Introduction to data assimilation and least squares methods

Eugenia Kalnay
and many friends

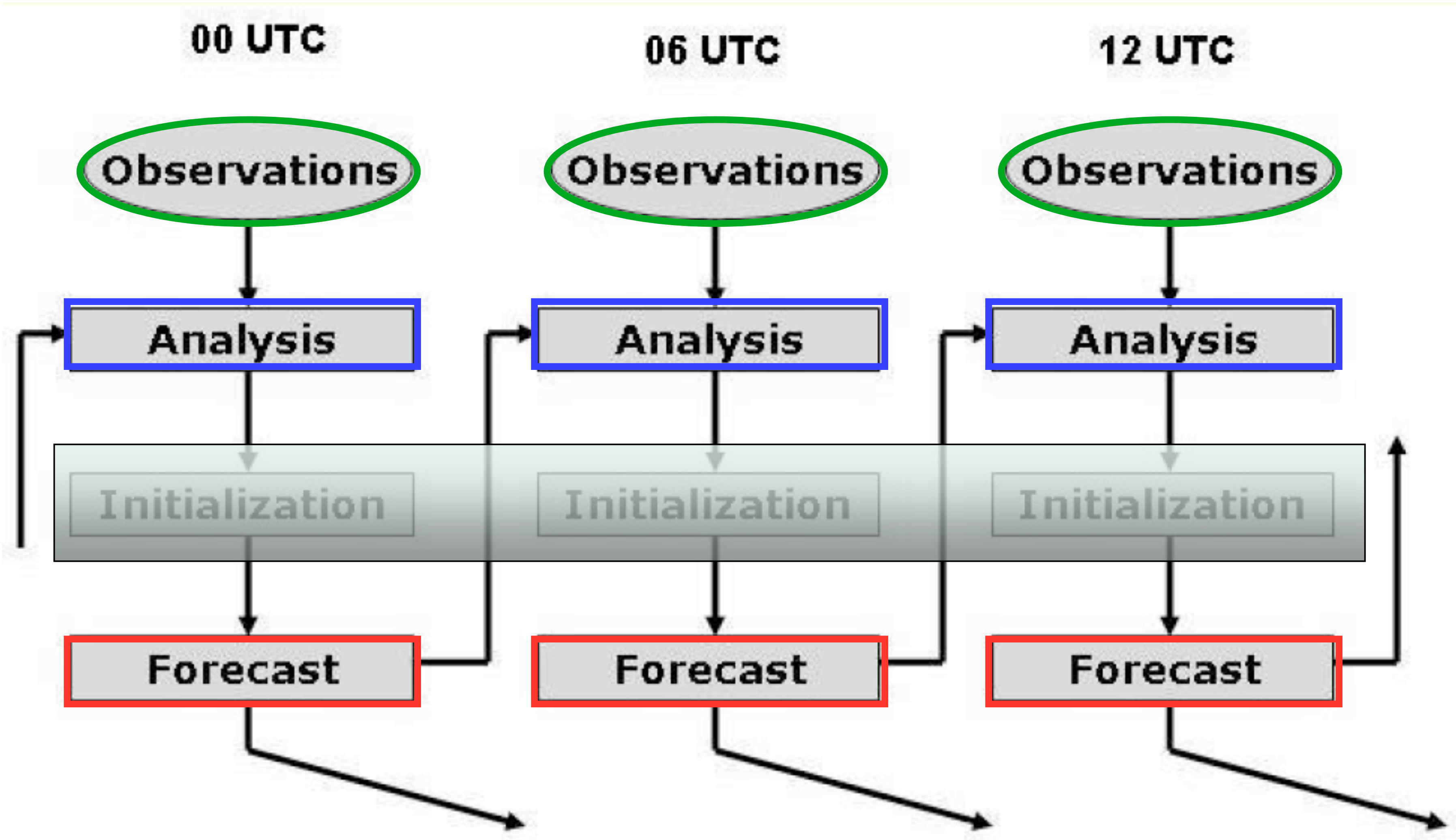
University of Maryland
October 2008 (part 1)

Contents (1)

- Forecasting the weather - we are really getting better!
- Why: Better obs? Better models? Better data assimilation? It's all three together!
- Intro to data assim: a toy scalar example 1, we measure with two thermometers, and we want an accurate temperature.
- Another toy example 2, we measure radiance but we want an accurate temperature: we derive OI/KF, 3D-Var, 4D-Var and EnKF for the toy model.

Contents (2)

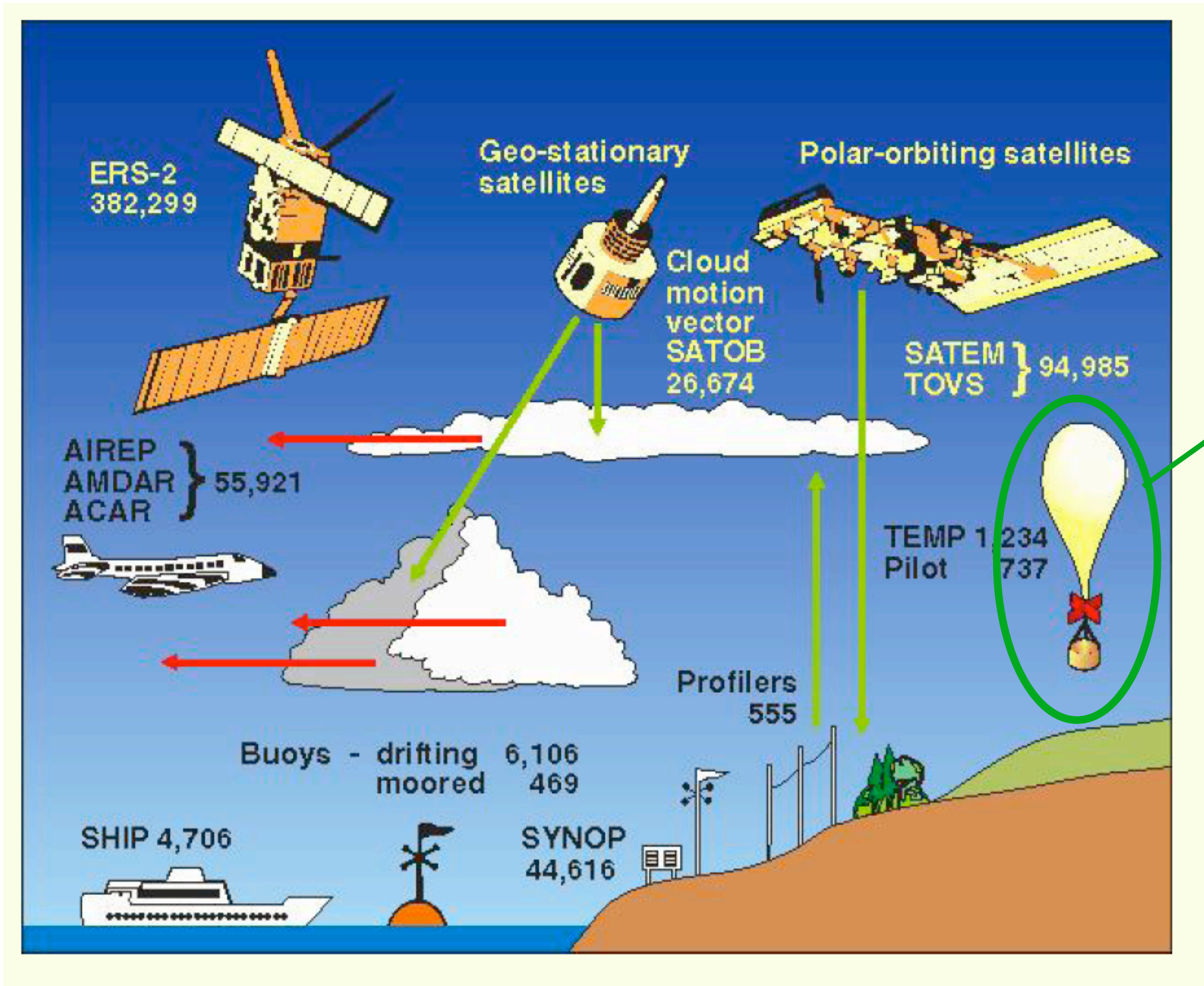
- Review of toy example 1
- Another toy example 2, we measure radiance but we want an accurate temperature:
- We derive OI/KF, 3D-Var, 4D-Var and EnKF for the toy model.
- Comparison of the toy and the real equations
- An example from JMA comparing 4D-Var and LETKF



Typical 6-hour analysis cycle.

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

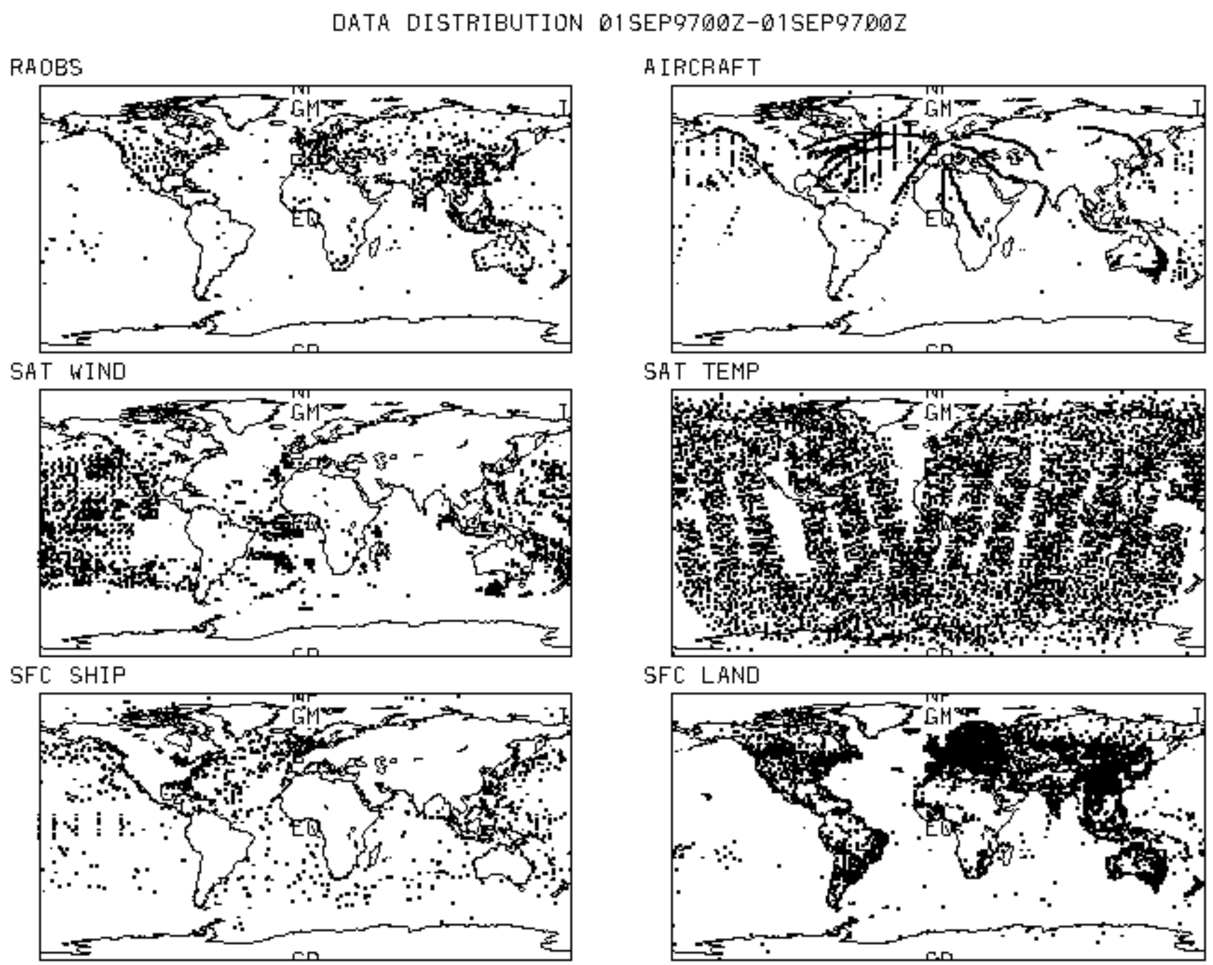
The observing system a few years ago…



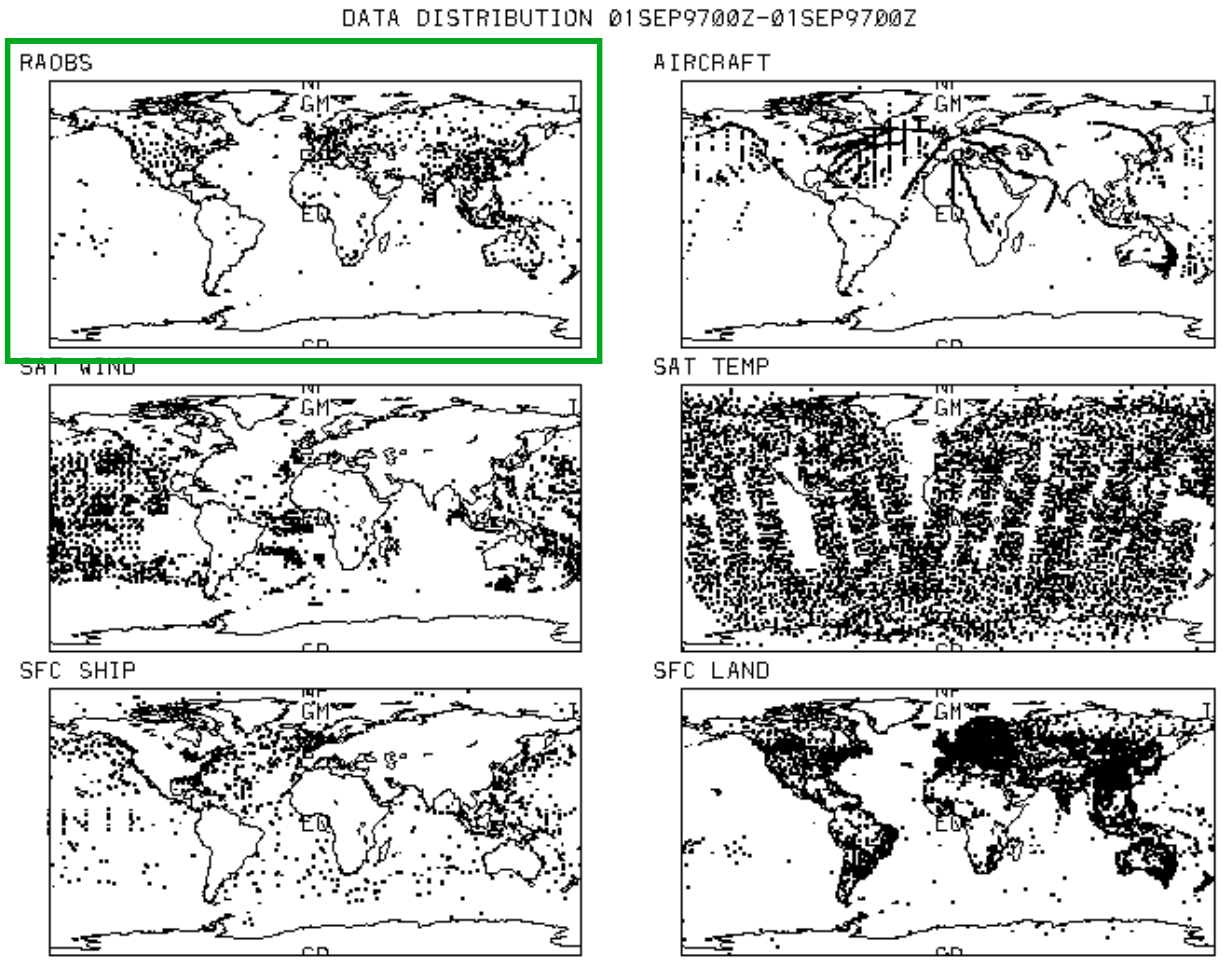
Before 1979 only raobs

Now we have even more satellite data…

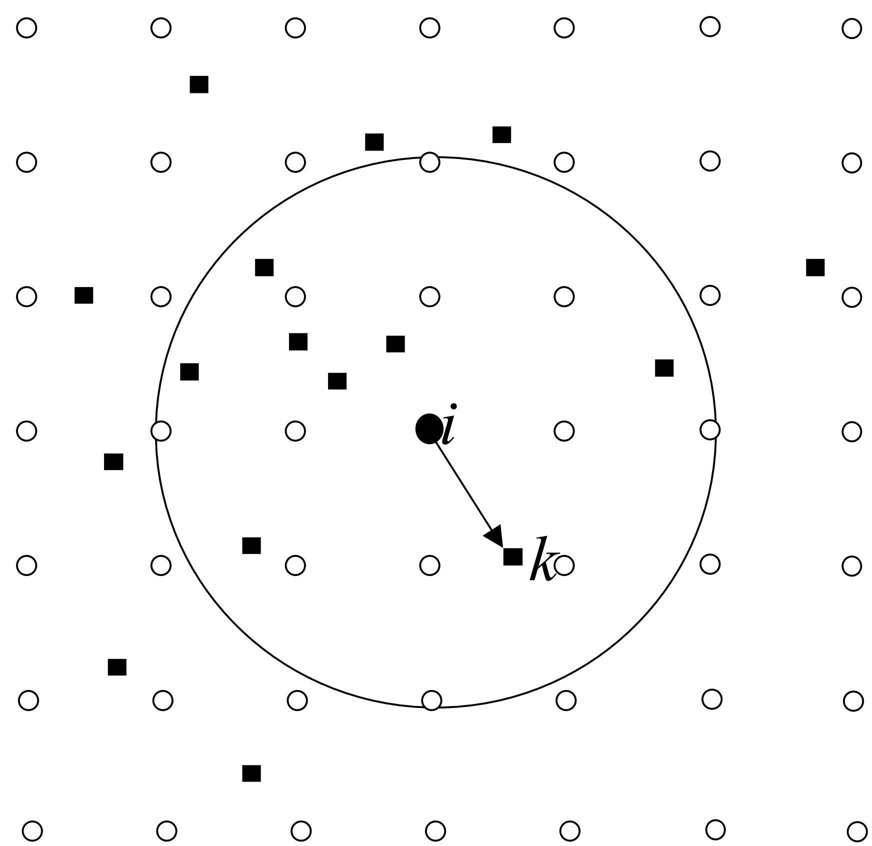
Typical distribution of the observing systems in a 6 hour period:
a real mess: different units, locations, times



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

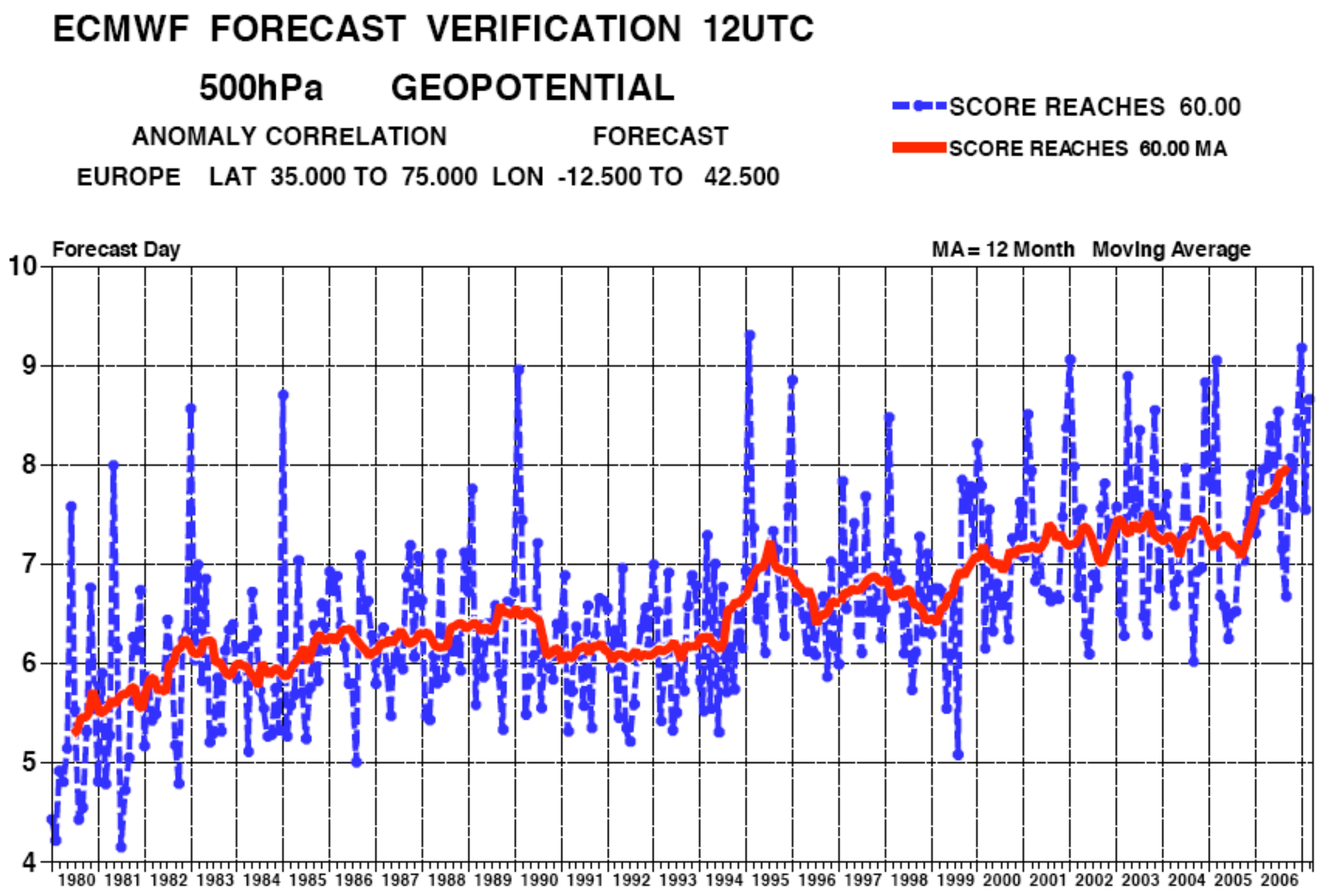


Model grid points (uniformly distributed) and observations (randomly distributed). For the grid point i only observations within a radius of influence may be considered

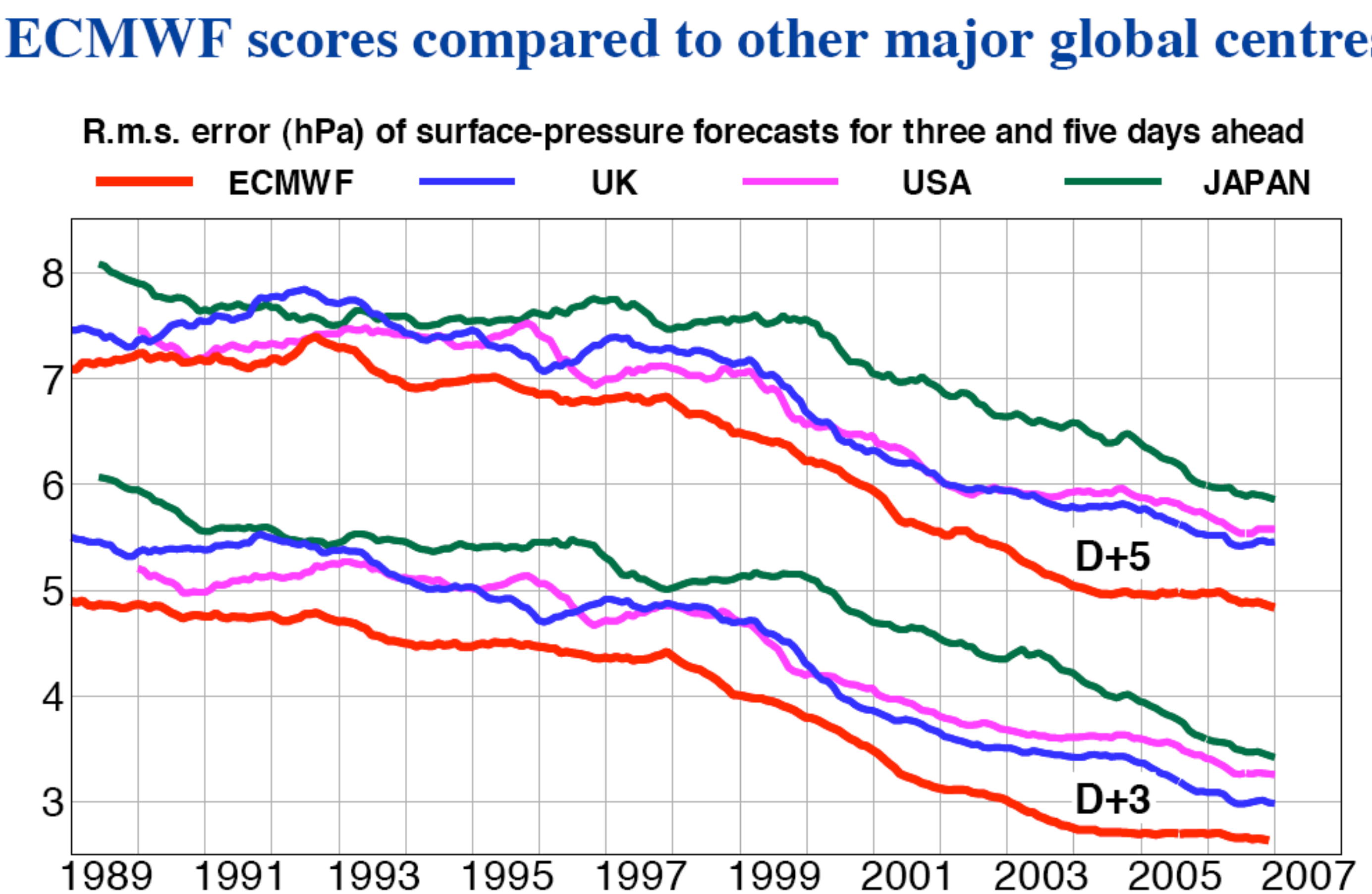


Some statistics of NWP…

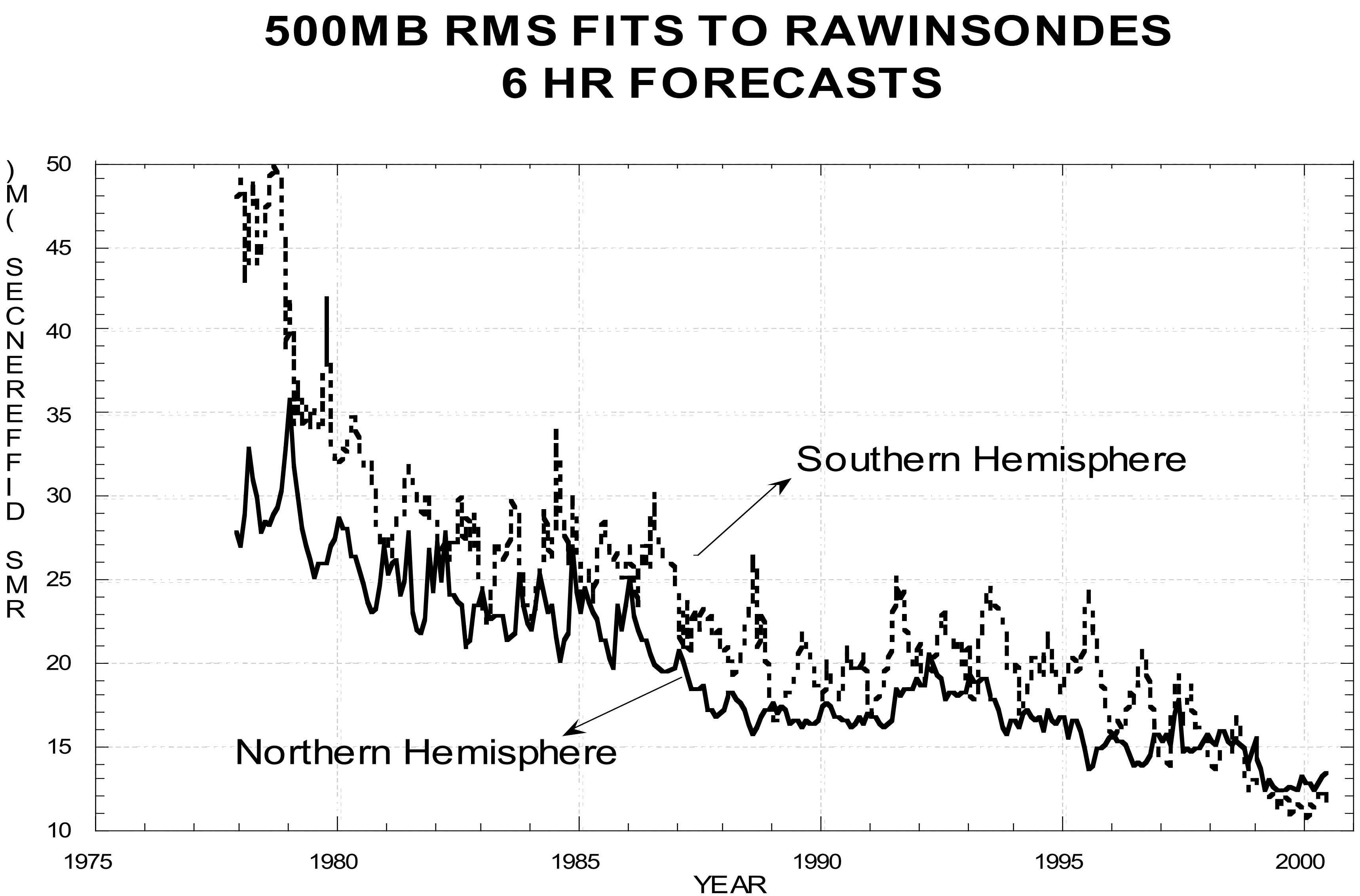
Permanent verifications of the forecast:



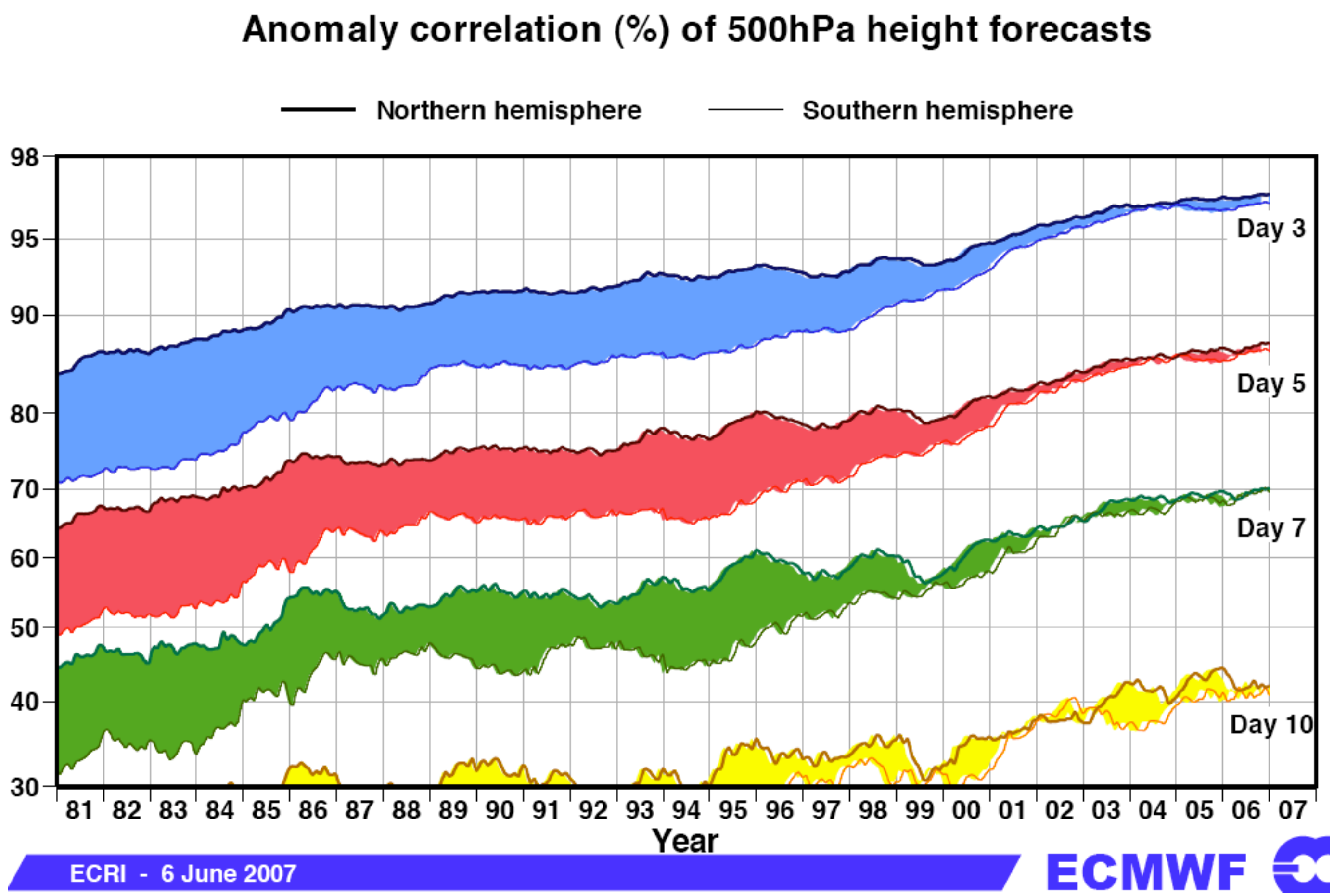
Some comparisons…



We are getting better… (NCEP observational increments)



Comparisons of Northern and Southern Hemispheres



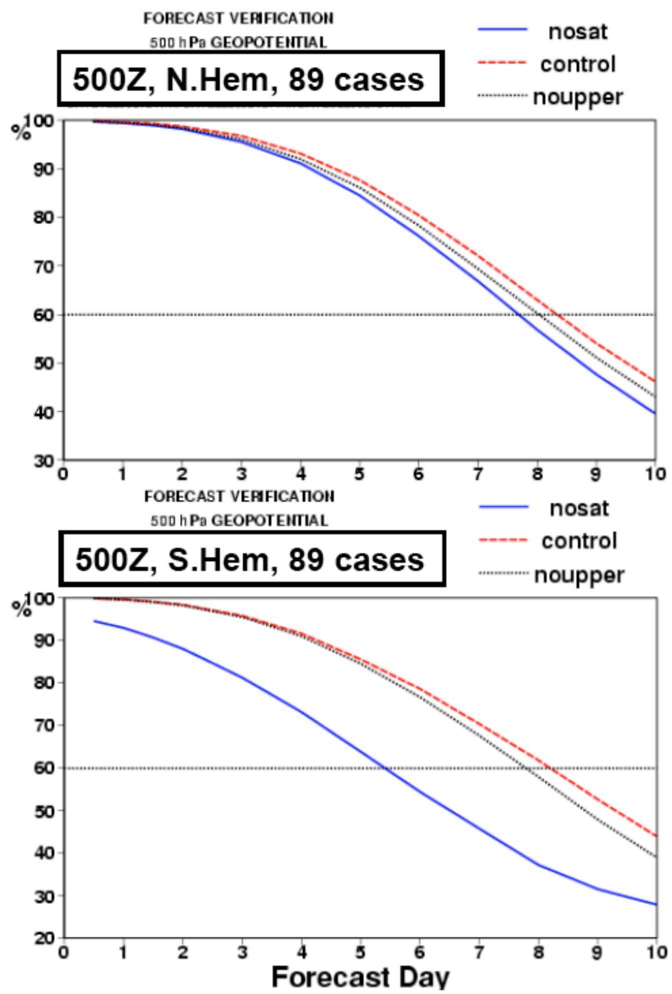
Satellite radiances are essential in the SH

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

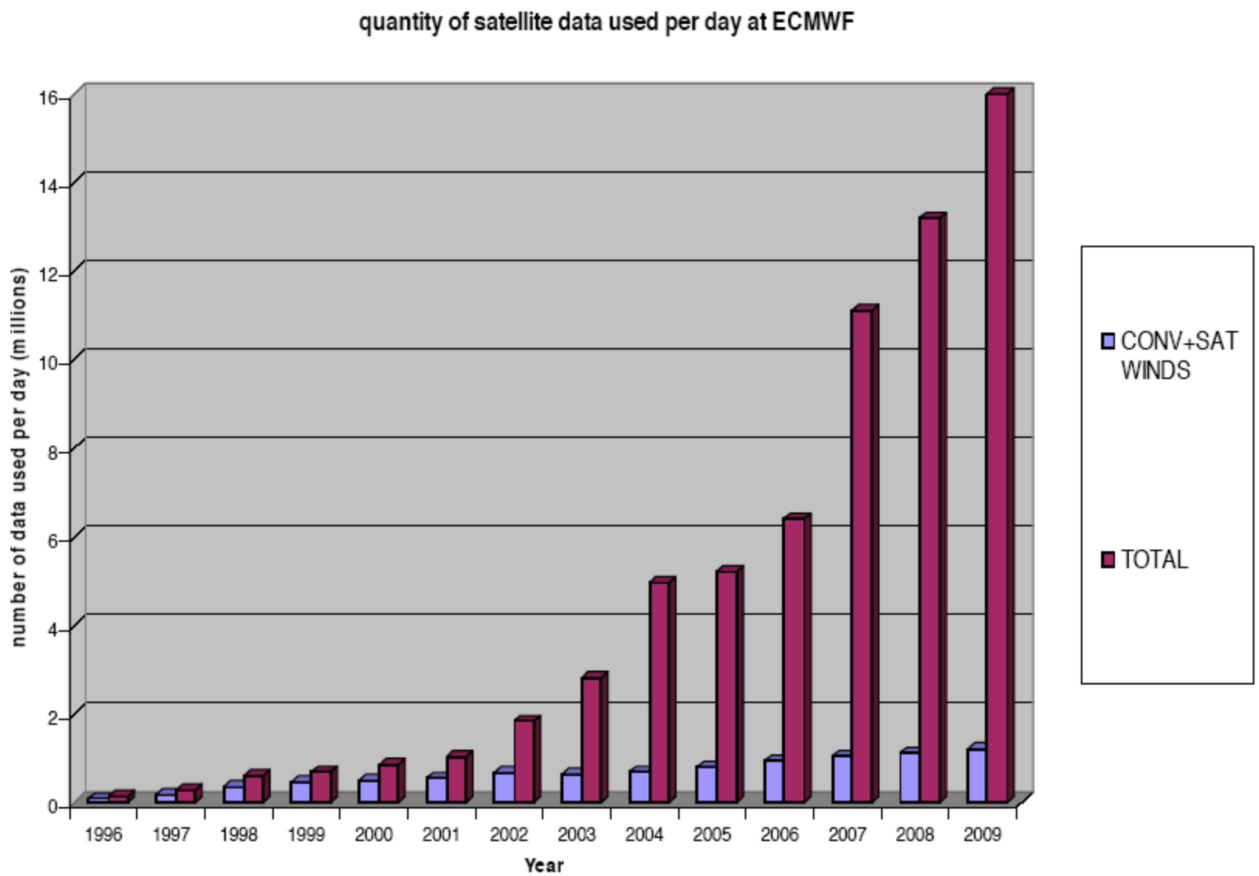
NoSAT= no satellite radiances or winds

Control= like operations

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



More and more satellite radiances…



Intro. to data assimilation: toy example 1

- We want to measure the temperature in this room, and we have two thermometers that measure with errors:

$$T_1 = T_t + \varepsilon_1$$
$$T_2 = T_t + \varepsilon_2$$

- We assume that the errors are unbiased:

$$\overline{\varepsilon_1} = \overline{\varepsilon_2} = 0$$

that we know their variances $\overline{\varepsilon_1^2} = \sigma_1^2 \quad \overline{\varepsilon_2^2} = \sigma_2^2$

and the errors of the two thermometers are uncorrelated: $\overline{\varepsilon_1 \varepsilon_2} = 0$

The question is: how can we estimate the true temperature optimally? We call this optimal estimate the "analysis of the temperature"

Intro. to data assimilation: toy example 1

- We try to estimate the analysis from a linear combination of the observations:

$$T_a = a_1 T_1 + a_2 T_2$$

and assume that the analysis errors are unbiased:

$$\overline{T_a} = \overline{T_t}$$

This implies that $a_1 + a_2 = 1$

Intro. to data assimilation: toy example 1

- We try to estimate the analysis from a linear combination of the observations:

$$T_a = a_1 T_1 + a_2 T_2$$

and assume that the analysis errors are unbiased:

$$\overline{T_a} = \overline{T_t}$$

This implies that $a_1 + a_2 = 1$

T_a will be the *best estimate* of T_t if the coefficients a_1, a_2 are chosen to minimize the mean squared error of T_a:

$$\sigma_a^2 = \overline{(T_a - T_t)^2} = \overline{[a_1(T_1 - T_t) + (1 - a_1)(T_2 - T_t)]^2}$$

Intro. to data assimilation: toy example 1

- Replacing $a_2 = 1 - a_1$

the minimization of σ_a^2 with respect to a_1 gives

$$\sigma_a^2 = \overline{(T_a - T_t)^2} = \overline{[a_1(T_1 - T_t) + (1 - a_1)(T_2 - T_t)]^2}$$

Intro. to data assimilation: toy example 1

- Replacing $a_2 = 1 - a_1$

the minimization of σ_a^2 with respect to a_1 gives

$$\sigma_a^2 = \overline{(T_a - T_t)^2} = \overline{[a_1(T_1 - T_t) + (1 - a_1)(T_2 - T_t)]^2}$$

$$\frac{\partial \sigma_a^2}{\partial a_1} = 0 \Longrightarrow \qquad a_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad a_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

or

$$a_1 = \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2} \quad a_2 = \frac{1/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

The first formula says that the weight of obs 1 is given by the variance of obs 2 divided by the total error.

The second formula says that the weights of the observations are proportional to the "precision" or accuracy of the measurements (defined as the inverse of the variances of the observational errors).

Intro. to data assim: toy example 1 summary

Two measurements and an optimal linear combination (analysis):

$$T_a = a_1 T_1 + a_2 T_2$$

Optimal coefficients (min $\sigma_a{}^2$)

$$a_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad a_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

or

$$a_1 = \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2} \quad a_2 = \frac{1/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

Replacing, we get $\sigma_a^2 = \dfrac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} < \sigma_1^2, \sigma_2^2$ or $\dfrac{1}{\sigma_a{}^2} = \dfrac{1}{\sigma_1{}^2} + \dfrac{1}{\sigma_2{}^2}$

Intro. to data assim: toy example 1 summary

Two measurements and an optimal linear combination (analysis):

$$T_a = a_1 T_1 + a_2 T_2$$

Optimal coefficients (min $\sigma_a^{\ 2}$)

$$a_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad a_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

or

$$a_1 = \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2} \quad a_2 = \frac{1/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

Replacing, we get $\sigma_a^2 = \dfrac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$, or $\dfrac{1}{\sigma_a^{\ 2}} = \dfrac{1}{\sigma_1^{\ 2}} + \dfrac{1}{\sigma_2^{\ 2}}$

Intro. to data assim: toy example 1 summary

Two measurements and an optimal linear combination (analysis):

$$T_a = a_1 T_1 + a_2 T_2$$

Optimal coefficients (min $\sigma_a{}^2$)

$$a_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad a_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

or

$$a_1 = \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2} \quad a_2 = \frac{1/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

Replacing, we get $\sigma_a^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$ or $\frac{1}{\sigma_a{}^2} = \frac{1}{\sigma_1{}^2} + \frac{1}{\sigma_2{}^2}$

Now assume that $T_1 = T_b$ (forecast) and $T_2 = T_o$ (observation). Then

$$T_a = a_1 T_b + a_2 T_o = T_b + a_2 (T_o - T_b)$$

Intro. to data assim: toy example 1 summary

Two measurements and an optimal linear combination (analysis):

$$T_a = a_1 T_1 + a_2 T_2$$

Optimal coefficients (min $\sigma_a^{\,2}$)

$$a_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad a_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \quad \text{or} \quad a_1 = \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2} \quad a_2 = \frac{1/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

Replacing, we get $\sigma_a^2 = \dfrac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$ or $\dfrac{1}{\sigma_a^{\,2}} = \dfrac{1}{\sigma_1^{\,2}} + \dfrac{1}{\sigma_2^{\,2}}$

Now assume that $T_1 = T_b$ (forecast) and $T_2 = T_o$ (observation). Then

$$T_a = a_1 T_b + a_2 T_o = T_b + a_2 (T_o - T_b)$$

or

$$T_a = T_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}(T_o - T_b) = T_b + w(T_o - T_b)$$

This is the form that is always used in analyses…

Intro. to data assim: toy example 1 summary

A forecast and an observation optimally combined (analysis):

$$T_a = T_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}(T_o - T_b) \quad \text{with} \quad \frac{1}{\sigma_a^2} = \frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}$$

If *the statistics of the errors are exact, and if the coefficients are optimal, then the "precision" of the analysis (defined as the inverse of the variance) is the sum of the precisions of the measurements.*

Now we are going to see a second toy example of data assimilation including remote sensing.

The importance of these toy examples is that the equations are identical to those obtained with big models and many obs.

Intro. to remote sensing and data assimilation: toy example 2

- Assume we have an object, a stone in space
- We want to estimate its temperature *T* (°K) accurately but we measure the radiance *y* (W/m²) that it emits. We have an *obs. model, e.g.*:

$$y = h(T) \sim \sigma T^4$$

Intro. to remote sensing and data assimilation: toy example 2

- Assume we have an object, a stone in space
- We want to estimate its temperature *T* (°K) accurately but we measure the radiance *y* (W/m²) that it emits. We have an *obs. model,* *e.g.*: $y = h(T) \sim \sigma T^4$

- We also have a *forecast model* for the temperature

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

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

Intro. to remote sensing and data assimilation: toy example 2

- Assume we have an object, a stone in space
- We want to estimate its temperature *T* (°K) accurately but we measure the radiance *y* (W/m²) that it emits. We have an *obs. model,* e.g.: $y = h(T) \sim \sigma T^4$
- 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: toy example 2

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

- 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 (OI/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:

$$y_o - h(T_b)$$

The final formula used has the same form as $T_a = T_b + \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}(T_o - T_b)$

$$T_a = T_b + w(y_o - h(T_b))$$

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

Where does *H* come from?

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 Gaussian

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., $h(T) = \sigma T^4$

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

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

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

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

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 $\sigma_o^2 >> \sigma_b^2 H^2$ $\quad T_a \approx T_b \approx T_t$

If $\sigma_o^2 << \sigma_b^2 H^2$ $\quad T_a \approx T_b + \frac{1}{H}\left[h(T_t) - h(T_b)\right] \approx T_t$

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

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

Subtracting T_t from both sides we obtain

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

Squaring the analysis error and averaging over many cases, we obtain

$$\sigma_a^2 = (1 - wH)\sigma_b^2$$

which can also be written as $$\frac{1}{\sigma_a^2} = \left(\frac{1}{\sigma_b^2} + \frac{H^2}{\sigma_o^2}\right)$$

Toy temperature data assimilation, measure radiance

Summary for OI/KF (sequential):

$$T_a = T_b + w(y_o - h(T_b))$$ analysis

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

The analysis error is computed from

$$\sigma_a^2 = (1 - wH)\sigma_b^2$$

which can also be written as

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

analysis precision=
forecast precision + observation precision

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:

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

Summary part 1

- Data assimilation methods have contributed much to the improvements in NWP.
- A toy example is easy to understand, and the equations are the same as in a realistic huge system
- Observation operator: model variables => observed variables
- We assume no bias, no error correlation
- Analysis = forecast +optimal weight x (innovation)
- Optimal weight = forecast error variance/total error variance
- Precision = 1/error variance
- Analysis precision=forecast precis. + obs. precis.