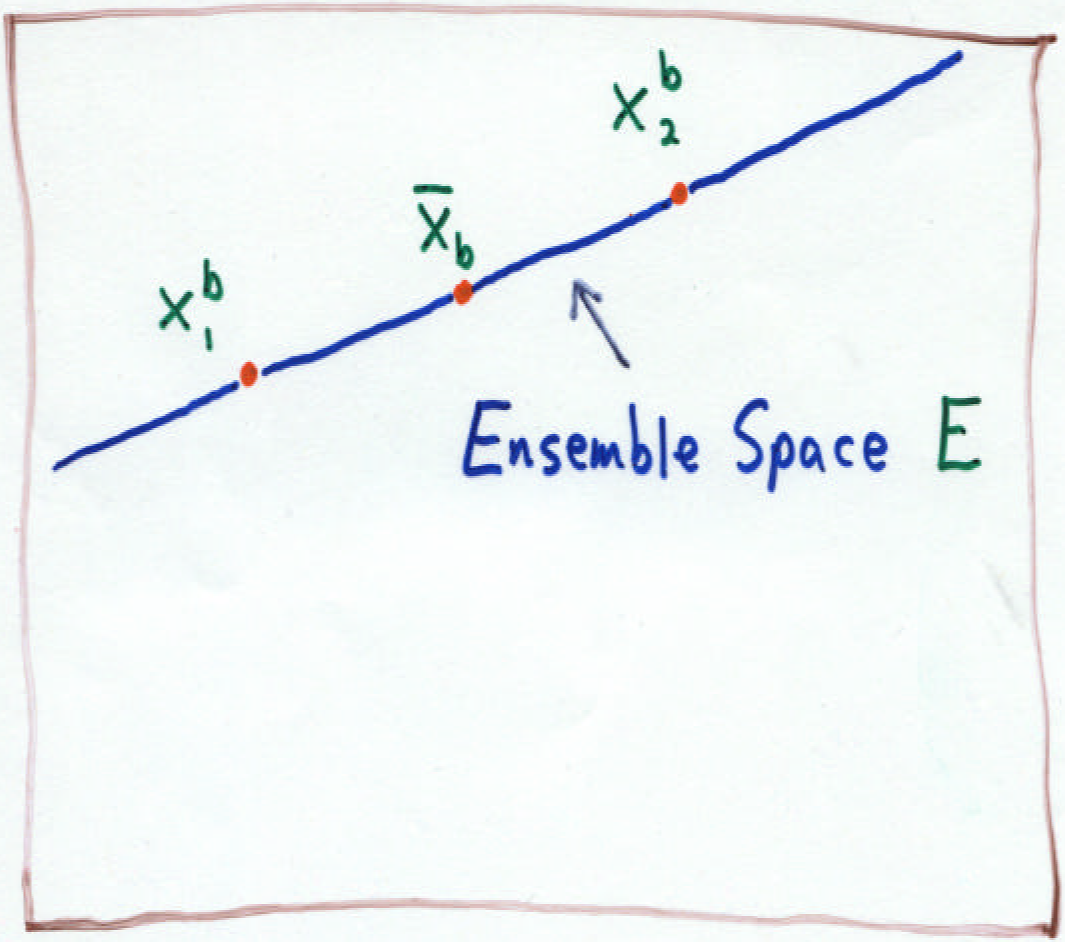
- Minimizing $J(x)$ corresponds to a maximum likelihood estimate of x.

- 3DVar [Lorenc 1986; Parrish + Derber 1992] replaces $\bar{x}^b$ and P^b by a single forecast from the previous analysis and a time-independent background covariance matrix.

BAD NEWS, GOOD NEWS

- Bad news: The background distribution is confined to the space spanned by the ensemble states ⇒ the analysis distribution is confined to the same space.
- Good news: We get to perform data assimilation in a low-dimensional space.
- An ensemble Kalman filter asks:
  Which linear combination of the ensemble states best fits the data?

ENSEMBLE SPACE



Forecast Model State Space

WEIGHT SPACE

- Let X be the matrix with columns

$$(k-1)^{-\frac{1}{2}}\left(x_i^b - \bar{x}^b\right).$$

Then $P^b = X X^T$.

- Every $x \in E$ can be written

$$x = T(w) = \bar{x}^b + X w$$

for some $w \in \mathbb{R}^k$ (weight space).

- Thus, we use the background ensemble perturbations as a "basis" for E [Bishop, Etherton, Majumdar 2001].

NEW AND IMPROVED COST FUNCTION

- Let w have Gaussian background distribution with mean 0 and covariance I.
- Then $T(w)$ is Gaussian with mean $\bar{x}^b$ and covariance $XX^T = P^b$.
- We use the cost function

$$J(w) = w^T w + (y^o - H(T(w)))^T R^{-1} (y^o - H(T(w))).$$

- If H is nonlinear, we can numerically minimize $J(w)$ on $\mathbb{R}^k$, or...

LINEARIZATION OF H ON E

- Evaluate H on each background ensemble state and interpolate [Houtekamer & Mitchell 2001; Anderson 2001].
- Let $y_i^b = H(x_i^b)$, and form the mean $\bar{y}^b$ and perturbation matrix Y of this observation ensemble as we did for the bg ensemble.
- Approximate

$$H(T(w)) = H(\bar{x}^b + Xw) \approx \bar{y}^b + Yw.$$

- Then

$$J(w) \approx w^T w + (y^o - \bar{y}^b - Yw)^T R^{-1} (y^o - \bar{y}^b - Yw).$$

THE ANALYSIS

- The analysis mean $\bar{w}^a$ and covariance A are

$$A = (I + Y^T R^{-1} Y)^{-1}$$

$$\bar{w}^a = A Y^T R^{-1} (y^o - \bar{y}^b).$$

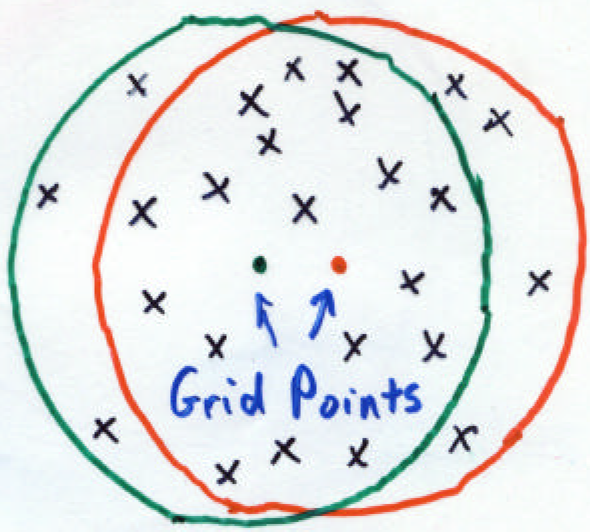
- We obtain A by inverting a k by k matrix with no small eigenvalues.
- Add to $\bar{w}^a$ the columns of $W = ((k-1)A)^{\frac{1}{2}}$ to obtain the analysis ensemble $w^a_1, \ldots, w^a_k$.
- Map to forecast model space: $x^a_i = T(w^a_i)$.
- Any W for which $WW^T = (k-1)A$ would do, but our choice minimizes the distance between the background & analysis ensembles for a suitable metric [OH et al. 2002, 2004].

LOCALIZATION

- For spatially extended systems, it's important to prevent observations from affecting the analysis state at distant locations because:
  - Small ensemble size will produce spurious long-range correlations [Houtekamer + Mitchell 1998; Hamill, Whitaker, Snyder 2001].
  - The linear combination of ensemble states that best fits the data in one region may be quite different from the best linear combination in another region.

LOCALIZATION IN LETKF

- Perform a separate analysis at each model grid point, using only observations from a surrounding local region.

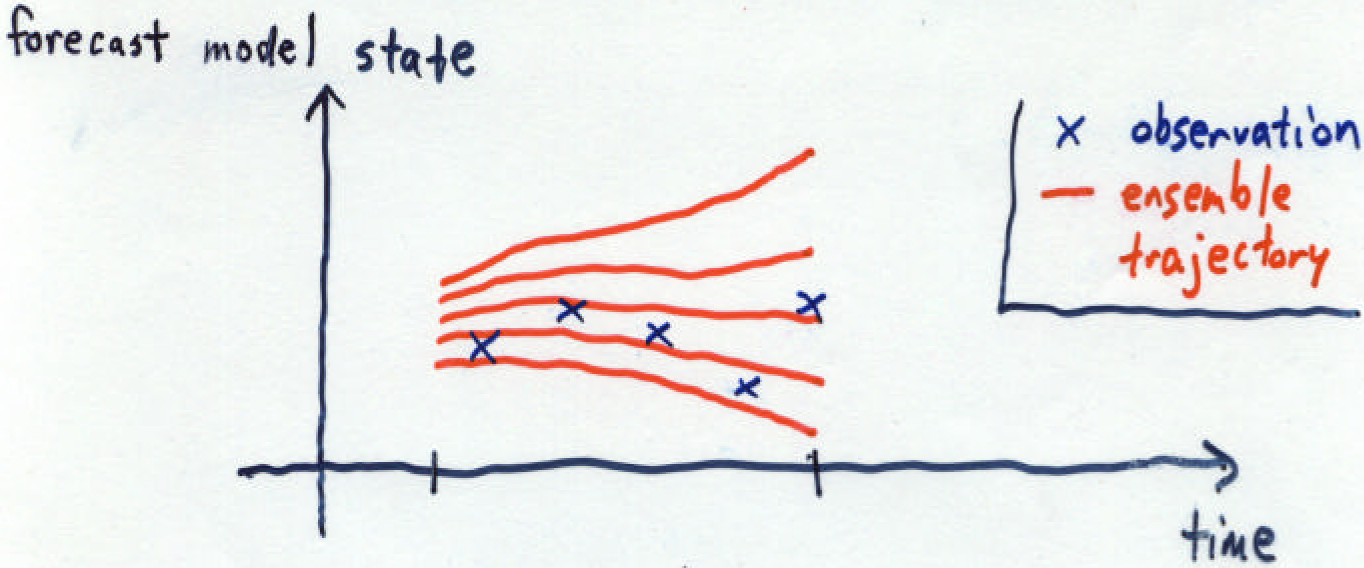


- Each local analysis is independent and can be done in parallel.

4DLETKF

- Suppose the previous analysis was at time t_0, and new observations are taken at times $t_1, t_2, \ldots, t_n$.
- Which linear combination of the background ensemble trajectories $x_i^b(t)$ best fits the data $y^o(t_1), \ldots, y^o(t_n)$?
- Form the observation ensemble whose ith member is $(y_i^b(t_1), \ldots, y_i^b(t_n))$ and proceed as before to get $w_1^a, \ldots, w_k^a$ for each model grid point.
- Map these weights to $x_1^a(t_n), \ldots, x_k^a(t_n)$.
- Simplifies 4DLEKF of [Hunt et al. 2004].

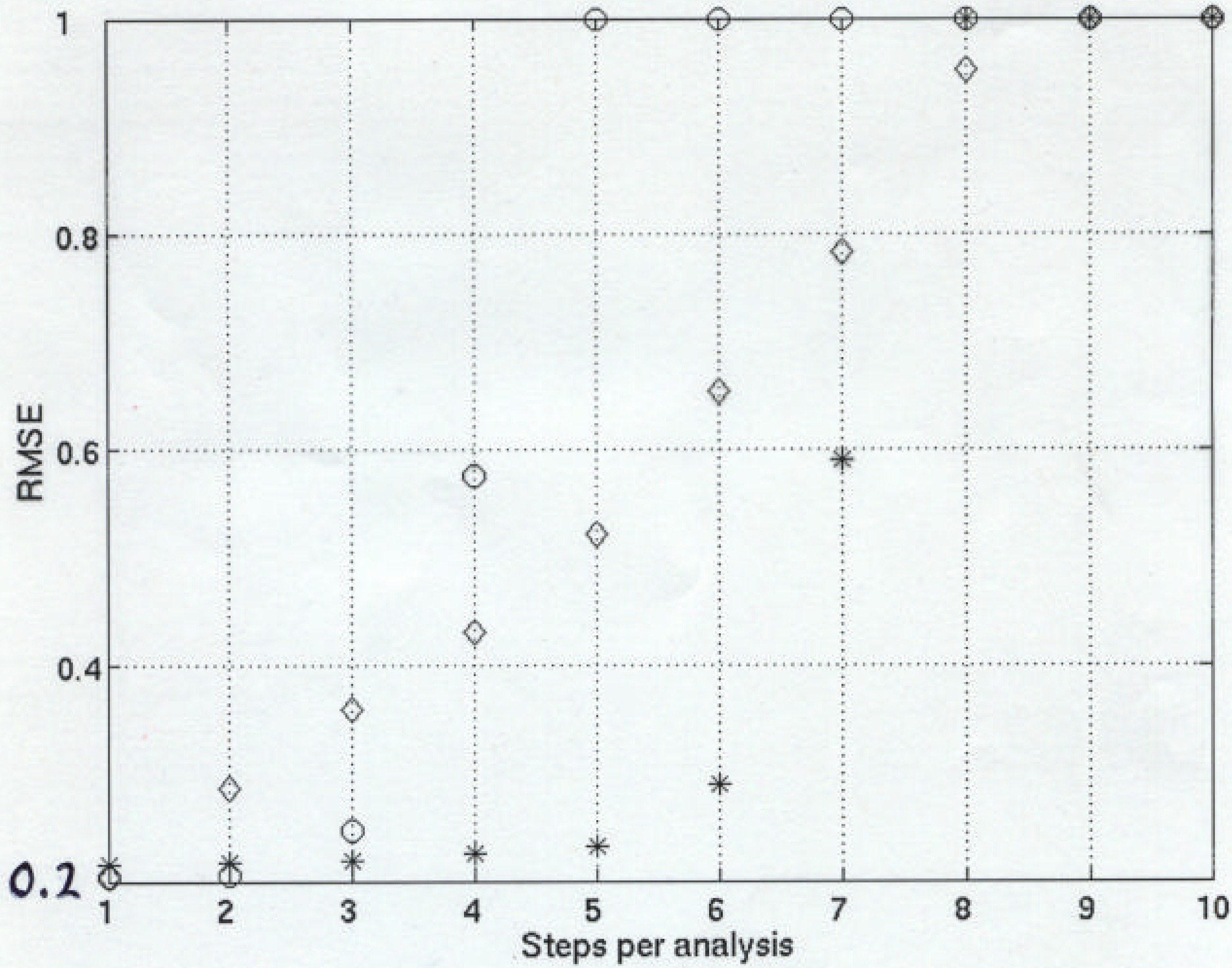
ENSEMBLE 4D-VAR: SCHEMATIC



- For linear dynamics and observations, and a perfect model, this approach yields the same result as applying the ensemble Kalman filter at each observation time.

RESULTS FOR LORENZ-96 MODEL

- Ring of 40 nodes, use observations from ≤6 nodes away, 10 member ensemble.
- Perfect model scenario, obs. at all nodes.



◇ LETKF w/ low variance inflation

○ 4DLETKF " " " "

* 4DLETKF " high " "

CONCLUSIONS

- LETKF provides a simplified framework for ensemble data assimilation.
- LETKF is fast: produces comparable analyses to our earlier LEKF on the NCEP global forecast model in $\sim\frac{1}{3}$ the time [E. Kostelich, I. Szunyogh].
- 4DLETKF may give significant improvement in an operational setting.
- Assimilate locally, forecast globally.

http://keck2.umd.edu/weather/