WRF-Var Overview

Hans Huang

Acknowledge:

NCAR/MMM Division Staff

USWRP, NSF-OPP, NCAR Data Assimilation Initiative

US Air Force Weather Agency, Korean Meteorological Administration

Taiwanese Central Weather Bureau, Civil Aeronautics Administration

AirDat

Outline of Talk

1) Introduction to data assimilation.
2) Basics of modern data assimilation.
3) Demonstration with a simple system.
4) WRF-Var.
5) Important issues.

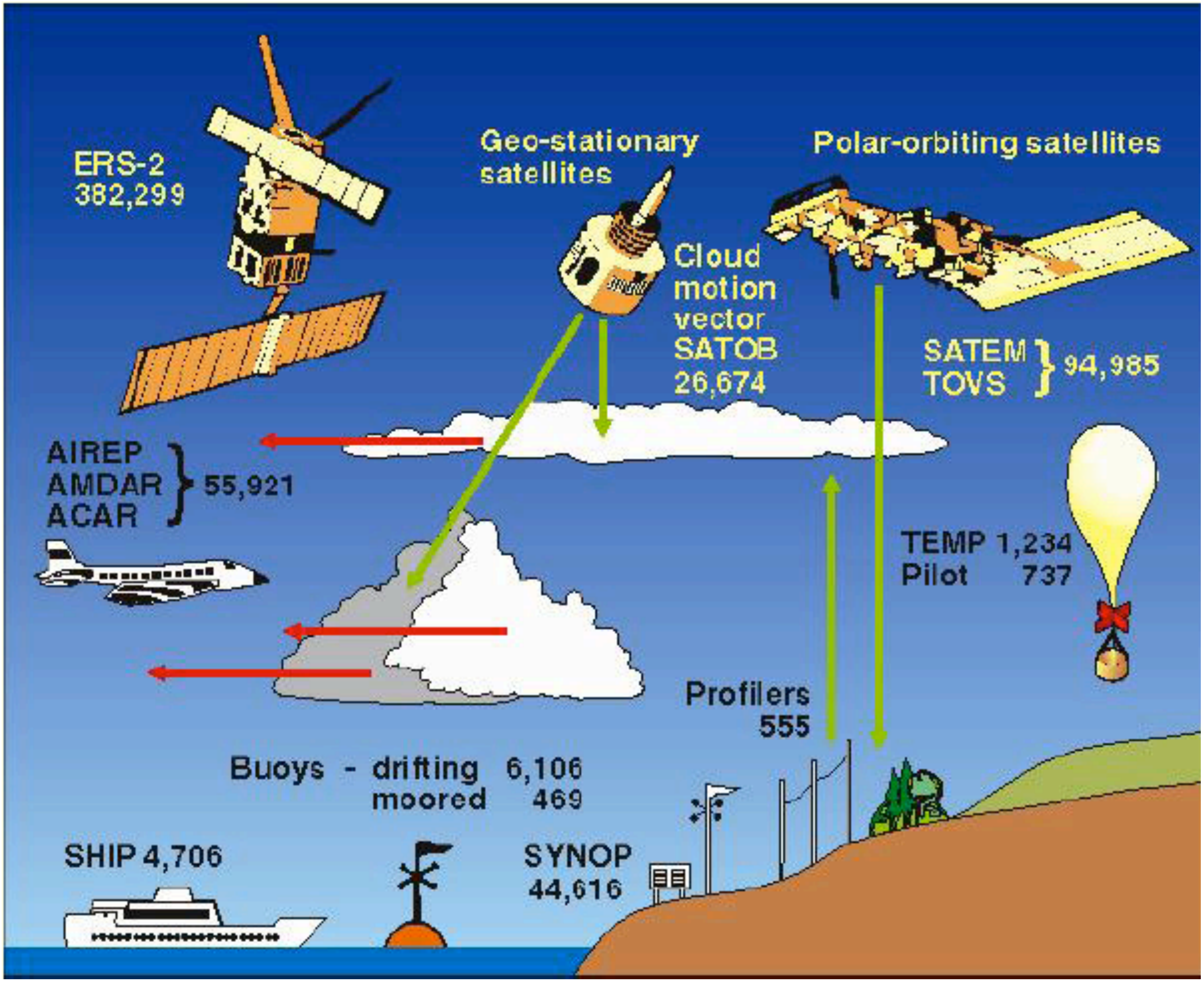
1. Introduction to data assimilation

Modern weather forecast (Bjerknes,1904)

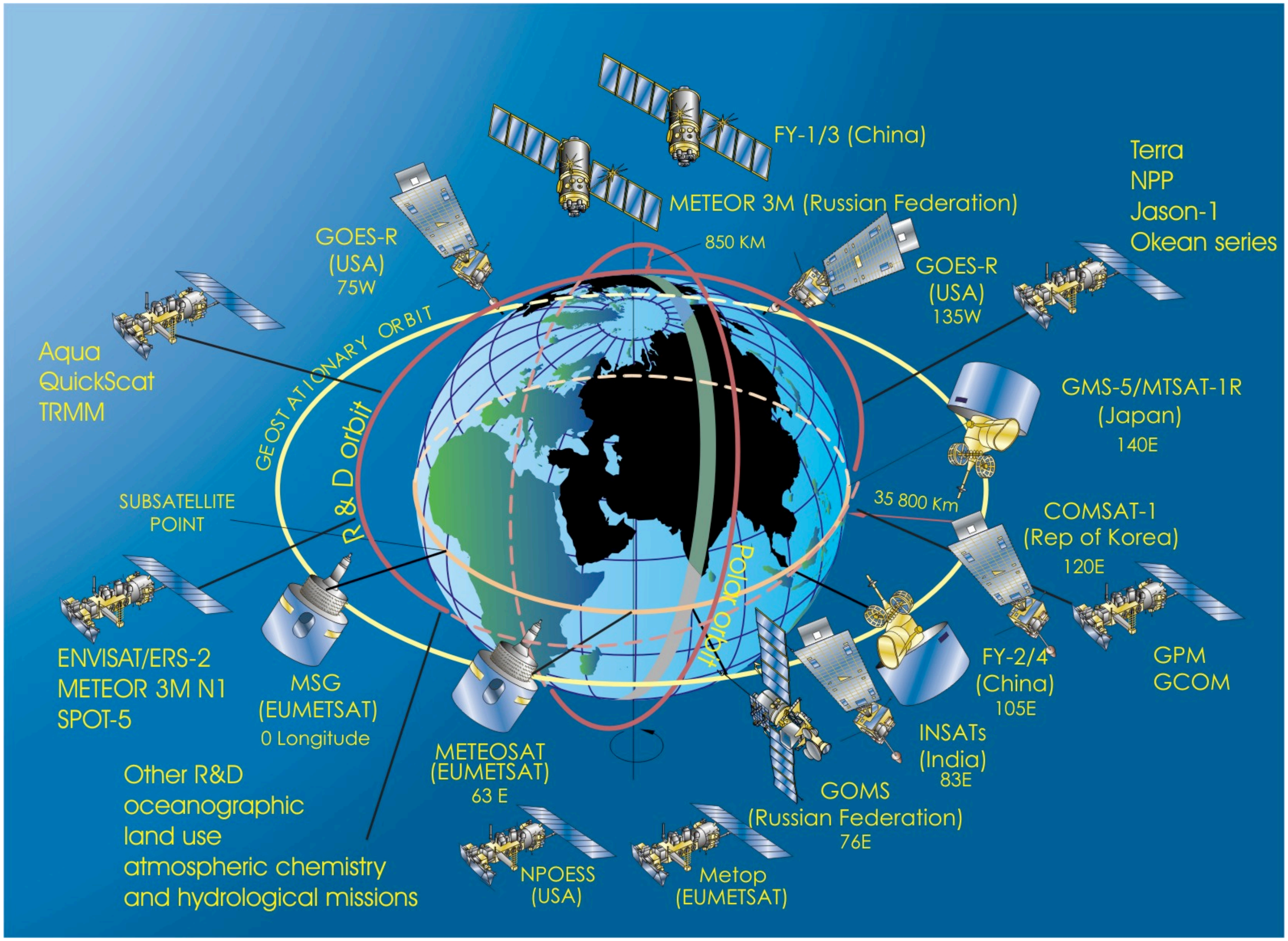
- A sufficiently accurate knowledge of the state of the atmosphere at the initial time
- A sufficiently accurate knowledge of the laws according to which one state of the atmosphere develops from another.

Vilhelm Bjerknes (1862–1951)

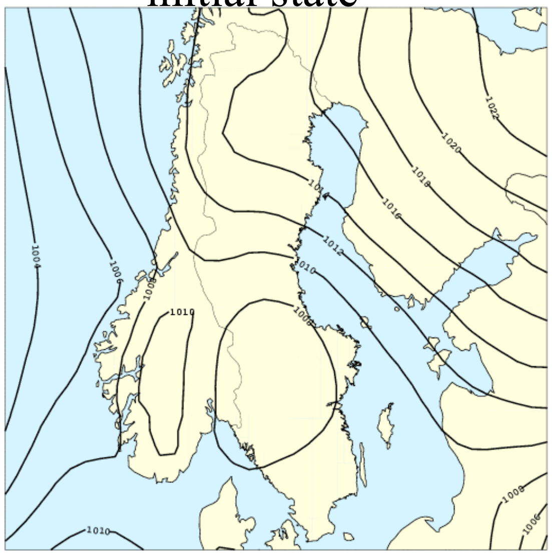
- Analysis: using observations and other information, we can specify the atmospheric state at a given initial time: “Today’s Weather”
- Forecast: using the equations, we can calculate how this state will change over time: “Tomorrow’s Weather”



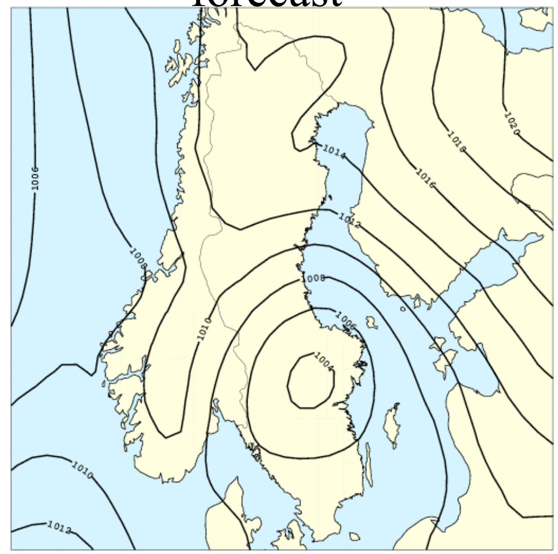
ERS-2
382,299
Geo-stationary
satellites
Polar-orbiting satellites
Cloud
motion
vector
SATOB
26,674
SATEM
TOVS
94,985
AIREP
AMDAR
ACAR
55,921
TEMP 1,234
Pilot 737
Profilers
555
Buoys - drifting 6,106
moored 469
SHIP 4,706
SYNOP
44,616


FY-1/3 (China)
METEOR 3M (Russian Federation)
Terra
NPP
Jason-1
Okean series
GOES-R
(USA)
75W
850 KM
GOES-R
(USA)
135W
Aqua
QuickScat
TRMM
GEOSTATIONARY ORBIT
R & D orbit
GMS-5/MTSAT-1R
(Japan)
140E
SUBSATELLITE
POINT
35 800 Km
COMSAT-1
(Rep of Korea)
120E
Polar orbit
ENVISAT/ERS-2
METEOR 3M N1
SPOT-5
MSG
(EUMETSAT)
0 Longitude
FY-2/4
(China)
105E
GPM
GCOM
METEOSAT
(EUMETSAT)
63 E
INSATs
(India)
83E
Other R&D
oceanographic
land use
atmospheric chemistry
and hydrological missions
GOMS
(Russian Federation)
76E
NPOESS
(USA)
Metop
(EUMETSAT)

initial state



forecast



Model

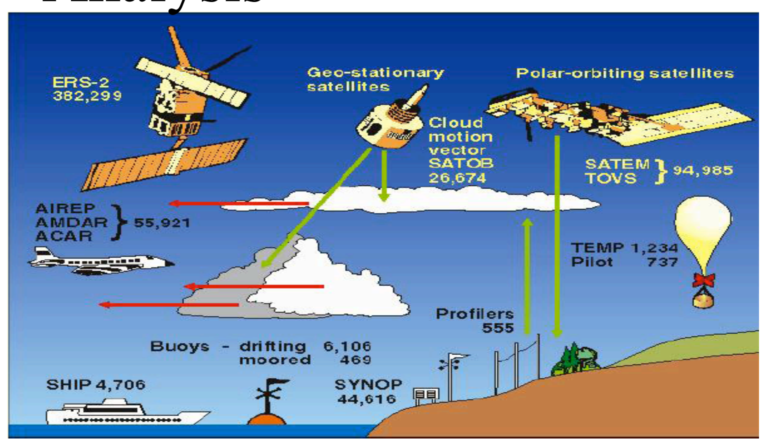
Analysis

Model state

x, $\sim 10^7$

Observations

y^o, $\sim 10^5$-10^6



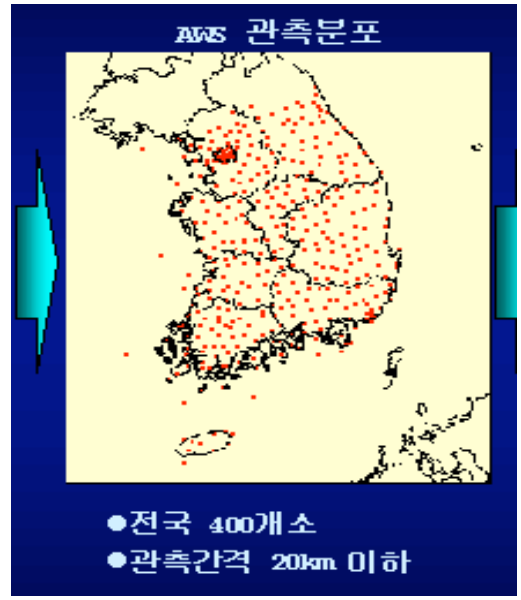
Need For Data Assimilation in NWP

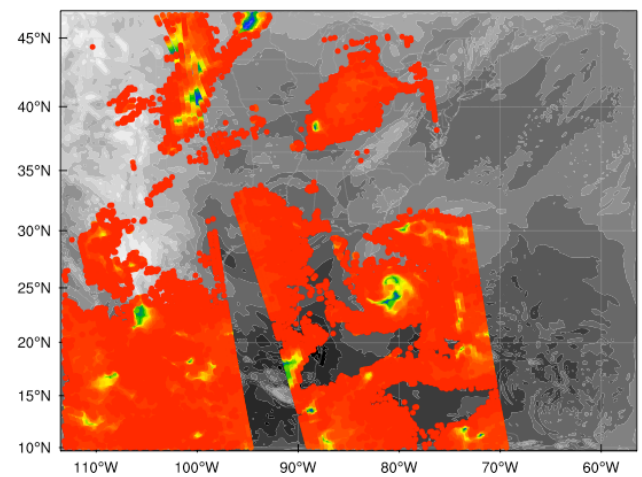
Fact: There are never enough good observations!!

- Consider NWP model:
  - Typical global model – 425 * 325 * 50 = 7 million gridpts.
  - Minimum number of variables = 6 *(u, v, w, T, p, q)*.
  - Number of degrees of freedom = 41.4 million.

- Typical number of observations = few x 10^6 but:
  - Inhomogeneous distribution of data.
  - Observations not always in sensitive areas.
  - Observations have errors.

- Solutions:
  - Use sophisticated (variational/ensemble) techniques.
  - Use previous forecast to propagate past observations.
  - Use approximate physical balance relationships.
  - More/better observations!





Assimilation methods

- Empirical methods
  - Successive Correction Method (SCM)
  - Nudging
  - Physical Initialisation (PI), Latent Heat Nudging (LHN)
- Statistical methods
  - Optimal Interpolation (OI)
  - 3-Dimensional VARiational data assimilation (3DVAR)
  - 4-Dimensional VARiational data assimilation (4DVAR)
- Advanced methods
  - Extended Kalman Filter (EKF)
  - Ensemble Kalman Filter (EnFK)

2. Basics of modern data assimilation

The analysis problem for a given time

Consider a scalar x.

The background (normally a short-range forecast):

$$x^b = x^t + b.$$

The observation:

$$x^r = x^t + r.$$

The error statistics are assumed to be known:

$$\begin{aligned}
< b > &= 0, \quad mean\ error\ (unbiased), \\
< r > &= 0, \quad mean\ error\ (unbiased), \\
< b^2 > &= B, \quad background\ error\ variance, \\
< r^2 > &= R, \quad observation\ error\ variance, \\
< br > &= 0, \quad nocorrelation\ between\ b\ and\ r,
\end{aligned}$$

where $< . >$ is ensemble average.

BLUE: the Best Linear Unbiased Estimate

The analysis: $x^a = x^t + a$.
Search for the best estimate: $x^a = \alpha x^b + \beta x^r$

Substitute the definitions, we have:

$$\alpha + \beta = 1.$$

$$< a >= 0$$

The variance:

$$A =< a^2 >= B - 2\beta B + \beta^2 (B + R)$$

To determine β: $\frac{dA}{d\beta} = -2B + 2\beta(B + R) = 0$
we have

$$\beta = \frac{B}{B + R}.$$

$$A^{-1} = B^{-1} + R^{-1}$$

The analysis: $x^a = x^b + \frac{B}{B+R}\left(x^r - x^b\right).$

"3DVAR"

The analysis is obtained by minimizing the cost function J, defined as:

$$J = \frac{1}{2}\left(x - x^b\right)^T B^{-1}\left(x - x^b\right) + \frac{1}{2}\left(x - x^r\right)^T R^{-1}\left(x - x^r\right).$$

The gradient of J with respect to x:

$$J' = B^{-1}\left(x - x^b\right) + R^{-1}\left(x - x^r\right).$$

At the minimum, $J' = 0$, we have:

$$x^a = x^b + \frac{B}{B + R}\left(x^r - x^b\right),$$

the same as BLUE.

Sequential data assimilation (I)

$$
\begin{array}{lcl}
True\ states & : & \ldots,\ x^t_{i-1},\ x^t_i,\ x^t_{i+1},\ \ldots \\
 & & \\
Observations & : & \ldots,\ x^r_{i-1},\ x^r_i,\ x^r_{i+1},\ \ldots \\
Forecasts & : & \ldots,\ x^f_{i-1},\ x^f_i,\ x^f_{i+1},\ \ldots \\
 & & \\
Analyses & : & \ldots,\ x^a_{i-1},\ x^a_i,\ x^a_{i+1},\ \ldots
\end{array}
$$

Sequential data assimilation (II)

Forecast model:

$$x_{i+1}^t = M(x_i^t) + q_i,$$

where q_i is the model error.

As q_i is unknown and x_i^a is the best estimate of x_i^t, the forecast model usually takes the form:

$$x_{i+1}^f = M(x_i^a).$$

OI (and 3DVAR):

$$x_i^a = x_i^f + \frac{B}{B+R}\left(x_i^r - x_i^f\right).$$

Sequential data assimilation (III)

4DVAR

4DVAR analysis is obtained by minimizing the cost function J, defined as:

$$J(x_i) = \frac{1}{2}\left(x_i - x_i^f\right)^T B^{-1}\left(x_i - x_i^f\right)$$
$$+\frac{1}{2}\sum_{k=0}^{K}\left[M_{k-1}(x_i) - x_{i+k}^r\right]^T R^{-1}\left[M_{k-1}(x_i) - x_{i+k}^r\right]$$

where, K is the assimilation window and

$$\begin{aligned} M_{-1}(x_i) &= x_i \\ M_0(x_i) &= M(x_i) \\ M_{k-1}(x_i) &= \underbrace{M(M(\ldots M}_{k}(x_i)\ldots)) \end{aligned}$$

Sequential data assimilation (IV)

4DVAR (continue)

The gradient of J with respect to x:

$$J^{'} = B^{-1}\left(x_i - x_i^f\right) + \sum_{k=0}^{K}\prod_{j=0}^{k-1}\mathbf{M}_{i+j}^{T}R^{-1}\left[M_{k-1}\left(x_i\right) - x_{i+k}^{r}\right]$$

where, $\mathbf{M}_{i+j}^{T}$ is the adjoint model of M at time step $i+j$.

Sequential data assimilation (V)

Extended Kalman Filters:

True states: $x_{i+1}^t = M(x_i^t) + q_i$

Model states: $x_{i+1}^f = M(x_i^a)$

Forecast error: $x_{i+1}^f - x_{i+1}^t = M(x_i^a) - M(x_i^t) - q_i$

A major assumption in KF:

$$M(x_i^a) - M(x_i^t) \approx \mathbf{M}_i(x_i^a - x_i^t)$$

Sequential data assimilation (VI)

Extended Kalman Filters (continue):

Forecast error covariance matrix:

$$
\begin{aligned}
P_{i+1}^{f} &= \left\langle \left(x_{i+1}^{f} - x_{i+1}^{t}\right)\left(x_{i+1}^{f} - x_{i+1}^{t}\right)^{T} \right\rangle \\
&\approx \mathbf{M}_i \left\langle \left(x_i^a - x_i^t\right)\left(x_i^a - x_i^t\right)^T \right\rangle \mathbf{M}_i^T + \left\langle q_i q_i^T \right\rangle \\
&= \mathbf{M}_i P_i^a \mathbf{M}_i^T + Q_i
\end{aligned}
$$

Sequential data assimilation (VII)

Extended Kalman Filters (continue):

For the analysis step:

$$K_i = P_i^f \left(P_i^f + R\right)^{-1}$$

$$x_i^a = x_i^f + K_i(x_i^r - x_i^f)$$

$$P_i^a = (I - K_i)P_i^f$$

For the forecast step:

$$x_{i+1}^f = M(x_i^a)$$

$$P_{i+1}^f = \mathbf{M}_i P_i^a \mathbf{M}_i^T + Q_i$$

Sequential data assimilation (VIII)

From scalar to vector:

Number of grid points $\approx 10^7$:

$$x \rightarrow \mathbf{x}$$

$$x^b \rightarrow \mathbf{x^b}$$

Dimension of $\mathbf{B}$, $\mathbf{P} \approx 10^7 \times 10^7$.

Number of observations, 10^6:

$$x^r \rightarrow \mathbf{y^o}$$

$$x - x^r \rightarrow H(\mathbf{x}) - \mathbf{y^o}$$

Dimension of $\mathbf{R} \approx 10^6 \times 10^6$.

Sequential data assimilation (IX)

OI:

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{B}\mathbf{H}^T \left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1} \left[\mathbf{y}^{\mathbf{o}} - H(\mathbf{x}_i^f)\right]$$

$$\mathbf{x}_{i+1}^f = M(\mathbf{x}_i^a)$$

4DVAR:

$$J(\mathbf{x}_i) = \frac{1}{2}\left(\mathbf{x}_i - \mathbf{x}_i^f\right)^T \mathbf{B}^{-1}\left(\mathbf{x}_i - \mathbf{x}_i^f\right)$$
$$+\frac{1}{2}\sum_{k=0}^{K}\left[H\left(M_{k-1}\left(\mathbf{x}_i\right)\right) - \mathbf{y^o}_{i+k}\right]^T \mathbf{R}^{-1}\left[H\left(M_{k-1}\left(\mathbf{x}_i\right)\right) - \mathbf{y^o}_{i+k}\right]$$

$$J' = \mathbf{B}^{-1}\left(\mathbf{x}_i - \mathbf{x}_i^f\right) + \sum_{k=0}^{K}\prod_{j=0}^{k-1}\mathbf{M}_{i+j}^T\mathbf{H}^T\mathbf{R}^{-1}\left[H\left(M_{k-1}\left(\mathbf{x}_i\right)\right) - \mathbf{y^o}_{i+k}\right]$$

Sequential data assimilation (X)

The Extended Kalman Filter:

For the analysis step i:

$$\mathbf{K}_i = \mathbf{P}_i^f \mathbf{H}_i^T \left(\mathbf{H}_i \mathbf{P}_i^f \mathbf{H}_i^T + \mathbf{R} \right)^{-1}$$

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}_i \left[\mathbf{y}^\mathbf{o} - H(\mathbf{x}_i^f) \right]$$

$$\mathbf{P}_i^a = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i) \mathbf{P}_i^f$$

For the forecast step, from i to $i+1$:

$$\mathbf{x}_{i+1}^f = M(\mathbf{x}_i^a)$$

$$\mathbf{P}_{i+1}^f = \mathbf{M}_i \mathbf{P}_i^a \mathbf{M}_i^T + \mathbf{Q}_i$$

Issues on data assimilation

- Observations $\mathbf{y}^o$
- Observation operator H
- Observation errors $\mathbf{R}$
- Background $\mathbf{x}^b$
- Size of $\mathbf{B}$: statistical model and tuning
- $\mathbf{M}$ and $\mathbf{M}^T$: development and validity
- Minimization algorithm (Quasi-Newton; Conjugate Gradient; ...)

- Model errors $\mathbf{Q}$
- Size of $\mathbf{P}^f$ and $\mathbf{P}^a$: simplifications

3. Demonstration with a simple system

The Lorenz 1964 equation

Nonlinear equation (NL):

$$x_{i+1} = ax_i - x_i^2 = M\left(x_i\right).$$

Depending on a, three types of solution are found: steady state; limited cycle; chaotic.

Tangent Linear equations (TL):

$$x_{i+1}^{tl} = \left(a - 2x_i^{bs}\right) x_i^{tl} = \mathbf{M}_i x_i^{tl}.$$

(linearized around basic state x_i^{bs})
Adjoint equation (AD):

$$x_i^{ad} = \left(a - 2x_i^{bs}\right) x_{i+1}^{ad} = \mathbf{M}_i^T x_{i+1}^{ad}.$$

Note here for this simple case we have

$$\mathbf{M}_i = \mathbf{M}_i^T = \left(a - 2x_i^{bs}\right).$$

Issues on data assimilation
for the system based on the Lorenz 64 eqation

- Observation operator $H = \mathbf{H} = \mathbf{H}^T = 1$
- Estimate of the true states - generated (with model error!)
- Observation error $x^o - x^t = \sigma_o G$
- Size of $\mathbf{B}$ is 1, but $B = \sigma_b^2$ still needs attention
- Model errors $\mathbf{Q}$: $q_i = \sigma_m G$ when "true" states are generated; but we assume $Q = B/4$
- Size of $\mathbf{P}^f$ or $\mathbf{P}^a$ is 1.
- $\mathbf{M}$ and $\mathbf{M}^T$: no e ort in development but their validity is still a major problem
  (model, assimilation window length, ...)
- Gaussian statistics

Too simple?

Try: Lorenz63 model

Try: 1-D advection equation

Try: 2-D shallow water equation

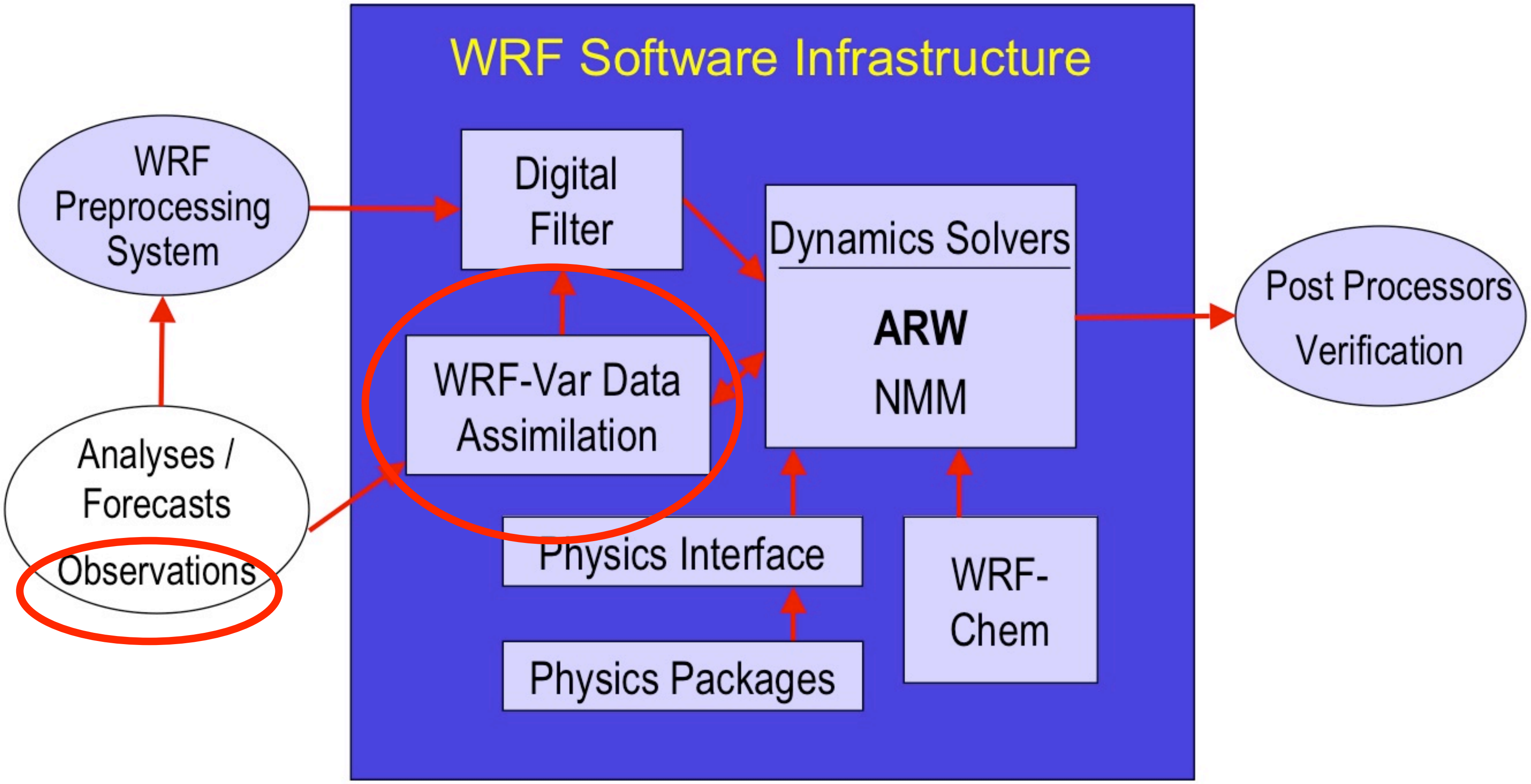
…

Try: ARW!!!

4. WRF-Var

…WRF-Var is a **Var**iational data assimilation system built within the software framework of **WRF**, used for application in both research and operational environments….

WRF Modeling System



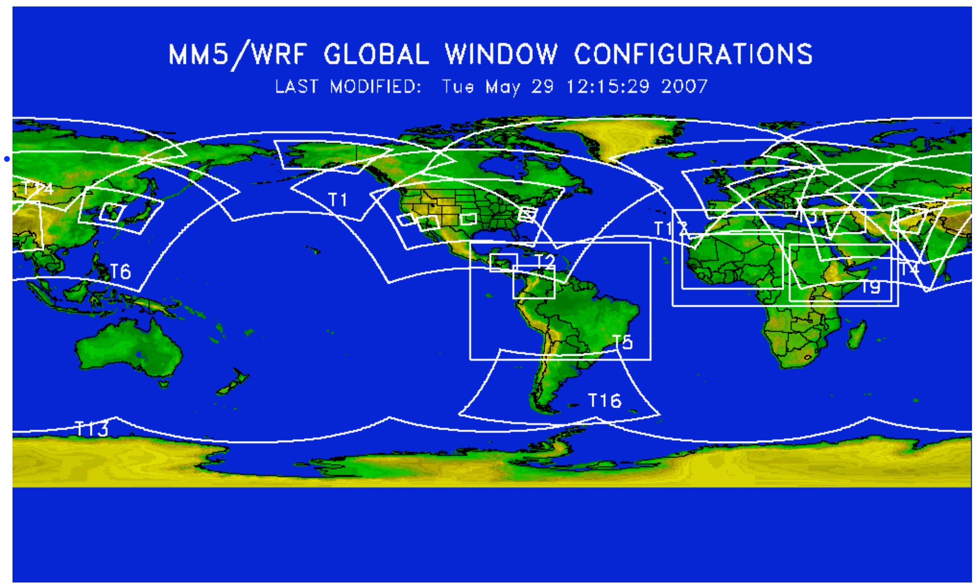
ARW = Advanced Research WRF (NCAR) Core

NMM = Nonhydrostatic Mesoscale Model (NCEP) Core

WRF-Var (WRFDA) Data Assimilation Overview

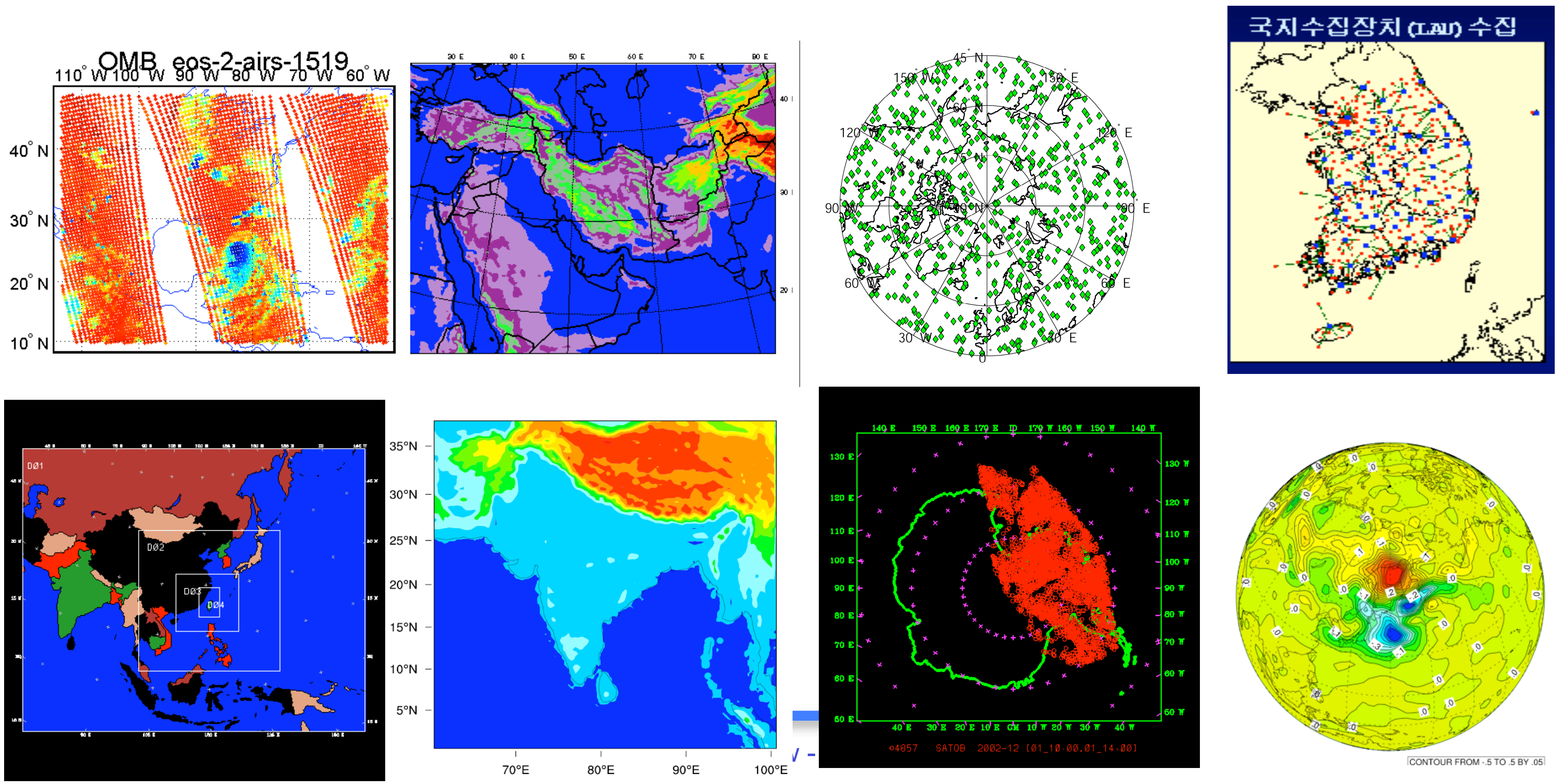
- **Goal:** Community WRF DA system for regional/global, research/operations, and deterministic/probabilistic applications.
- **Techniques:**
  - 3D-Var
  - 4D-Var (regional)
  - Ensemble DA,
  - Hybrid Variational/Ensemble DA.
- **Model:** WRF (ARW and NMM)
- **Support:**
  - NCAR/ESSL/MMM/DAG
  - NCAR/RAL/JNT/DATC
- **Observations:** Conv.+Sat.+Radar

AFWA Theaters:



The WRF-Var Program

- NCAR staff: 23FTE, ~12 projects.
- Non-NCAR collaborators (AFWA, KMA, CWB, BMB, etc): ~10FTE.
- Community users: ~30 (more in 4000 general WRF downloads?).



Future Plans

General Goals:

- Unified, multi-technique WRF DA system.
- Retain flexibility for research, multi-applications.
- Leverage international WRF community efforts.

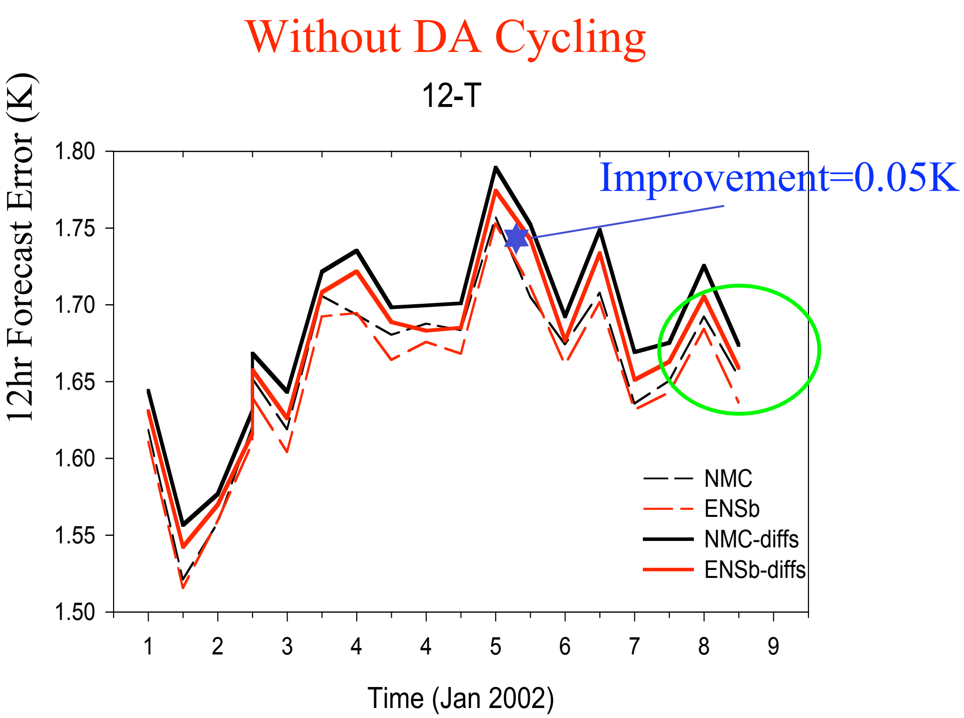
WRF-Var Development (MMM Division):

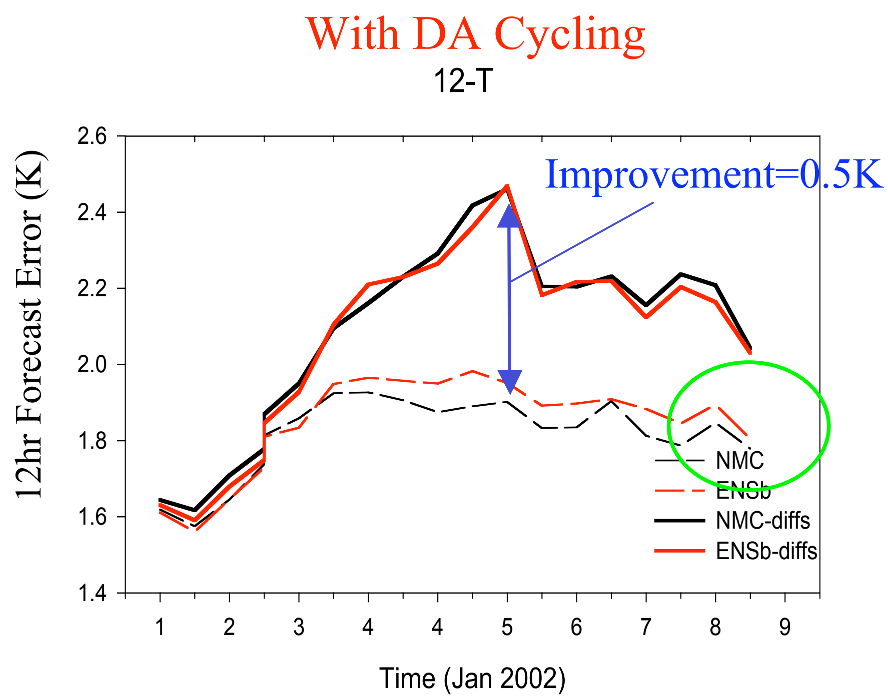
- 4D-Var (additional physics, optimization).
- Sensitivities tools (adjoint, ensemble, etc.).
- EnKF within WRF-Var -> WRFDA.
- Instrument-specific radiance QC, bias correction, etc.

Data Assimilation Testbed Center (DATC):

- Technique intercomparison: 3/4D-Var, EnKF, Hybrid
- Obs. impact: AIRS, TMI, SSMI/S, METOP.
- New Regional testbeds: US, India, Arctic, Tropics.

Importance of Data Assimilation For General WRF Development/Testing





Warning: Cycling with insufficient observations leads to degradation (1.8K vs. 1.7K)

Experiment (Mi-Seon Lee, KMA)

5. Important issues covered by this tutorial

- y^o observations - collection, quality control, bias correction, thinning, …
- H observation operator, including the tangent linear operator $\mathbf{H}$ and the adjoint operator $\mathbf{H}^T$.
- M forecast model, including the tangent linear model $\mathbf{M}$ and adjoint model $\mathbf{M}^T$.
- $\mathbf{B}$ background error covariance (~ $10^7 \mathrm{x} 10^7$).
- $\mathbf{R}$ observation error covariance which includes the representative error ($10^6 \mathrm{x} 10^6$).
- Minimization algorithm
- $\mathbf{P}^a$ and $\mathbf{P}^f$ analysis and forecast error covariances