# Fundamentals of Data Assimilation

Kayo Ide
University of Maryland

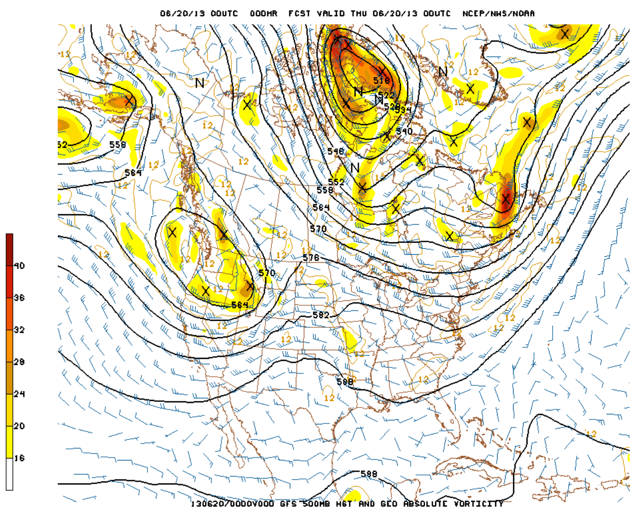
2013 Joint DTC-EMC-JCSDA GSI Tutorial

# Objectives of Data Assimilation from NWP Viewpoint

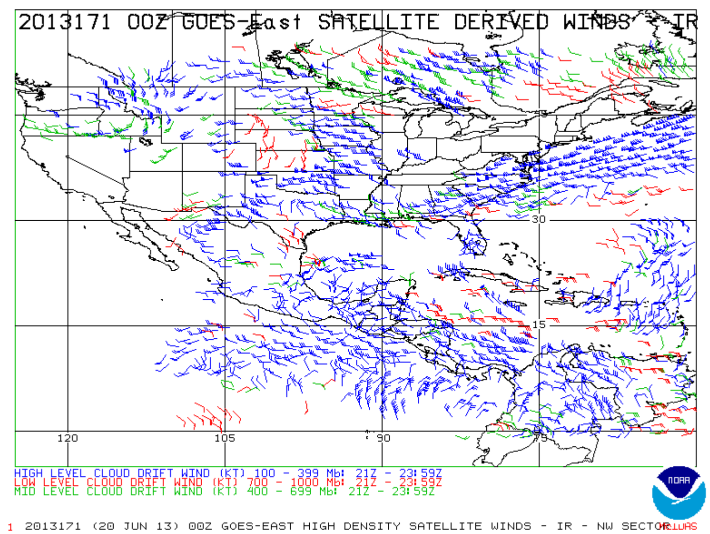
- Primary objectives of data assimilation:
  - Estimation of the current state
  - Forecast of the future state

  by periodically integrating information from

1. Computational (forecast) model




2. Observations (in real-time)




- DA is a growing area:

  - Many variants, extensions, applications etc exist & are being developed

# Objectives of This Lecture

- Basic ideas of data assimilation
- Unifying perspectives of currently popular & practical approaches
- Understanding of the variational formulation
  - Cost function for minimization(=optimization)

    $J(\mathbf{x}) = J^b(\mathbf{x}) + J^o(\mathbf{x})$ [+ $J^c(\mathbf{x})$]

    $J^b(\mathbf{x}) = 1/2(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b)$ : *b*ackground cost function

    $J^o(\mathbf{x}) = 1/2(\mathbf{y}^o - \mathbf{h}(\mathbf{x}))^T \mathbf{R}^1(\mathbf{y}^o - \mathbf{h}(\mathbf{x}))$ : *o*bservation cost function

    $J^c(\mathbf{x})$ : *c*onstraint cost function

    Where ($\mathbf{x}$, $\mathbf{B}$): model state vector & background error covariance matrix

    ($\mathbf{y}$, $\mathbf{R}$): observation vector& observation error covariance matrix

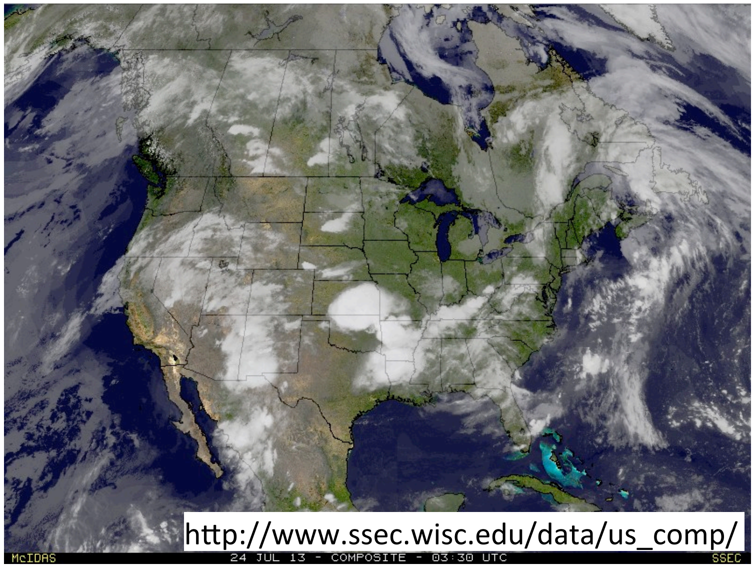
    $\mathbf{y}=\mathbf{h}(\mathbf{x})$ : forward model / observation operator
  - With emphasis on
    - Role of **B**
    - Flexibility of variational approach

# Outline

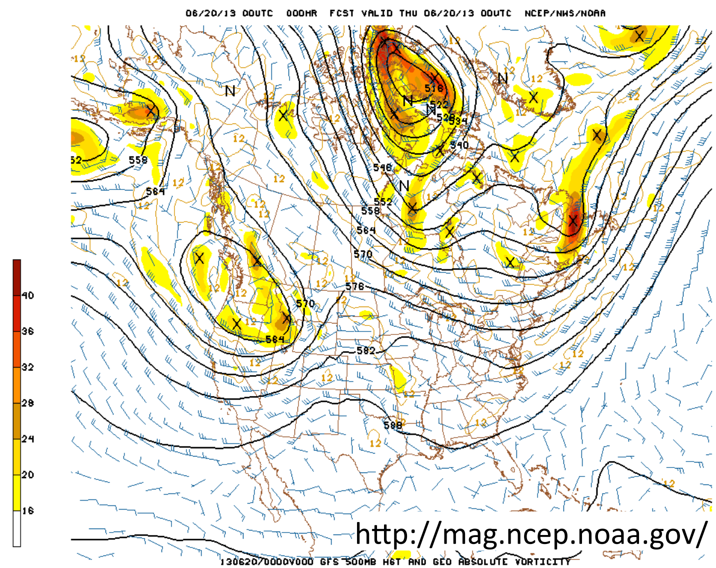
- Objectives
- Background
  - State **x** & Observation **y** - also **m**(**x**) & **h**(**x**)
  - Probability $p(\mathbf{x})$ - also $p(\mathbf{x}|\mathbf{y})$ & $p(\mathbf{y}|\mathbf{x})$
  - Data assimilation perspectives
- 3D Method - 3 dimensions in space
  - OI = Optimal Interpolation
  - 3DVar =Variational
- 3D to 4D - 4th dimension is time
  - EKF/EnKF= Extended/Ensemble Kalman filter
  - FGAT = First Guess at Appropriate Time
  - 4DVar
  - Hybrid = between Var and EnKF - Current operational system
- Concluding remarks

# State **x** and $\mathbf{x}^t$ ?

- Real atmosphere




- Atmospheric modeling




"true state" $\mathbf{x}^t_k$ :

target of data assimilation

[= consistent projection of real state onto **x**]

→ representation of $\mathbf{x}^t$ has uncertainty

→ Probability $p(\mathbf{x})$

Computational model (from time $t_{k-1}$ to $t_k$)

$\mathbf{x}_k = \mathrm{m}_{k,k-1}(\mathbf{x}_{k-1})$

**m**: model

**x**: $N$-dim spatially discretized vector of atmospheric variables ($N$~$10^7$)

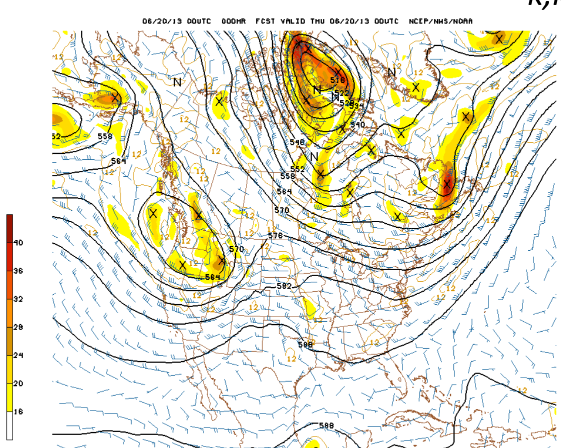
# Computational Model: $\mathbf{x}_k = \mathbf{m}_{k,k-1}(\mathbf{x}_{k-1})$

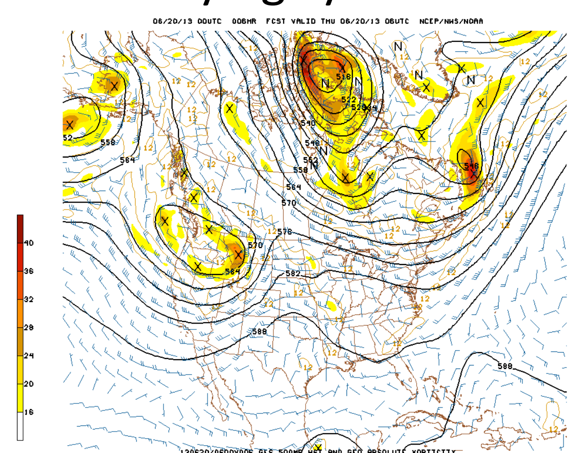
- Variables: $\mathbf{x}$ is high-dimensional state vector $N$~O($10^{6-7}$)

  Ex: GFS

$$\mathbf{x} = \begin{pmatrix} \mathbf{T} \\ \mathbf{q} \\ \mathbf{D} \\ \zeta \\ \mathbf{p}_s \\ \vdots \end{pmatrix} = \begin{pmatrix} \text{Temperature} \\ \text{moisture} \\ \text{divergence} \\ \text{vorticity} \\ \text{surface pressure} \\ \vdots \end{pmatrix}_{\text{All grid points}}$$

- Forecast model:
  - Model $\mathbf{m}_{k,k-1}$ : flow dependent, complex, and nonlinear
  - Initial condition(IC) $\mathbf{x}_{k-1}$: accurate representation of the current condition
  - Skill: $\mathbf{m}_{k,k-1}$ , $\mathbf{x}_{k-1}$, and underlying system itself





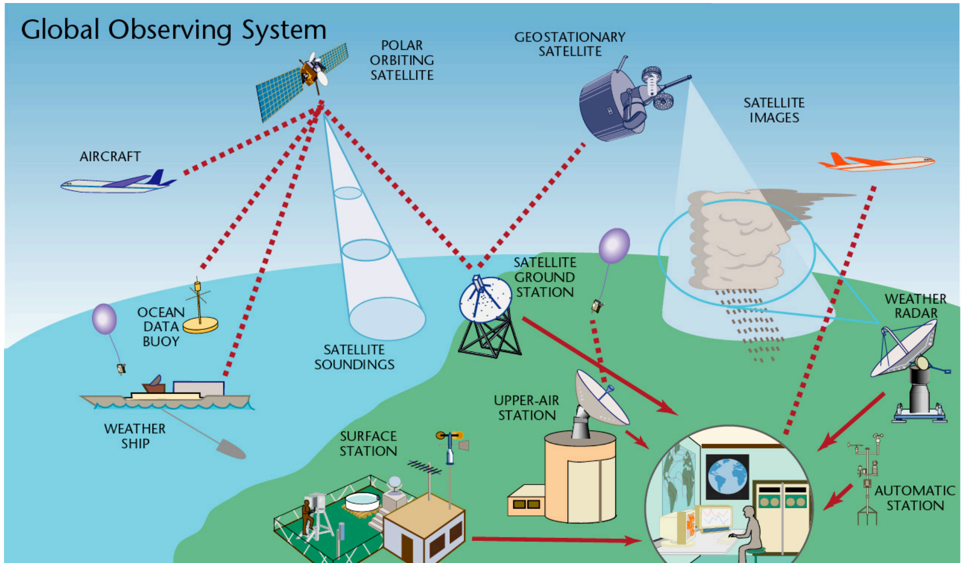
GFS $\mathbf{x}_{k-1}$ at =06/23/2013 0UTC → $\mathbf{x}_k = m_{k,k-1}(\mathbf{x}_{k-1})$ at $t_k = t_{k-1}$ +6hr

# Observation & Forward Model : y=h(x)

- Observation vector **y**: Sampling of (real) atmospheric state
  - Ex: GSI

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_T \\ \mathbf{y}_q \\ \mathbf{y}_{\text{wind}} \\ \mathbf{y}_{ps} \\ \vdots \end{pmatrix} = \begin{pmatrix} \text{Temperature} \\ \text{moisture} \\ \text{wind field} \\ \text{surface pressure} \\ \vdots \end{pmatrix}$$

  - Platforms
    - In situ
    - Remote sensing




http://www.wmo.int/pages/prog/www/OSY/GOS.html

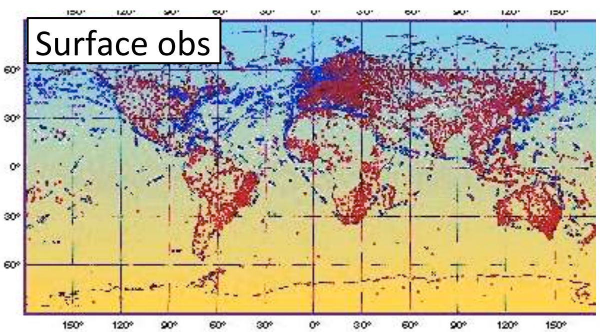
- Forward model **h**(**x**): Relationship between **x** and **y**
  - Can be simple or complex

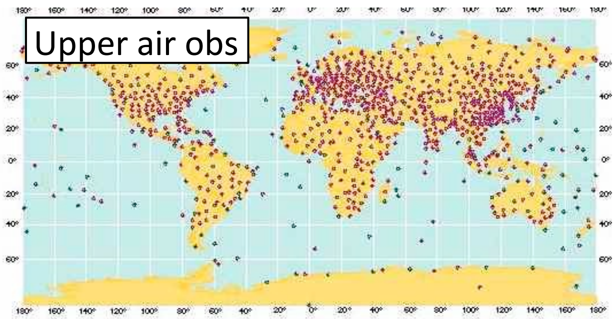
# Uncertainty in Observations

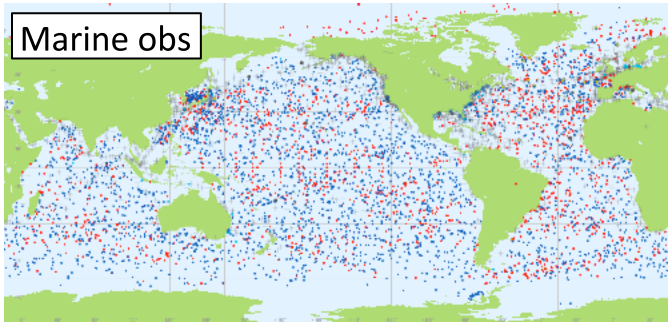
- Characteristics of $\mathbf{y}^o$
  - Heterogeneous, asynchronous, & noisy sampling of the evolving state $\mathbf{x}^t$

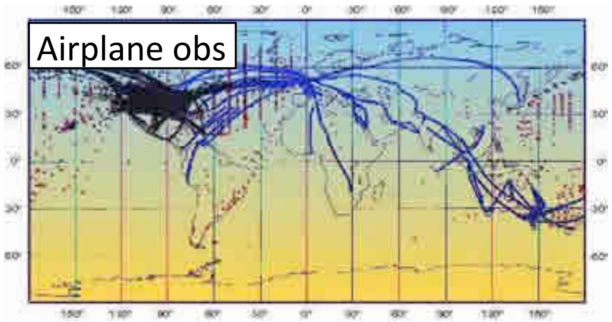
  → Uncertainty in the representation $\mathbf{y}=\mathbf{h}(\mathbf{x})$ [ $\mathbf{y}^t=\mathbf{h}(\mathbf{x}^t)$? ]

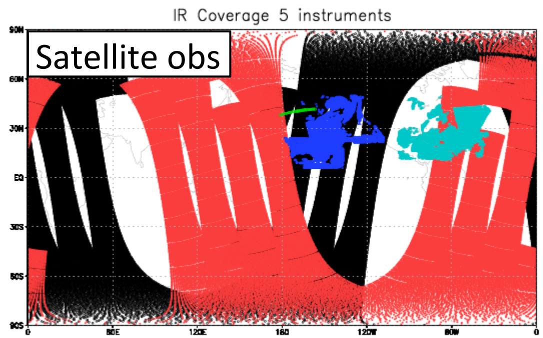
  ➔ Modeling of observation likelihood: $p(\mathbf{y}|\mathbf{x})$

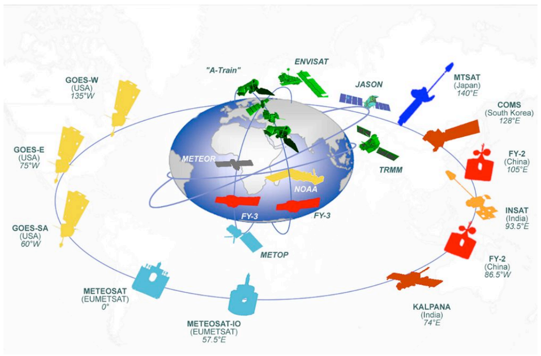


http://www.wmo.int/pages/prog/www/OSY/GOS.html

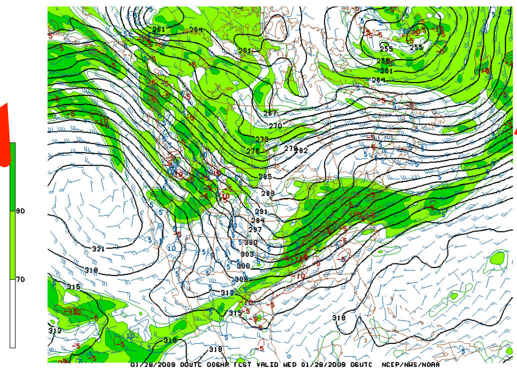
# Schematic of Data Assimilation: Assimilation Window

- Data assimilation is a method that iterates the cycle over a window:

6hr for Operational System

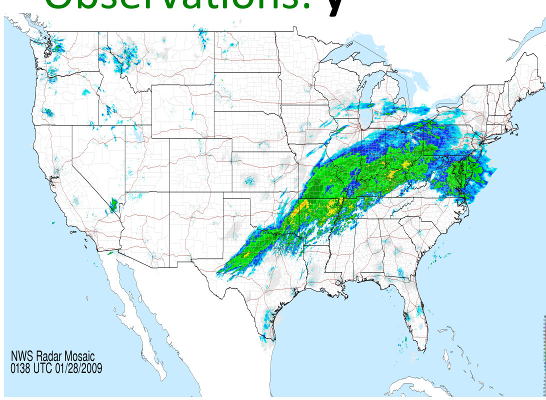
- Model forecast: $\mathbf{x}^b$

6h forecast (700mb) for 1.28.09 06UTC



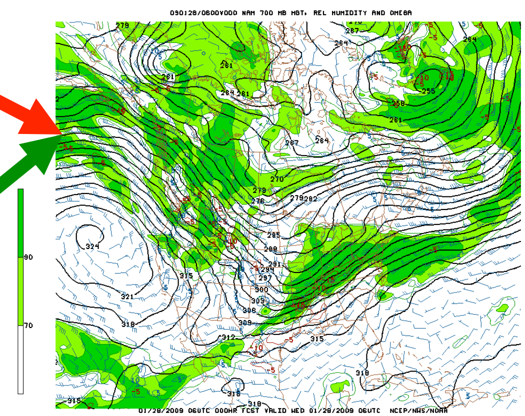
$p(\mathbf{x})$
[prior]

- Observations: $\mathbf{y}^o$



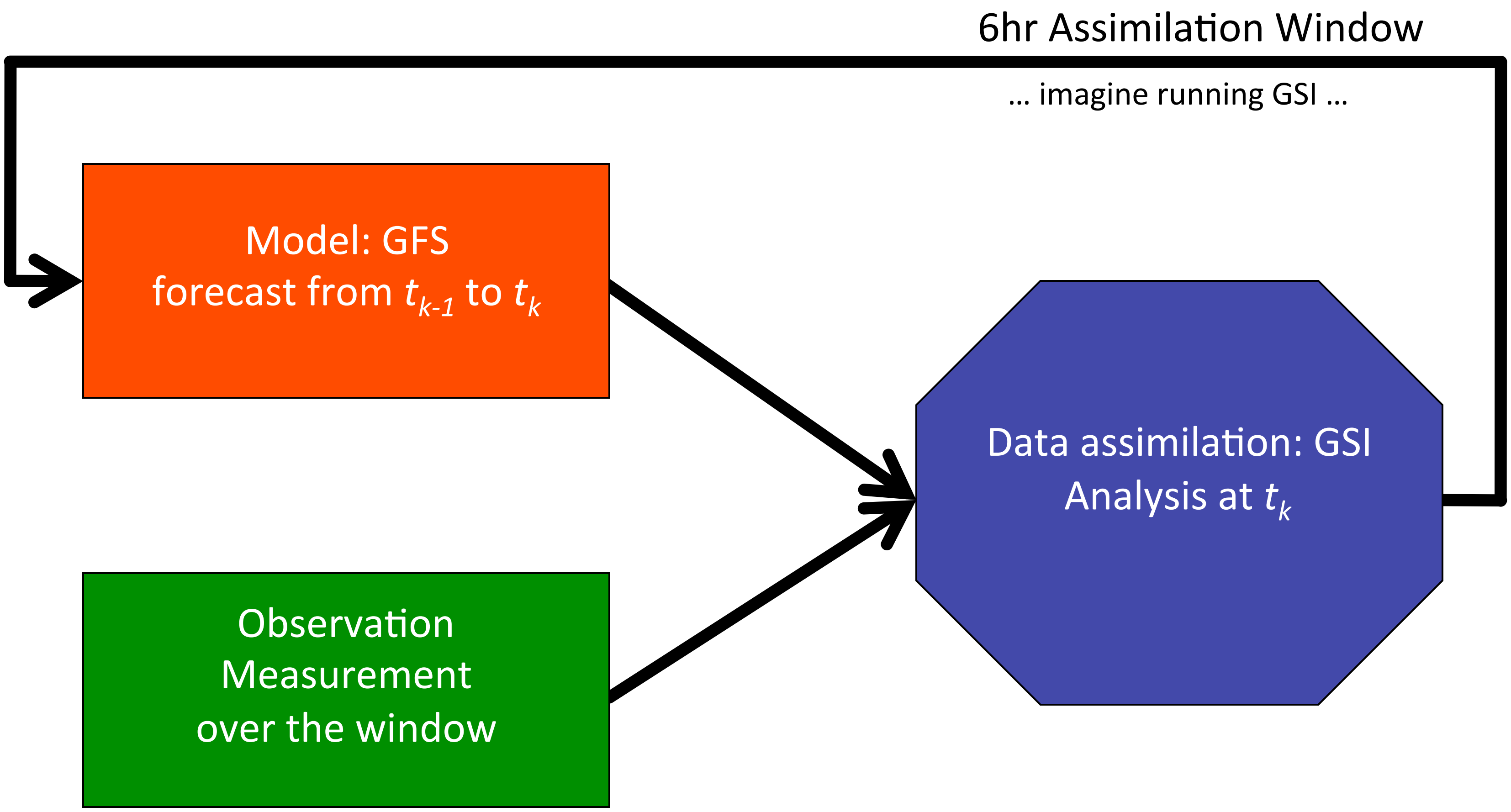

$p(\mathbf{y}|\mathbf{x})$
[obs likelihood]

- Analyzed model state: $\mathbf{x}^a$

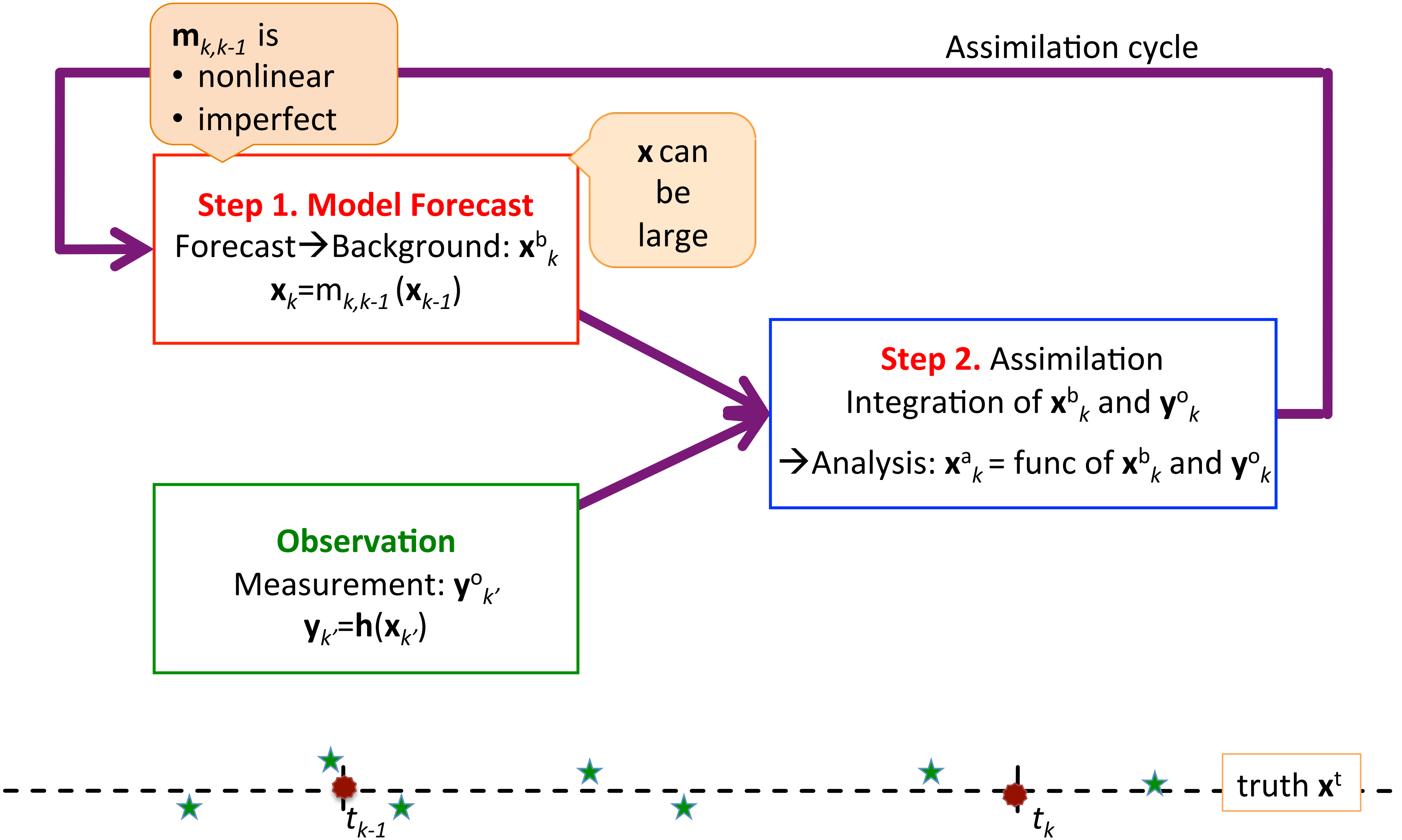


$p(\mathbf{x}|\mathbf{y})$
[posterior]

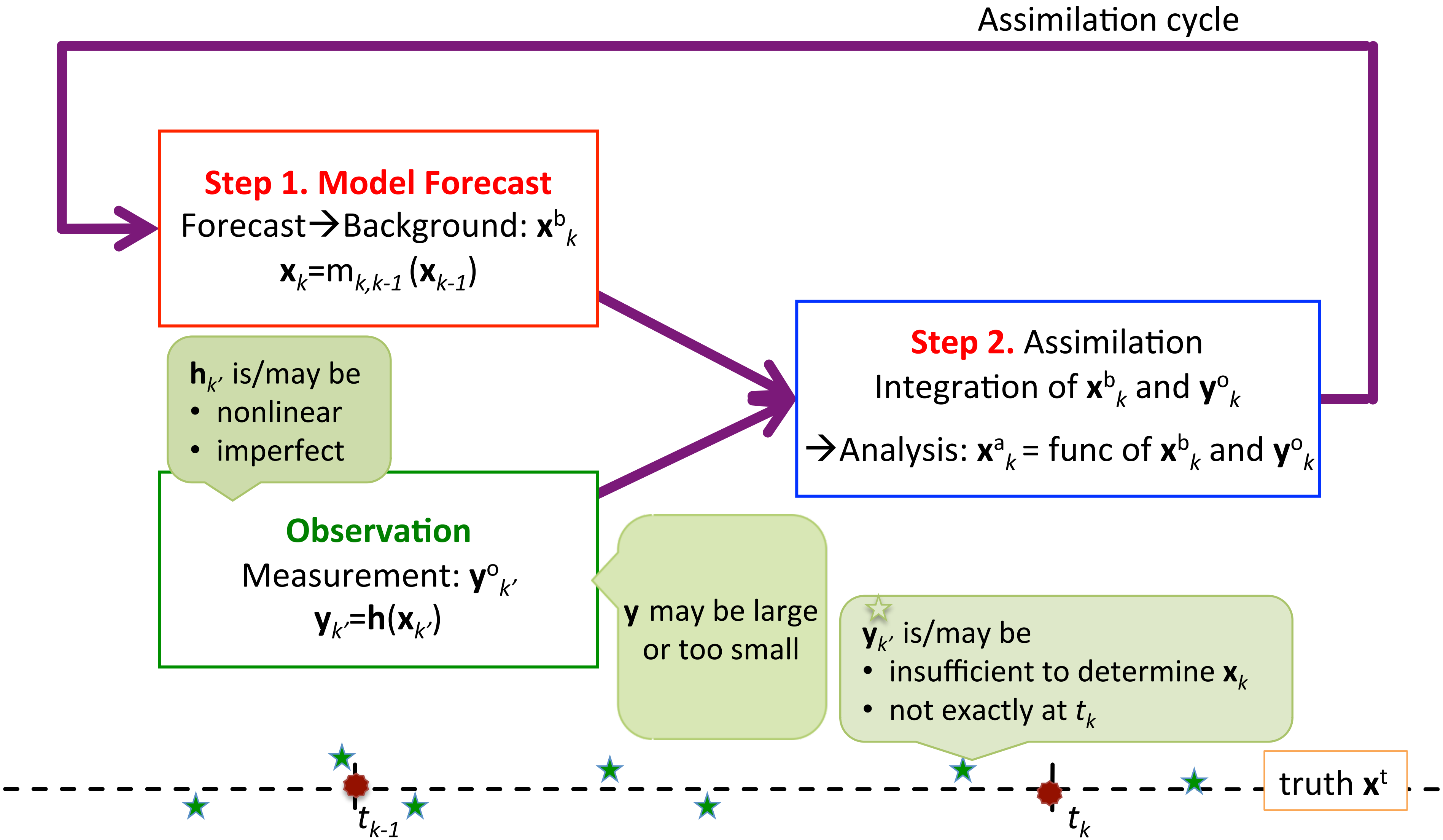
# Elements of Data Assimilation



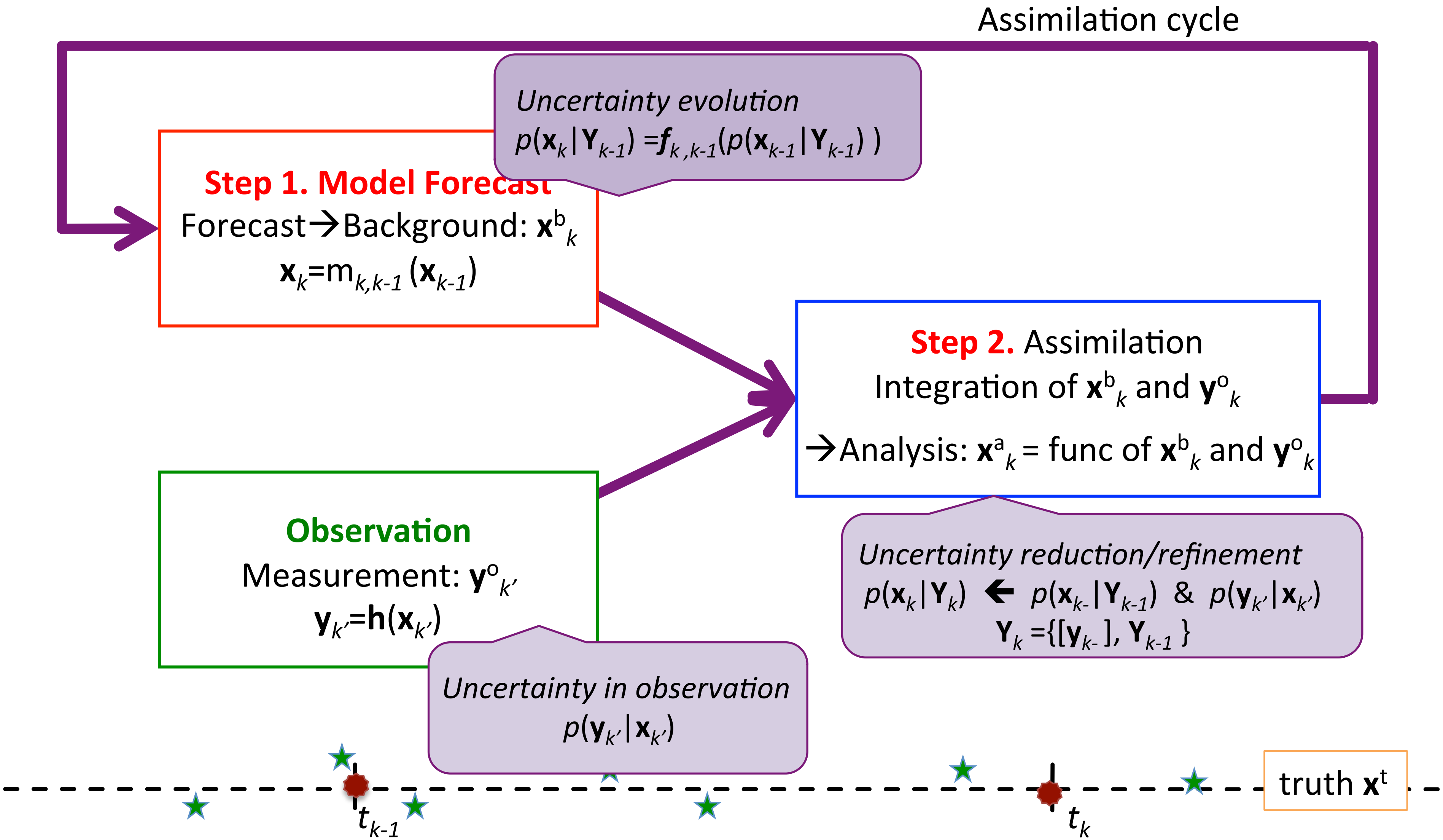
# Challenges of Data Assimilation: Model



# Challenges of Data Assimilation: Observation



# Data Assimilation Cycle With Probabilistic View



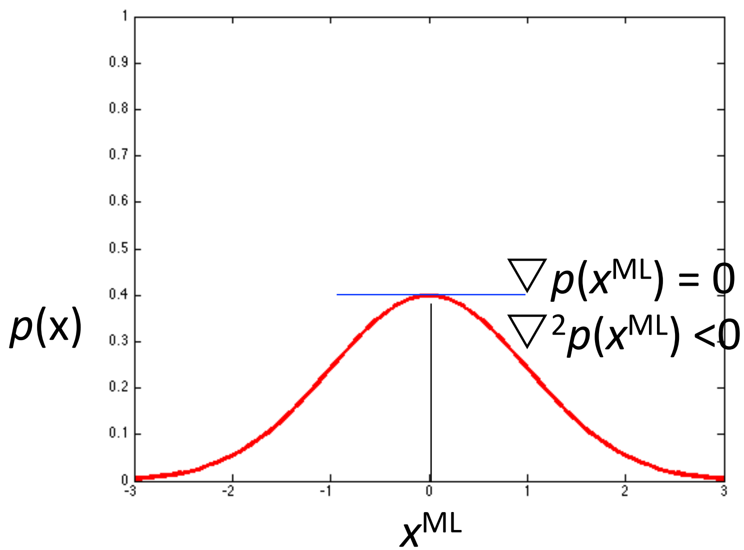
# Probability $p(\mathbf{x})$ & Data Assimilation Perspective I

- Probability $p(\mathbf{x})$ ~ Likelihood (between 0 and 1) that value being $\mathbf{x}$

- Perspective: How to choose a "good" $\mathbf{x}$ given $p(\mathbf{x})$

  → Maximum Likelihood (ML): $\mathbf{x}^{ML}$ that is most likely

  - Conditions at $\mathbf{x}=\mathbf{x}^{ML}$ that maximizes $p(\mathbf{x})$
    - Extreme (=gradient $\mathbf{0}$): $\nabla p(\mathbf{x}) = \mathbf{0}$
    - Maximum (~convex) : $\nabla^2 p(\mathbf{x})$ =(semi-)negative definite




<u>Basic axioms</u>

- $p(\mathbf{x})>0$ for all $\mathbf{x}$
- $\int p(\mathbf{x})d\mathbf{x}=1$

# Expectation Based on Probability $p(\mathbf{x})$

- Expectation: $E[f(\mathbf{x})] = \int f(\mathbf{x}) p(\mathbf{x}) d\mathbf{x}$
  - Mean $\bar{\mathbf{x}}$ = expected value of $\mathbf{x}$

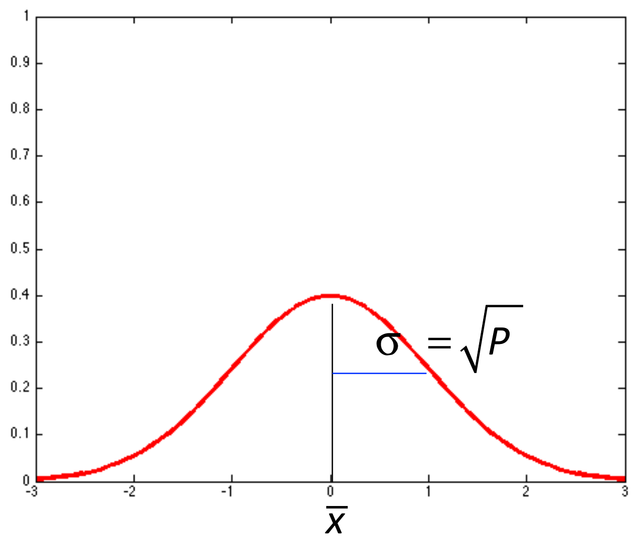
$$\bar{x}_n = E[x_n] = \int x_n p(\mathbf{x}) d\mathbf{x}$$

  - Variance = (standard deviation)$^2$

    = Uncertainty around $\bar{\mathbf{x}}$

$$P_{nn} = E[(x_n - \bar{x}_n)^2] = \int (x_n - \bar{x}_n)^2 p(\mathbf{x}) d\mathbf{x}$$

$$\sigma_n = \sqrt{P_{nn}}$$




$$\bar{\mathbf{x}} = E[\mathbf{x}] = \int \mathbf{x} p(\mathbf{x}) d\mathbf{x} = \begin{pmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_N \end{pmatrix}$$

Cross-variable:

- Covariance

$$P_{in} = P_{ni} = E[(x_i - \bar{x}_i)(x_n - \bar{x}_n)] = \int (x_i - \bar{x}_i)(x_n - \bar{x}_n) p(\mathbf{x}) d\mathbf{x}$$

- Covariance matrix

$$\mathbf{P} = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T] = \int (\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T p(\mathbf{x}) d\mathbf{x}$$

$$= \begin{pmatrix} P_{11} & & \cdots & & P_{1N} \\ & \ddots & & P_{ni} & \\ & & & \vdots & \\ \vdots & P_{in} & \cdots & P_{nn} & \vdots \\ & & & \ddots & \\ P_{N1} & & \cdots & & P_{NN} \end{pmatrix}$$

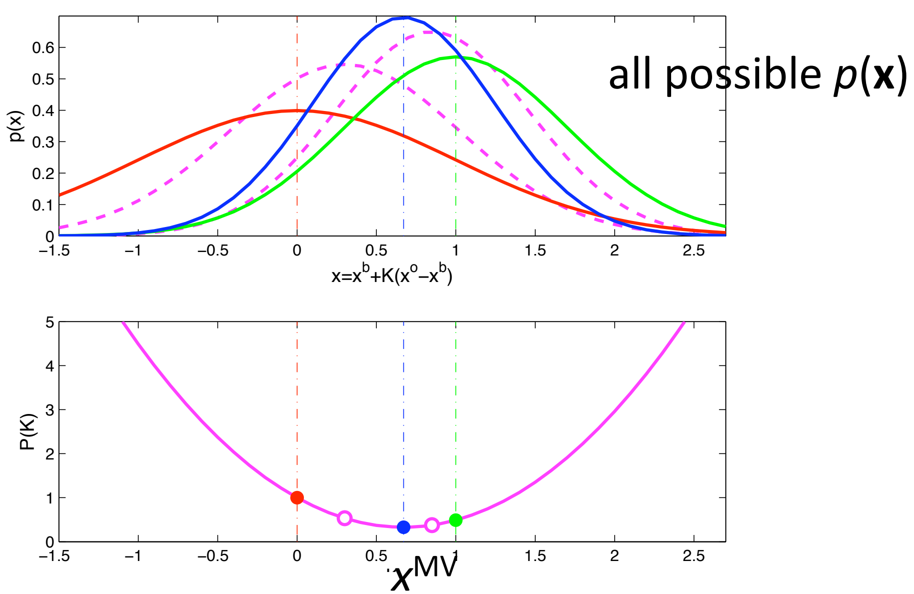
# Probability $p(\mathbf{x})$ & Data Assimilation Perspective II

- Total uncertainty ~ total variance associated with $p(\mathbf{x})$: $\sum_{n=1}^{N} P_{nn} = tr\mathbf{P} = \sigma^2$

- Perspective: How to choose a "good" **x**

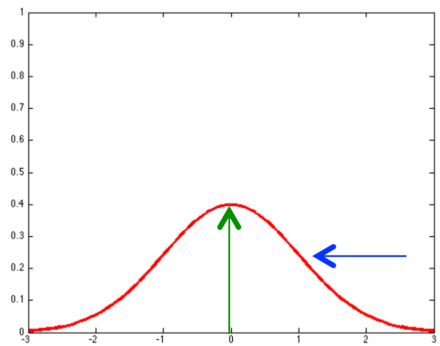
  → Minimum Variance (MV): $\mathbf{x}^{MV}$ with the least the risk

  - Conditions
    - Corresponding $p^{MV}(\mathbf{x})$ has the minimum variance among all $p(\mathbf{x})$
    - $\mathbf{x}^{MV}$ is the expectation (mean) associated with $p^{MV}(\mathbf{x})$



# Maximum Likelihood, Minimum Variance, and Data Assimilation

- Relationship between $\mathbf{x}^{ML}$ and $\mathbf{x}^{MV}$



$$\int p(\mathbf{x})d\mathbf{x} = 1$$

$$p_G(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{P}|^{1/2}}\exp\{-J_G(\mathbf{x})\}$$

$$J_G(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\overline{\mathbf{x}})^T\mathbf{P}^{-1}(\mathbf{x}-\overline{\mathbf{x}})$$

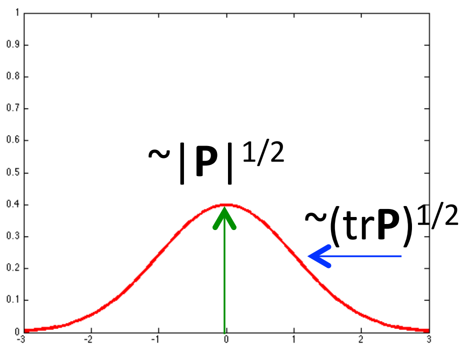
If probability density function (pdf) is Gaussian $p_G(\mathbf{x})$ & obs is linear,

then $\mathbf{x}^{ML} = \mathbf{x}^{MV}$

- Practical data assimilation
  - Estimation of $\overline{\mathbf{x}}$ and $\mathbf{P}$
  - For large $N$, not easy to completely estimate $\mathbf{P}$
    - Statistical: 3D methods
    - Dynamical
      - Tangent linear model: Extended Kalman filter / 4DVar
      - Monte Carlo: Ensemble Kalman Filter

# Gaussian $p_G(\mathbf{x})$; 1D and 2D

- 1D

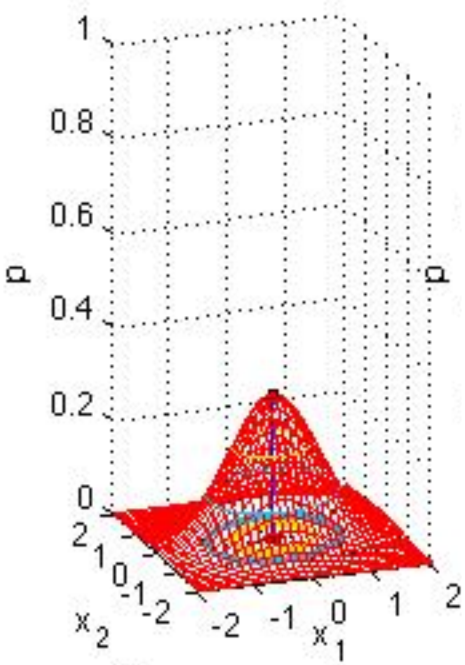


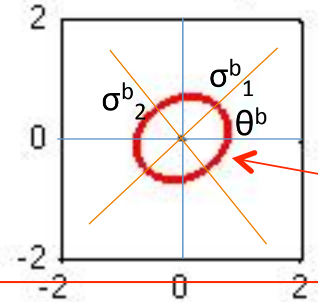

$$p_G(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{P}|^{1/2}} \exp\{-J_G(\Delta\mathbf{x})\}$$

$$\mathbf{x} = \overline{\mathbf{x}} + \Delta\mathbf{x}$$

$$J_G(\Delta\mathbf{x}) = \frac{1}{2}(\Delta\mathbf{x})^T \mathbf{P}^{-1}(\Delta\mathbf{x})$$

- 2D




**P** gives size of uncertainty as "ellipsoid"

18

# Ensemble Approach to Represent $p(\mathbf{x})$

- Ensemble
  - Members $\mathbf{X} = \{\mathbf{x}^{(m)}\} \quad = \{\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(M)}\}$
  - Spread $\Delta\mathbf{X} = \{\mathbf{x}^{(m)} - \overline{\mathbf{x}}\} = \{\mathbf{x}^{(1)} - \overline{\mathbf{x}},\ldots,\mathbf{x}^{(M)} - \overline{\mathbf{x}}\}$
  - Mean

$$\overline{x}_n = \frac{1}{M}\sum_{m=1}^{M} x_n^{(m)}$$

$$\overline{\mathbf{x}} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{x}^{(m)}$$

  - Covariance

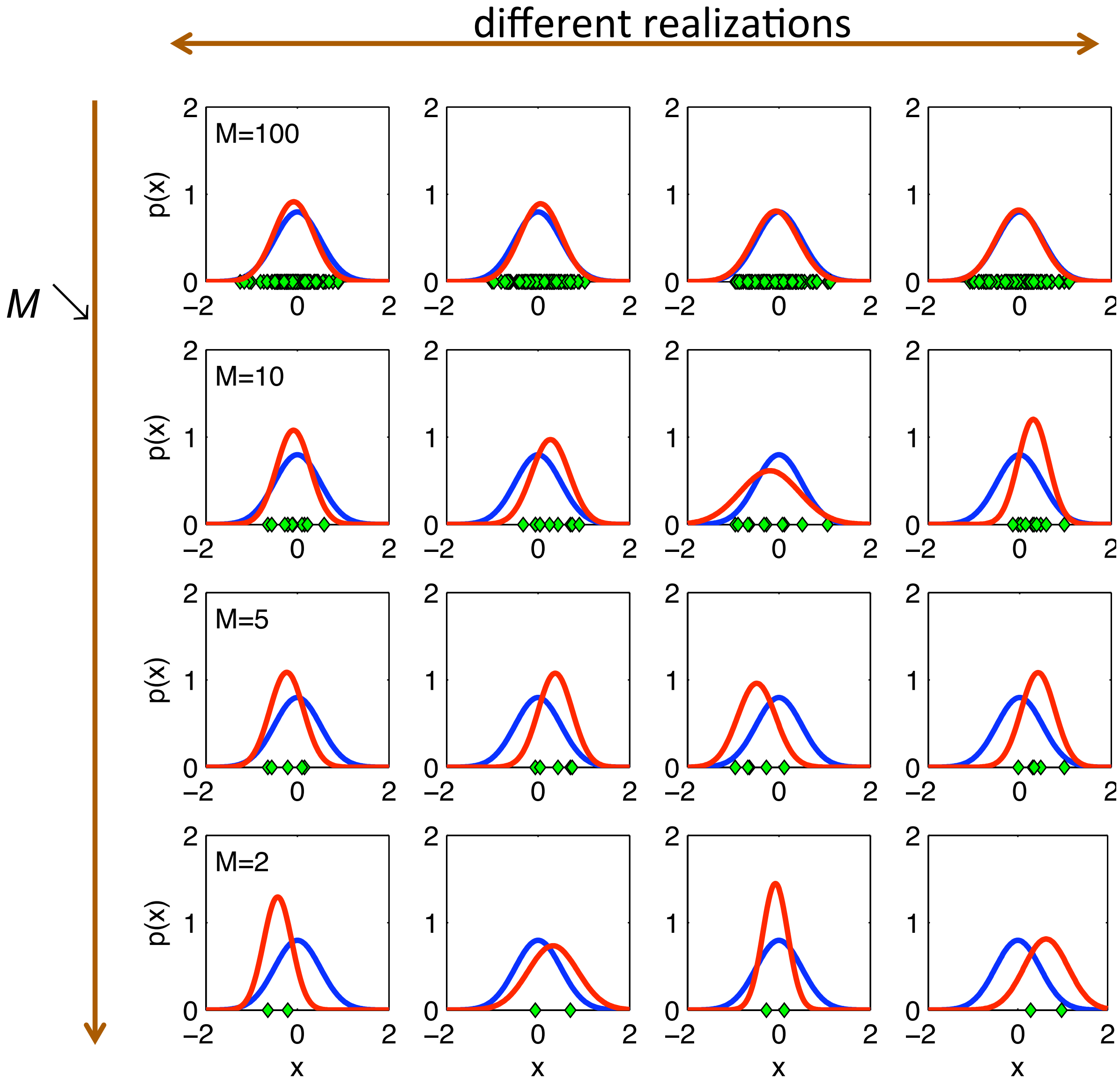
$$P_{nn} = \frac{1}{M-1}\sum_{m=1}^{M}(x_n^{(m)} - \overline{x}_n)^2$$

$$P_{in} = P_{ni} = \frac{1}{M-1}\sum_{m=1}^{M}(x_i^{(m)} - \overline{x}_i)(x_n^{(m)} - \overline{x}_n)$$

$$\mathbf{P} = \frac{1}{M-1}(\Delta\mathbf{X})(\Delta\mathbf{X})^T$$

- Issues
  - Sampling of by ensemble can be lousy, especially for
    - Small $M$
    - Small $P_{in}$
  - Rank of **P** is at most $M$-1
  - There infinitely many $\Delta\mathbf{X}$ that have the same $\mathbf{P}=(1/M-1)\Delta\mathbf{X}(\Delta\mathbf{X})^{\mathrm{T}}$

# $p(\mathbf{x})$ Sampling & Reconstruction by Ensemble: 1D




Assuming Gaussian pdfs

$$p(x) = \frac{1}{(2\pi)^{1/2} P^{1/2}} \exp\left(-\frac{1}{2}\frac{(x-\bar{x})^2}{P}\right)$$

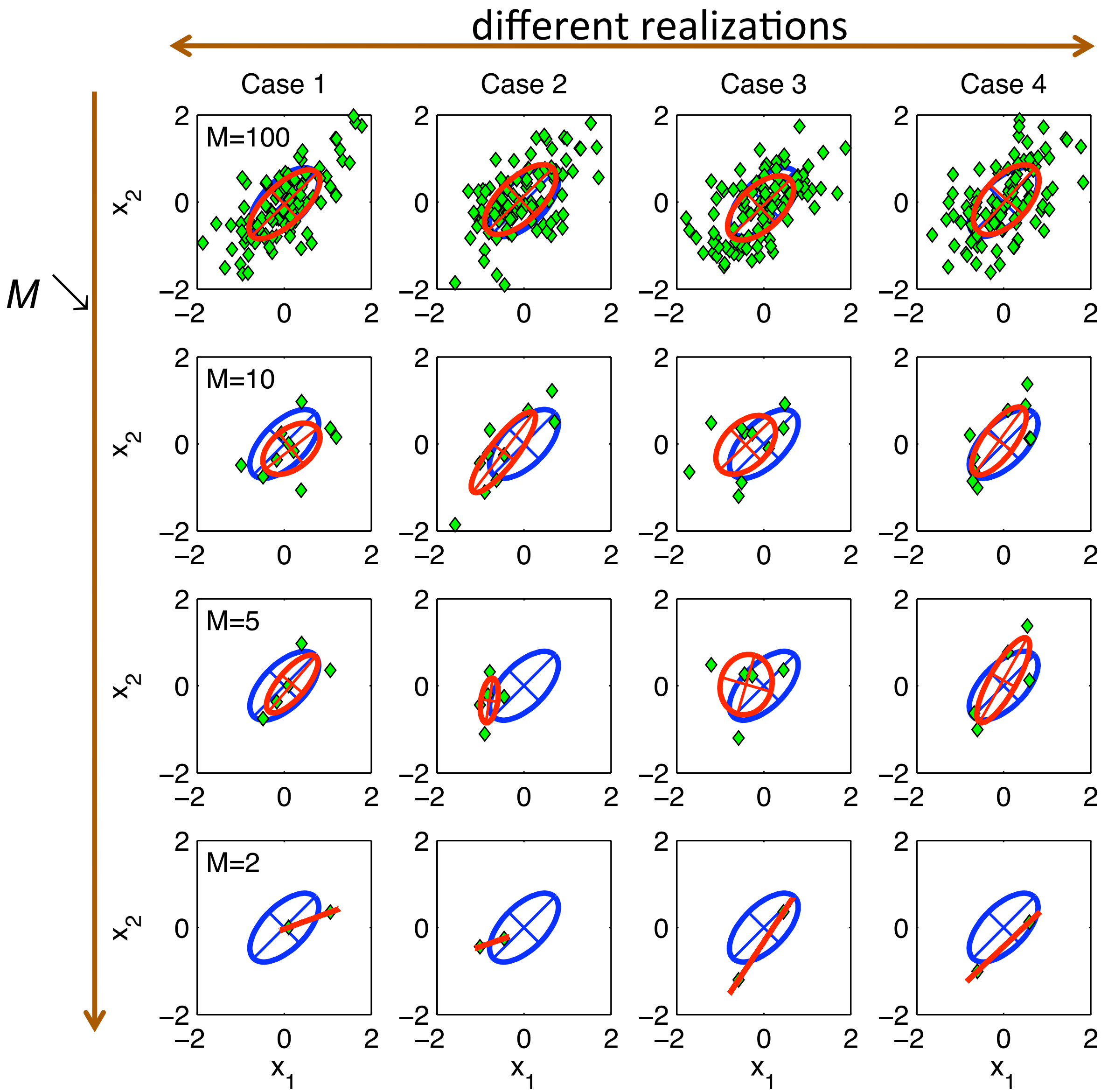
$p(\mathbf{x}) \Rightarrow \{\mathbf{x}_m\}$

- orig. $p$(x) by $(\bar{x}, P) = (0, 0.5)$
- *M* sample* from *p*(x)
- Reconstructed $p^*$(x) by $(\bar{x}^*, P^*)$

$\{\mathbf{x}_m\} \Rightarrow p^*(\mathbf{x})$

- If sampling is well-done, then $p^*$(x)~$p$(x).
- 'Fitness' of $p^*$(x) to $p^*$(x) vary case by case particularly for small *M*.
- All cases, N<M.

# $p(\mathbf{x})$ Sampling & Reconstruction by Ensemble: 2D




Assuming Gaussian pdfs

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}\,|\,\mathbf{P}\,|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{P}^{-1}(\mathbf{x}-\bar{\mathbf{x}})\right)$$

$$p(\mathbf{x}) \Rightarrow \{\mathbf{x}_m\}$$

— orig. $p(\mathbf{x})$ by $(\bar{\mathbf{x}},\mathbf{P})$

◆ $M$ sample from $p(\mathbf{x})$

— Reconstructed $p^*(\mathbf{x})$ by $(\bar{\mathbf{x}}^*,\mathbf{P}^*)$

$$\{\mathbf{x}_m\} \Rightarrow p^*(\mathbf{x})$$

- If sampling is well-done, then $p^*(x) \sim p(x)$.
- 'Fitness' of $p^*(x)$ to $p^*(x)$ vary case by case particularly for small $M$.
- All cases, N≤M.

# Outline

- Objectives
- Background
  - State **x** & Observation **y** - also **m**(**x**) & **h**(**x**)
  - Probability $p(\mathbf{x})$ - also $p(\mathbf{x}|\mathbf{y})$ & $p(\mathbf{y}|\mathbf{x})$
  - Data assimilation perspectives
- 3D Method - 3 dimensions in space
  - OI = Optimal Interpolation
  - 3DVar =Variational
- 3D to 4D - $4^{th}$ dimension is time
  - EKF/EnKF= Extended/Ensemble Kalman filter
  - FGAT = First Guess at Appropriate Time
  - 4DVar
  - Hybrid = between Var and EnKF - Current operational system
- Concluding remarks

# Data Assimilation With Probabilistic View in 3D

*Prior*
$p(\mathbf{x}) \sim N(\mathbf{x}^b, \mathbf{B})$
**B** : Prescribed

**Step 1. Model Forecast**
Forecast→Background: $\mathbf{x}^b$

**Step 2.** Assimilation
Integration of $\mathbf{x}^b_k$ and $\mathbf{y}^o_k$
→Analysis: $\mathbf{x}^a_k$ = func of $\mathbf{x}^b_k$ and $\mathbf{y}^o_k$

**Observation**
Measurement: $\mathbf{y}^o$

*Observation Likelihood*
$p(\mathbf{y}|\mathbf{x}) \sim N(\mathbf{y}^o, \mathbf{R})$
**R** : Given

3D Method" computes $\mathbf{x}^a$
- 3DVar (ML)
  - by maximizing likelihood in posterior $p(\mathbf{x})$
- Optimal Interpolation (MV)
  - by minimizing tr**P** of posterior $p(\mathbf{x})$

# Perspectives of Data Assimilation

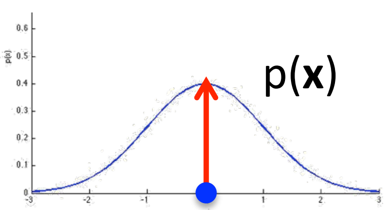
- Two main perspectives of practical data assimilation & hybrid approach

**Variational Approach:**

Least square estimation

[maximum likelihood]

- 3D-Var (3 dim in space)
- 4D-Var ($4^{th}$ dim is time)

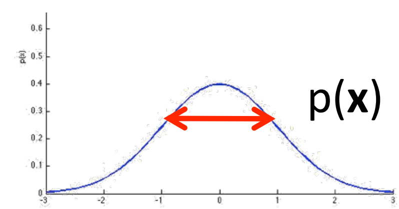



**Sequential (KF) Approach:**

Minimum Variance estimate

[least uncertainty]

- Optimal Interpolation (OI)
- (Extended / Ensemble) **Kalman Filter**




Hybrid

- Data assimilation is fast growing domain
  - Advanced methods are being developed as interdisciplinary science
  - A variety of extension & applications exist and are under development

# OI

- Approach: MV using Best Linear Unbiased Estimation (BLUE)
  - Idea: Determine $\mathbf{x}$ as linear combination of the two information

    $$\mathbf{x} = \mathbf{G}\,\mathbf{x}^b + \mathbf{K}\,\mathbf{y}^o = \mathbf{x}^t + \boldsymbol{\varepsilon} \qquad \text{with } \boldsymbol{\varepsilon} \sim (0, \mathbf{P}) \quad \text{or} \quad \mathbf{P} = E[\boldsymbol{\varepsilon}(\boldsymbol{\varepsilon})^T]$$

    such that resulting $\mathbf{x}^{OI}$ has
    - Minimum variance (least risk) min: $tr\mathbf{P} = \Sigma_n^N P_{nn}$
    - No bias: $E[\boldsymbol{\varepsilon}] = \mathbf{0}$
  - Mathematical problem: Determine $\mathbf{G}$ and $\mathbf{K}$ as such.
  - Statistical property: $\mathbf{x}^a = \mathbf{x}^{OI}$ has less risk (=is better) than $\mathbf{x}^b$ or $\mathbf{y}^o$

- Solution
  - Analytical:

    $$\mathbf{x}^a = \mathbf{x}^{OI} = \mathbf{x}^b + \mathbf{K}^{OI}(\mathbf{y}^o - \mathbf{h}(\mathbf{x}^b))$$

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

    Although not required

    $$\mathbf{P}^a = \mathbf{P}^{OI} = (\mathbf{I} - \mathbf{K}^{OI}\mathbf{H})\mathbf{B}$$

  - Computational:

    $$\mathbf{x}^a = \mathbf{x}^{OI} = \mathbf{x}^b + \mathbf{B}\mathbf{H}^T\mathbf{z}$$

    where $\mathbf{z}$ is solution to $(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}\mathbf{z} = \mathbf{d}$ $\qquad \mathbf{d} = \mathbf{y}^o - \mathbf{h}(\mathbf{x}^b)$

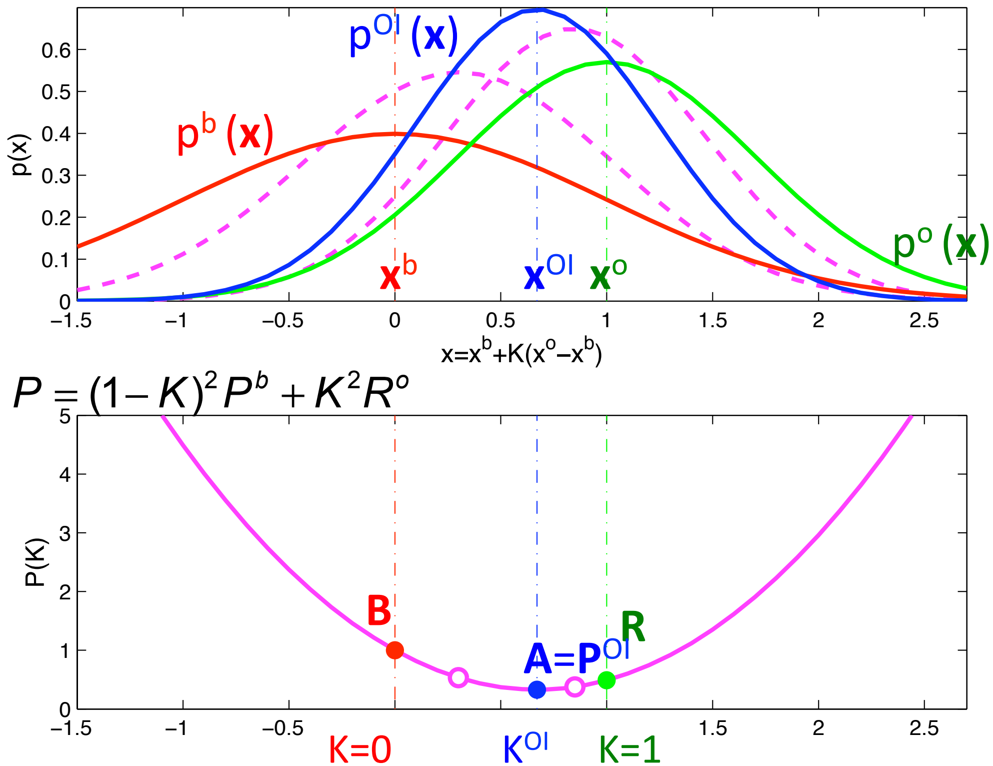
# OI: 1D Example

- Background $(x^b, B)$
- Observation $(x^o, R)$ with y=x

→Analysis $x^a = x^{OI} = (1-K^{OI})\, x^b + K^{OI}\, x^o$

$trP^{OI} = (1-K^{OI})\, B = R/(B+R)$

$$K^{OI} = \frac{B}{R+B} = \frac{(B/R)}{1+(B/R)} \rightarrow \begin{cases} 0 & \text{as } B/R \rightarrow 0 \\ 1 & \text{as } B/R \rightarrow \infty \end{cases}$$



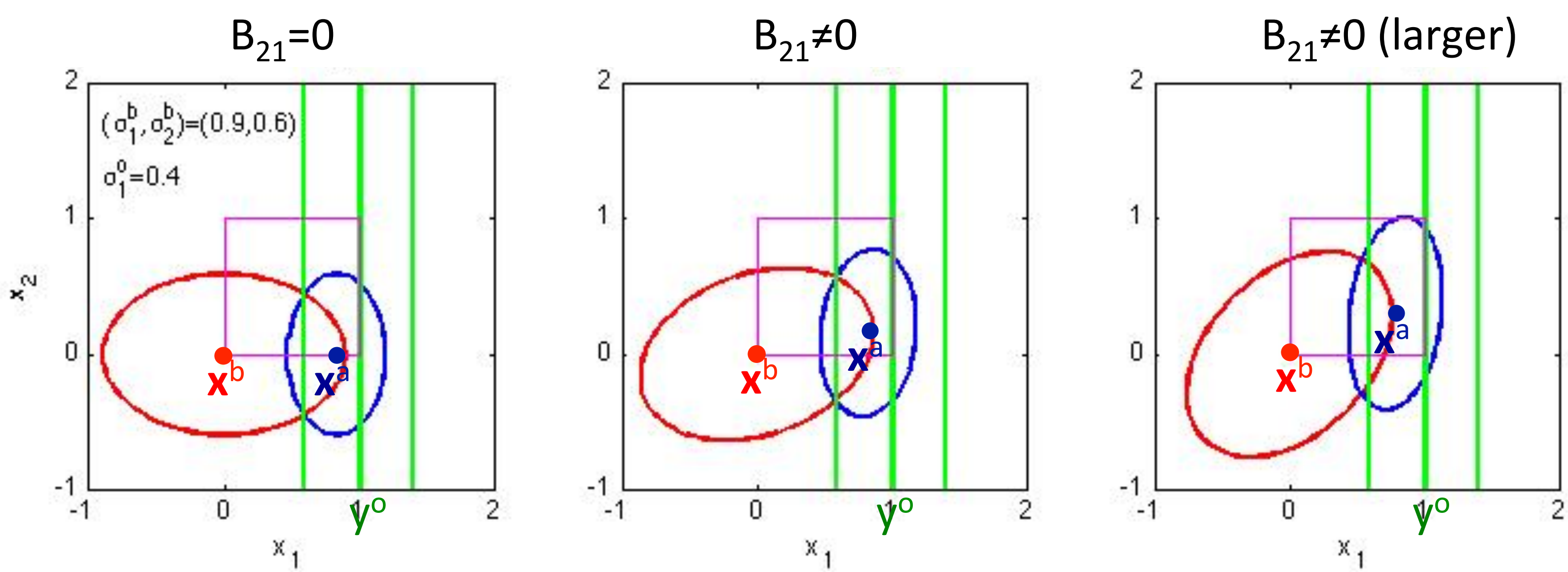
# OI: 2D Example with Effect of Correlation $B_{21}$

- Background $\quad \mathbf{x}^b = \begin{pmatrix} x_1^b \\ x_2^b \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$
- Observation $\quad y^o = x^o{}_1, \quad R \quad$ with $\mathbf{Hx} = [1,0]\mathbf{x}$

→Analysis $\quad \mathbf{x}^{OI} = \mathbf{x}^b + \Delta\mathbf{x}^{OI}$

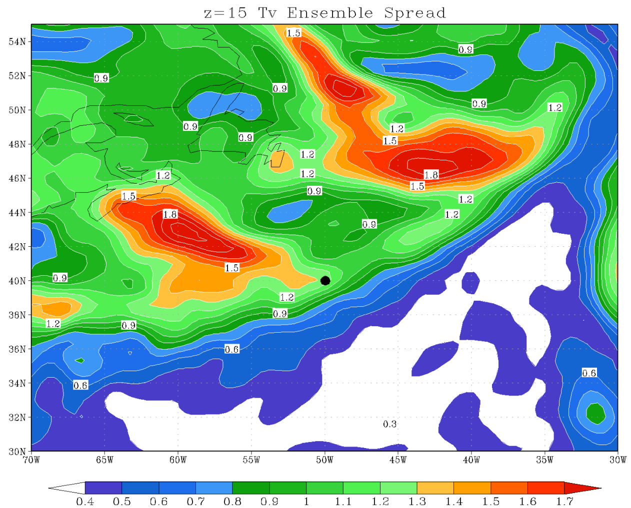
$$\Delta\mathbf{x}^{OI} = \begin{pmatrix} \Delta x_1^{OI} \\ \Delta x_2^{OI} \end{pmatrix} = \frac{1}{B_{11}+R} \begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix} (x_1^o - x_1^b)$$

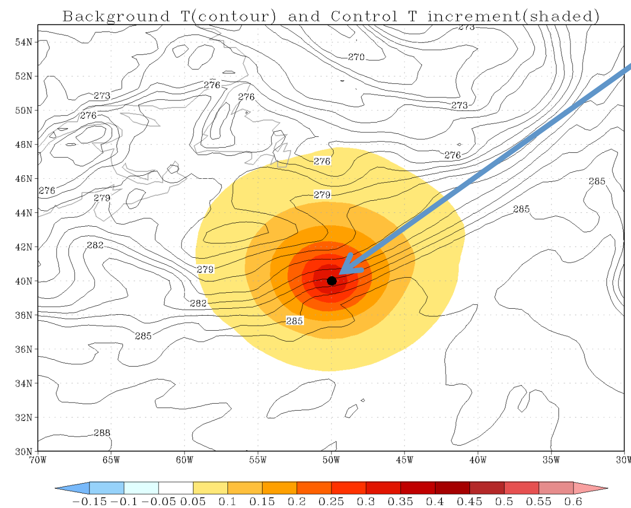
$\Delta\,\mathbf{x}^{OI}$ and $\mathbf{BH}^T$ has the same form



# Optimal Interpolation (OI): GSI Example

- Effect of correlation: **B**→ $\triangle x^{OI}{}_n$ (single obs for $x_l$) with GSI




$x_l$ position

3DVar [=OI]

$$\begin{pmatrix} \Delta x_1^{OI} \\ \vdots \\ \Delta x_l^{OI} \\ \vdots \\ \Delta x_N^{OI} \end{pmatrix} = \frac{1}{R + B_{l\,l}} \begin{pmatrix} B_{1\,l} \\ \vdots \\ B_{l\,l} \\ \vdots \\ B_{N\,l} \end{pmatrix} (y^o - x^b)$$

**B** determines the quality of $\Delta \mathbf{x}^{OI}$

Single 850mb Tv observation (1K O-F, 1K error) – Courtesy of D. Kleist

# Perspectives of Data Assimilation

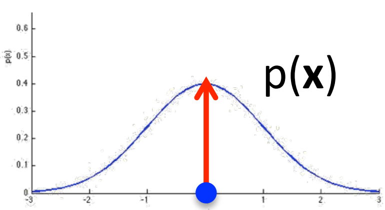
- Two main perspectives of practical data assimilation & hybrid approach

**Variational Approach:**

<u>Least square estimation</u>

[maximum likelihood]

– 3D-Var (3 dim in space)
– 4D-Var ($4^{th}$ dim is time)

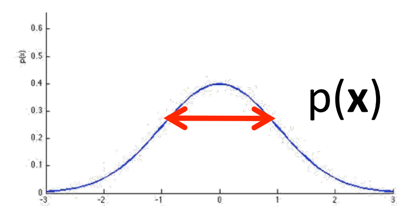



**Sequential (KF) Approach:**

Minimum Variance estimate

[least uncertainty]

– Optimal Interpolation (OI)
– (Extended / Ensemble) **Kalman Filter**




Hybrid

- Data assimilation is fast growing domain
  - Advanced methods are being developed as interdisciplinary science
  - A variety of extension & applications exist and are under development

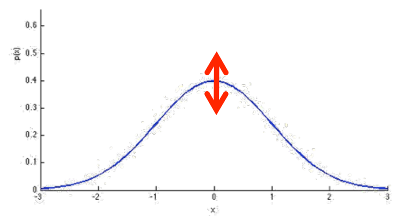
# 3DVar: Setting

Baye's theorem for Maximum Likelihood (ML / MLL)

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}$$

← $p(\mathbf{x},\mathbf{y})=p(\mathbf{x}|\mathbf{y})p(\mathbf{y})=p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$

- Goal: find x that maximizes $p(\mathbf{x}|\mathbf{y})$ obtained by maximum (log)likelihood.



- Base PDFs in RHS.
  - Background (prior $\mathbf{x}$): $\boldsymbol{\varepsilon}^b = \mathbf{x}^b - \mathbf{x}^t$ is Gaussian

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\mathbf{B}|^{1/2}} \exp\left\{-J^b(\mathbf{x})\right\} \qquad J^b(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b)$$

  - Observation (conditional $\mathbf{y}$): $\boldsymbol{\varepsilon}^o = \mathbf{y}^o - \mathbf{y}^t$ is Gaussian

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi)^{L/2}|\mathbf{R}|^{1/2}} \exp\left\{-J^o(\mathbf{x})\right\} \qquad J^o(\mathbf{x}) = \frac{1}{2}(\mathbf{y}^o-\mathbf{h}(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}^o-\mathbf{h}(\mathbf{x}))$$

# 3DVar: Formulation

- Maximum Likelihood (ML) by Bayes' theorem

$$p(\mathbf{x}|\mathbf{y}) = \frac{\exp\{-J(\mathbf{x})\}}{(2\pi)^{(N+L)/2}|\mathbf{R}|^{1/2}|\mathbf{B}|^{1/2}\ p(\mathbf{y})}$$

= Minimum of the Cost function

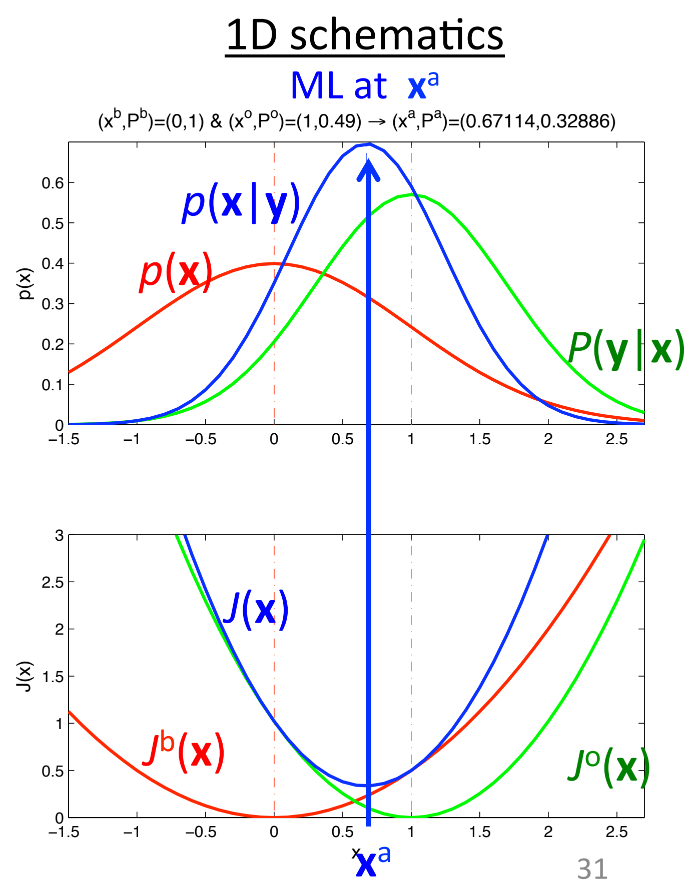
$J(\mathbf{x}) = J^b(\mathbf{x}) + J^o(\mathbf{x})$

$J^b(\mathbf{x}) = (1/2)\ (\mathbf{x}-\mathbf{x}^b)^T\ \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}^b)$

$J^o(\mathbf{x}) = (1/2)\ (\mathbf{y}^o-\mathbf{h}(\mathbf{x}))^T\ \mathbf{R}^{-1}(\mathbf{y}^o-\mathbf{h}(\mathbf{x}))$

Conditions

- $\nabla J(\mathbf{x}) = 0$
- $\nabla^2 J(\mathbf{x})$ = semi-positive definite




- For linear observation **y=Hx**, analytical solution

$\mathbf{x}^a = \mathbf{x}^{OI} \qquad = \mathbf{x}^{3DVar}$

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

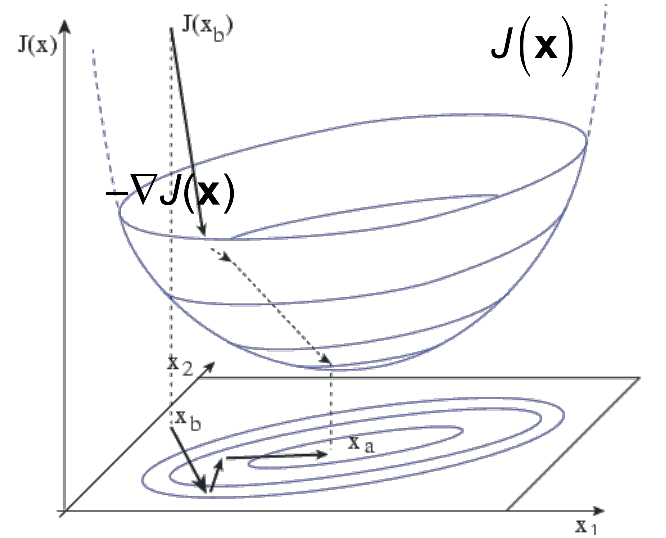
31

# 3DVar: Computational Approach

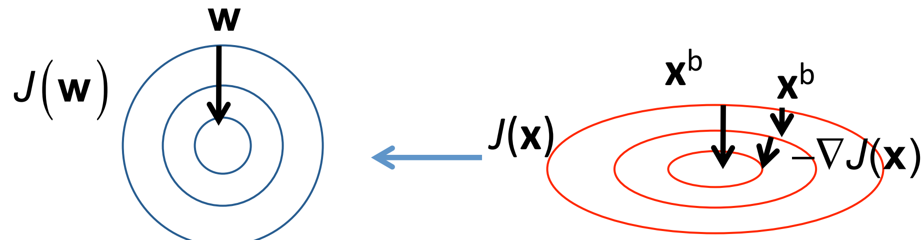
- Computational algorithm: Given
  - Cost function $J(\mathbf{x})$
  - Gradient function $\nabla J(\mathbf{x})$
  - Initial "guess" $\mathbf{x}^b$

  Using algorithms such as
  - Conjugate gradient
  - Quasi-Newton method

ECMWF Technical Note




- Efficiency
  - Pre-conditioning
    - Change of coordinate **x** to **w**




    - (In)homogeneity around the minimum: Controlled by the Hessian of $J(\mathbf{x})$
  - For nonlinear forward model for observation,
    - Inner loop by incremental 3DVar (linearization of nonlinear model)
    - Occasional outer loop with full nonlinear forward model

# 3DVar: Additional Topics

- Flexibility of $J(\mathbf{x}) = J^b(\mathbf{x}) + J^o(\mathbf{x})$
  - Quadratic, under the Gaussian assumption for the error distribution but can be other functional form

    → Variational quality control for observations
  - Can add constraint : $J(\mathbf{x}) = J^b(\mathbf{x}) + J^o(\mathbf{x}) + J^c(\mathbf{x})$

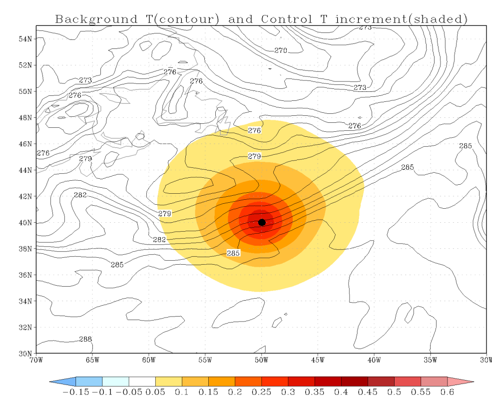
    → Reduction of unwanted fast gravity waves
  - Can use variable transformation (weak constraint)

    → Improved use of observations & preservation of dynamic balance

- Challenges: Modeling of static **B** (same for OI)
  - Cross-variable correlations

    ←Variable transformation for better representation
  - Mostly homogenous

    (can add some flow dependence but not much)



# Outline

- Objectives
- Background
  - State **x** & Observation **y** - also **m**(**x**) & **h**(**x**)
  - Probability $p(\mathbf{x})$ - also $p(\mathbf{x}|\mathbf{y})$ & $p(\mathbf{y}|\mathbf{x})$
  - Data assimilation perspectives
- 3D Method - 3 dimensions in space
  - OI = Optimal Interpolation
  - 3DVar =Variational
- 3D to 4D - $4^{th}$ dimension is time
  - EKF/EnKF= Extended/Ensemble Kalman filter -- Dynamics
  - FGAT = First Guess at Appropriate Time -- Obs
  - 4DVar
  - Hybrid = between Var and EnKF - Current operational system
- Concluding remarks

# Kalman Filters: Extension to Sequential Methods

- By explicitly targeting **P**, the analytical form of $\mathbf{P}^{OI}$ (=**A**) is also available.

- OI itself doesn't require the computation of $\mathbf{P}^{OI}$.

  $\mathbf{x}^b_k \rightarrow \mathbf{x}^a_k$ [**B** is static, no need for **A**]

- Extended/Ensemble Kalman filter makes use of **A**,

  $(\mathbf{x}^b_k, \mathbf{B}_k) \rightarrow (\mathbf{x}^a_k, \mathbf{A}_k)$ [$\mathbf{A}_k$ is obtained along with $\mathbf{x}^a_k$]

  to estimate $\mathbf{B}_{k+1}$ in the next assimilation cycle

  $(\mathbf{x}^a_k, \mathbf{A}_k) \rightarrow (\mathbf{x}^b_{k+1}, \mathbf{B}_{k+1})$ [Model needs to forecast $\mathbf{A}_k$ to $\mathbf{B}_{k+1}$]

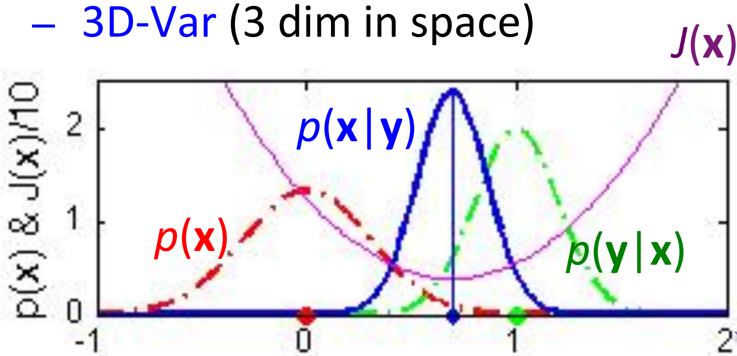
# “4D”ness in Dynamic Forecast: Extended Kalman Filter

**Variational Approach:**

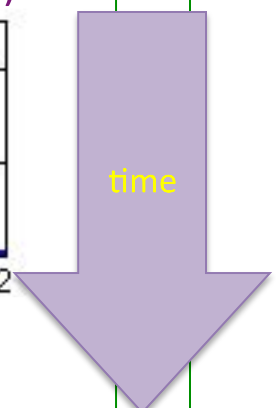
Least square estimation

[maximum likelihood]

- 3D-Var (3 dim in space)




- 4D-Var (4th dim is time)

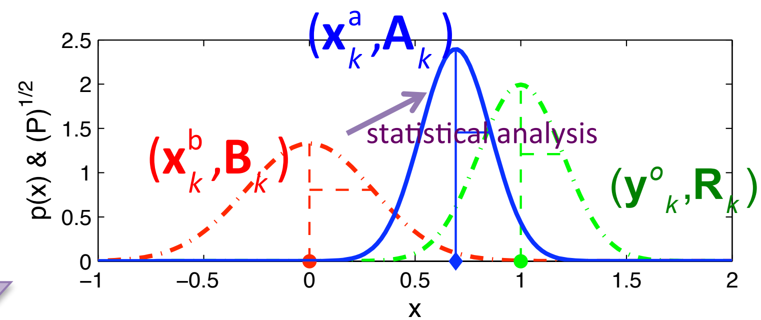



**“Sequential” Approach:**

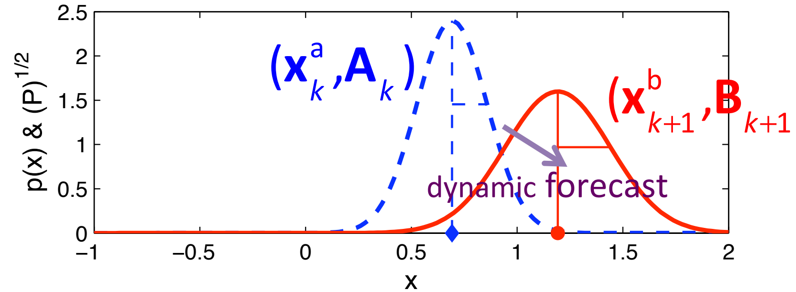
Minimum Variance estimate

[least uncertainty]

- Optimal Interpolation (OI)




- Extended Kalman Filter




**x** forecast: Nonlinear model

**A**->**B** forecast: Tangent linear model (TLM)

# Sequential Approach: Extended Kalman Filter

- Basic concept for dynamic estimation of $\mathbf{A}_{k-1}$ to $\mathbf{B}_k$ by perturbation growth
- Extended Kalman Filter (EKF)
  - Error covariance $\mathbf{P}_k = E[\boldsymbol{\varepsilon}_k (\boldsymbol{\varepsilon}_k)^T]$ evolution using Tangent Linear Model (TLM)

$$\mathbf{P}_k = \mathbf{M}_{k,k-1}\mathbf{P}_{k-1} (\mathbf{M}_{k,k-1})^T$$

$$[\boldsymbol{\varepsilon}_k = \mathbf{M}_{k,k-1}\boldsymbol{\varepsilon}_{k-1}]$$

  - Formulation

Step 1. Forecast

$$\mathbf{x}_k^{\mathrm{b}} = \mathbf{m}_{k,k-1}(\mathbf{x}_{k-1}^{\mathrm{a}})$$

$$\mathbf{B}_k = \mathbf{M}_{k,k-1}\mathbf{A}_k\mathbf{M}_{k,k-1}^T \quad [+\mathbf{Q}_k]$$

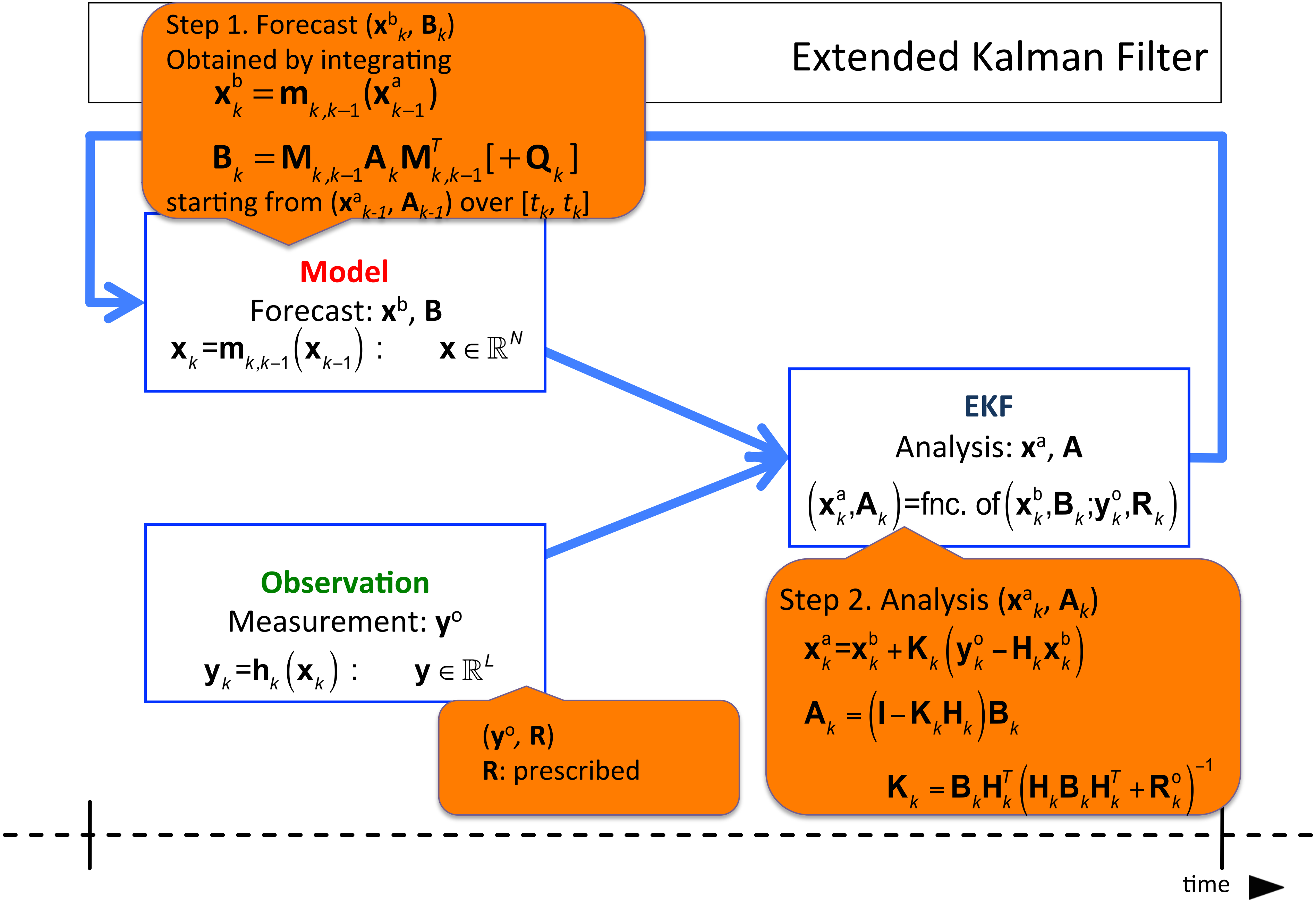
Step 2. Analysis

$$\mathbf{x}_k^{\mathrm{a}} = \mathbf{x}_k^{\mathrm{b}} + \mathbf{K}_k\left(\mathbf{y}_k^{\mathrm{o}} - \mathbf{H}_k\mathbf{x}_k^{\mathrm{b}}\right)$$

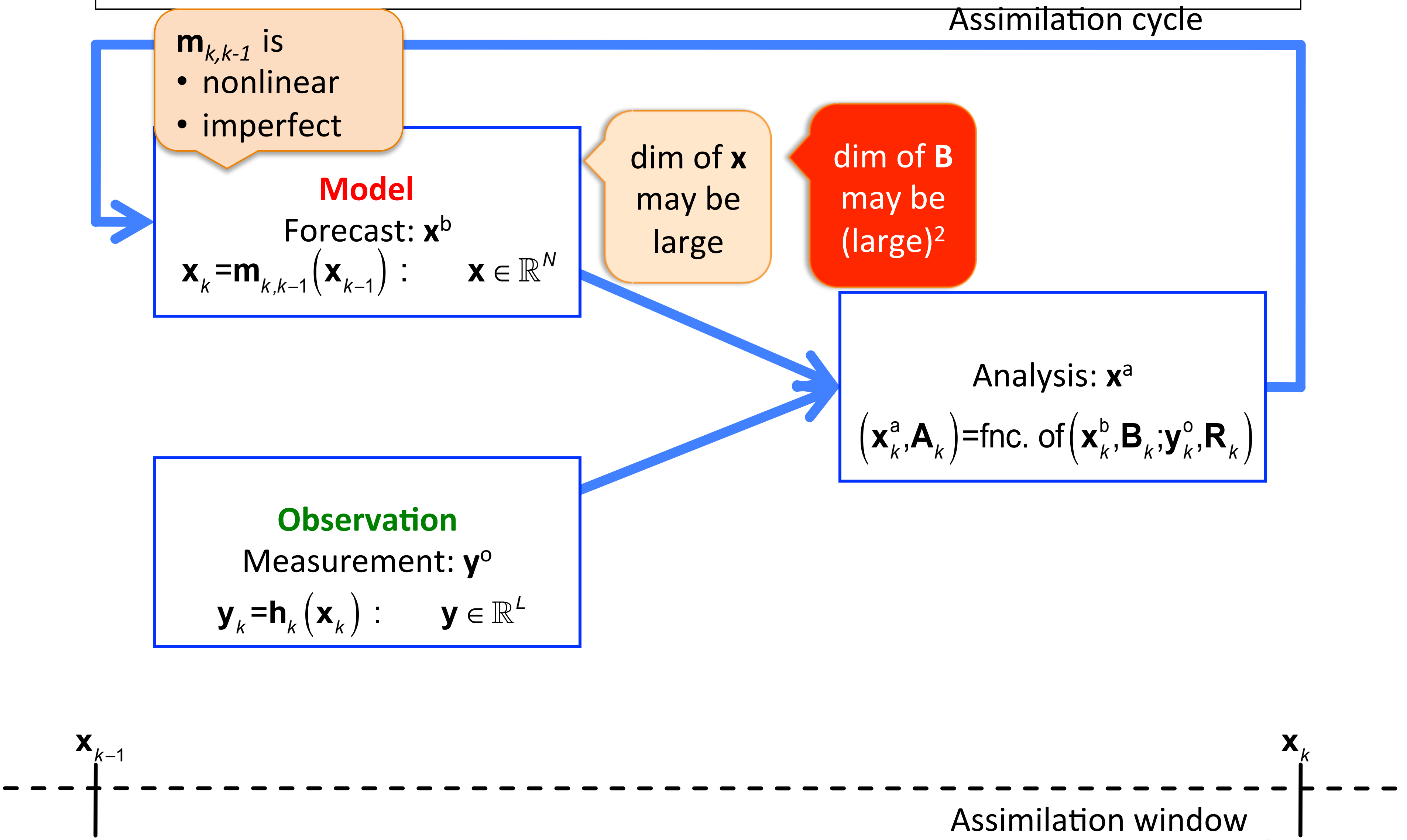
$$\mathbf{A}_k = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{B}_k$$

$$\mathbf{K}_k = \mathbf{B}_k\mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{B}_k\mathbf{H}_k^T + \mathbf{R}_k^{\mathrm{o}}\right)^{-1}$$

# Extended Kalman Filter



# Extended Kalman Filter: Challenges



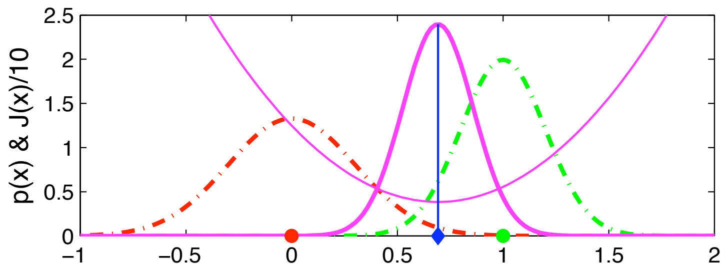
# "4D"ness in Dynamic Forecast: Ensemble Kalman Filter (EnKF)

**Variational Approach:**

Least square estimation

[maximum likelihood]

– 3D-Var (3 dim in space)



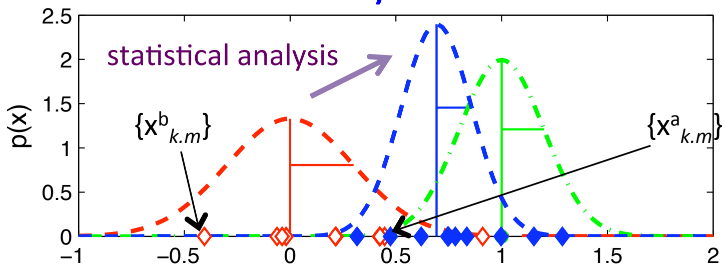
– 4D-Var ($4^{th}$ dim is time)

**"Ensemble Kalman Filter (EnKF)" Approach :**

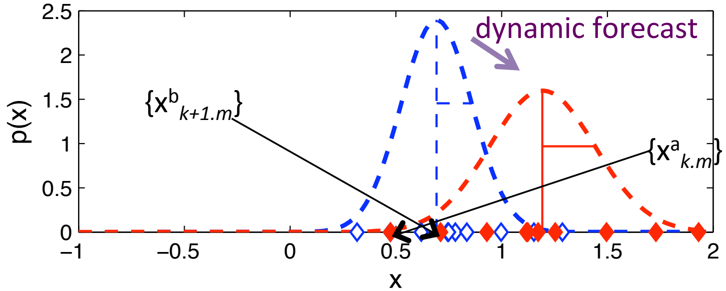
Minimum Variance estimate

use $\{\mathbf{x}_{k.m}\}$ to infer/approx. $p(\mathbf{x},t_k)$

– Ensemble analysis




– Ensemble forecast




$$\{\mathbf{x}_{k.m}\}=(\mathbf{x}_{k.1},\ldots,\mathbf{x}_{k.M}) \in \mathbb{R}^{N\times M}$$

$k$: time $t_k$ $M$: ensemble size

# Ensemble Kalman Filter

Step 1. Forecast $\{\mathbf{x}^{b(m)}{}_k\}$

Obtained by integrating

$$\mathbf{x}_k^{b(m)} = \mathbf{m}_{k,k-1}(\mathbf{x}_{k-1}^{a(m)})$$

starting from $\mathbf{x}^{a(m)}{}_{k-1}$ over $[t_k, t_k]$

**Model**

Forecast: $\mathbf{x}^b$

$$\mathbf{x}_k = \mathbf{m}_{k,k-1}\left(\mathbf{x}_{k-1}\right) : \qquad \mathbf{x} \in \mathbb{R}^N$$

**EnKF**

Analysis: $\mathbf{x}^a$

$$\left(\mathbf{x}_k^a, \mathbf{A}_k\right) = \text{fnc. of} \left(\mathbf{x}_k^b, \mathbf{B}_k; \mathbf{y}_k^o, \mathbf{R}_k\right)$$

**Observation**

Measurement: $\mathbf{y}^o$

$$\mathbf{y}_k = \mathbf{h}_k\left(\mathbf{x}_k\right) : \qquad \mathbf{y} \in \mathbb{R}^L$$

$(\mathbf{y}^o, \mathbf{R})$

**R**: prescribed

Step 2. Analysis $(\mathbf{x}^{a(m)}{}_k\}$ that achieves

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k\left(\mathbf{y}_k^o - \mathbf{H}_k\mathbf{x}_k^b\right)$$

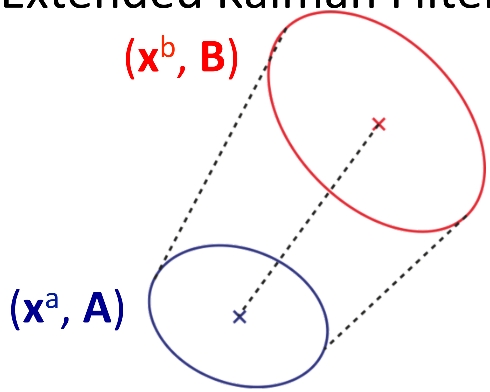
$$\mathbf{A}_k = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{B}_k$$

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

time

# Two Main Branches of EnKF: Analysis Processes at a Fixed $t_k$

- Extended Kalman Filter



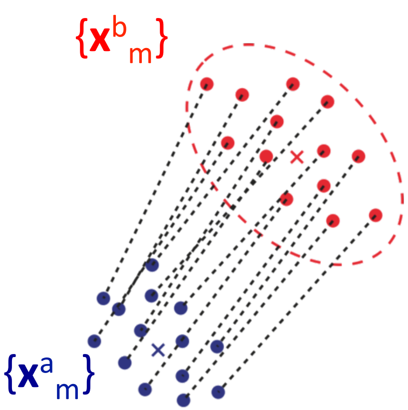

Given $\mathbf{x}^b$, $\mathbf{B}$, $\mathbf{y}^o$, $\mathbf{R}$, $\mathbf{h}$ / $\mathbf{H}$

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

Obtain $\mathbf{x}^a = \mathbf{x}^b + \Delta\mathbf{x}^a$ ; $\Delta\mathbf{x}^a = \mathbf{K}\,(\mathbf{y}^o - \mathbf{h}(\mathbf{x}^b))$

& $\mathbf{A} = (\mathbf{I} - \mathbf{K}\,\mathbf{H})\mathbf{B}$

- Stochastic EnKF approach
  - Perturbed Observation (PO)
    (Houtekamer & Mitchell, MWR, 1998)
    (Burgers et al, MWR, 1998)

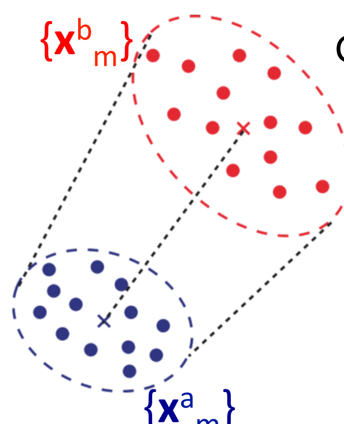



Given $\{\mathbf{x}^b{}_m\}$, $\mathbf{y}^o$, $\mathbf{R}$, $\mathbf{h}$ / $\mathbf{H}$

Apply $\mathbf{K}$ to $\{\mathbf{x}^b{}_m\}$

Obtain $\{\mathbf{x}^a{}_m\}$

- Square-Root EnKF approach
  - Serial Ensemble Square Root Filter (EnSRF)
    (Whitaker & Hamill, MWR, 2002)
  - [Local] ensemble transform KF ([L]ETKF)
    (Hunt et al, Physica D, 2007)




Given $\{\mathbf{x}^b{}_m\}$, $\mathbf{y}^o$, $\mathbf{R}$, $\mathbf{h}$ (/ $\mathbf{H}$)

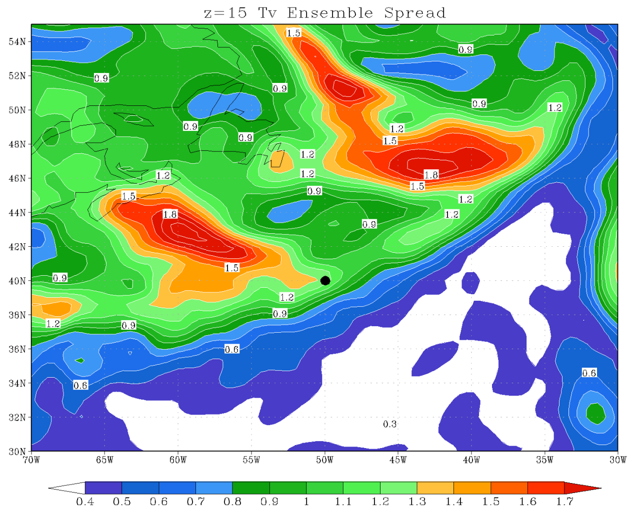
Apply $\mathbf{K}$ to $\mathbf{x}^b$, $\mathbf{B}$

Obtain $\mathbf{x}^a$, $\mathbf{A}$

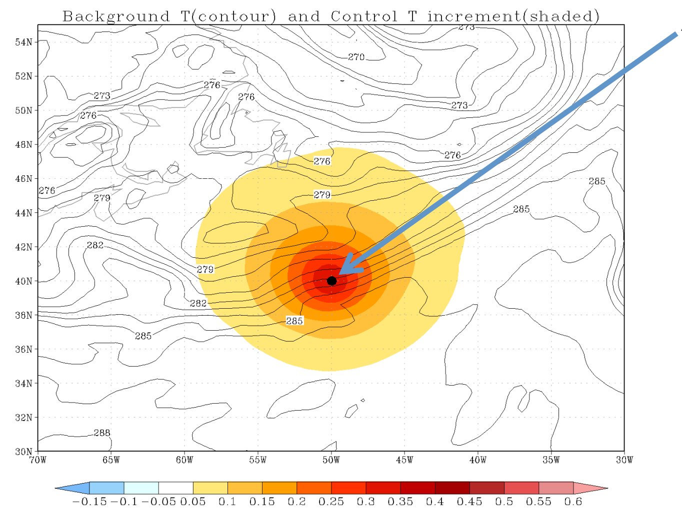
& hence $\{\mathbf{x}^a{}_m\}$

# 3DVar/OI vs EnKF: GSI Example

- Effect of correlation: $\boldsymbol{B} \Rightarrow \triangle x_n$ (single obs for $x_l$) with GSI

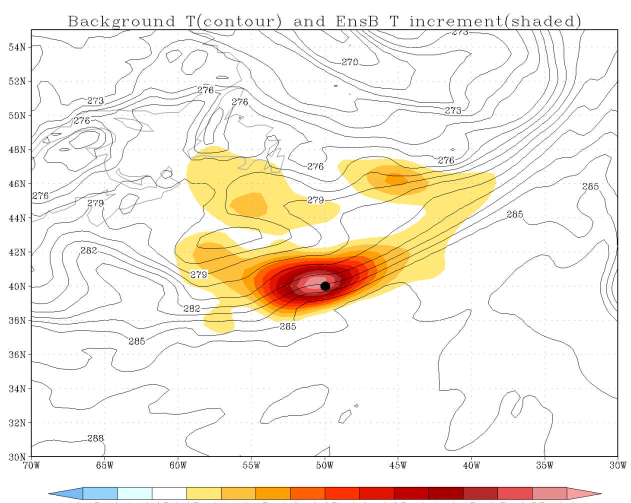



*$x_l$ position*




3DVAR

EnKF




$$\begin{pmatrix} \Delta x_1^a \\ \vdots \\ \Delta x_l^a \\ \vdots \\ \Delta x_N^a \end{pmatrix} = \frac{1}{R + B_{l\,l}} \begin{pmatrix} B_{1\,l} \\ \vdots \\ B_{l\,l} \\ \vdots \\ B_{N\,l} \end{pmatrix} (y^o - x^b)$$

**B** determines the quality of $\Delta\mathbf{x}^a$

Single 850mb Tv observation (1K O-F, 1K error) – Courtesy of D. Kleist

# Var Representation of EnKF: **L** and ρ

- 3DVAR: Static $\mathbf{B}_{3DVAR}$

$$J_{3DVAR}(\Delta\mathbf{x}) = J^{b}_{3DVAR}(\Delta\mathbf{x}) + J^{o}(\Delta\mathbf{x}) \; ; \qquad \Delta\mathbf{x} = \mathbf{x} - \mathbf{x}^{b} \quad \& \quad \mathbf{d} = \mathbf{y}^{o} - \mathbf{h}(\mathbf{x}^{b})$$

$$J^{b}_{3DVAR}(\Delta\mathbf{x}) = \frac{1}{2}(\Delta\mathbf{x})^{T}(\mathbf{B}_{3DVAR})^{-1}(\Delta\mathbf{x});$$

$$J^{o}(\Delta\mathbf{x}) = \frac{1}{2}(\mathbf{d}^{o} - \mathbf{H}\Delta\mathbf{x})^{T}(\mathbf{R}^{o})^{-1}(\mathbf{d}^{o} - \mathbf{H}\Delta\mathbf{x})$$

- EnKF flow evolving covariance estimation

$$J_{EnKF}(\alpha) = J^{b}_{EnKF}(\alpha) + J^{o}_{EnKF}(\alpha) \; ;$$

$$J^{b}_{EnKF}(\alpha) = \frac{1}{2\rho}(\alpha)^{T}(\mathbf{L})^{-1}(\alpha); \quad J^{o}_{EnKF}(\alpha) = J^{o}(\Delta\mathbf{x})$$

$$\Delta\mathbf{x} = \sum_{m=1}^{M} \alpha^{(m)} \circ \Delta\mathbf{x}^{(m)}$$

$\alpha^{(m)}$: ~local weight of individual members, scaled by $(M\text{-}1)^{1/2}$

**L**: Localization matrix on extended control

ρ: inflation of background covariance matrix

$\Delta\mathbf{x}^{(m)}$ *:* ensemble perturbations

# Full 4D Data Assimilation: Problem Setting

- Goal: Over a window $[t_{k-1}, t_k]=[t_{k,0}, t_{k,I}]$,

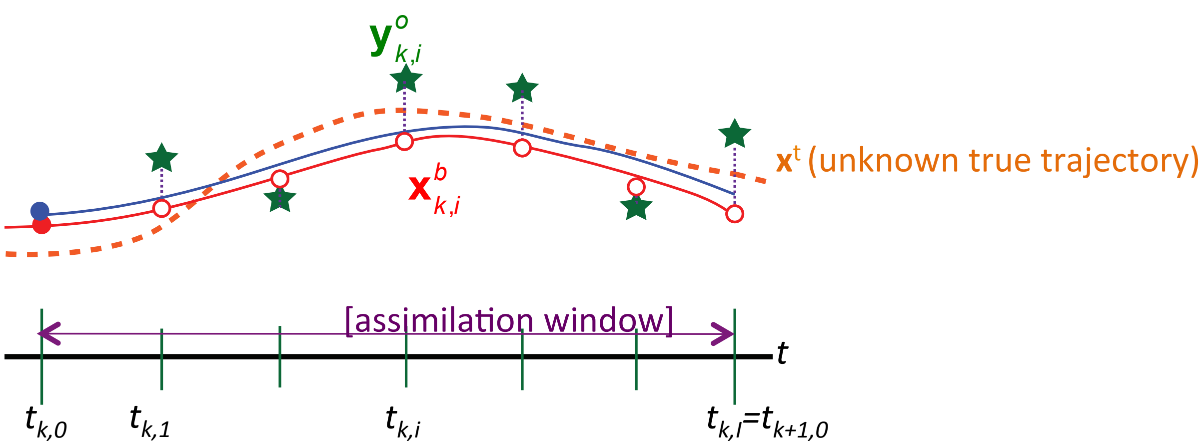
  Obtain the 'best possible' estimate of the unknown true state $\mathbf{x}^t_{k-1}$ ($R^N$) from two information resources

  1. Background by model forecast $\mathbf{x}^b_{k,i}$($R^N$)

     $$\mathbf{x}_{k,i} = \mathbf{m}_k(\mathbf{x}_{k,0})$$

  2. Asynchronous Observations $\mathbf{y}^o_{k,i}$ ($R^L$)

     $$\mathbf{y}_{k,i} = \mathbf{h}_{k,i}(\mathbf{x}_{k,i})$$



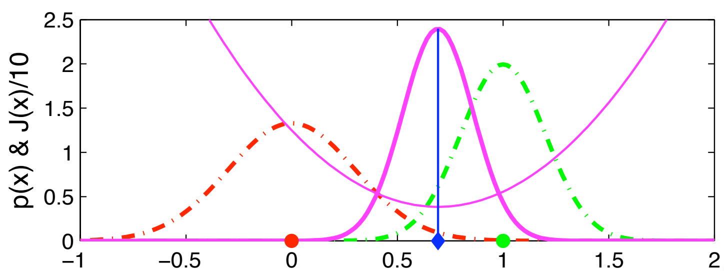
# "4D"ness

**Variational Approach:**

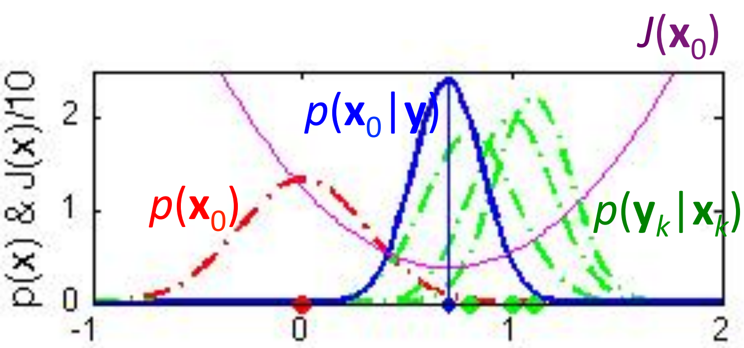
Least square estimation

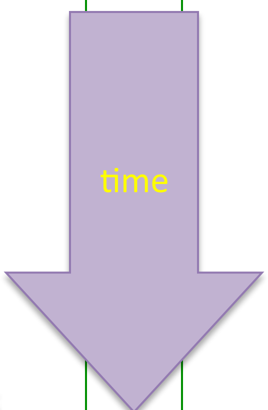
[maximum likelihood]

– 3D-Var (3 dim in space)




– 4D-Var ($4^{th}$ dim is time)



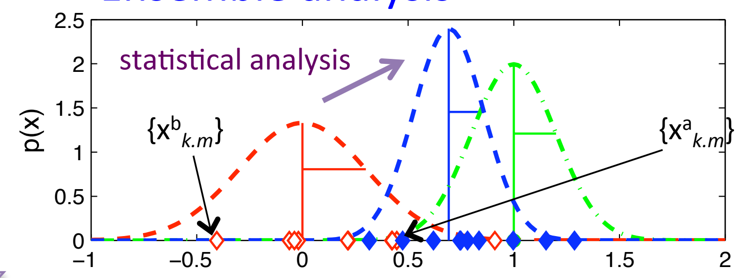




**"Ensemble Kalman Filter (EnKF)" Approach :**

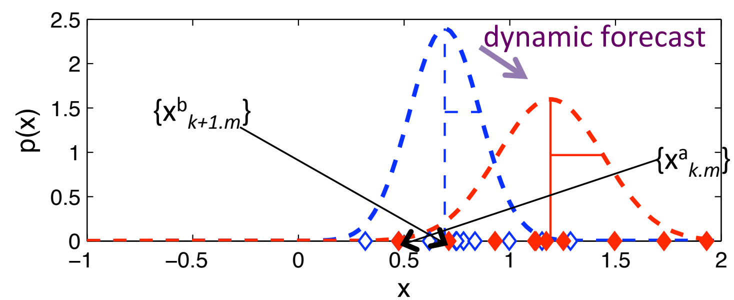
Minimum Variance estimate

use $\{\mathbf{x}_{k.m}\}$ to infer/approx. $p(\mathbf{x},t_k)$

– Ensemble analysis




– Ensemble forecast




$$\{\mathbf{x}_{k.m}\}=(\mathbf{x}_{k.1},\ldots,\mathbf{x}_{k.M})\in\mathbb{R}^{N\times M}$$

$k$: time $t_k$ $M$: ensemble size

# 3DVar Approach to 4D (Asynchronous) Observation: FGAT

- FGAT (First Guess at Appropriate Time)
  - Basic idea: Incorporate the time distribution of $\mathbf{y}^o_{k,i}$ of incremental 3D-Var while keeping the control variable $\Delta\mathbf{x}_k$ at $t_I$

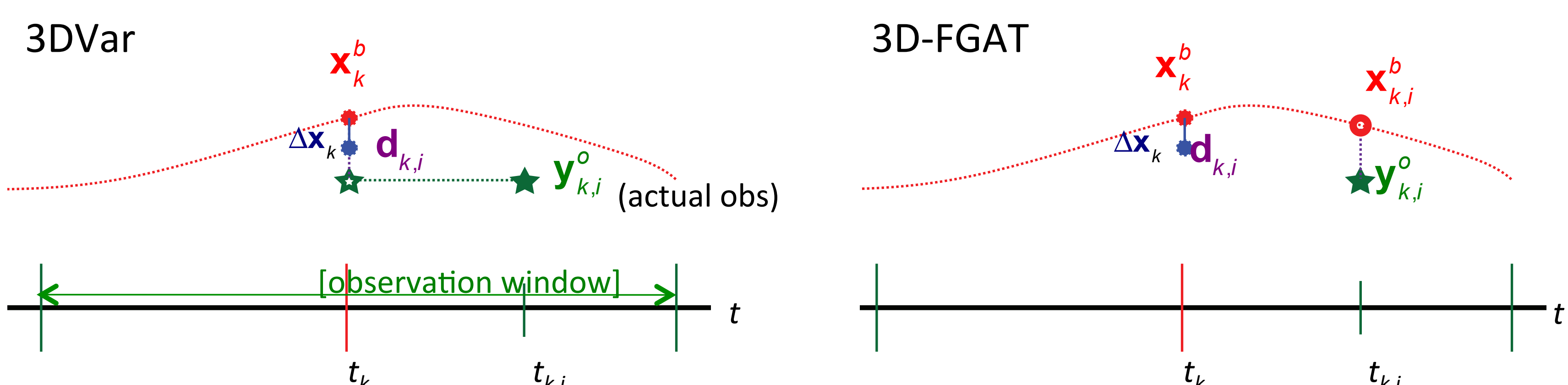
$$J_{3DFGAT}(\Delta\mathbf{x}) = J^b_{3DVAR}(\Delta\mathbf{x}) + J^o_{FGAT}(\Delta\mathbf{x}) \; ; \qquad \Delta\mathbf{x} = \mathbf{x} - \mathbf{x}^b \quad \& \quad \mathbf{d} = \mathbf{y}_{k,i}{}^o - \mathbf{h}(\mathbf{x}_{k,i}{}^b)$$

$$J^b_{3DVAR}(\Delta\mathbf{x}) = \frac{1}{2}(\Delta\mathbf{x})^T \mathbf{B}^{-1}(\Delta\mathbf{x});$$

$$J^o_{FGAT}(\Delta\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{I} (\mathbf{d}_{k,i} - \mathbf{H}_{k,i}\Delta\mathbf{x})^T \mathbf{R}_{k,i}{}^{-1}(\mathbf{d}_{k,i} - \mathbf{H}_{k,i}\Delta\mathbf{x})$$

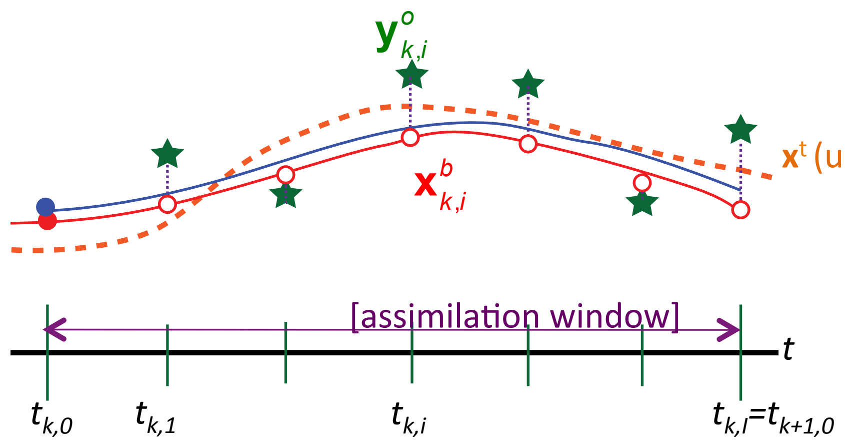
$$\mathbf{d}^{ob}_{i,k} = \mathbf{y}^o_{i,k} - \mathbf{h}_{i,k}(\mathbf{x}^b_{i,k})$$

If $\mathbf{x}^b_{i,k}$ is available (stored)



# Variational Approach to "4D"ness in Dynamics: 4DVar

- Addressing "4D"- ness
  - 4D-Var is the extension of 3D-Var to the time domain and forward model takes into the account the dynamical evolution the model state
  - Two obvious and related advantages of 4D-Var over 3D-Var are:
    - Better representation of the temporal distribution of the observations, like FGAT
    - Time evolution of the model state




- In the 4D-Var, the observation window and the forecast window (in the simplest form) coincide, which we call the assimilation cycle.
  - One assimilation cycle at a time → drop the window index *k*, use only *i*.

# 4DVar: Formulation

- Standard 4D-Var:
  - Cost function

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \sum_{i=1}^{I} \frac{1}{2}(\mathbf{y}_i^o - \mathbf{h}_i(\mathbf{x}_i))^T (\mathbf{R}_i)^{-1}(\mathbf{y}_i^o - \mathbf{h}_i(\mathbf{x}_i))$$

$$= \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \sum_{i=1}^{I} \frac{1}{2}(\mathbf{y}_i^o - \mathbf{h}_i(\mathbf{m}_{i,0}(\mathbf{x}_0)))(\mathbf{R}_i)^{-1}(\mathbf{y}_i^o - \mathbf{h}_i(\mathbf{m}_{i,0}(\mathbf{x}_0)))$$

$$= J^b(\mathbf{x}_0) + J^o(\mathbf{x}_0)$$

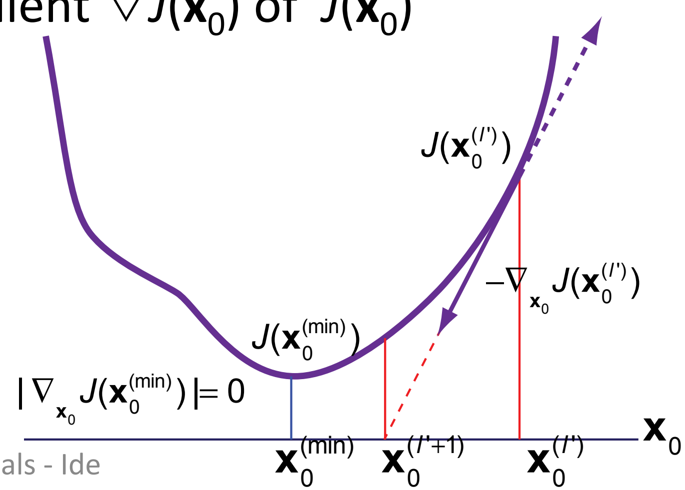
$$= J^b(\mathbf{x}_0) + \sum_{i=1}^{I} J_i^o(\mathbf{x}_0)$$

$$\nabla J(\mathbf{x}_0) = \nabla J^b(\mathbf{x}_0) + \sum_{i=1}^{I} \nabla J_i^o(\mathbf{x}_0)$$

$$\nabla J_i^o(\mathbf{x}_0) = -(\mathbf{H}_i \mathbf{M}_{i,0})^T (\mathbf{R}_i)^{-1}(\mathbf{y}_i^o - \mathbf{h}_i(\mathbf{m}_{i,0}(\mathbf{x}_0)))$$

  - Minimization algorithms require the gradient $\nabla J(\mathbf{x}_0)$ of $J(\mathbf{x}_0)$

    → Adjoint $\mathbf{M}_{i,0}$ of nonlinear model $\mathbf{m}_{i,0}$



# 4DVar: Additional Topics

- Flexibility of $J(\mathbf{x}) = J^b(\mathbf{x}) + J^o(\mathbf{x})$ $[+ J^c(\mathbf{x})]$
  - Similar to 3DVar, once minimization of $J(\mathbf{x})$ is in place.
    - Variational quality control
    - Constraints
    - Preconditioning
    - Outer loop (for nonlinearity in both dynamic and forward models)

- Challenges: Modeling of static **B** (similar to 3DVar)
  - 4DVar is not a sequential method: **B** is still static though evolves over the assimilation window, effectively just like EKF, i.e., $\mathbf{B}_{i,0} = \mathbf{M}_{i,0}\,\mathbf{B}\,\mathbf{M}_{i,0}^{\mathrm{T}}$
  - Computationally intensive: in particular TLM $\mathbf{M}_{i,0}$ and its adjoint $\mathbf{M}_{i,0}^{\mathrm{T}}$

# Integrated Approach: Hybrid

- Incremental 4DVar with Static $\mathbf{B}_f$

$$J(\Delta\mathbf{x}_0) = \frac{1}{2}(\Delta\mathbf{x}_0)^T\mathbf{B}^{-1}(\Delta\mathbf{x}_0) + \sum_{i=1}^{I} \frac{1}{2}(\mathbf{d}_i^o - \mathbf{H}_i\Delta\mathbf{x}_i)^T(\mathbf{R}_i)^{-1}(\mathbf{d}_i^o - \mathbf{H}_i\Delta\mathbf{x}_i)$$

- EnKF in Var formulation with dynamic $\mathbf{B}_e$

$$J_{\text{EnKF}}(\alpha) = \frac{1}{2\rho}\alpha^T\mathbf{L}^{-1}\alpha + \sum_{i=1}^{I} \frac{1}{2}(\mathbf{d}_i^o - \mathbf{H}_i\Delta\mathbf{x}_i)^T(\mathbf{R}_i)^{-1}(\mathbf{d}_i^o - \mathbf{H}_i\Delta\mathbf{x}_i)\ ;$$

$$\Delta\mathbf{x}_i = \sum_{m=1}^{M}\alpha^{(m)} \circ \Delta\mathbf{x}_i^{(m)}$$

- Hybrid: with $\mathbf{P}^b = \beta_f\mathbf{B}_f + \beta_e\mathbf{B}_e$

$$J_{\text{hybrid}}(\Delta\mathbf{x}_f, \alpha) = \beta_f\frac{1}{2}(\Delta\mathbf{x}_0)^T\mathbf{B}^{-1}(\Delta\mathbf{x}_0) + \beta_e\frac{1}{2\rho}\alpha^T\mathbf{L}^{-1}\alpha$$

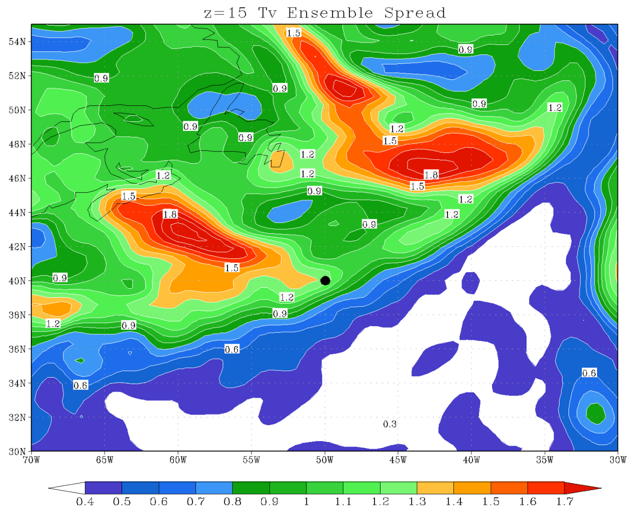
$$+\sum_{i=1}^{I} \frac{1}{2}(\mathbf{d}_i^o - \mathbf{H}_i\Delta\mathbf{x}_i)^T(\mathbf{R}_i)^{-1}(\mathbf{d}_i^o - \mathbf{H}_i\Delta\mathbf{x}_i)$$

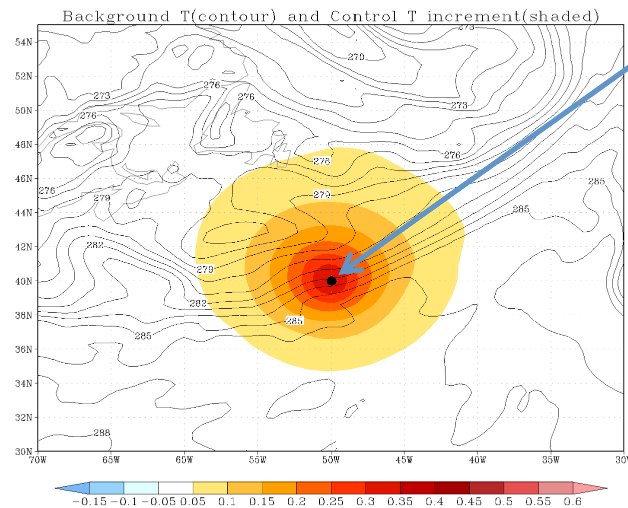
$$1/\beta_f + 1/\beta_e = 1$$

$$\Delta\mathbf{x}_i = \Delta\mathbf{x}_f + \sum_{m=1}^{M}\alpha^{(m)} \circ \Delta\mathbf{x}_i^{(m)}$$

# Hybrid: GSI Example

- Effect of correlation: $\mathbf{P}^b \Rightarrow \triangle \mathbf{x}_n$ (single obs for $x_l$) with GSI



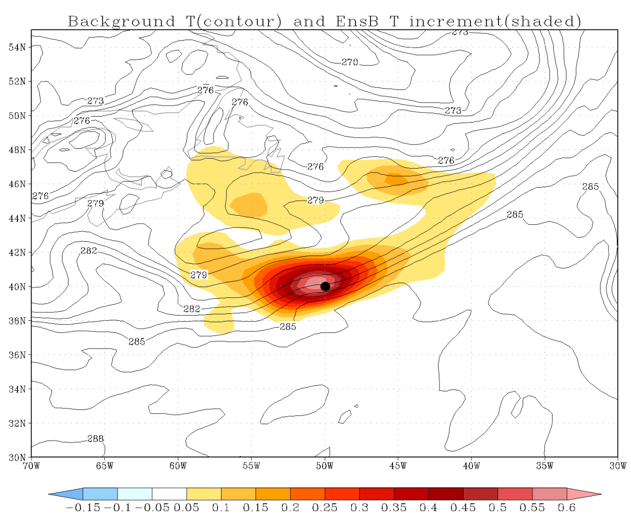

$x_l$ *position*

3DVAR

$\beta_e$=0
$\beta_f$=1

EnKF

$\beta_e$=1
$\beta_f$=0



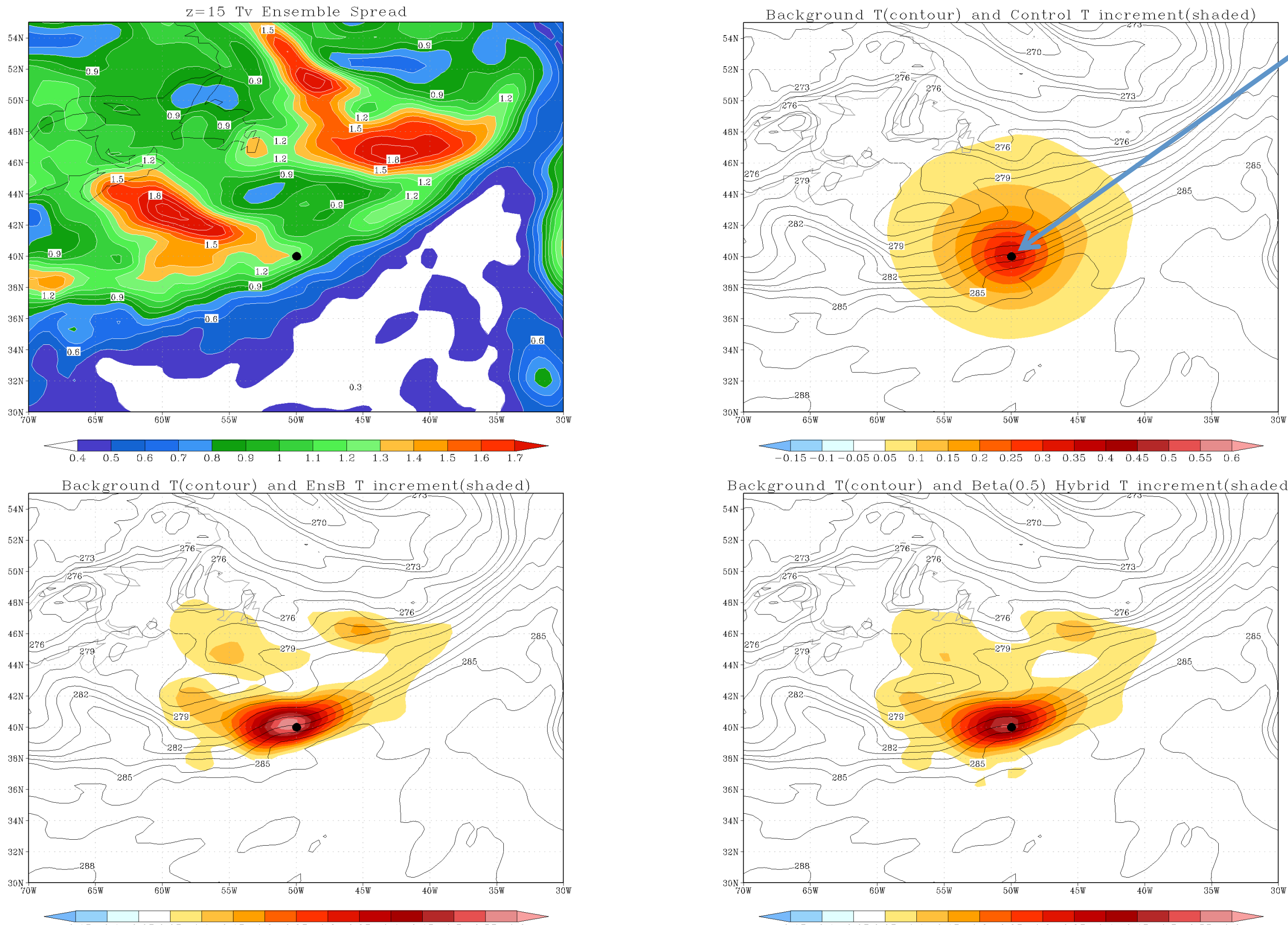

$$\Delta \mathbf{x}^a = \begin{pmatrix} P^b_{1l} / P^b_{ll} \\ \vdots \\ P^b_{ll} / P^b_{ll} \\ \vdots \\ P^b_{Nl} / P^b_{ll} \end{pmatrix} \frac{P^b_{ll}}{P^b_{ll} + R_{ll}} (y^o_l - x^b_l)$$

$\mathbf{P}^b = \beta_f \mathbf{B}_f + \beta_e \mathbf{B}_e$ : hybrid

Single 850mb Tv observation (1K O-F, 1K error) – Courtesy of D. Kleist

# Hybrid: GSI Example

- Effect of correlation: $\mathbf{P}^b \Rightarrow \triangle \mathbf{x}_n$ (single obs for $x_l$) with GSI




Single 850mb Tv observation (1K O-F, 1K error) – Courtesy of D. Kleist

# Summary

- Background ideas of data assimilation
- Unifying perspectives of the current approaches
- Understanding of the variational formulation
  - Cost function

    $J(\mathbf{x}) = J^b(\mathbf{x}) + J^o(\mathbf{x})$  $[+ J^c(\mathbf{x})]$

    $J^b(\mathbf{x}) = 1/2(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b)$ : *b*ackground cost function

    $J^o(\mathbf{x}) = 1/2(\mathbf{y}^o - \mathbf{h}(\mathbf{x}))^T \mathbf{R}^{1}(\mathbf{y}^o - \mathbf{h}(\mathbf{x}))$ : *o*bservation cost function

    $J^c(\mathbf{x})$ : *c*onstraint cost function

    Where ($\mathbf{x}$, $\mathbf{B}$): model state vector & background error covariance matrix

    ($\mathbf{y}$, $\mathbf{R}$): observation vector& observation error covariance matrix

    $\mathbf{y}=\mathbf{h}(\mathbf{x})$ : forward model / observation operator
  - With emphasis on
    - Role of **B**
    - Flexibility of variational approach