# Introduction to data assimilation and least squares methods

Eugenia Kalnay
and many friends

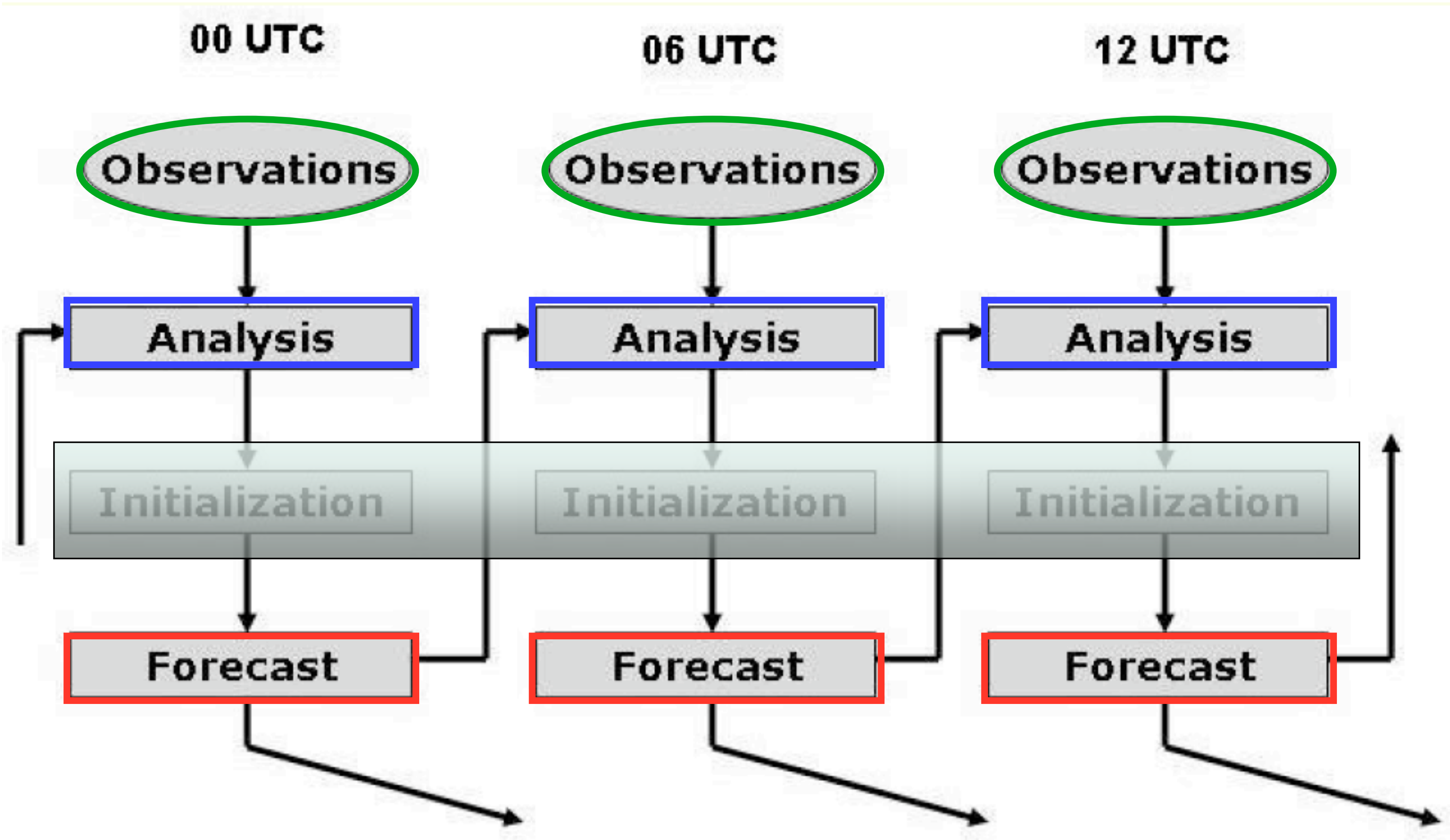
University of Maryland
October 2008 (Part 2)

# Contents (part 1)

- Forecasting the weather - we are really getting better!
- Why: Better obs? Better models? Better data assimilation?
- 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 will derive OI/KF, 3D-Var, 4D-Var and EnKF for the toy model.

# Contents (part 2)

- Review of toy example 1
- Another toy example 2, we measure radiance but we want an accurate temperature:
- We derive Optimal Interpolation/Kalman Filter (sequential algorithms) for the toy model.
- 3D-Var, 4D-Var (variational algorithms) and EnKF for the toy model.
- We compare the toy equations 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")

# 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 $T$ ($^oK$) 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: 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

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

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

# 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 (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 is very similar to that in toy model 1:

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

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

Recall that $T_a = T_b + w(y_o - h(T_b)) = T_b + w(\varepsilon_o - H\varepsilon_b)$

So that, subtracting the truth, $\varepsilon_a = \varepsilon_b + w(\varepsilon_o - H\varepsilon_b)$

# Toy temperature data assimilation, measure radiance

Summary for Optimal Interpolation/Kalman Filter (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 obtained from squaring $\varepsilon_a = \varepsilon_b + w\left[\varepsilon_o - H\varepsilon_b\right]$

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

It 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 $J$:

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

# 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 $J$:

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

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 $J$

# 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:*

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

The cost function 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:*

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

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

# 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 $J$:

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

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

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

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

Minimizing the cost function maximizes the likelihood of the estimate of 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 $J$.

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

To find the minimum we use an incremental approach: find $T_a - T_b$:

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

So that from $\partial J / \partial(T_a - T_b) = 0$ 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 $J$.

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

To find the minimum we use an incremental approach: find $T_a - T_b$:

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

So that from $\partial J / \partial (T_a - T_b) = 0$ 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: $y_o - h(T_b)$

3D-Var: $T_a$ minimizes the distance to both the background and the observations

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

3D-Var solution

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

# Toy temperature data assimilation, variational approach

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

Innovation: $y_o - h(T_b)$

3D-Var: $T_a$ minimizes the distance to both the background and the observations

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

3D-Var solution $T_a = T_b + w(y_o - h(T_b))$ with $w_{3D-Var} = \left(\sigma_b^{-2} + H\sigma_o^{-2}H\right)^{-1} H\sigma_o^{-2} = \sigma_a^2 H\sigma_o^{-2}$

This variational solution looks different but is the **same** as the one obtained before with Kalman filter (a sequential approach, like Optimal Interpolation, Lorenc 86)):

KF/OI solution $T_a = T_b + w(y_o - h(T_b))$ with $w_{OI} = \sigma_b^2 H(\sigma_o^2 + \sigma_b^2 H^2)^{-1}$

# Toy temperature data assimilation, variational approach

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

Innovation: $y_o - h(T_b)$

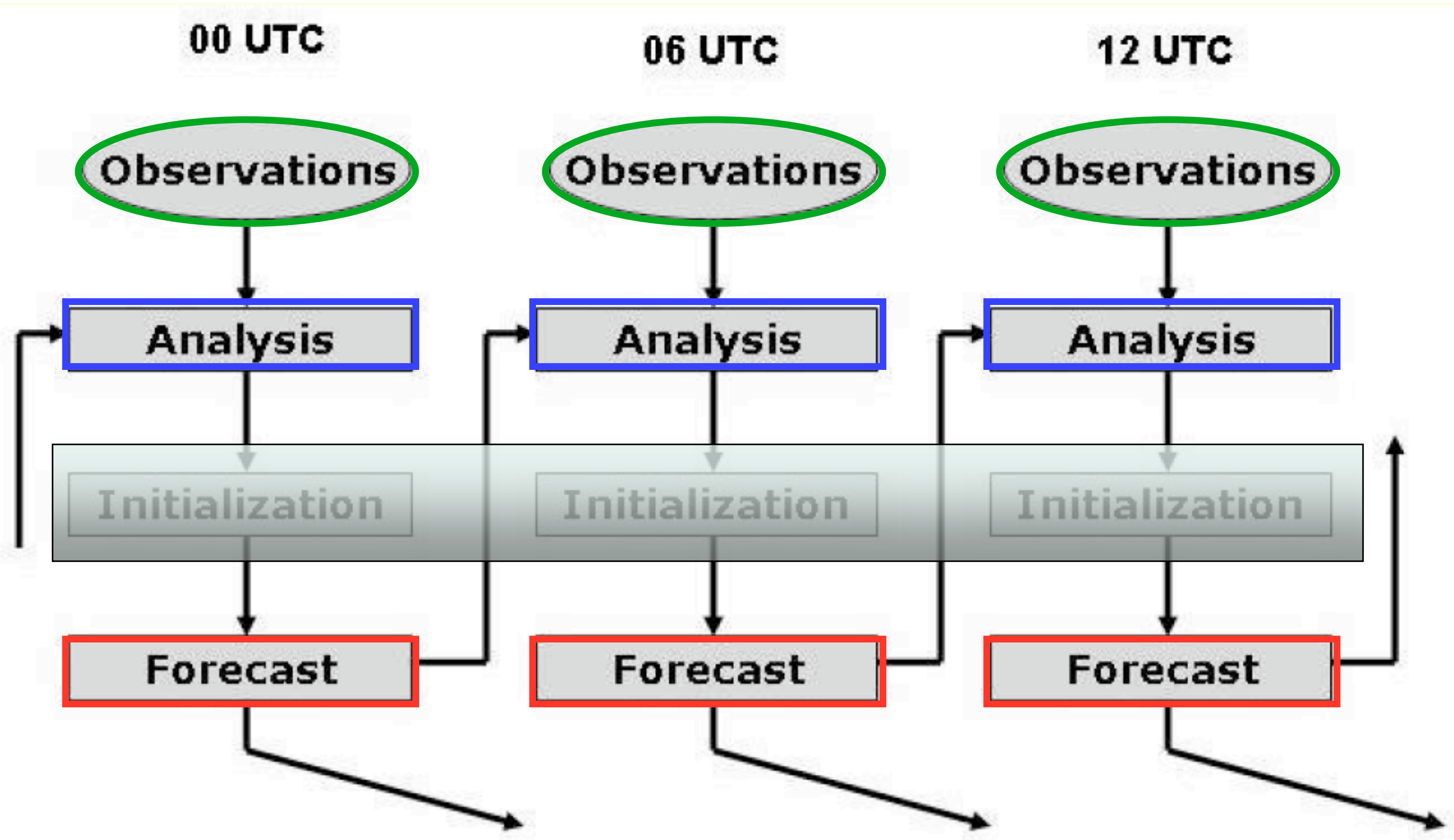
3D-Var: $T_a$ minimizes the distance to both the background and the observations $2J_{\min} = \frac{(T_a - T_b)^2}{\sigma_b^2} + \frac{(h(T_a) - y_o)^2}{\sigma_o^2}$

3D-Var solution $T_a = T_b + w(y_o - h(T_b))$ with $w_{3D-Var} = \left(\sigma_b^{-2} + H\sigma_o^{-2}H\right)^{-1} H\sigma_o^{-2} = \sigma_a^2 H\sigma_o^{-2}$

This variational solution is the **<u>same</u>** as the one obtained before with Kalman filter (a sequential approach, like Optimal Interpolation, Lorenc 86)):

KF/OI solution $T_a = T_b + w(y_o - h(T_b))$ with $w_{OI} = \sigma_b^2 H(\sigma_o^2 + \sigma_b^2 H^2)^{-1}$

Show that the 3d-Var and the OI/KF weights are the same: both methods find the same optimal solution!




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

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

## Summary of toy Analysis Cycle (for a scalar)

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

Interpretation…

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

## Summary of toy Analysis Cycle (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 Analysis Cycle (for a scalar)

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

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

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

"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 Analysis Cycle (for a scalar)

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

"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."

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

Summary of toy Analysis Cycle (for a scalar)

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 $10^7$-$10^9$!

# 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 $10^7$-$10^9$!

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

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

and their error variances by error covariances:

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

# 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 $10^7$-$10^9$!

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

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

and their error variances by error covariances:

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

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

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

"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 and its adjoint to forecast **B**"

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

# 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 <u>horrendously</u> expensive that it makes Kalman Filter <u>completely unfeasible</u>".

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

# 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"

## 3D-Var

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

Distance to forecast

Distance to observations

at the analysis time

## 4D-Var

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

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)

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

And the new ensemble perturbations are given by (transform)

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

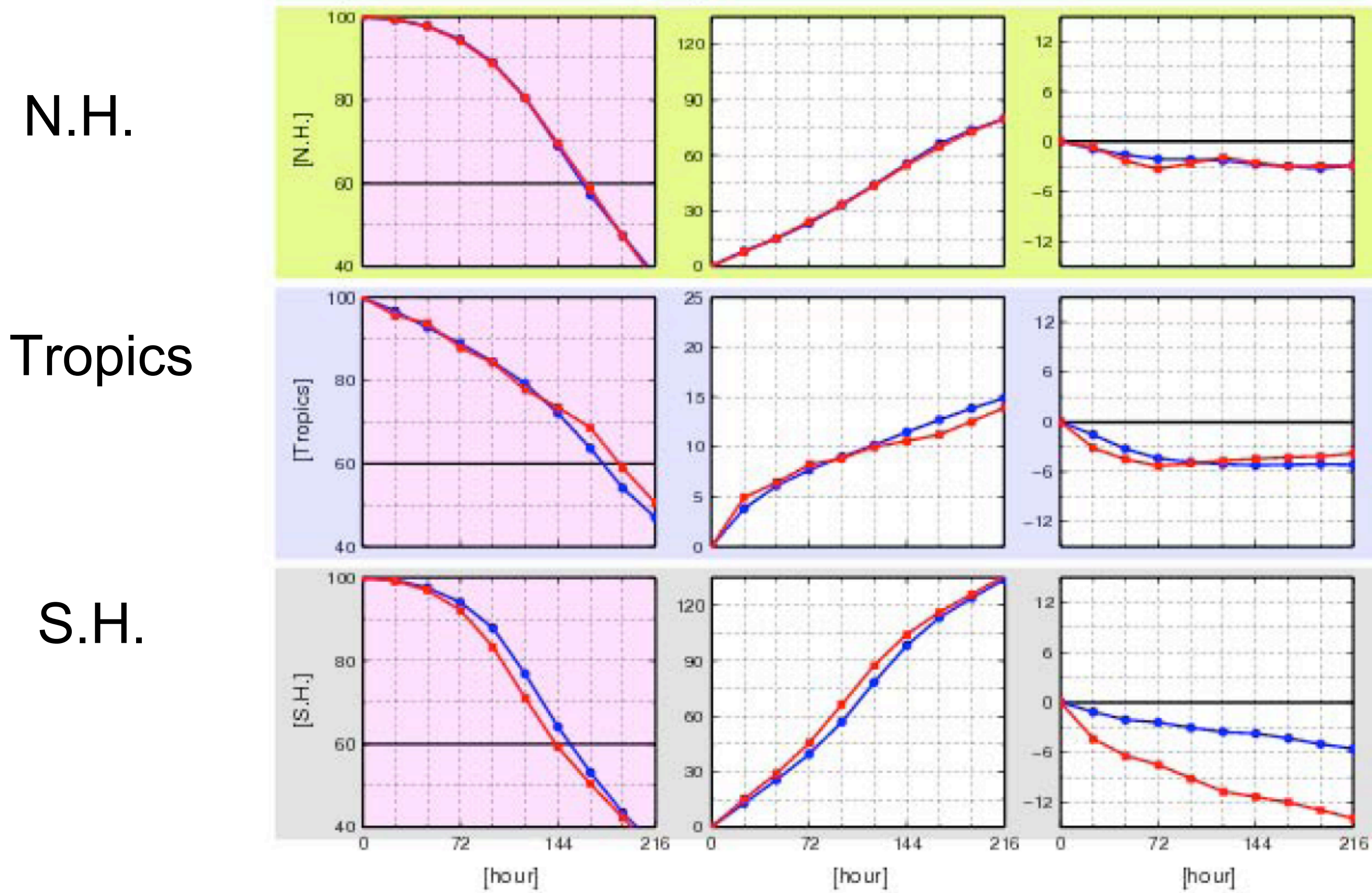
# Comparisons of 4-D Var and LETKF at JMA

T. Miyoshi and Y. Sato

- 4D-Var and EnKF are the two advanced, feasible methods
- http://4dvarenkf.cima.fcen.uba.ar/ Workshop in Buenos Aires
- At JMA, Takemasa Miyoshi has performed 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




RMS error against analysis

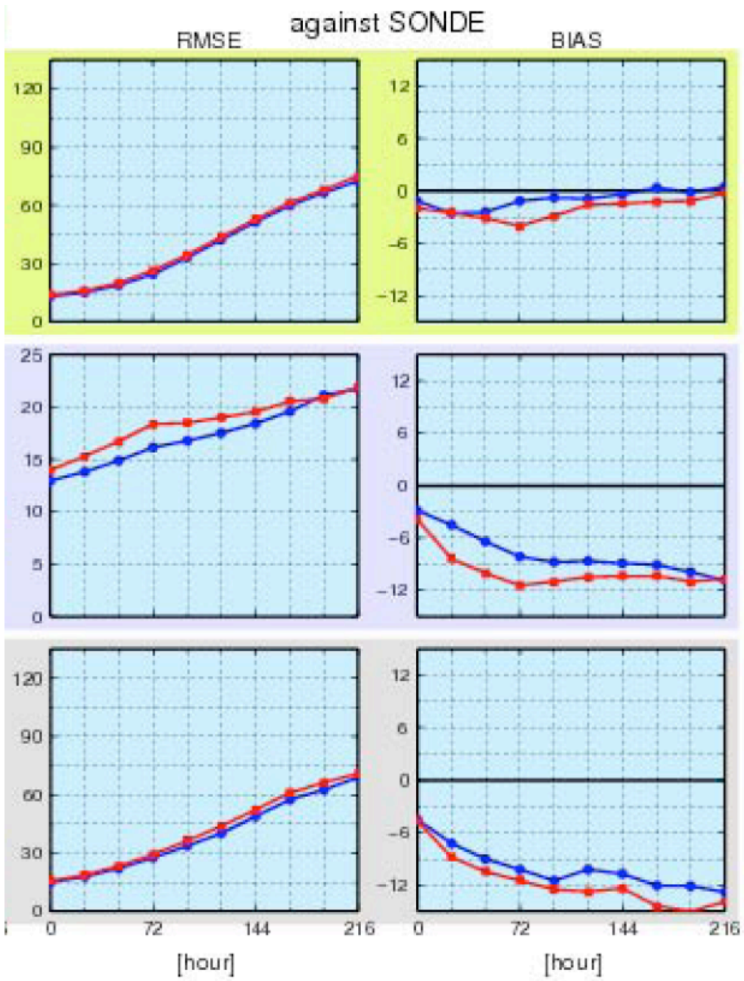
# Comparison of LETKF and 4D-Var at JMA

T. Miyoshi and Y. Sato

N.H.

Tropics

S.H.




Verifying against Rawinsondes

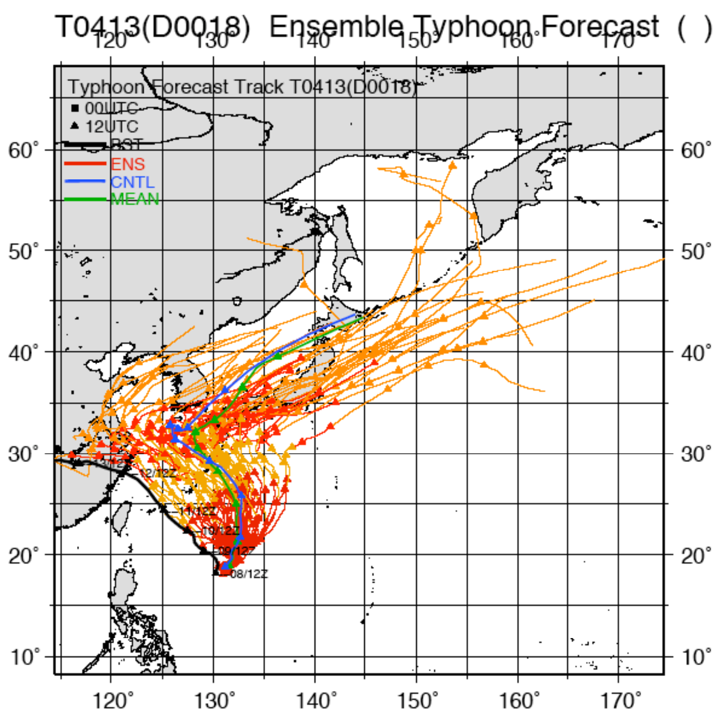
RMS error

Bias

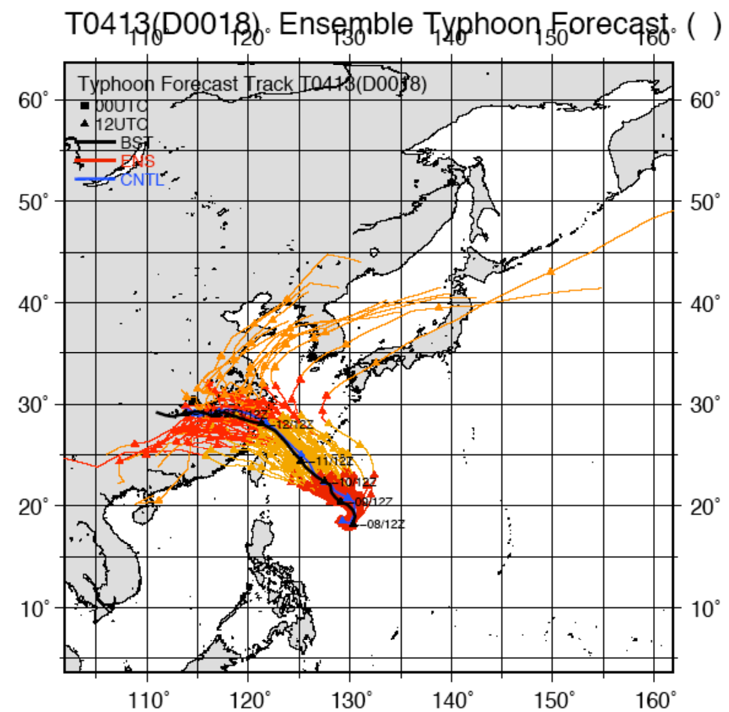
# Comparison of 4-D Var and LETKF at JMA

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

T. Miyoshi and Y. Sato




operational

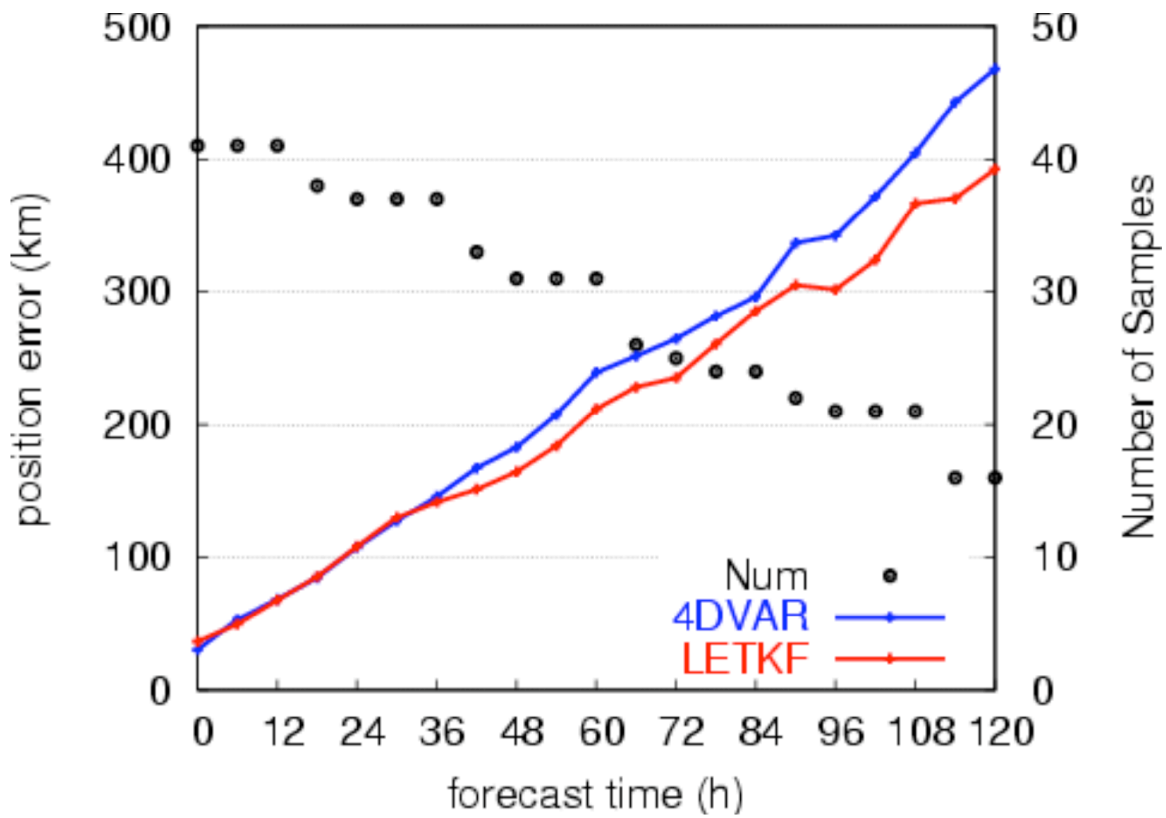



LETKF

## 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

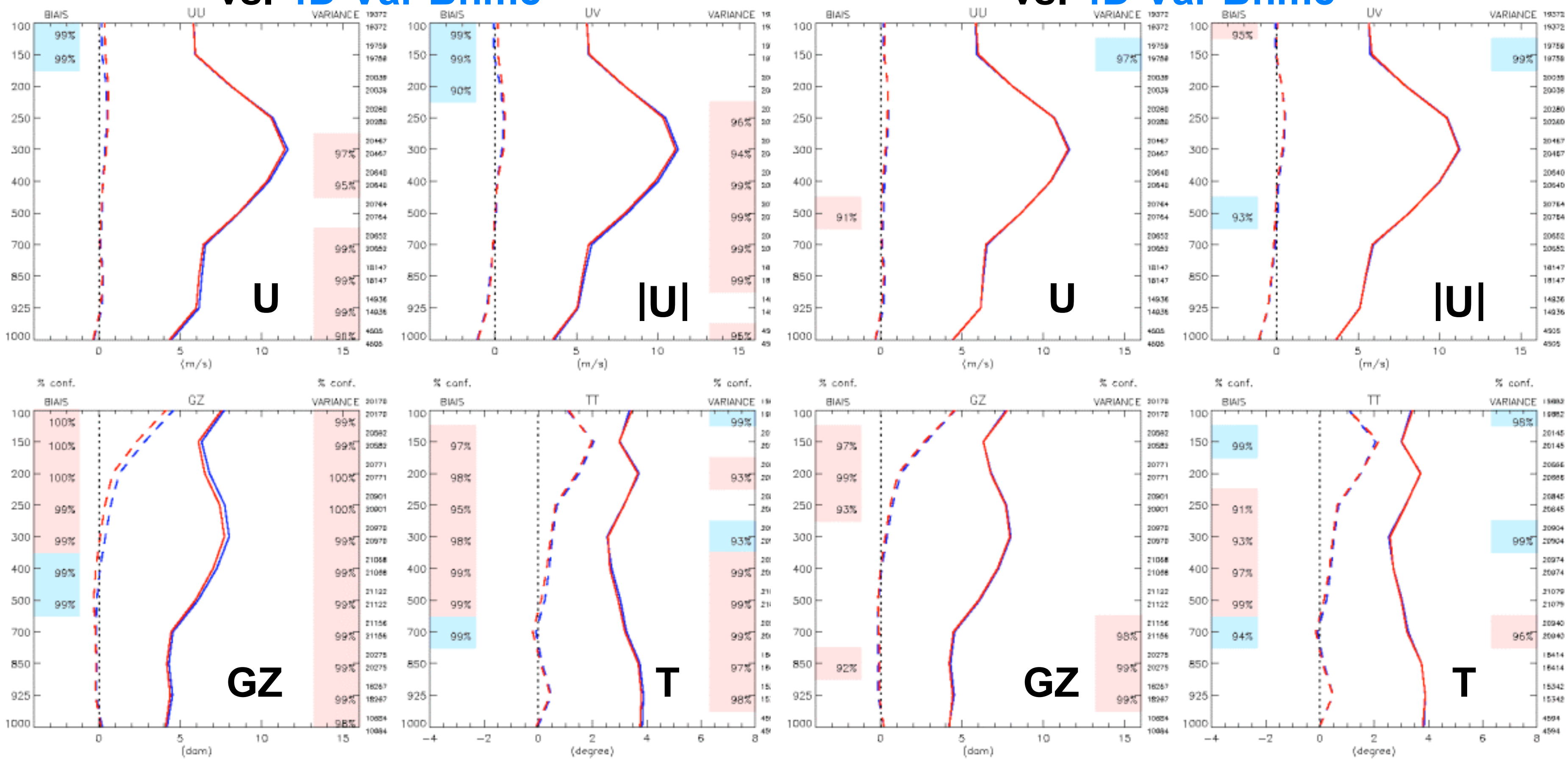
LETKF

# Buehner et al., 2008: Forecast Results – 120h NH

**EnKF mean analysis vs. 4D-Var Bnmc**

**4D-Var Benkf vs. 4D-Var Bnmc**

**HYBRID**

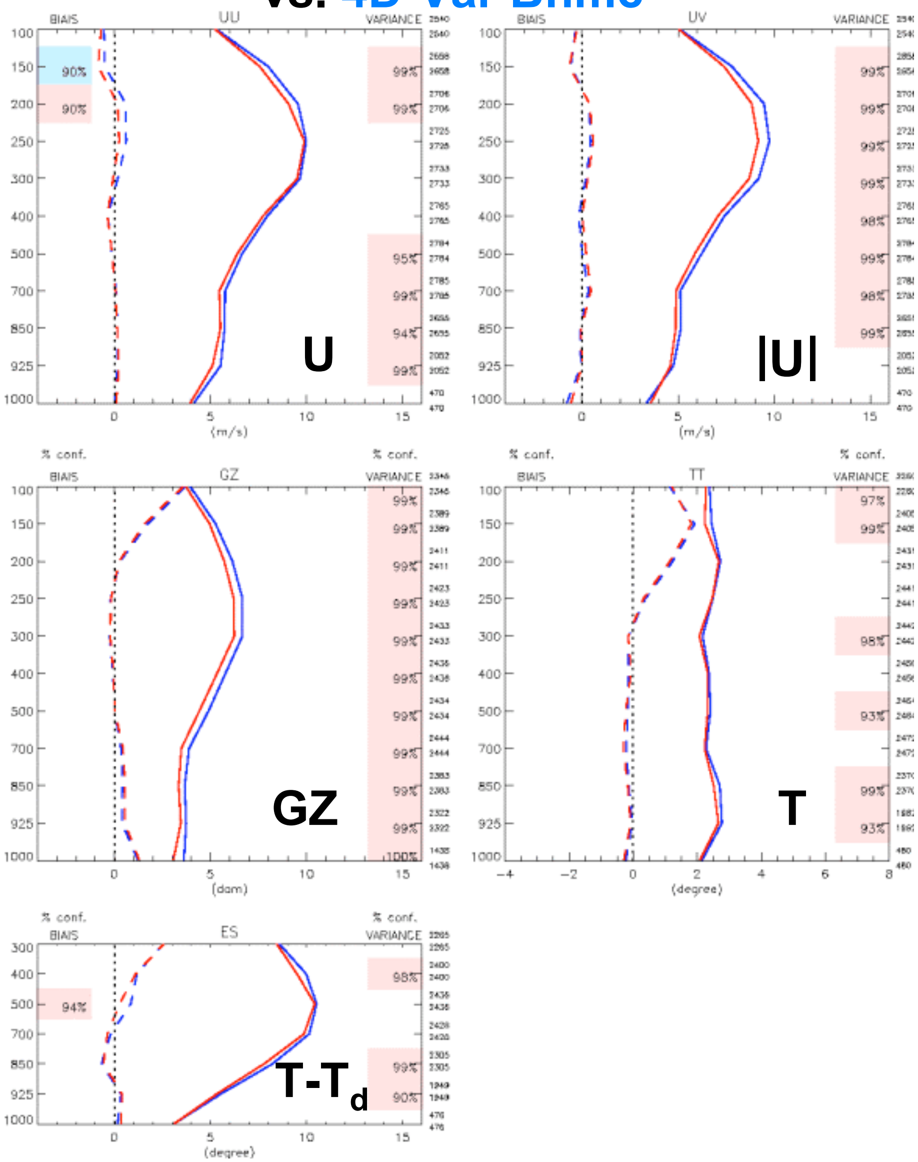




**stddev & bias relative to radiosondes**

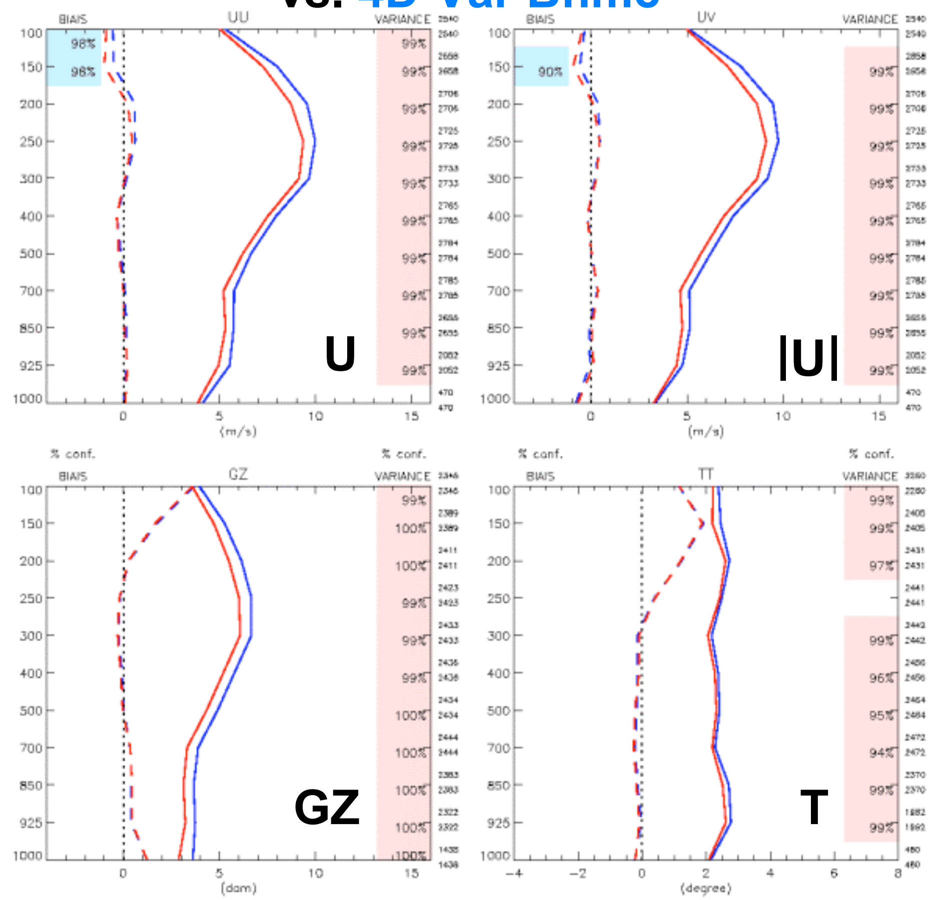
# Buehner et al., 2008: Forecast Results – 120h SH

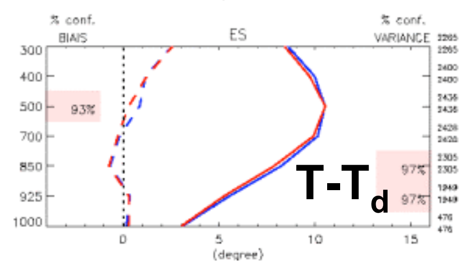
**EnKF mean analysis vs. 4D-Var Bnmc**




**4D-Var Benkf vs. 4D-Var Bnmc**

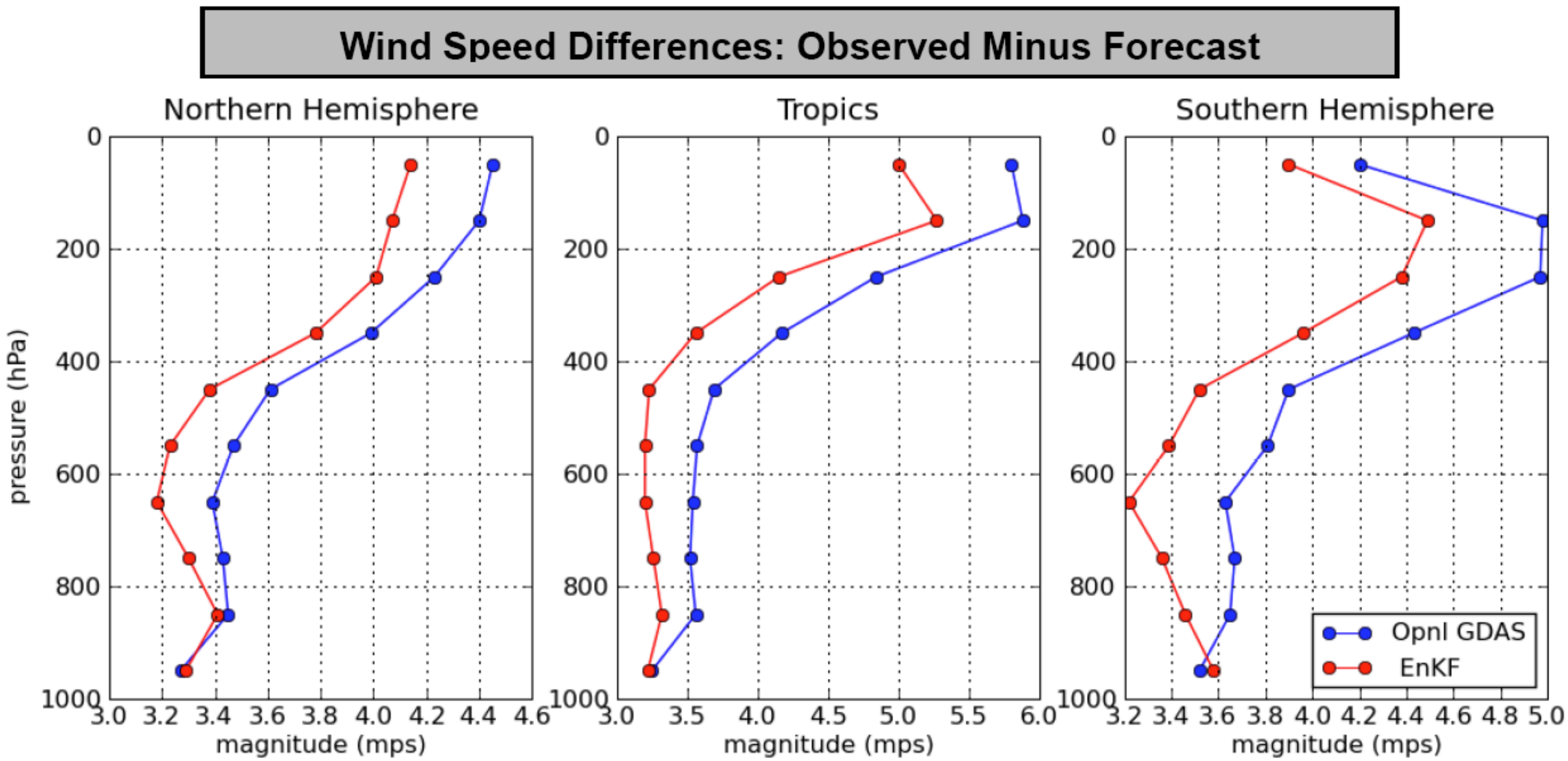
**HYBRID**




**stddev & bias relative to radiosondes**

# Whitaker: Comparison of T190, 64 members EnKF with T382 operational GSI, same observations




Vertical profiles of the RMS difference between six hour forecasts and in-situ observations for the period 2007120700 – 2008010718. Observations are aggregated in 100 hPa layers. The red curve is for the ensemble mean of the experimental 64-member T190 EnKF system, and the blue curve is for the T382 GSI-based GDAS system operational in December 2007.

# 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? Tellus 2007
  - Papers posted in “Weather Chaos UMD”
  - Workshop in Buenos Aires Nov ’08: http://4dvarenkf.cima.fcen.uba.ar/