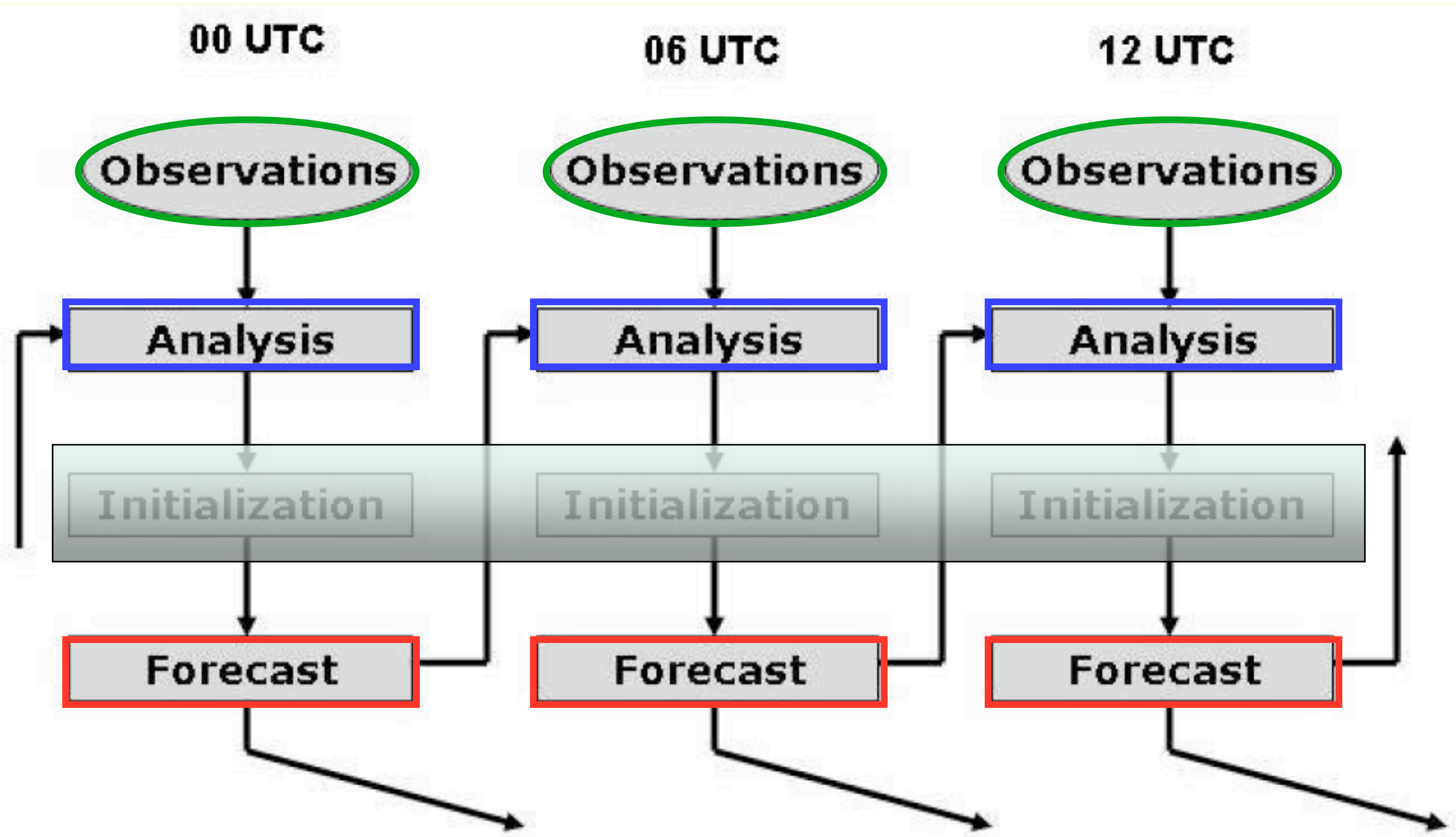
Introduction to data assimilation and least squares methods

Eugenia Kalnay
and many friends

University of Maryland

Contents

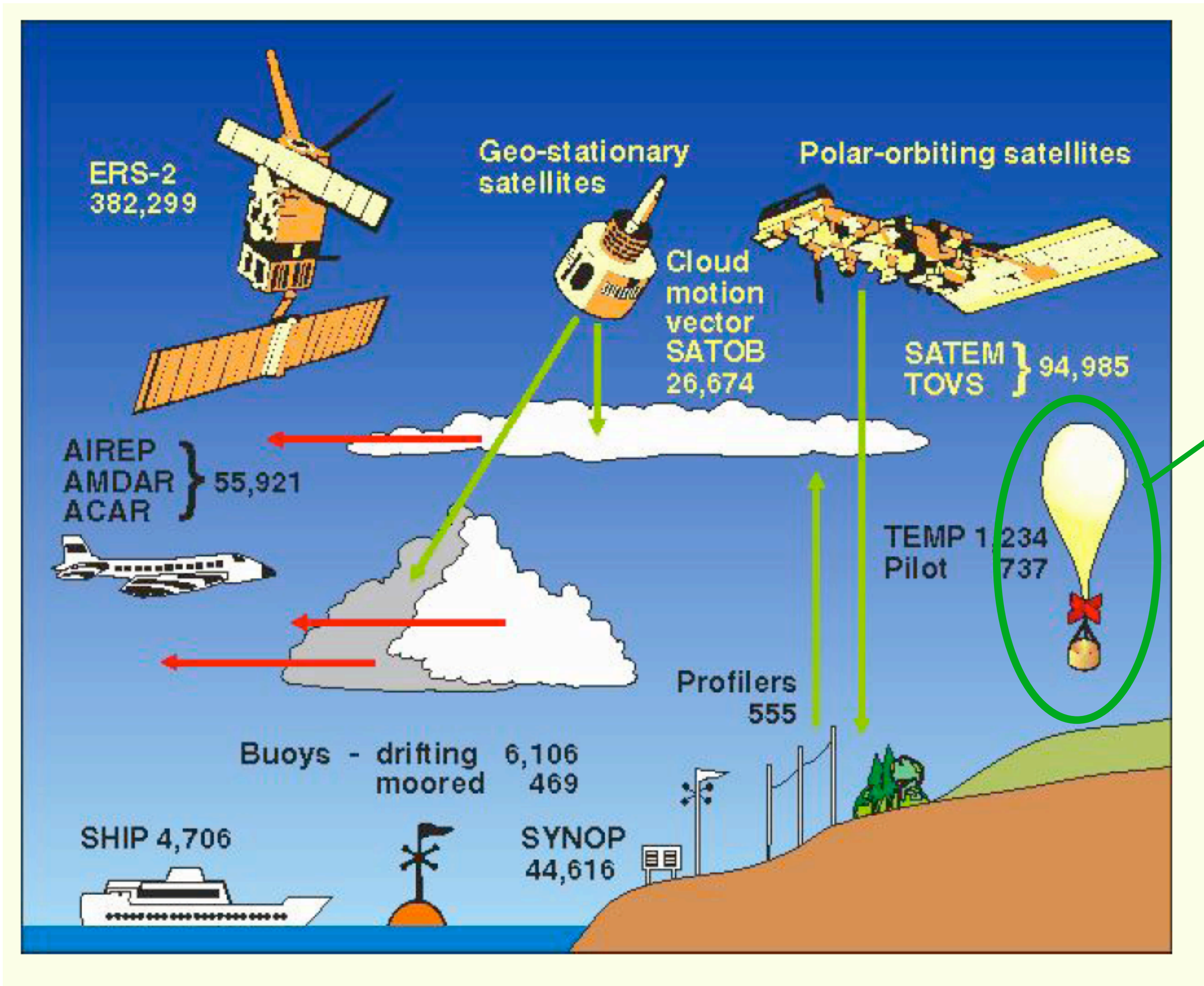
- Forecasting the weather - we are really getting better…
- Why: Better obs? Better models? Better data assimilation?
- Intro to data assim: a toy example, we measure radiance but we want an accurate temperature
- Comparison of the toy and the real equations
- An example from JMA comparing 4D-Var and LETKF (a type of Ensemble Kalman Filter)



Typical 6-hour analysis cycle.

Bayes interpretation: a forecast (the "prior"), is combined with the new observations, to create the Analysis (IC) (the "posterior")

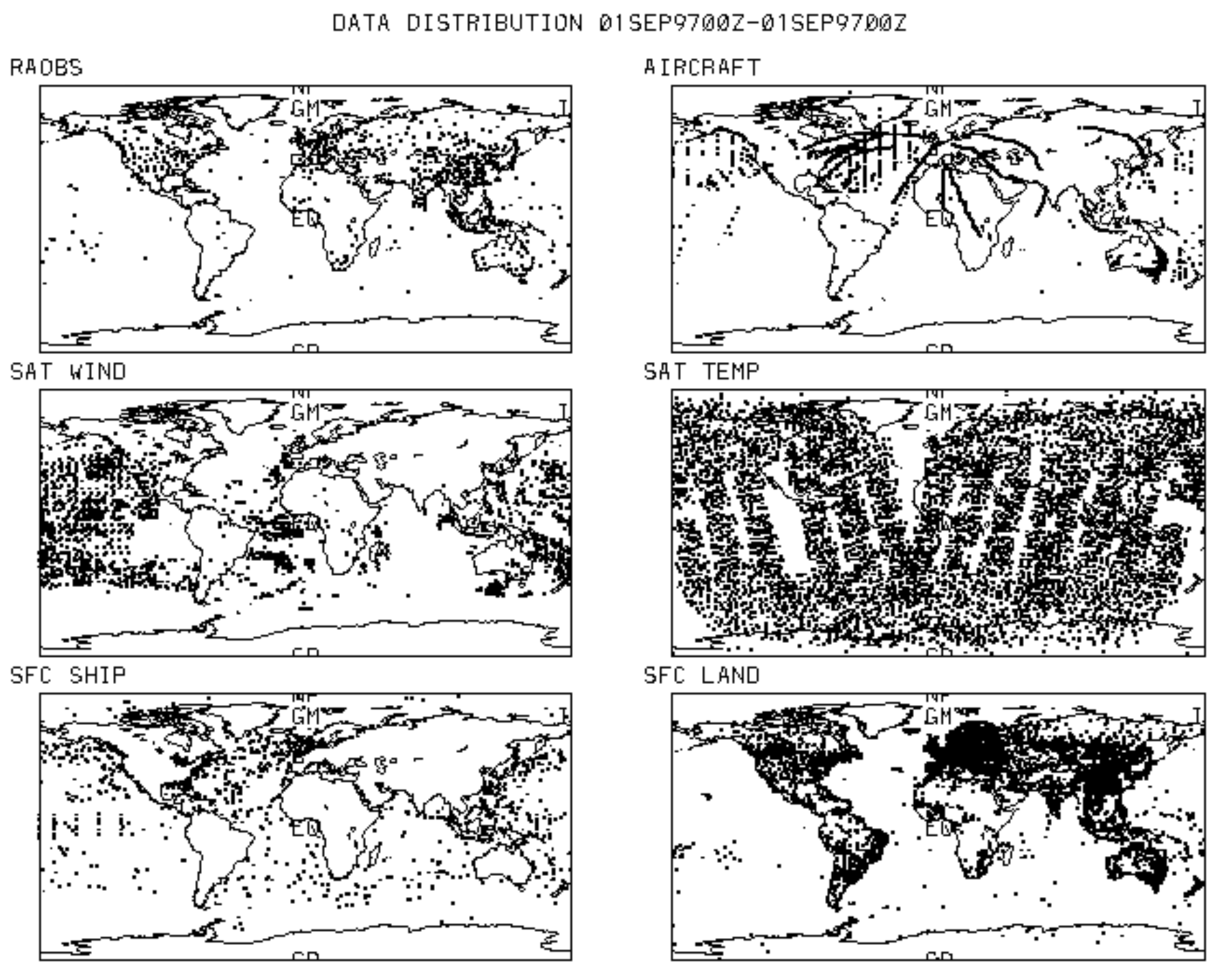
The observing system a few years ago…



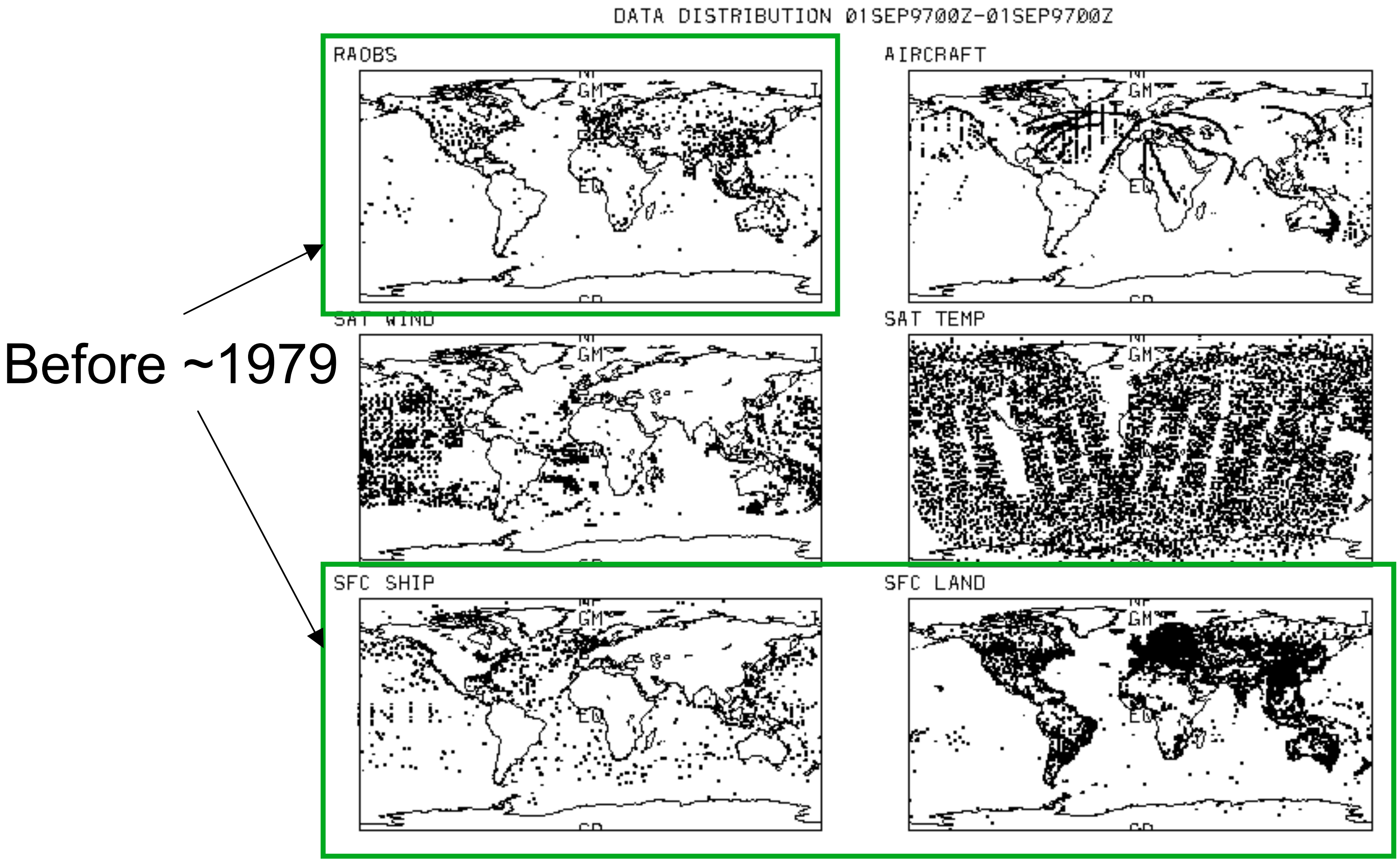
Before 1979 only raobs

Now we have even more satellite data…

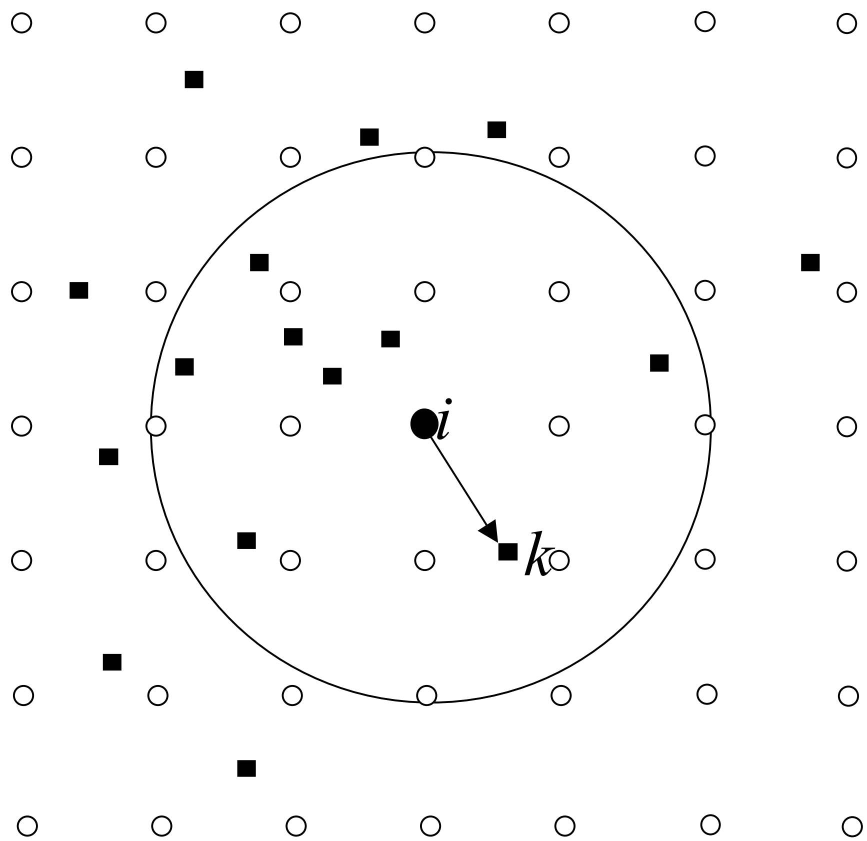
Typical distribution of the observing systems in a 6 hour period:
a real mess: different units, locations, times



Typical distribution of the observing systems in a 6 hour period:
a real mess: different units, locations, times

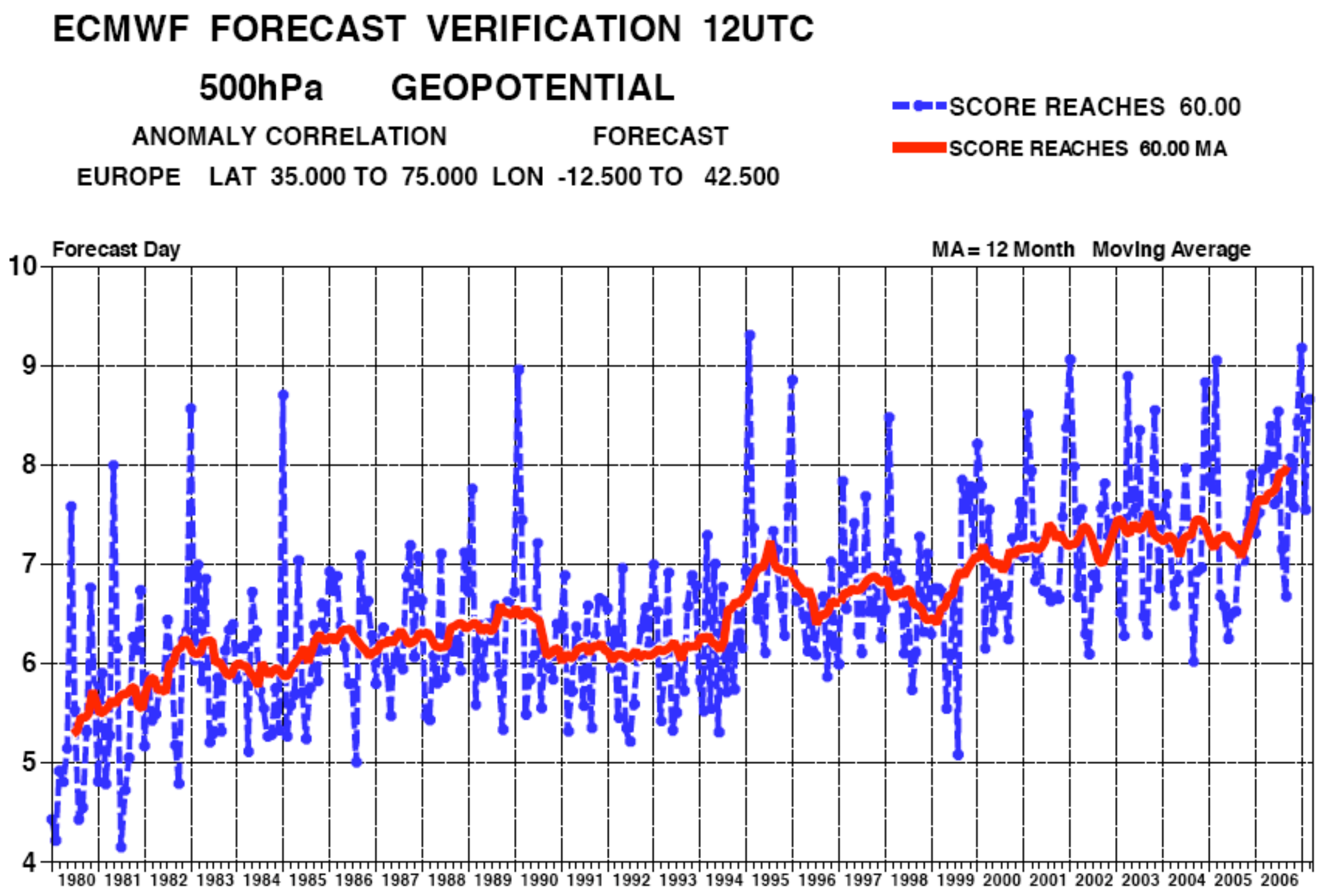


Model grid points (uniformly distributed) and observations (randomly distributed). In a local approach only observations within a radius of influence may be considered

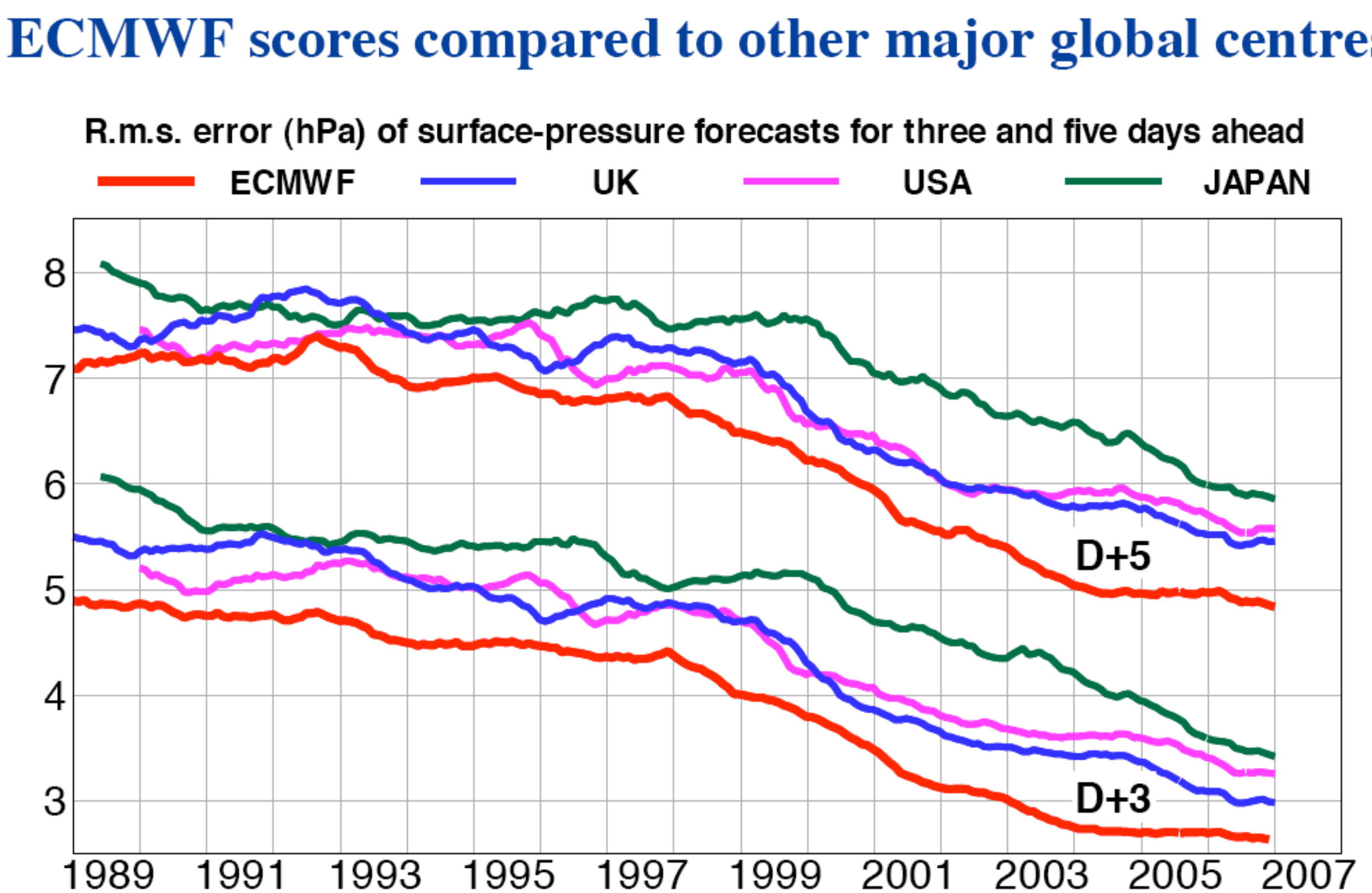


Some statistics of NWP…

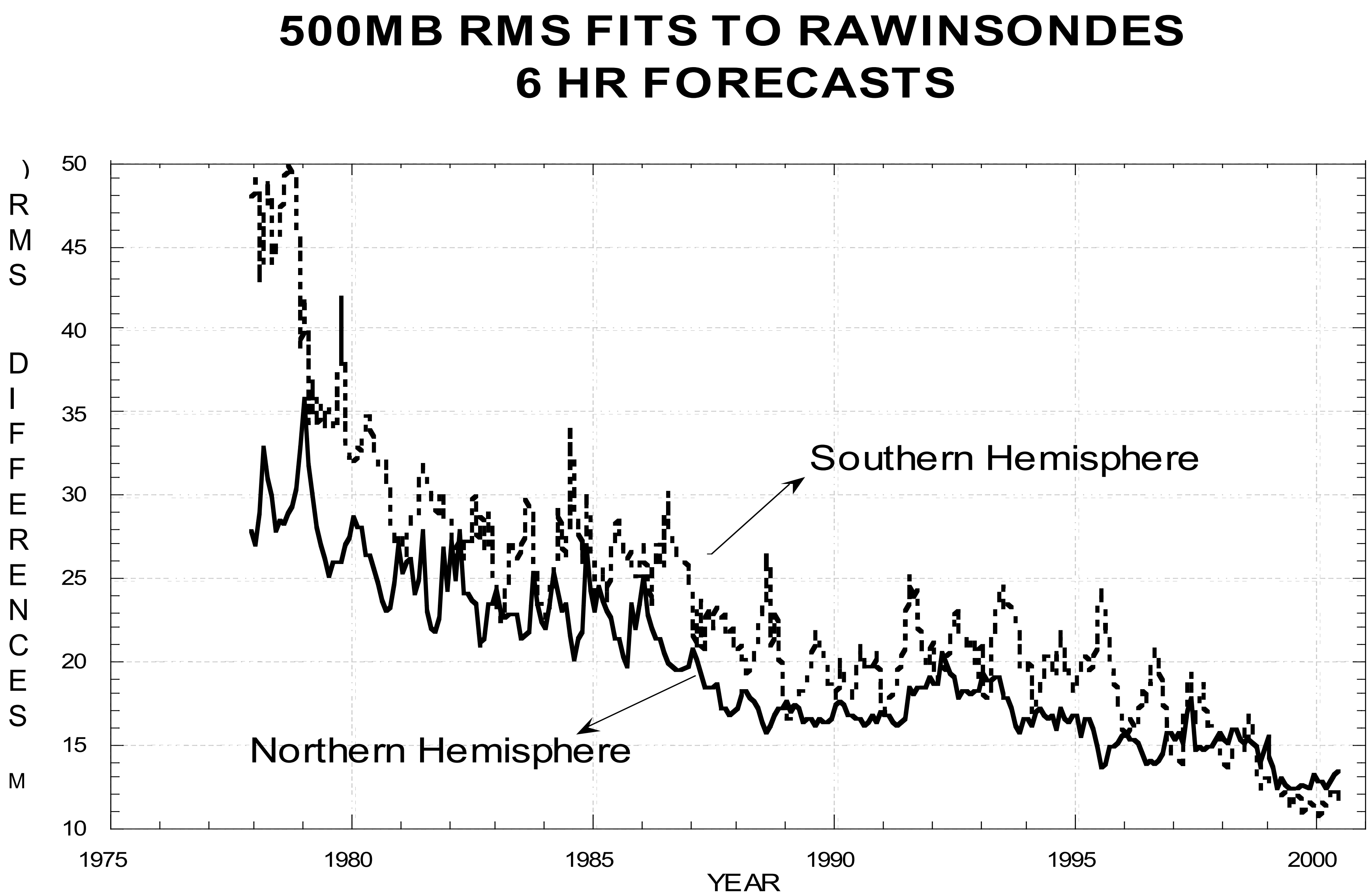
Permanent verifications of the forecasts



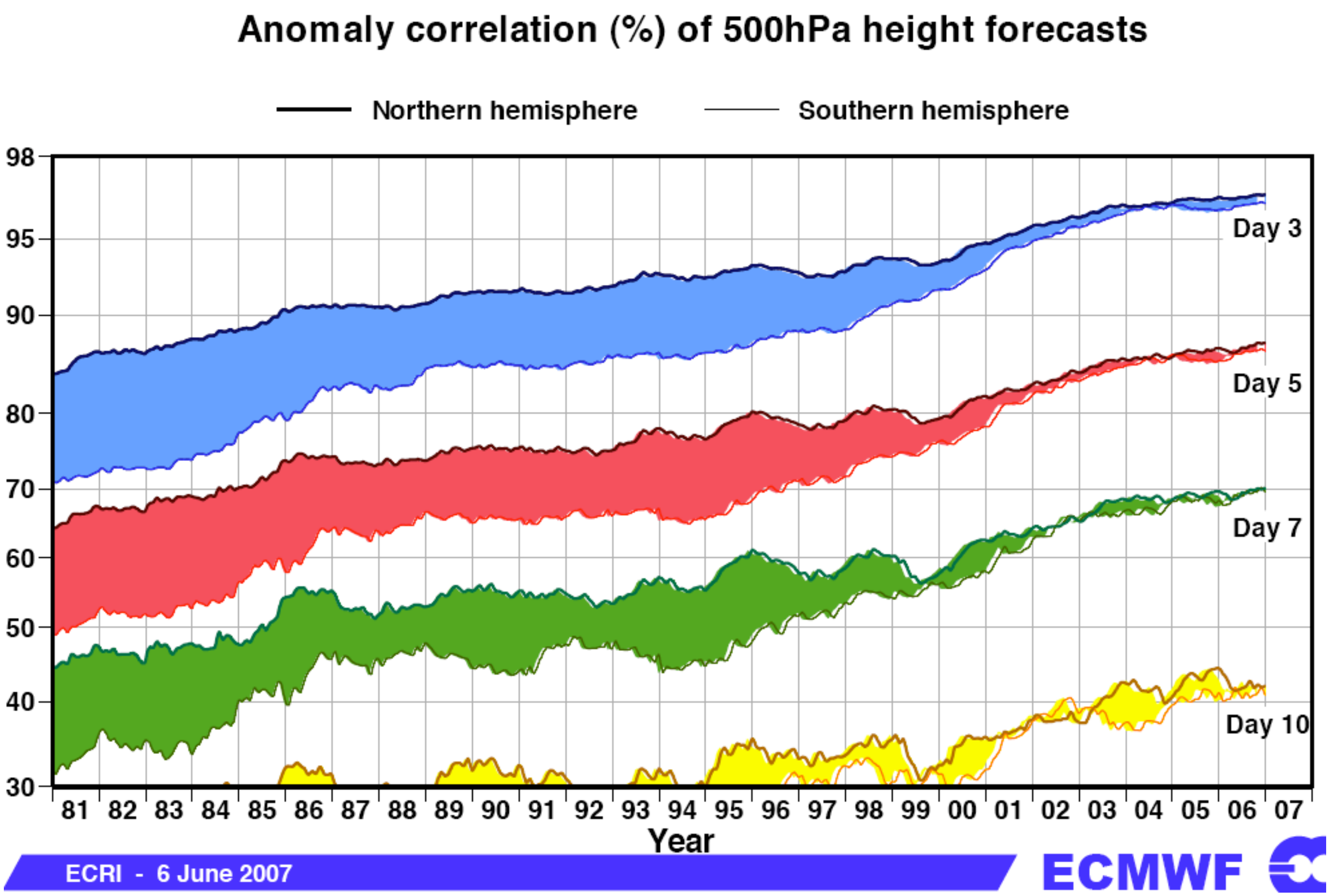
Some comparisons…



We are getting better… (NCEP observational increments)



Comparisons of Northern and Southern Hemispheres



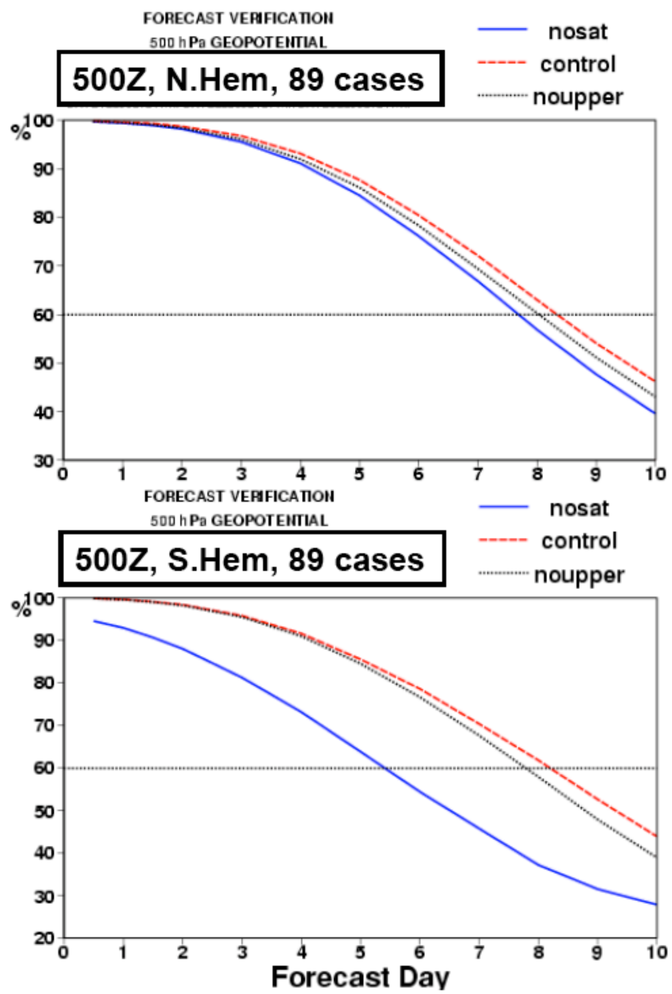
Satellite radiances are essential in the SH

Observing System Experiments (ECMWF - G. Kelly et al.)

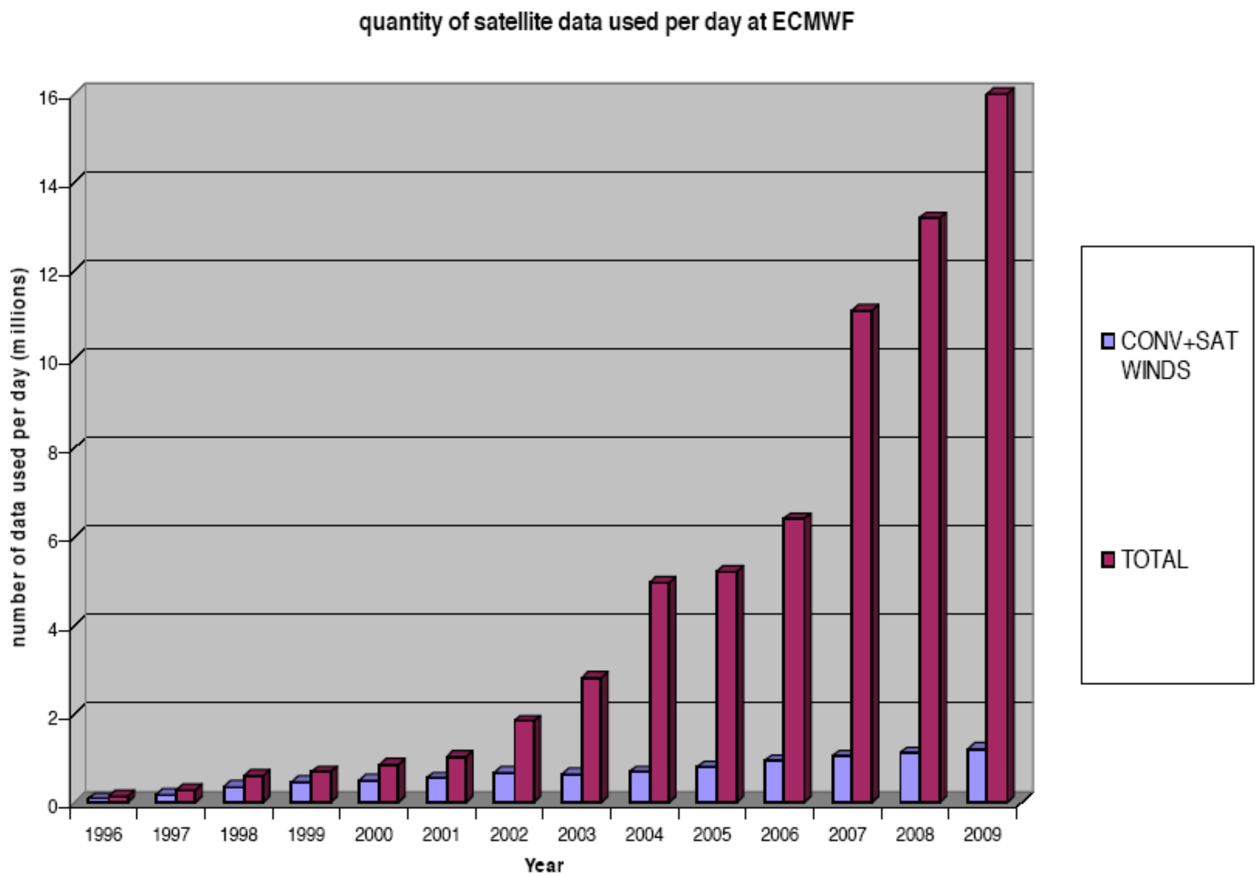
NoSAT= no satellite radiances or winds

Control= like operations

NoUpper=no radiosondes, no pilot winds, no wind profilers



More and more satellite radiances…



Intro. to remote sensing and data assimilation: a toy example

- Assume we have an object, a stone in space
- We want to estimate its temperature T (^oK) accurately but we measure the radiance y (W/m^2) that it emits. We have an *obs. model,* *e.g.*: $y = h(T) \sim \sigma T^4$

Intro. to remote sensing and data assimilation: a toy example

- Assume we have an object, a stone in space
- We want to estimate its temperature T (^{o}K) accurately but we measure the radiance y (W/m^2) that it emits. We have an *obs. model,* e.g.: $y = h(T) \sim \sigma T^4$
- (The *observational model* is also called *forward model*)

We also have a *forecast model* for the temperature

$$T(t_{i+1}) = m\left[T(t_i)\right];$$

$$\text{e.g., } T(t_{i+1}) = T(t_i) + \frac{\Delta t}{C}\left[\text{SW heating-LW cooling}\right]$$

Intro. to remote sensing and data assimilation: a toy example

- Assume we have an object, a stone in space
- We want to estimate its temperature *T* (^{o}K) accurately but we measure the radiance *y* (W/m^2) that it emits. We have an *obs. model,* e.g.: $y = h(T) \sim \sigma T^4$
- (The *observational model* is also called *forward model*)

We also have a *forecast model* for the temperature

$$T(t_{i+1}) = m\left[T(t_i)\right];$$

$$\text{e.g., } T(t_{i+1}) = T(t_i) + \Delta t\left[\text{SW heating-LW cooling}\right]$$

- We will derive the data assim eqs (KF and Var) for this toy system (easy to understand!)

Intro. to remote sensing and data assimilation: a toy example

- Assume we have an object, a stone in space
- We want to estimate its temperature *T* (ºK) accurately but we measure the radiance *y* (W/m^2) that it emits. We have an *obs. model,* *e.g.*: $y = h(T) \sim \sigma T^4$
- (The *observational model* is also called *forward model*)
- We also have a *forecast model* for the temperature

$$T(t_{i+1}) = m\left[T(t_i)\right];$$

$$\text{e.g., } T(t_{i+1}) = T(t_i) + \Delta t\left[\text{SW heating-LW cooling}\right]$$

- We will derive the data assim eqs (KF and Var) for this toy system (easy to understand!)
- Will compare the toy and the real huge vector/matrix equations: they are exactly the same!

Toy temperature data assimilation, measure radiance

We have a forecast T_b (prior) and a radiance obs $y_o = h(T_t) + \varepsilon_0$

The new information (or innovation) is the observational increment:

$$y_o - h(T_b)$$

Toy temperature data assimilation, measure radiance

We have a forecast T_b (prior) and a radiance obs $y_o = h(T_t) + \varepsilon_0$

The new information (or innovation) is the observational increment:

$$y_o - h(T_b)$$

We assume that the obs. and model errors are unbiased, Gaussian and uncorrelated

The innovation can be written in terms of errors:

$$y_o - h(T_b) = h(T_t) + \varepsilon_0 - h(T_b) = \varepsilon_0 + h(T_t) - h(T_b) = \varepsilon_0 - H\varepsilon_b$$

where $H = \partial h / \partial T$ includes changes of units and observation model nonlinearity, e.g.,

$$h(T) \sim \sigma T^4, \partial h / \partial T \sim 4\sigma T^3$$

Toy temperature data assimilation, measure radiance

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

$$y_o - h(T_b) = \varepsilon_0 - H\varepsilon_b$$

Toy temperature data assimilation, measure radiance

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

$$y_o - h(T_b) = \varepsilon_0 - H\varepsilon_b$$

From an OI/KF (sequential) point of view:

$$T_a = T_b + w(y_o - h(T_b)) = T_b + w(\varepsilon_0 - H\varepsilon_b)$$

or

$$\varepsilon_a = \varepsilon_b + w(\varepsilon_0 - H\varepsilon_b)$$

Here w is a weight, and we want to find the optimal weight

Now, the analysis error variance (over many cases) is

$$\overline{\varepsilon_a^2} = \sigma_a^2$$

Toy temperature data assimilation, measure radiance

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

$$y_o - h(T_b) = \varepsilon_0 - H\varepsilon_b$$

From an OI/KF (sequential) point of view:

$$T_a = T_b + w(y_o - h(T_b)) = T_b + w(\varepsilon_0 - H\varepsilon_b)$$

or $$\varepsilon_a = \varepsilon_b + w(\varepsilon_0 - H\varepsilon_b)$$

In OI/KF we choose w to minimize the analysis error: $\overline{\varepsilon_a^2} = \sigma_a^2$

By taking ε_a^2 and averaging in time we can compute: $$\sigma_a^2 = \sigma_b^2 + w^2(\sigma_o^2 + H\sigma_b^2 H) - 2w\sigma_b^2 H$$

assuming that $\varepsilon_b, \varepsilon_0$ are uncorrelated

Toy temperature data assimilation, measure radiance

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

$$y_o - h(T_b) = \varepsilon_0 - H\varepsilon_b$$

From an OI/KF (sequential) point of view:

$$T_a = T_b + w(y_o - h(T_b)) = T_b + w(\varepsilon_0 - H\varepsilon_b)$$

or

$$\varepsilon_a = \varepsilon_b + w(\varepsilon_0 - H\varepsilon_b)$$

In OI/KF we choose w to minimize the analysis error: $\overline{\varepsilon_a^2} = \sigma_a^2$

$$\sigma_a^2 = \sigma_b^2 + w^2(\sigma_o^2 + H\sigma_b^2 H) - 2w\sigma_b^2 H$$

From $\dfrac{\partial \sigma_a^2}{\partial w} = 0$ we obtain $w = \sigma_b^2 H(\sigma_o^2 + H\sigma_b^2 H)^{-1}$

do it!

Toy temperature data assimilation, measure radiance

Repeat: from an OI/KF point of view the analysis (posterior) is:

$$T_a = T_b + w(y_o - h(T_b)) = T_b + w(\varepsilon_0 - H\varepsilon_b)$$

with $$w = \sigma_b^2 H(\sigma_o^2 + \sigma_b^2 H^2)^{-1}$$

Note that the scaled weight wH is between 0 and 1

If $\sigma_o^2 >> \sigma_b^2 H^2$ $\qquad T_a \approx T_b$

If $\sigma_o^2 << \sigma_b^2 H^2$ $\qquad T_a \approx w y_o$

The analysis interpolates between the background and the observation, giving more weights to smaller error variances.

Toy temperature data assimilation, variational approach

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

Innovation: $y_o - h(T_b)$

From a 3D-Var point of view, we want to find a T_a that minimizes the cost function J:

$$2J = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$$

This analysis temperature T_a is closest to both the forecast T_b and the observation y_o and maximizes the likelihood of $T_a \sim T_{truth}$ given the information we have.

It is easier to find the analysis *increment* T_a-T_b that minimizes the cost function J rather than the analysis T_a (this is called *incremental method*)

Toy temperature data assimilation, variational approach

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

Innovation: $y_o - h(T_b)$

From a 3D-Var point of view, we want to find a T_a that minimizes the cost function J:

$$2J_{\min} = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$$

The cost function J comes from a maximum likelihood analysis:

Toy temperature data assimilation, variational approach

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

Innovation: $y_o - h(T_b)$

From a 3D-Var point of view, we want to find a T_a that minimizes the cost function J:

$$2J_{\min} = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$$

Likelihood of T_{truth} given T_b: $\frac{1}{\sqrt{2\pi}\sigma_b} \exp\left[-\frac{T_{truth} - T_b}{2\sigma_b^2} \right]$

Toy temperature data assimilation, variational approach

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

Innovation: $y_o - h(T_b)$

From a 3D-Var point of view, we want to find a T_a that minimizes the cost function J:

$$2J_{\min} = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$$

Likelihood of T_{truth} given T_b: $\dfrac{1}{\sqrt{2\pi}\sigma_b} \exp\left[-\dfrac{\left(T_{truth} - T_b\right)^2}{2\sigma_b^2} \right]$

Likelihood of $h(T_{truth})$ given y_o: $\dfrac{1}{\sqrt{2\pi}\sigma_o} \exp\left[-\dfrac{\left(h(T_{truth}) - y_o\right)^2}{2\sigma_o^2} \right]$

Toy temperature data assimilation, variational approach

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

Innovation: $y_o - h(T_b)$

From a 3D-Var point of view, we want to find a T_a that minimizes the cost function J:

$$2J_{\min} = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$$

Likelihood of T_{truth} given T_b: $\dfrac{1}{\sqrt{2\pi}\sigma_b} \exp\left[-\dfrac{\left(T_{truth} - T_b\right)^2}{2\sigma_b^2} \right]$

Likelihood of $h(T_{truth})$ given y_o: $\dfrac{1}{\sqrt{2\pi}\sigma_o} \exp\left[-\dfrac{\left(h(T_{truth}) - y_o\right)^2}{2\sigma_o^2} \right]$

Joint likelihood of T_{truth}: $\dfrac{1}{2\pi\sigma_b\sigma_o} \exp\left[-\dfrac{\left(T_{truth} - T_b\right)^2}{2\sigma_b^2} - \dfrac{\left(h(T_{truth}) - y_o\right)^2}{2\sigma_o^2} \right]$

Minimizing the cost function maximizes the likelihood of the truth

Toy temperature data assimilation, variational approach

Again, we have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

Innovation: $y_o - h(T_b)$

From a 3D-Var point of view, we want to find $(T_a - T_b)$ that minimizes the cost function J.

This maximizes the likelihood of $T_a \sim T_{truth}$ given both T_b and y_o

$$2J_{\min} = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$$

Toy temperature data assimilation, variational approach

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

Innovation: $y_o - h(T_b)$

From a 3D-Var point of view, we want to find $(T_a - T_b)$ that minimizes the cost function J:

$$2J_{\min} = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$$

Now $h(T_a) - y_o = h(T_b) - y_o + H(T_a - T_b)$

So that from $\partial 2J / \partial (T_a - T_b) = 0$ we get

$$(T_a - T_b)\left(\frac{1}{\sigma_b^2} + \frac{H^2}{\sigma_o^2}\right) = (T_a - T_b)\frac{1}{\sigma_a^2} = H\frac{(y_o - h(T_b))}{\sigma_o^2}$$

Toy temperature data assimilation, variational approach

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

Innovation: $y_o - h(T_b)$

From a 3D-Var point of view, we want to find $(T_a - T_b)$ that minimizes the cost function J:

$$2J_{\min} = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$$

Now $h(T_a) - y_o = h(T_b) - y_o + H(T_a - T_b)$

So that from $\partial 2J / \partial (T_a - T_b) = 0$ we get

$$(T_a - T_b)\left(\frac{1}{\sigma_b^2} + \frac{H^2}{\sigma_o^2}\right) = (T_a - T_b)\frac{1}{\sigma_a^2} = H\frac{(y_o - h(T_b))}{\sigma_o^2}$$

or $T_a = T_b + w\left(y_o - h(T_b)\right)$ where now

$$w = \left(\sigma_b^{-2} + H\sigma_o^{-2}H\right)^{-1} H\sigma_o^{-2} = \sigma_a^2 H\sigma_o^{-2}$$

Toy temperature data assimilation, variational approach

We have a forecast T_b and a radiance obs $y_o = h(T_t) + \varepsilon_0$

Innovation: $y_o - h(T_b)$

From a 3D-Var point of view, we want to find $(T_a - T_b)$ that minimizes the cost function J:

$$2J_{\min} = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$$

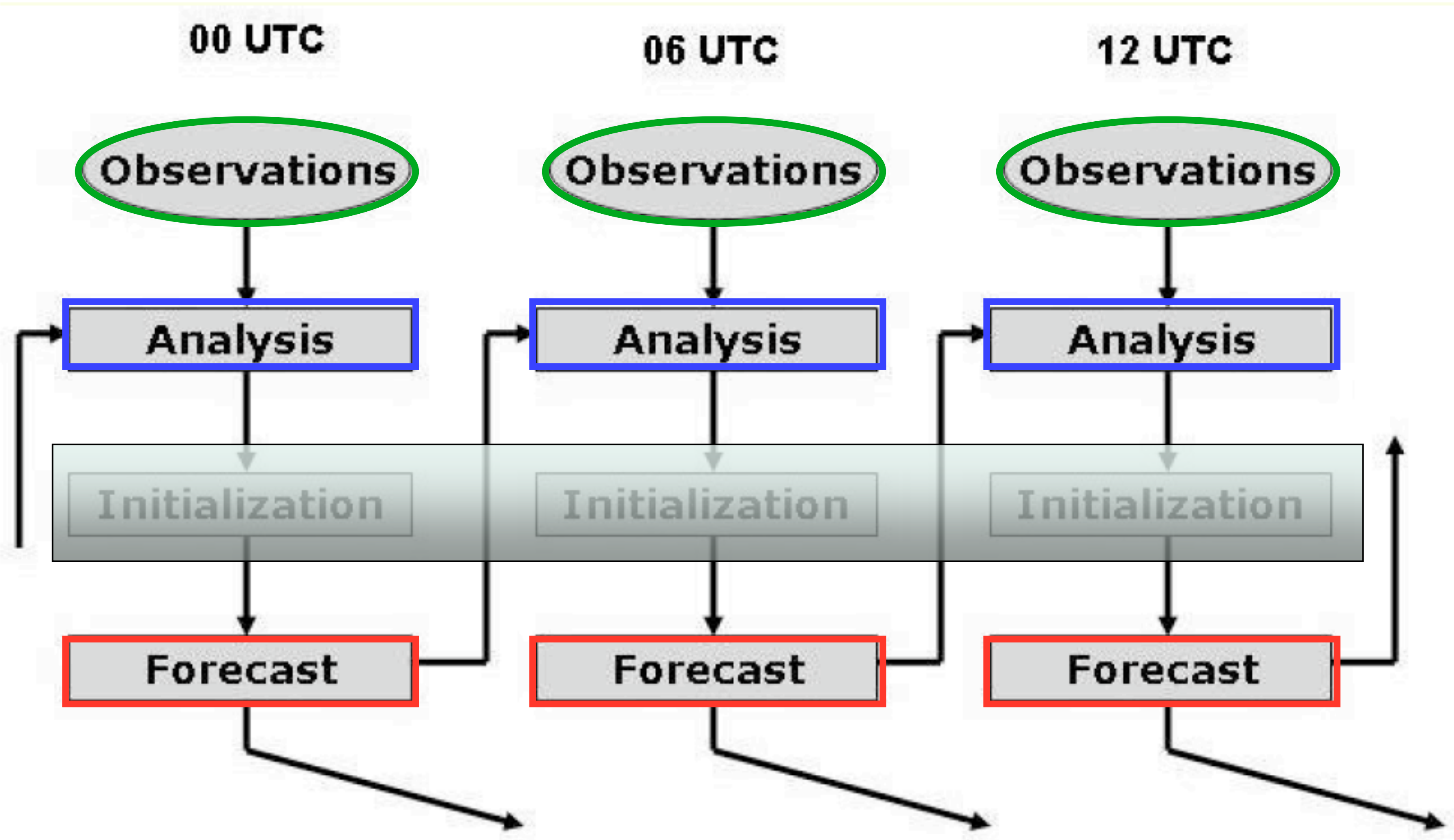
$$T_a = T_b + w(y_o - h(T_b)) = T_b + w(\varepsilon_0 - H\varepsilon_b)$$ where

$$w = \left(\sigma_b^{-2} + H\sigma_o^{-2}H\right)^{-1} H\sigma_o^{-2} = \sigma_a^2 H\sigma_o^{-2}$$

This variational solution <u>is the same</u> as the one obtained with Kalman filter (sequential approach, like Optimal Interpolation):

with $$w = \sigma_b^2 H(\sigma_o^2 + \sigma_b^2 H^2)^{-1}$$

(Show the w's are the same!)



Typical 6-hour analysis cycle.

Forecast phase, followed by Analysis phase

Toy temperature analysis cycle (Kalman Filter)

Forecasting phase, from t_i to t_{i+1}: $$T_b(t_{i+1}) = m\left[T_a(t_i)\right]$$

Forecast error: $$\varepsilon_b(t_{i+1}) = T_b(t_{i+1}) - T_t(t_{i+1}) =$$
$$m\left[T_a(t_i)\right] - m\left[T_t(t_i)\right] + \varepsilon_m(t_{i+1}) = M\varepsilon_a(t_i) + \varepsilon_m(t_{i+1})$$

So that we can predict the forecast error variance

$$\sigma_b^2(t_{i+1}) = M^2\sigma_a^2(t_i) + Q_i; \quad Q_i = \overline{\varepsilon_m^2(t_{i+1})}$$

(The forecast error variance comes from the analysis and model errors)

Toy temperature analysis cycle (Kalman Filter)

Forecasting phase, from t_i to t_{i+1}: $T_b(t_{i+1}) = m\left[T_a(t_i)\right]$

Forecast error: $\varepsilon_b(t_{i+1}) = T_b(t_{i+1}) - T_t(t_{i+1}) =$

$$m\left[T_a(t_i)\right] - m\left[T_t(t_i)\right] + \varepsilon_m(t_{i+1}) = M\varepsilon_a(t_i) + \varepsilon_m(t_{i+1})$$

So that we can predict the forecast error variance

$$\sigma_b^2(t_{i+1}) = M^2\sigma_a^2(t_i) + Q_i; \quad Q_i = \overline{\varepsilon_m^2(t_{i+1})}$$

(The forecast error variance comes from the analysis and model errors)

Now we can compute the optimal weight (KF or Var, whichever form is more convenient, since they are equivalent):

$$w = \sigma_b^2 H(\sigma_o^2 + H\sigma_b^2 H)^{-1} = \left(\sigma_b^{-2} + H\sigma_o^{-2}H\right)^{-1} H\sigma_o^{-2}$$

Toy temperature analysis cycle (Kalman Filter)

Analysis phase: we use the new observation $y_o(t_{i+1})$

compute the new observational increment $y_o(t_{i+1}) - h\left(T_b(t_{i+1})\right)$

and the new analysis:

$$T_a(t_{i+1}) = T_b(t_{i+1}) + w_{i+1}\left[y_o(t_{i+1}) - h\left(T_b(t_{i+1})\right)\right]$$

We also need the compute the new analysis error variance:

from $\sigma_a^{-2} = \sigma_b^{-2} + H\sigma_o^{-2}H$

we get $\sigma_a^2(t_{i+1}) = \left(\dfrac{\sigma_o^2\sigma_b^2}{\sigma_o^2 + H^2\sigma_b^2}\right)_{i+1} = (1 - w_{i+1}H)\sigma_{b\,i+1}^2 < \sigma_{b\,i+1}^2$

now we can advance to the next cycle $t_{i+2}, t_{i+3}, \ldots$

Summary of toy system equations (for a scalar)

$$T_b(t_{i+1}) = m\left[T_a(t_i)\right] \qquad \sigma_b^2(t_{i+1}) = M^2\left[\sigma_a^2(t_i)\right] + Q$$

$$M = \partial m / \partial T$$

Interpretation…

"We use the model to forecast T_b and to update the forecast error variance from t_i to t_{i+1}"

Q: model deficiencies error covariance

Summary of toy system equations (for a scalar)

$$T_b(t_{i+1}) = m\left[T_a(t_i)\right] \qquad \sigma_b^2(t_{i+1}) = M^2\left[\sigma_a^2(t_i)\right] \qquad M = \partial m / \partial T$$

"We use the model to forecast T_b and to update the forecast error variance from t_i to t_{i+1} "

At t_{i+1} $\quad T_a = T_b + w\left[y_o - h\left(T_b\right)\right]$

"The analysis is obtained by adding to the background the innovation (difference between the observation and the first guess) multiplied by the optimal weight:

Summary of toy system equations (for a scalar)

$$T_b(t_{i+1}) = m\left[T_a(t_i)\right] \quad \sigma_b^2(t_{i+1}) = M^2\left[\sigma_a^2(t_i)\right] + Q \qquad M = \partial m / \partial T$$

"We use the model to forecast T_b and to update the forecast error variance from t_i to t_{i+1} "

At t_{i+1} $\quad T_a = T_b + w\left[y_o - h\left(T_b\right)\right]$

"The analysis is obtained by adding to the background the innovation (difference between the observation and the first guess) multiplied by the optimal weight:

$$w = \sigma_b^2 H(\sigma_o^2 + H\sigma_b^2 H)^{-1}$$

"The optimal weight is the background error variance divided by the sum of the observation and the background error variance. $H = \partial h / \partial T$ ensures that the magnitudes and units are correct."

Summary of toy system equations (cont.)

$$w = \sigma_b^2 H(\sigma_o^2 + H\sigma_b^2 H)^{-1}$$

"The optimal weight is the background error variance divided by the total variance (sum of the observation and the background error variance). $H = \partial h / \partial T$ ensures that the magnitudes and units are correct."

Note that the larger the background error variance, the larger the correction to the first guess.

Summary of toy system equations (cont.)

The analysis error variance is given by

$$\sigma_a^2 = \left(\frac{\sigma_o^2 \sigma_b^2}{\sigma_o^2 + H^2 \sigma_b^2} \right) = (1 - wH)\sigma_b^2$$

"The analysis error variance is reduced from the background error by a factor (1 - scaled optimal weight)"

Summary of toy system equations (cont.)

The analysis error variance is given by

$$\sigma_a^2 = \left(\frac{\sigma_o^2 \sigma_b^2}{\sigma_o^2 + H^2 \sigma_b^2} \right) = (1 - wH)\sigma_b^2$$

"The analysis error variance is reduced from the background error by a factor (1 - scaled optimal weight)"

This can also be written as

$$\sigma_a^{-2} = \left(\sigma_b^{-2} + \sigma_o^{-2} H^2 \right)$$

"The analysis precision is given by the sum of the background and observation precisions"

Equations for toy and real huge systems

These statements are important because they hold true for data assimilation systems in very large multidimensional problems (e.g., NWP).

Instead of model, analysis and observational scalars, we have 3-dimensional vectors of sizes of the order of 10^7-10^8

Equations for toy and real huge systems

These statements are important because they hold true for data assimilation systems in very large multidimensional problems (e.g., NWP).

Instead of model, analysis and observational scalars, we have 3-dimensional vectors of sizes of the order of 10^7-10^8

We have to replace scalars (obs, forecasts) by vectors

$$T_b \to \mathbf{x}_b; \quad T_a \to \mathbf{x}_a; \quad y_o \to \mathbf{y}_o;$$

and their error variances by error covariances:

$$\sigma_b^2 \to \mathbf{B}; \quad \sigma_a^2 \to \mathbf{A}; \quad \sigma_o^2 \to \mathbf{R};$$

Interpretation of the NWP system of equations

"We use the model to forecast from t_i to t_{i+1} "

$$\mathbf{x}_b(t_{i+1}) = M\left[\mathbf{x}_a(t_i)\right]$$

At t_{i+1} $\quad \mathbf{x}_a = \mathbf{x}_b + \mathbf{K}\left[\mathbf{y}_o - H\left(\mathbf{x}_b\right)\right]$

Interpretation of the NWP system of equations

"We use the model to forecast from t_i to t_{i+1} "

$$\mathbf{x}_b(t_{i+1}) = M\left[\mathbf{x}_a(t_i)\right]$$

At t_{i+1} $\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}\left[\mathbf{y}_o - H\left(\mathbf{x}_b\right)\right]$

"The analysis is obtained by adding to the background the innovation (difference between the observation and the first guess) multiplied by the optimal Kalman gain (weight) matrix"

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

Interpretation of the NWP system of equations

"We use the model to forecast from t_i to t_{i+1} "

$$\mathbf{x}_b(t_{i+1}) = M\left[\mathbf{x}_a(t_i)\right]$$

At t_{i+1} $\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}\left[\mathbf{y}_o - H\left(\mathbf{x}_b\right)\right]$

"The analysis is obtained by adding to the background the innovation (difference between the observation and the first guess) multiplied by the optimal Kalman gain (weight) matrix"

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

"The optimal weight is the background error covariance divided by the sum of the observation and the background error covariance. $\mathbf{H} = \partial H / \partial \mathbf{x}$ ensures that the magnitudes and units are correct. The larger the background error variance, the larger the correction to the first guess."

Interpretation of the NWP system of equations

Forecast phase:

"We use the model to forecast from t_i to t_{i+1} "

$$\mathbf{x}_b(t_{i+1}) = M\left[\mathbf{x}_a(t_i)\right]$$

Interpretation of the NWP system of equations

Forecast phase:

"We use the model to forecast from t_i to t_{i+1} "

$$\mathbf{x}_b(t_{i+1}) = M\left[\mathbf{x}_a(t_i)\right]$$

"We use the linear tangent model **M** and its adjoint **M***^{T}* to forecast **B** (plus model errors covariance **Q**)"

$$\mathbf{B}(t_{i+1}) = \mathbf{M}\left[\mathbf{A}(t_i)\right]\mathbf{M}^T + \mathbf{Q}$$

Interpretation of the NWP system of equations

Forecast phase:

"We use the model to forecast from t_i to t_{i+1} "

$$\mathbf{x}_b(t_{i+1}) = M\left[\mathbf{x}_a(t_i)\right]$$

"We use the linear tangent model and its adjoint to forecast **B**"

$$\mathbf{B}(t_{i+1}) = \mathbf{M}\left[\mathbf{A}(t_i)\right]\mathbf{M}^T$$

"However, this step is so horrendously expensive that it makes Kalman Filter computationally unfeasible for NWP".

"Ensemble Kalman Filter solves this problem by estimating **B** using an ensemble of forecasts."

Interpretation of the NWP system of equations

Forecast phase:

"We use the model to forecast from t_i to t_{i+1} "

$$\mathbf{x}_b(t_{i+1}) = M\left[\mathbf{x}_a(t_i)\right]$$

$$\mathbf{B}(t_{i+1}) = \mathbf{M}\left[\mathbf{A}(t_i)\right]\mathbf{M}^T$$ "This is too expensive!"

"Ensemble Kalman Filter solves this problem by estimating **B** using an ensemble of K~50-100 forecasts."

$$\mathbf{x}_b^k(t_{i+1}) = M\left[\mathbf{x}_a^k(t_i)\right],\ k = 1,2,...K$$ ensemble of forecasts

$$\mathbf{B}(t_{i+1}) = \frac{1}{K-1}\sum_{k=1}^{K}(\mathbf{x}_b^k - \overline{\mathbf{x}}_b)*(\mathbf{x}_b^k - \overline{\mathbf{x}}_b)^T = \frac{1}{K-1}\mathbf{X}_b\mathbf{X}_b^T$$

Summary of NWP equations (cont.)

The analysis error covariance is given by

$$\mathbf{A} = \left(\mathbf{I} - \mathbf{KH}\right)\mathbf{B}$$

“The analysis covariance is reduced from the background covariance by a factor (**I** - scaled optimal gain)”

Summary of NWP equations (cont.)

The analysis error covariance is given by

$$\mathbf{A} = \left(\mathbf{I} - \mathbf{KH}\right)\mathbf{B}$$

"The analysis covariance is reduced from the background covariance by a factor (**I** - scaled optimal gain)"

This can also be written as

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}$$

"The analysis precision is given by the sum of the background and observation precisions"

Summary of NWP equations (cont.)

The analysis error covariance is given by

$$\mathbf{A} = \left(\mathbf{I} - \mathbf{KH}\right)\mathbf{B}$$

"The analysis covariance is reduced from the background covariance by a factor (**I** - scaled optimal gain)"

This can also be written as

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H}$$

"The analysis precision is given by the sum of the background and observation precisions"

$$\mathbf{K} = \mathbf{BH}^{T}(\mathbf{R} + \mathbf{HBH}^{T})^{-1} = (\mathbf{B}^{-1} + \mathbf{H}^{T}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{T}\mathbf{R}^{-1}$$

"The variational approach and the sequential approach are solving the same problem, with the same **K**, but only KF (or EnKF) provide an estimate of the analysis error covariance"

Variational: 3D-Var

$$J = \min \frac{1}{2}[(\mathbf{x}^a - \mathbf{x}^b)^{\mathbf{T}} \mathbf{B}^{-1} (\mathbf{x}^a - \mathbf{x}^b) + (H\mathbf{x}^a - \mathbf{y})^{\mathbf{T}} \mathbf{R}^{-1} (H\mathbf{x}^a - \mathbf{y})]$$

Distance to forecast

Distance to observations

at the analysis time

4D-Var

$$J = \min \frac{1}{2}[(\mathbf{x}_0 - \mathbf{x}_0^b)^{\mathbf{T}} \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \sum_{i=1}^{s} (H\mathbf{x}_i - \mathbf{y}_i)^{\mathbf{T}} \mathbf{R}_i^{-1} (H\mathbf{x}_i - \mathbf{y}_i)]$$

Distance to background at the initial time

Distance to observations in a **time window interval t_0-t_1**

Control variable $\mathbf{x}(t_0)$

Analysis $\mathbf{x}(t_1) = M[\mathbf{x}(t_0)]$

It seems like a simple change, but it is not! (e.g., adjoint)

What is B? It should be tuned…

Ensemble Transform Kalman Filter (EnKF)

Forecast step:

$$\mathbf{x}^b_{n,k} = M_n\left(\mathbf{x}^a_{n-1,k}\right)$$

$$\mathbf{B}_n = \frac{1}{K-1}\mathbf{X}^b_n\mathbf{X}^{bT}_n, \textit{ where } \mathbf{X}^b_n = \left[\mathbf{x}^b_{n,1} - \overline{\mathbf{x}}^b_n; \ldots, \mathbf{x}^b_{n,K} - \overline{\mathbf{x}}^b_n\right]$$

Analysis step:

$$\mathbf{x}^a_n = \mathbf{x}^b_n + \mathbf{K}_n(\mathbf{y}_n - H\overline{\mathbf{x}}^b_n); \mathbf{K}_n = \mathbf{B}_n\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{B}_n\mathbf{H}^T)^{-1}$$

The new analysis error covariance in the ensemble space is (Hunt et al. 2007)

$$\tilde{\mathbf{A}}_n = \left[(K-1)\mathbf{I} + (\mathbf{H}\mathbf{X}^b_n)^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{X}^b_n)\right]^{-1}$$

And the new ensemble perturbations are given by a matrix transform:

$$\mathbf{X}^a_n = \mathbf{X}^b_n\left[(K-1)\tilde{\mathbf{A}}_n\right]^{1/2}$$

Comparison of 4-D Var and LETKF at JMA

T. Miyoshi and Y. Sato

- 4D-Var and EnKF are the two advanced, feasible methods
  - There will be a workshop on them in Buenos Aires (Nov'08)!!!
- In Ensemble Kalman Filter the background error covariance **B** is approximated and advanced in time with an ensemble of K forecasts. In the subspace of the ensemble, **B=I** so that matrix inversions are efficient.
- So far, comparisons show EnKF is slightly better than 3D-Var, but there has not been enough time to develop tunings
- At JMA, Takemasa Miyoshi has been performing comparisons of the Local Ensemble Transform Kalman Filter (Hunt et al., 2007) with their operational 4D-Var
- Comparisons are made for August 2004

Comparison of 4D-Var and LETKF at JMA

T. Miyoshi and Y. Sato

N.H.

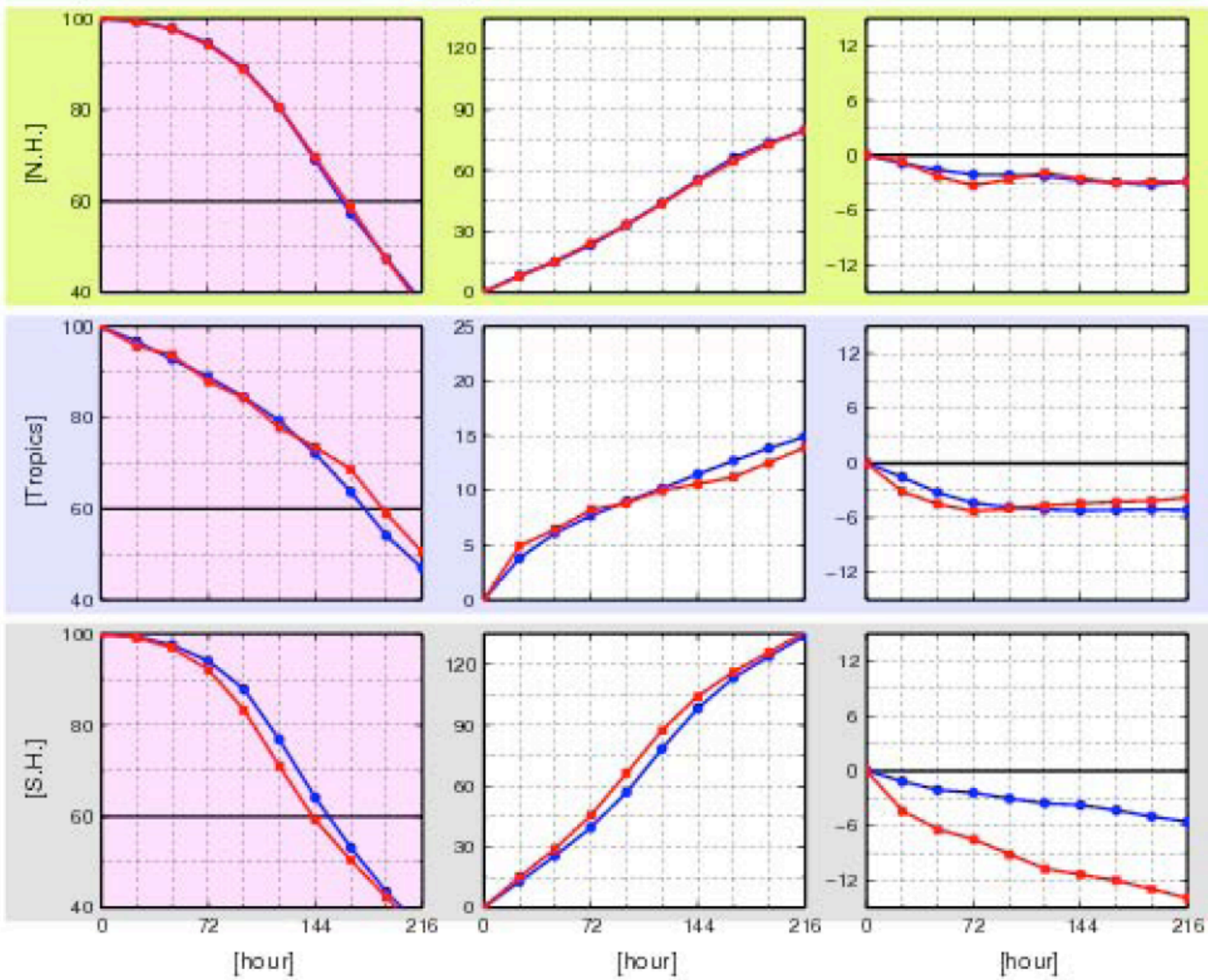
better

Tropics

better

S.H.

worse



RMS error against analysis

Comparison of LETKF and 4D-Var at JMA

T. Miyoshi and Y. Sato

N.H.

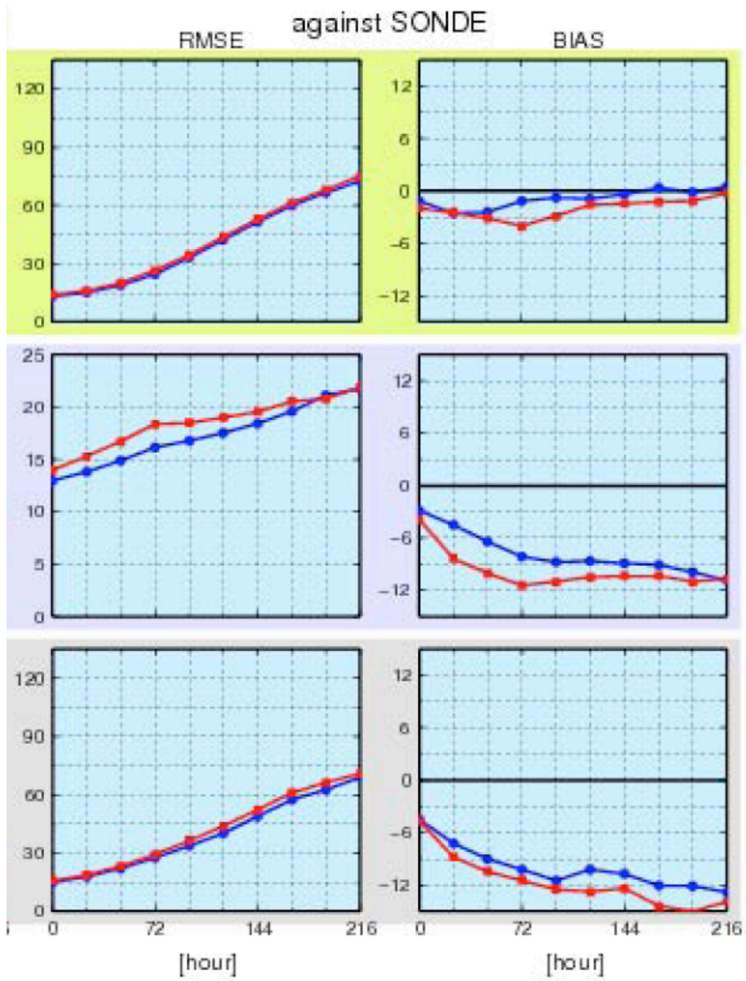
same

Tropics

worse

S.H.

worse



RMS error

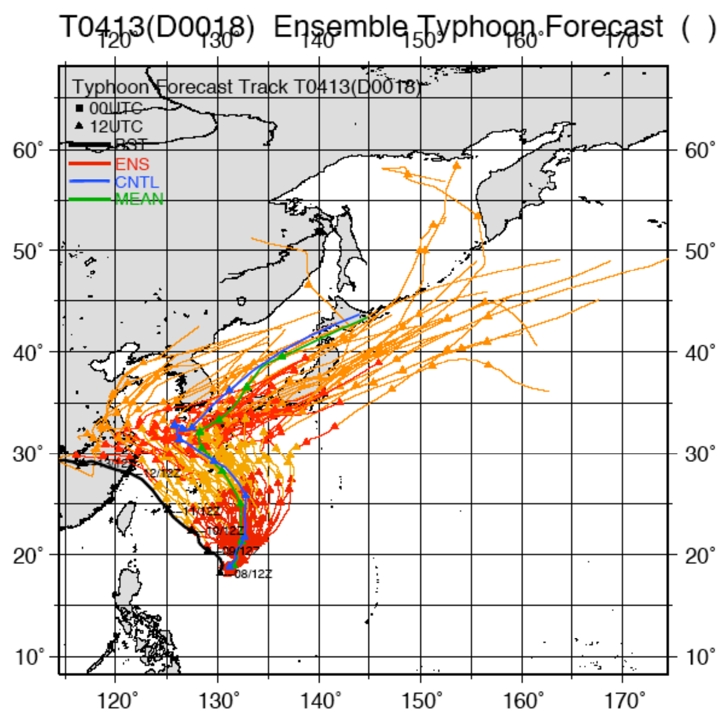
Bias

Verifying against Rawinsondes

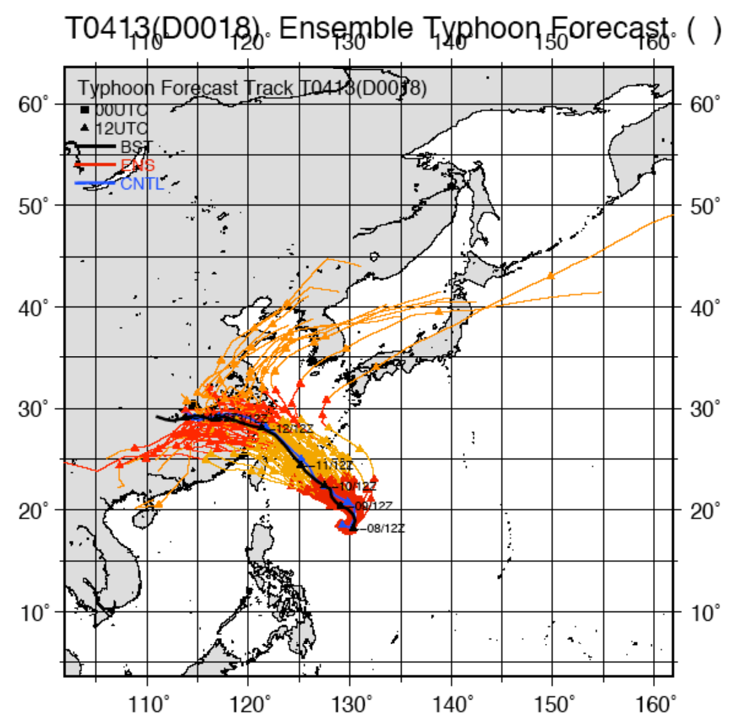
Comparison of 4-D Var and LETKF at JMA

18th typhoon in 2004, IC 12Z 8 August 2004

T. Miyoshi and Y. Sato



operational

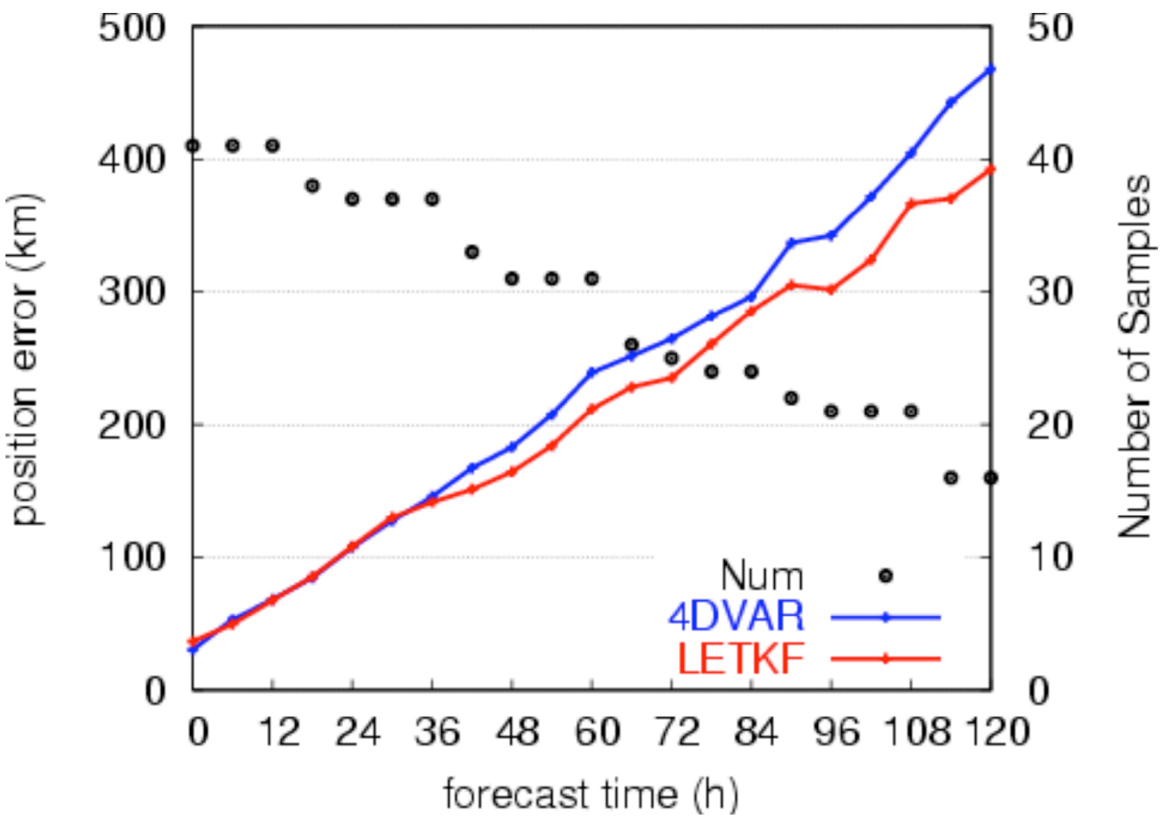


LETKF

Comparison of 4-D Var and LETKF at JMA

RMS error statistics for all typhoons in August 2004

T. Miyoshi and Y. Sato



Operational 4D-Var

LETKF

Summary

- Data assimilation methods have contributed much to the improvements in NWP.
- A toy example is easy to understand, and the equations are the same for a realistic system
- Kalman Filter (too costly) and 4D-Var (complicated) solve the same problem (if model is linear and we use long assimilation windows)
- Ensemble Kalman Filter is feasible and simple
- It is starting to catch up with operational 4D-Var
- EnKF can also estimate observational errors online
- Important problems: estimate and correct model errors & obs. errors, optimal obs. types and locations, tuning additive/multiplicative inflation, parameters estimation,…
  - Tellus: 4D-Var or EnKF? In press
  - Papers posted in “Weather Chaos UMD”, Hunt et al, Szunyogh et al
  - Workshop in Buenos Aires Nov ’08