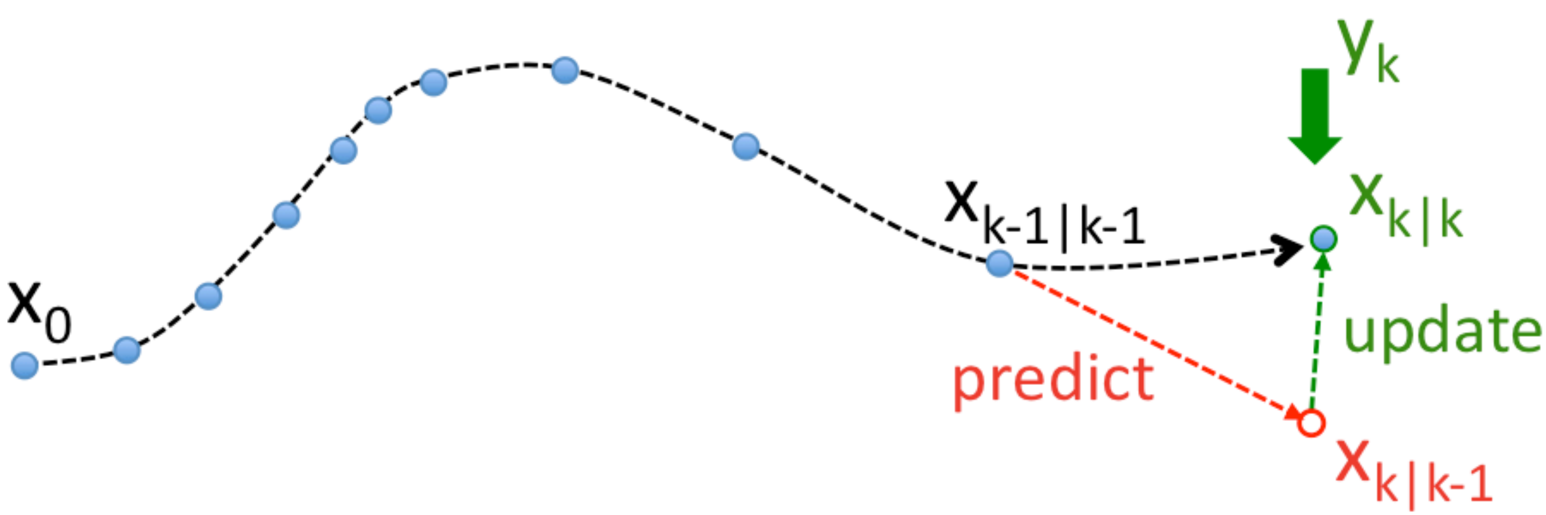
How do we solve it and what does the solution look like?

KF/PFs offer solutions to dynamical systems, nonlinear in general, using prediction and update as data becomes available. Tracking in time or space offers an ideal framework for studying KF/PF.



## The Model

Consider the discrete, linear system,

$$\mathbf{x}_{k+1} = \mathbf{M}_k \mathbf{x}_k + \mathbf{w}_k, \quad k = 0, 1, 2, \ldots, \tag{1}$$

where

- $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector at time $t_k$
- $\mathbf{M}_k \in \mathbb{R}^{n \times n}$ is the state transition matrix (mapping from time $t_k$ to $t_{k+1}$) or model
- $\{\mathbf{w}_k \in \mathbb{R}^n; k = 0, 1, 2, \ldots\}$ is a white, Gaussian sequence, with $\mathbf{w}_k \sim N(\mathbf{0}, \mathbf{Q}_k)$, often referred to as model error
- $\mathbf{Q}_k \in \mathbb{R}^{n \times n}$ is a symmetric positive definite covariance matrix (known as the model error covariance matrix).



Some of the following slides are from: Sarah Dance, University of Reading

## The Observations

We also have discrete, linear observations that satisfy

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \quad k = 1, 2, 3, \ldots, \tag{2}$$

where

- $\mathbf{y}_k \in \mathbb{R}^p$ is the vector of actual measurements or observations at time $t_k$
- $\mathbf{H}_k \in \mathbb{R}^{n \times p}$ is the observation operator. Note that this is not in general a square matrix.
- $\{\mathbf{v}_k \in \mathbb{R}^p; k = 1, 2, \ldots\}$ is a white, Gaussian sequence, with $\mathbf{v}_k \sim N(\mathbf{0}, \mathbf{R}_k)$, often referred to as observation error.
- $\mathbf{R}_k \in \mathbb{R}^{p \times p}$ is a symmetric positive definite covariance matrix (known as the observation error covariance matrix).

We assume that the initial state, $\mathbf{x}_0$ and the noise vectors at each step, $\{\mathbf{w}_k\}$, $\{\mathbf{v}_k\}$, are assumed mutually independent.

# Summary of the Kalman filter

Prediction step

Mean update: $\widehat{\mathbf{x}}_{k+1|k} = \mathbf{M}_k \widehat{\mathbf{x}}_{k|k}$

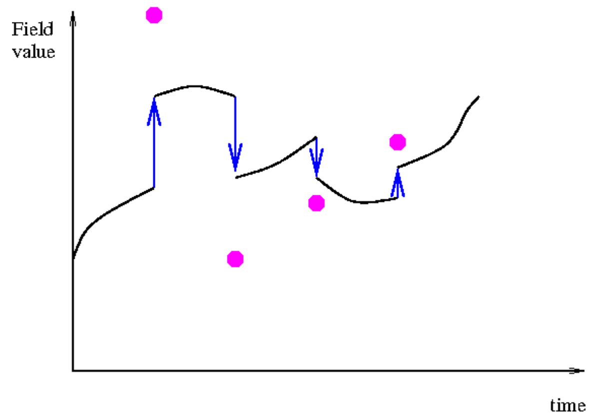
Covariance update: $\mathbf{P}_{k+1|k} = \mathbf{M}_k \mathbf{P}_{k|k} \mathbf{M}_k^T + \mathbf{Q}_k.$

Observation update step

Mean update: $\widehat{\mathbf{x}}_{k|k} = \widehat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k \widehat{\mathbf{x}}_{k|k-1})$

Kalman gain: $\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k)^{-1}$

Covariance update: $\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}.$



# CAGLARS HOMEWORK

1. Assume that we are trying to track a whale. We assume that the whale moves up and down relatively slowly and is going at a constant velocity radially. Show that state equation for the whale (*s*) can be given as below where *z, r*, and *v* are depth, range, and the speed of the whale, $v_z$ and $v_a$ are vertical depth errors and radial acceleration error terms. Δt is the time between measurement k and k-1.

$$\mathbf{s}_k = \mathbf{F}^{\mathbf{s}}_{k-1}\mathbf{s}_{k-1} + \mathbf{B}^{\mathbf{s}}_{k-1}\mathbf{v}^{\mathbf{s}}{}_{k-1}$$

$$\begin{bmatrix} z_s \\ r_s \\ v_s \end{bmatrix}_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_s \\ r_s \\ v_s \end{bmatrix}_{k-1} + \begin{bmatrix} 1 & 0 \\ 0 & \frac{\Delta t^2}{2} \\ 0 & \Delta t \end{bmatrix} \begin{bmatrix} \mathrm{v}_{z_s} \\ \mathrm{v}_{a_s} \end{bmatrix}_{k-1}$$

We use sonar that measures r and z at every k with zero mean Gaussian measurement noise values w:

$$y_k = \begin{bmatrix} z \\ r \end{bmatrix}_k + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_k$$

## Prediction step

We first derive the equation for one-step prediction of the mean using the state propagation model (1).

$$\begin{aligned}\widehat{\mathbf{x}}_{k+1|k} &= \mathbb{E}\left[\mathbf{x}_{k+1}|\mathbf{y}_1, \ldots \mathbf{y}_k\right], \\ &= \mathbb{E}\left[\overset{k+1|k}{\mathbf{M}}_k \mathbf{x}_k + \mathbf{w}_k\right], \\ &= \mathbf{M}_k \widehat{\mathbf{x}}_{k|k} \end{aligned} \quad (5)$$

mesurement

$$\mathbf{x}_{k+1|k} = \mathbf{M}\mathbf{x}_{k|k} + \mathbf{B}\mathbf{w}$$

$$\widehat{\mathbf{x}}_{k+1|k} = E[\mathbf{x}_{k+1|k} \mid \mathbf{y}_1, \ldots, \mathbf{y}_k] = \mathbf{M}\hat{\mathbf{x}}_{k|k}$$

# CAGLARS HOMEWORK

The one step prediction of the covariance is defined by,

$$\mathbf{P}_{k+1|k} = \mathbb{E}\left[(\mathbf{x}_{k+1} - \widehat{\mathbf{x}}_{k+1|k})(\mathbf{x}_{k+1} - \widehat{\mathbf{x}}_{k+1|k})^T|\mathbf{y}_1, \ldots \mathbf{y}_k\right]. \quad (6)$$

Exercise: Using the state propagation model, (1), and one-step prediction of the mean, (5), show that

$$\mathbf{P}_{k+1|k} = \mathbf{M}_k \mathbf{P}_{k|k} \mathbf{M}_k^T + \mathbf{Q}_k. \quad (7)$$

mesurement

$$\mathbf{x}_{k+1|k} = \mathbf{M}\mathbf{x}_{k|k} + \mathbf{B}\mathbf{w}$$

$$\mathbf{P}_{k+1|k} = E[(\mathbf{x}_{k+1|k} - \hat{\mathbf{x}}_{k+1|k})(\mathbf{x}_{k+1|k} - \hat{\mathbf{x}}_{k+1|k})^T \mid \mathbf{y}_1, \ldots, \mathbf{y}_k] = \mathbf{M}\mathbf{P}_{k|k}\mathbf{M}^T + \mathbf{B}\mathbf{Q}\mathbf{B}^T$$

# Bayesian Framework

m : model parameter vector (unknown parameters to be estimated)
d : data vector relating to m via an equation h(.)
d = h(m) + noise

Classical parameter estimation framework: Unknown but deterministic m
Bayesian parameter estimation framework: Unknown and random variable m

Bayes' Formula

$$p(m,d) = p(m \mid d)\,p(d) = p(d \mid m)\,p(m)$$

$$\overbrace{p(\mathbf{m}|\mathbf{d})}^{\text{POSTERIOR}} = \frac{p(\mathbf{d}|\mathbf{m})\,p(\mathbf{m})}{p(\mathbf{d})} = \frac{\overbrace{p(\mathbf{d}|\mathbf{m})}^{\text{LIKELIHOOD}}\,\overbrace{p(\mathbf{m})}^{\text{PRIOR}}}{\underbrace{\int p(\mathbf{d}|\mathbf{m})\,p(\mathbf{m})\,d\mathbf{m}}_{\text{EVIDENCE}}}$$

Bayes

# Sequential updates

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p(\mathbf{m})$$

Consider $\mathbf{d}$ consisting of two independent data set

$$p(\mathbf{d}_1,\mathbf{d}_2) = p(\mathbf{d}_1)p(\mathbf{d}_2)$$

$$\begin{aligned}
p(\mathbf{m}|\mathbf{d}) = p(\mathbf{m}|\mathbf{d}_1,\mathbf{d}_2) \\
&= \frac{p(\mathbf{d}_2|\mathbf{d}_1,\mathbf{m})p(\mathbf{m}|\mathbf{d}_1)}{p(\mathbf{d}_2|\mathbf{d}_1)} \\
&= \frac{p(\mathbf{d}_2|\mathbf{d}_1,\mathbf{m})}{p(\mathbf{d}_2)}\frac{p(\mathbf{d}_1|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d}_1)} \\
&\propto p(\mathbf{d}_2|\mathbf{m})p(\mathbf{d}_1|\mathbf{m})p(\mathbf{m})
\end{aligned}$$

Generalizing

$$p(\mathbf{m}|\mathbf{d}) \propto \prod_{i=1}^{N} p(\mathbf{d}_i|\mathbf{m})p(\mathbf{m})$$

**Thus, in principle with no measurement equation, you can update sequentially or just at once**

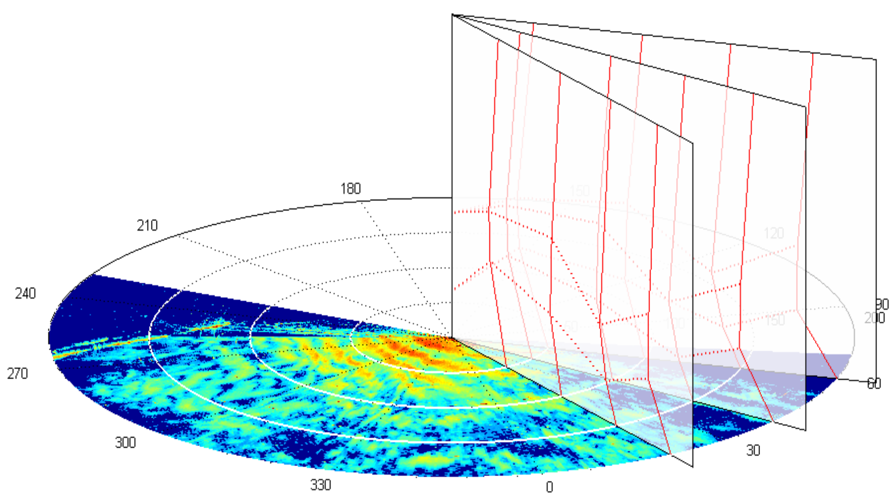
# Inversion, Filtering and Smoothing

$p(\mathbf{x}_t \mid \mathbf{y}_t)$: Inversion , Only observations at time $t$

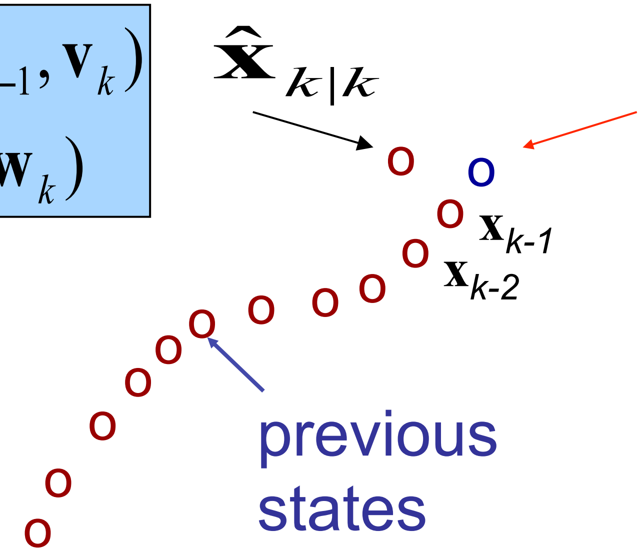
$p(\mathbf{x}_t \mid \mathbf{y}_{1:t})$: Filter , Observations from time $1{:}t$

$p(\mathbf{x}_t \mid \mathbf{y}_{1:T})$: Smoother, Observations from time $1{:}T$

$p(\mathbf{x}_t \mid \mathbf{y}_{1:T})$: Predictor, $t > T$ Observations from time $1{:}T$



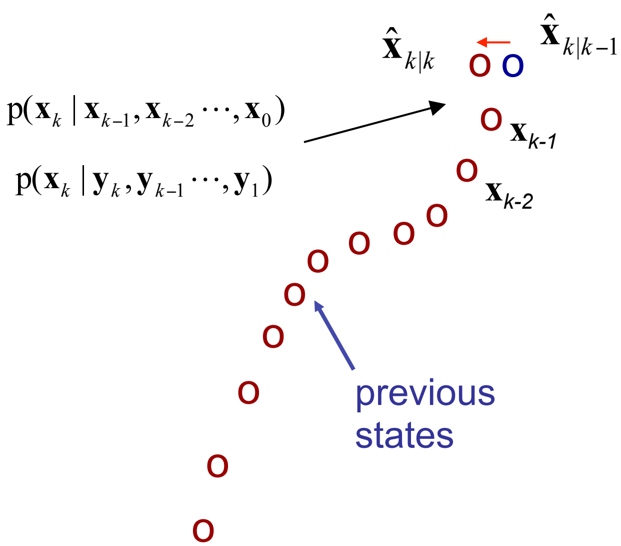

$$\mathbf{x}_k = f_{k-1}(\mathbf{x}_{k-1}, \mathbf{v}_k)$$
$$\mathbf{y}_k = h_k(\mathbf{x}_k, \mathbf{w}_k)$$



# A Single Kalman Iteration

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_k$$
$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k$$

$$\mathbf{x}_{k|k} \sim \mathcal{N}(\hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})$$




**PREDICT**

1. Predict the mean $\hat{\mathbf{x}}_{k|k-1}$ using previous history.

   $p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$

   $$\hat{\mathbf{x}}_{k|k-1} = \mathrm{E}\{\mathbf{x}_k \mid \mathbf{x}_{k-1}\} = \int \mathbf{x}_k \, p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) d\mathbf{x}_k$$

2. Predict the covariance $\mathbf{P}_{k|k-1}$ using previous history.

**UPDATE**

3. Correct/update the mean using new data $\mathbf{y}_k$

   $p(\mathbf{x}_k \mid \mathbf{Y}_k)$

   $$\hat{\mathbf{x}}_{k|k} = \mathrm{E}\{\mathbf{x}_k \mid \mathbf{Y}_k\} = \int \mathbf{x}_k \, p(\mathbf{x}_k \mid \mathbf{Y}_k) d\mathbf{x}_k$$

4. Correct/update the covariance $\mathbf{P}_{k|k}$ using $\mathbf{y}_k$

$$\cdots \Rightarrow p(\mathbf{x}_{k-1} \mid \mathbf{Y}_{k-1}) \Rightarrow p(\mathbf{x}_k \mid \mathbf{Y}_{k-1}) \Rightarrow p(\mathbf{x}_k \mid \mathbf{Y}_k) \Rightarrow \cdots$$

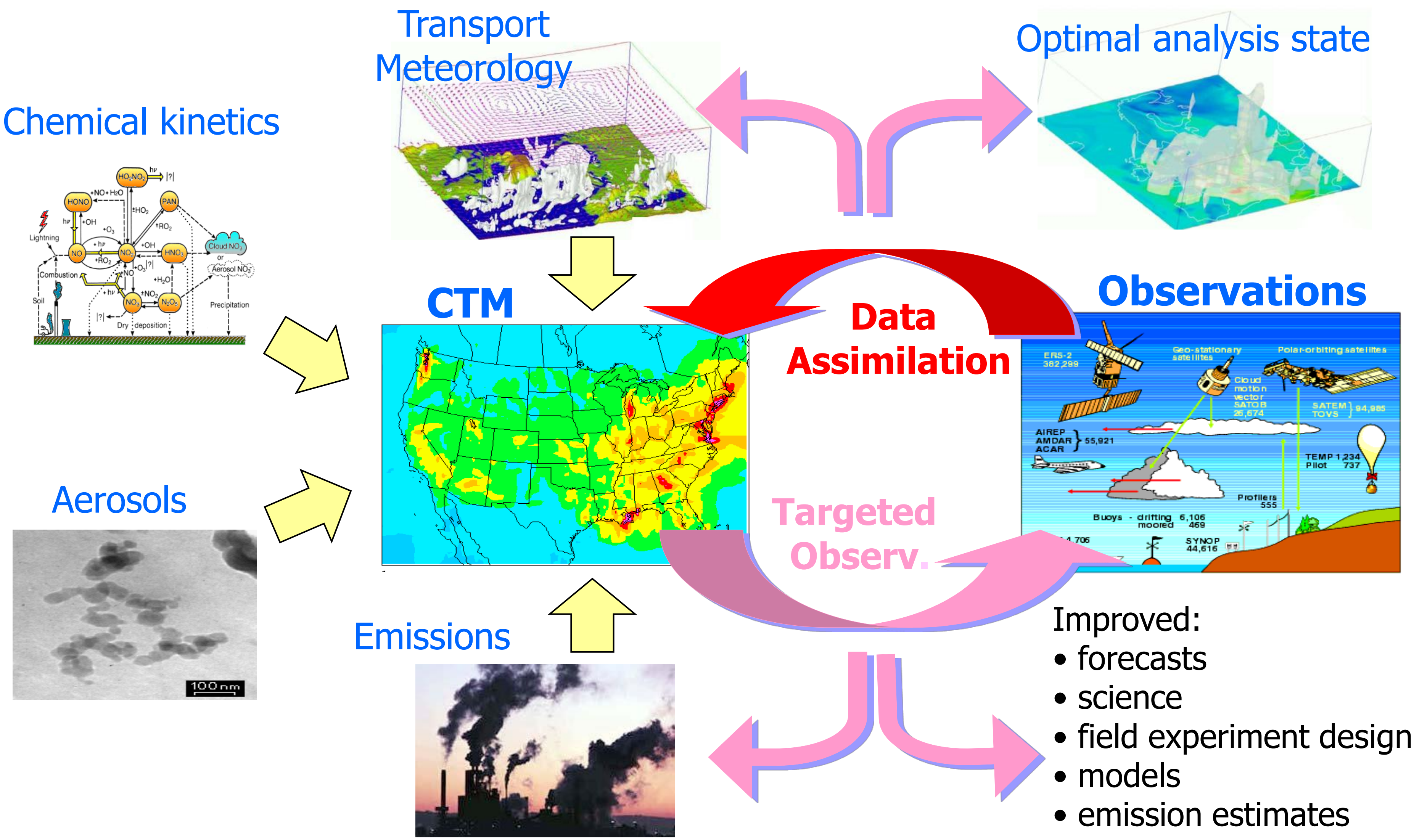
**PREDICTOR-CORRECTOR**

**DENSITY PROPAGATOR**

# Tutorial Lecture: Data Assimilation

Adrian Sandu
*Computational Science Laboratory*
Virginia Polytechnic Institute
and State University

# Data assimilation fuses information from (1) prior, (2) model, (3) observations to obtain consistent description of a physical system

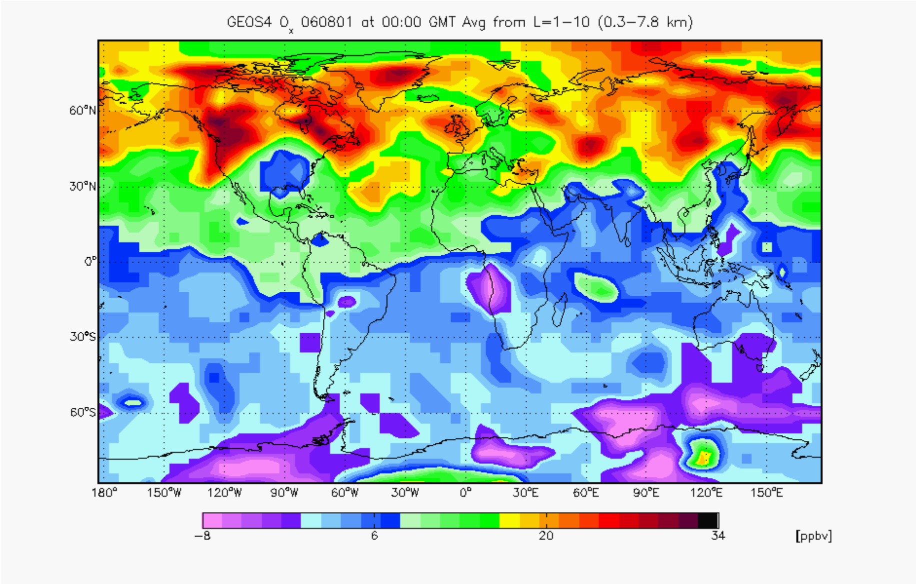


# Source of information #1: **the prior** encapsulates our current knowledge about the state of the system

- Background (prior) pdf: $\mathcal{P}^{\rm b}(\mathbf{x})$
- Current best estimate: background state $\mathbf{x}^{\rm b}$.
- Typical assumption:

$$\varepsilon^{\rm b} = \mathbf{x}^{\rm b} - \mathcal{S}(\mathbf{x}^{\rm true}) \in \mathcal{N}\left(\mathbf{0}, \mathbf{B}\right) \,.$$

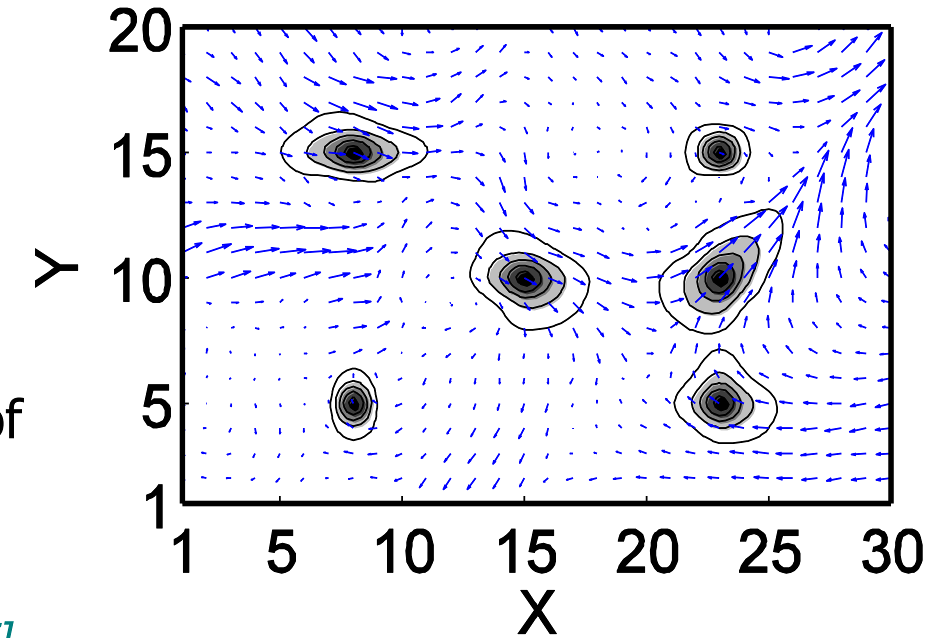
- With nonlinear models the normality assumption is difficult to justify, but is nevertheless used because of its convenience.



# Correct models of background (prior) errors are very important for data assimilation

- Background error representation determines the spread of information, and impacts the assimilation results
- Needs: high rank, capture dynamic dependencies, efficient computations
- Traditionally estimated empirically (NMC, Hollingsworth-Lonnberg)

1. Tensor products of 1d correlations, decreasing with distance (Singh et al, 2010)
2. Multilateral AR model (Constantinescu et al 2007)
3. Hybrid methods in the context of 4D-Var (Cheng et al, 2009)




*[Constantinescu and Sandu, 2007]*

# Source of information #2: **the model** encapsulates our knowledge about the physical laws that govern the evolution of the system

- The model evolves an initial state $\mathbf{x}_0 \in \mathbb{R}^n$ to future times

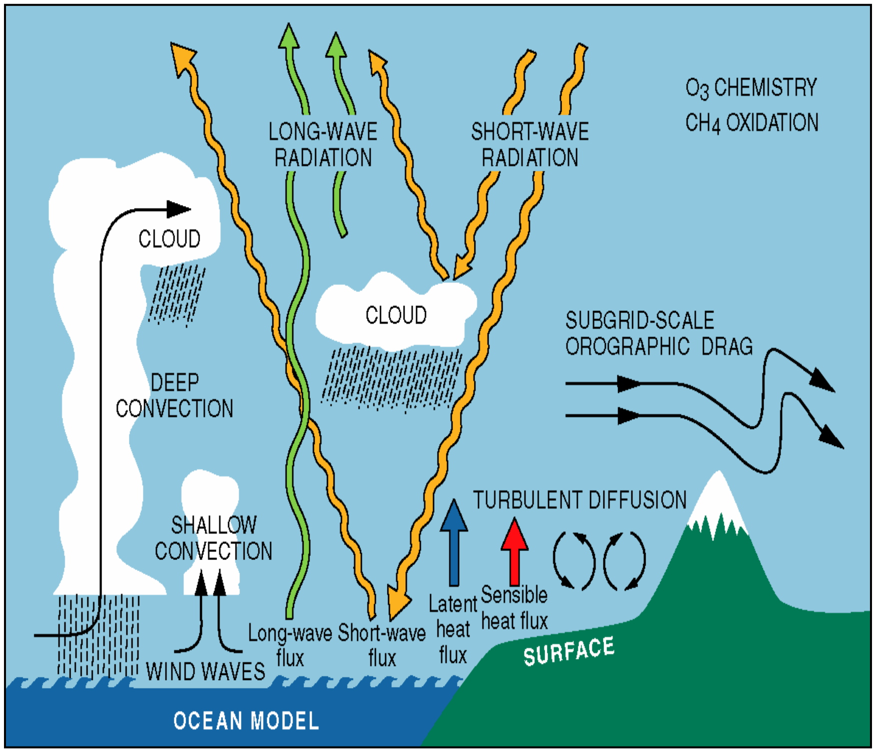
$$\mathbf{x}_i = \mathcal{M}_{t_0 \to t_i}\left(\mathbf{x}_0\right) .$$

- The model is imperfect

$$\mathcal{S}\left(\mathbf{x}_i^{\mathrm{true}}\right) = \mathcal{M}_{t_{i-1} \to t_i} \cdot \mathcal{S}\left(\mathbf{x}_{i-1}^{\mathrm{true}}\right) - \eta_i ,$$

where $\eta_i$ is the model error in step $i$.

How large are the models of interest? Typically O($10^8$) variables, and O(10) different physical processes




Picture: L. Isaksen (http://www.ecmwf.int)

# Source of information #3: **the observations** are sparse and noisy snapshots of reality

- Measurements $\mathbf{y}_i \in \mathbb{R}^m$ ($m \ll n$) taken at times $t_1, \ldots, t_N$

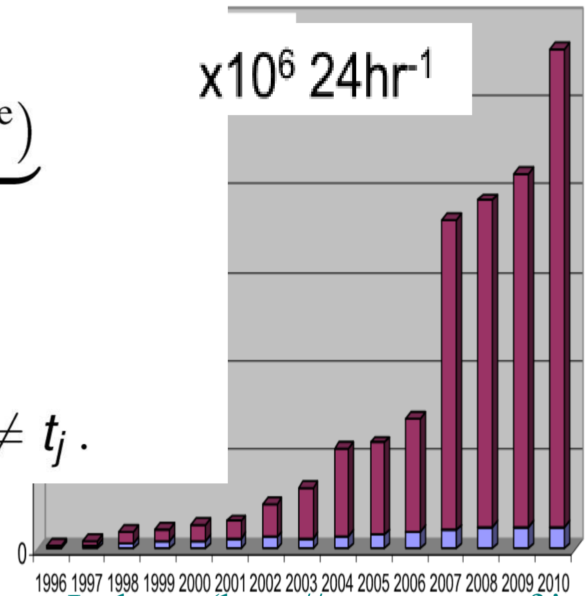
$$\mathbf{y}_i = \mathcal{H}^{\mathrm{t}}\left(\mathbf{x}_i^{\mathrm{true}}\right) - \varepsilon_i^{\mathrm{instrument}} = \mathcal{H}\left(\mathcal{S}(\mathbf{x}_i^{\mathrm{true}})\right) - \varepsilon_i^{\mathrm{obs}}, \quad i = 1, \cdots, N\,.$$

- Observation operators
  - $\mathcal{H}^{\mathrm{t}}$ : physical space $\rightarrow$ observation space, while
  - $\mathcal{H}$ : the model space $\rightarrow$ observation space.
- The *observation error*

$$\varepsilon_i^{\mathrm{obs}} = \underbrace{\varepsilon_i^{\mathrm{instrument}}}_{\text{instrument error}} + \underbrace{\mathcal{H}\left(\mathcal{S}(\mathbf{x}_i^{\mathrm{true}})\right) - \mathcal{H}^{\mathrm{t}}\left(\mathbf{x}_i^{\mathrm{true}}\right)}_{\text{representativeness error}}$$

- Typical assumptions:

$$\varepsilon_i^{\mathrm{obs}} \in \mathcal{N}\left(\mathbf{0}, \mathbf{R}_i\right)\ ; \quad \varepsilon_i^{\mathrm{obs}}\,,\, \varepsilon_j^{\mathrm{obs}} \quad \text{independent for} \quad t_i \neq t_j\,.$$

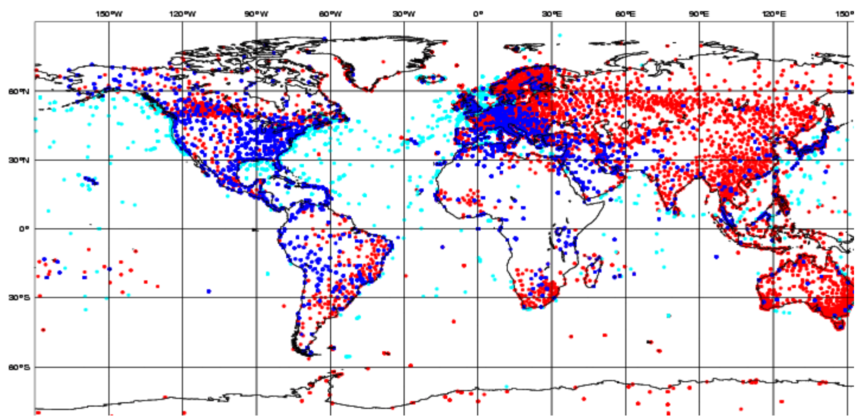


Lars Isaksen (http://www.ecmwf.int)

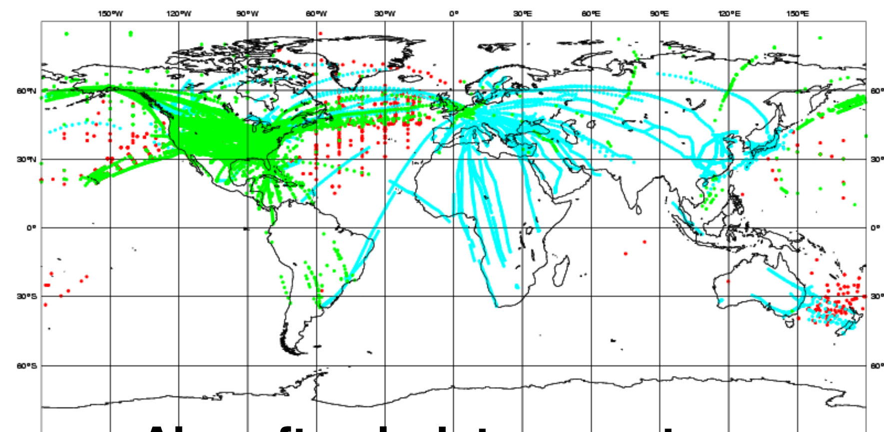
How many observations? ECMWF: $O(10^7)$

# Some conventional and remote data sources used at ECMWF for numerical weather prediction

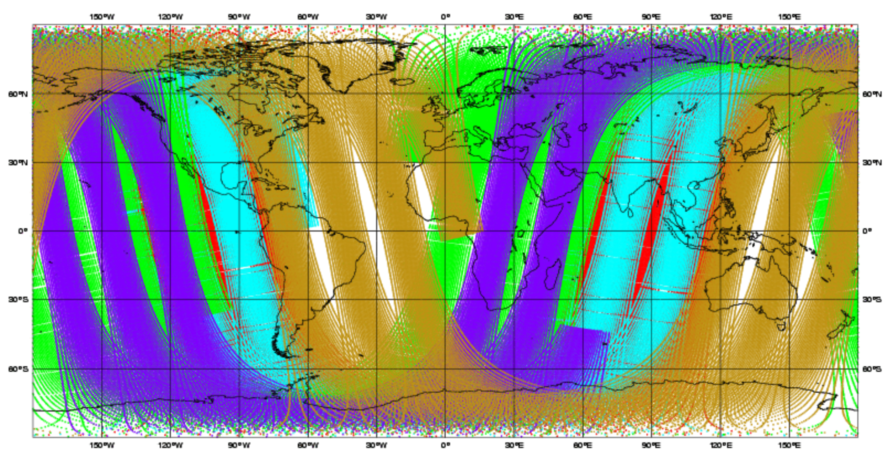
Lars Isaksen (http://www.ecmwf.int)



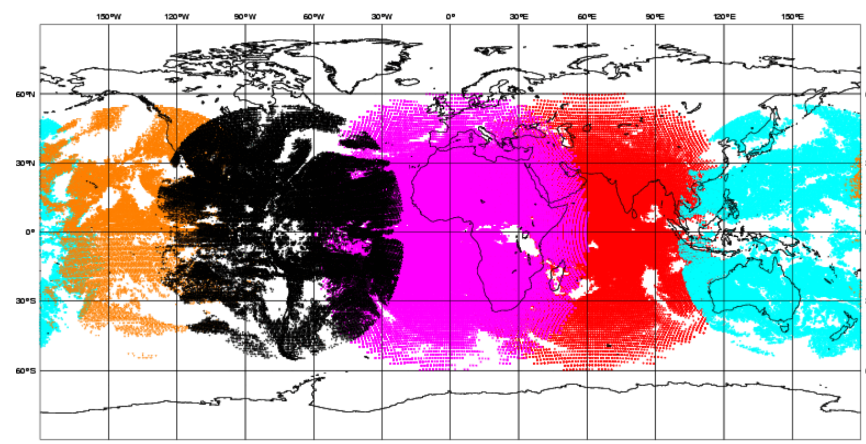
**SYNOP/METAR/SHIP: pres., wind, RH**



**Aircraft: wind, temperature**

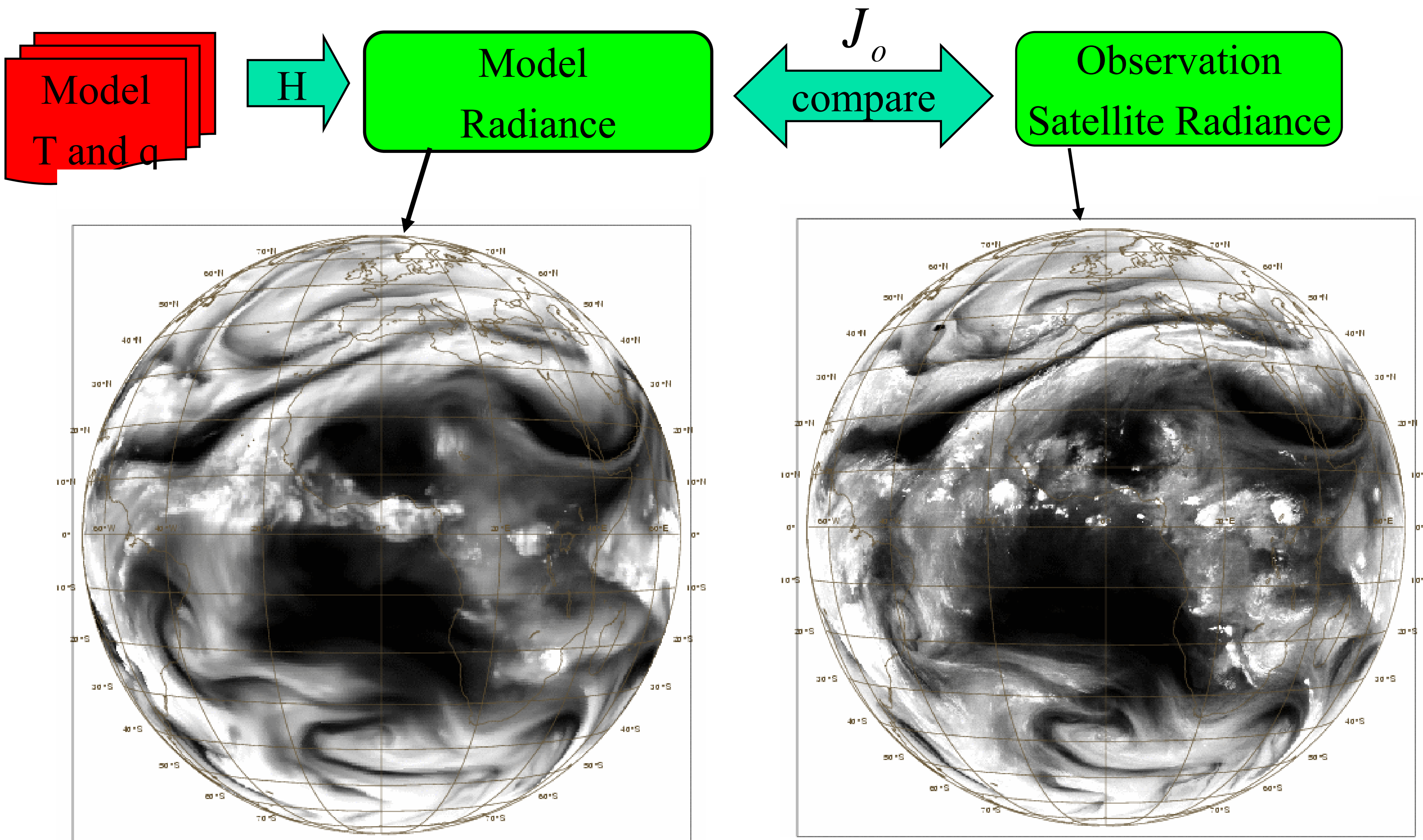


**13 Sounders: NOAA AMSU-A/B, HIRS, AIRS, …**



**Geostationary, 4 IR and 5 winds**

# To allow model-data comparison, **observation operators** map the model state space to observation space




Lars Isaksen (http://www.ecmwf.int)