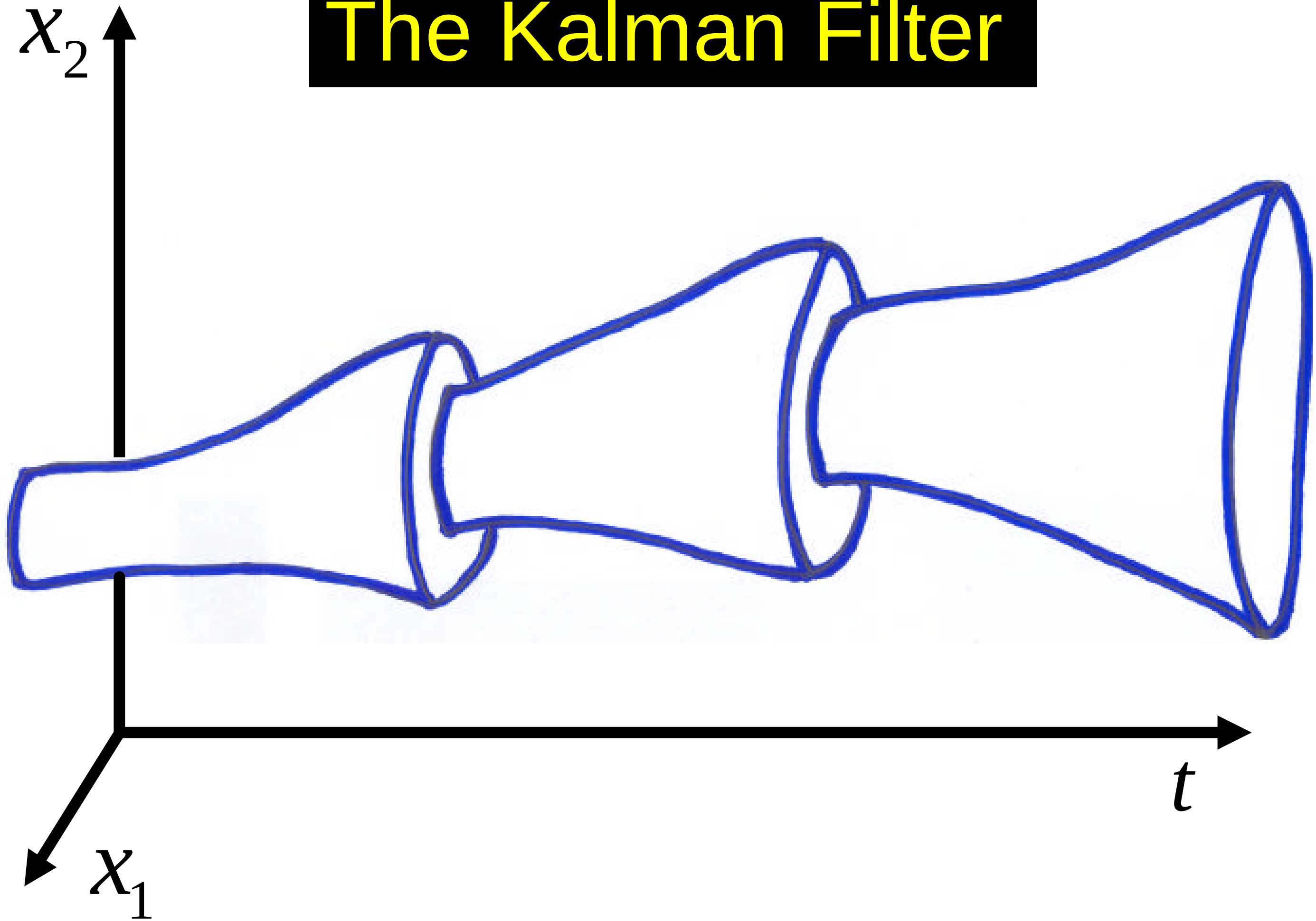
Some Useful References

- Atmospheric Data Analysis by R. Daley, Cambridge University Press.
- Atmospheric Modelling, Data Assimilation and Predictability by E. Kalnay, C.U.P.
- The Ocean Inverse Problem by C. Wunsch, C.U.P.
- Inverse Problem Theory by A. Tarantola, Elsevier.
- Inverse Problems in Atmospheric Constituent Transport by I.G. Enting, C.U.P.
- ECMWF Lecture Notes at www.ecmwf.int

The Kalman Filter



Extended Kalman Filter

- Assumes the model is *non-linear* and imperfect.
- The tangent linear model depends on the state and on time.
- Could be a “gold standard” for data assimilation, but very expensive to implement because of the very large dimension of the state space (~ 10^6 – 10^7 for NWP models).

Ensemble Kalman Filter

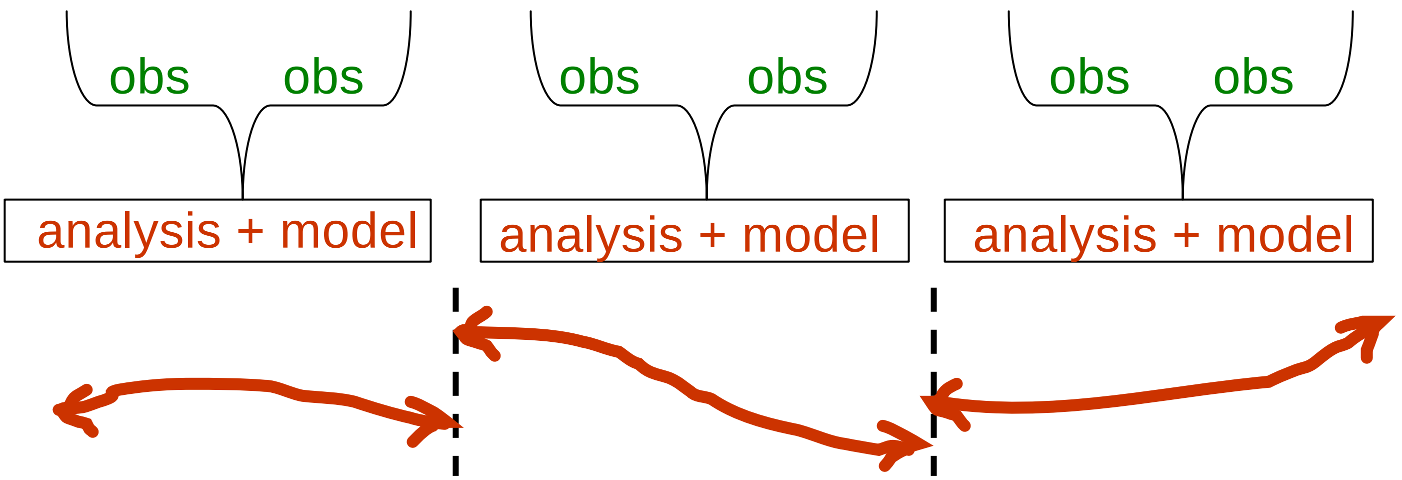
- Carry forecast error covariance matrix forward in time by using ensembles of forecasts:

$$\mathbf{B} \approx \frac{1}{K-1}\sum_{k\neq 1}^{K}(\mathbf{x}_k^f - <\mathbf{x}^f>)(\mathbf{x}_k^f - <\mathbf{x}^f>)^{\mathrm{T}}$$

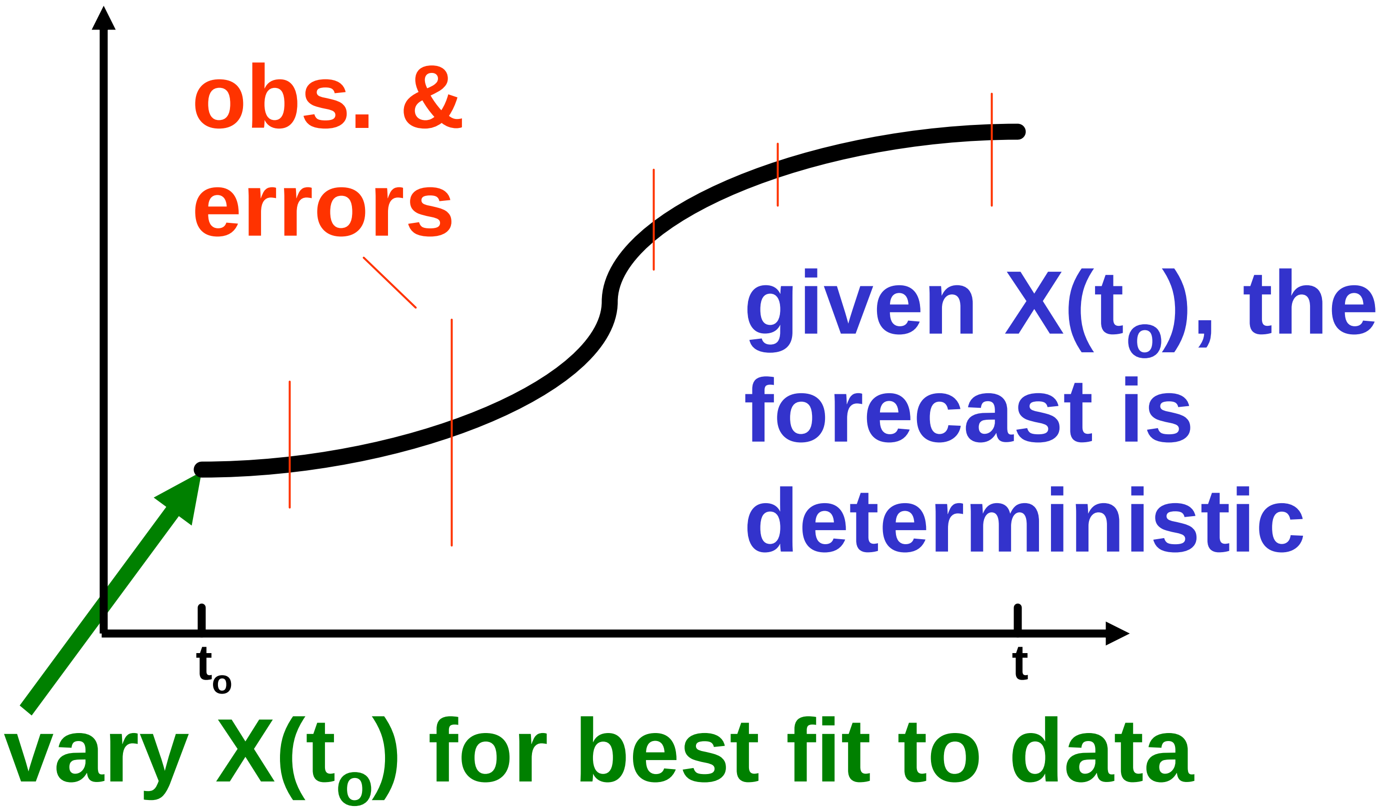
- Only ~ 10 + forecasts needed.
- Does not require computation of tangent linear model and its adjoint.
- Does not require linearization of evolution of forecast errors.
- Fits in neatly into ensemble forecasting.

4d-Variational Assimilation

Non-sequential Intermittent Assimilation



4D Variational Data Assimilation



4d-Variational Assimilation

$$J(\mathbf{x}(t_0)) = \frac{1}{2}\sum_{i=0}^{N}[\mathbf{y}_i - H(\mathbf{x}_i)]^{\mathrm{T}}\mathbf{R}_i^{-1}[\mathbf{y}_i - H(\mathbf{x}_i)]$$

$$+\frac{1}{2}[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)]^{\mathrm{T}}\mathbf{B}_0^{-1}[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)]$$

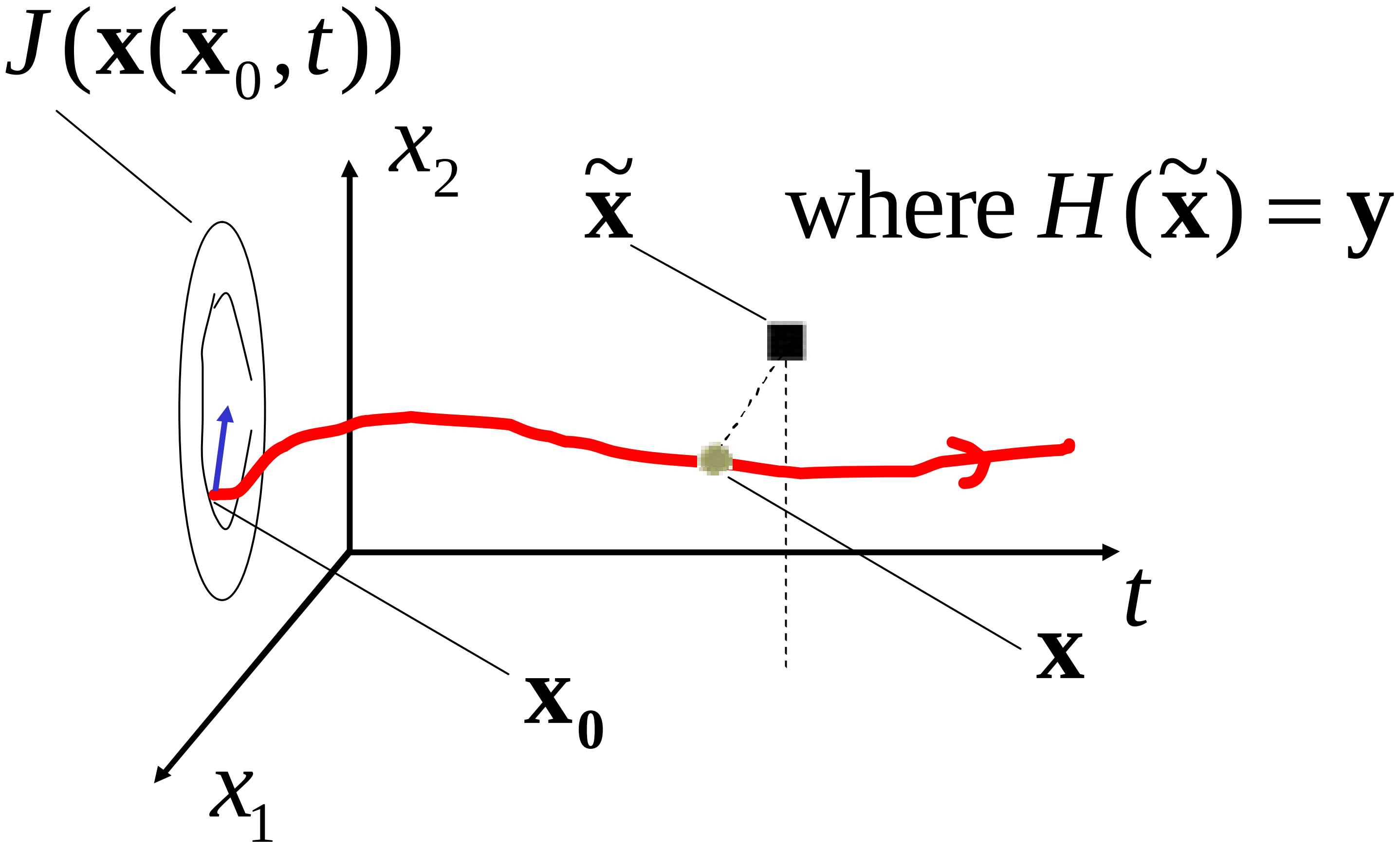
where $\mathbf{x}(t_i) = M_{0\rightarrow i}(\mathbf{x}(t_0))$ i.e. the model is treated as a strong constraint

Minimize the cost function by finding the gradient $\partial J / \mathbf{x}(t_0)$ ("Jacobian") with respect to the control variables in $\mathbf{x}(t_0)$

4d-VAR Continued

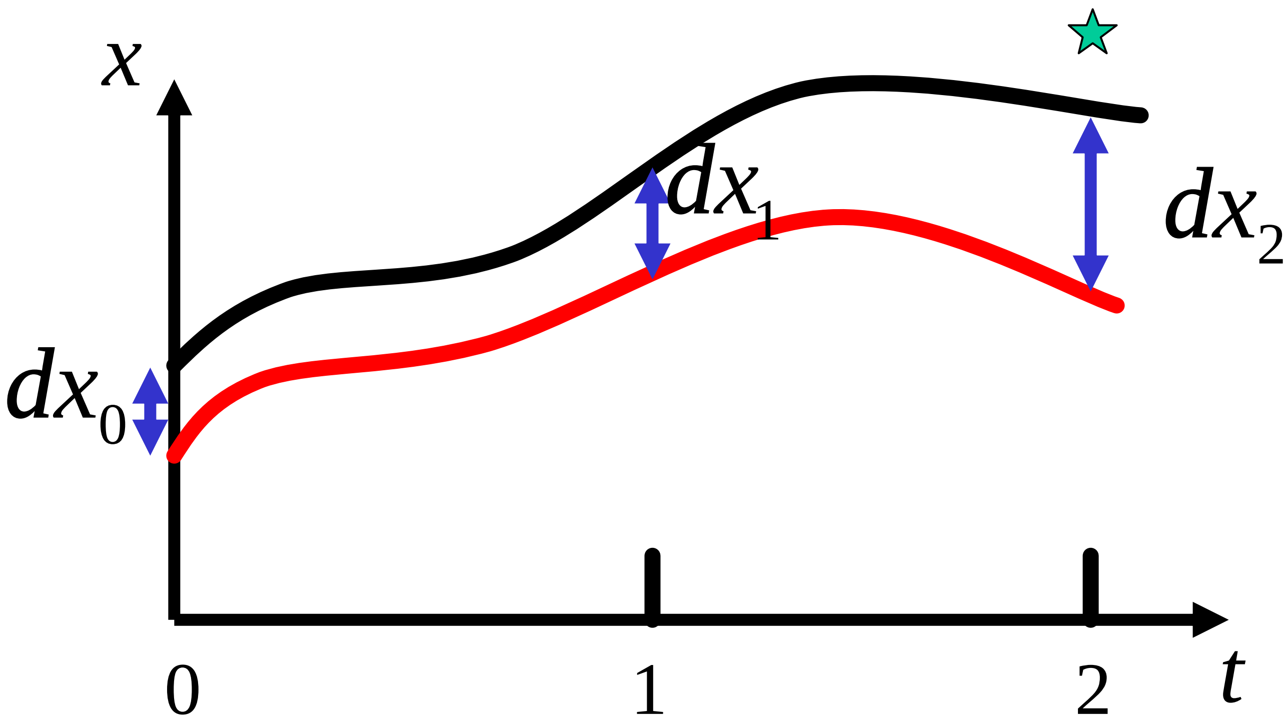
The 2nd term on the RHS of the cost function measures the distance to the background at the beginning of the interval. The term helps join up the sequence of optimal trajectories found by minimizing the cost function for the observations. The “analysis” is then the optimal trajectory in state space. Forecasts can be run from any point on the trajectory, e.g. from the middle.

4d-VAR For Single Observation at time t



The Time-Stepping Model
Scalar Case

$$\begin{tikzcd}
M(x_0) \arrow[r] & x_1 \arrow[dl] \\
M(x_1) \arrow[r] & x_2 \arrow[dl] \\
{} &
\end{tikzcd}$$



x
dx0
dx1
dx2
0
1
2
t

$$x_1 = M(x_0)$$

$$dx_1 = \left.\frac{\partial M}{\partial x}\right|_{x_0} dx_0$$

$$dx_2 = \left.\frac{\partial M}{\partial x}\right|_{x_1} \left.\frac{\partial M}{\partial x}\right|_{x_0} dx_0$$

$$x_2 = M(x_1)$$

$$dx_2 = \left.\frac{\partial M}{\partial x}\right|_{x_1} dx_1$$

$$\frac{\partial x_2}{\partial x_0} = \left.\frac{\partial M}{\partial x}\right|_{x_1} \left.\frac{\partial M}{\partial x}\right|_{x_0}$$

$$J(x_2(x_0))$$

$$\frac{\partial J}{\partial x_0} = \frac{\partial x_2}{\partial x_0} \frac{\partial J}{\partial x_2}$$

$$\boxed{\frac{\partial J}{\partial x_0} = \left.\frac{\partial M}{\partial x}\right|_{x_1} \left.\frac{\partial M}{\partial x}\right|_{x_0} \frac{\partial J}{\partial x_2}}$$

Vector Case

$$J = J(\mathbf{x}(\mathbf{x}_0))$$

Then
$$\frac{\partial J}{\partial \mathbf{x}_0} = \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}_0}\right)^{\mathrm{T}} \frac{\partial J}{\partial \mathbf{x}}$$

adjoint of the tangent-linear model

Transpose reverses the sequence of matrices in the chain rule, so the tangent linear model is run BACKWARDS in time.

A Useful Result

Let J have the following form: $J = \mathbf{z}^{\mathrm{T}}(\mathbf{x})\mathbf{A}\mathbf{z}(\mathbf{x})$

Then it can be shown that $\dfrac{\partial J}{\partial \mathbf{x}} = \left(\dfrac{\partial \mathbf{z}}{\partial \mathbf{x}}\right)^{\mathrm{T}} \mathbf{A}\mathbf{z}$

Combining these results: $\dfrac{\partial J}{\partial \mathbf{x}_0} = \left(\dfrac{\partial \mathbf{x}}{\partial \mathbf{x}_0}\right)^{\mathrm{T}} \left(\dfrac{\partial \mathbf{z}}{\partial \mathbf{x}}\right)^{\mathrm{T}} \mathbf{A}\mathbf{z}$

4d-VAR for Single Observation

$$J(\mathbf{x}(\mathbf{x}_0)) = \frac{1}{2}[\mathbf{y} - H(\mathbf{x}(\mathbf{x}_0))]^{\mathrm{T}} \mathbf{R}^{-1} [\mathbf{y} - H(\mathbf{x}(\mathbf{x}_0))]$$

$$\frac{\partial J}{\partial \mathbf{x}_0} = -\mathbf{L}_{0\to t}^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{R}^{-1} [\mathbf{y} - H(\mathbf{x}(\mathbf{x}_0))] \equiv -\mathbf{L}_{0\to t}^{\mathrm{T}} \mathbf{d}$$

$$\text{where} \quad \mathbf{L}_{0\to t}^{\mathrm{T}} = \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} \right)^{\mathrm{T}} = \left(\frac{\partial M_{0\to t}(\mathbf{x}_0)}{\partial \mathbf{x}_0} \right)^{\mathrm{T}},$$

$$\mathbf{L}_{0\to t}^{\mathrm{T}} = \mathbf{L}_{0\to t_1}^{\mathrm{T}} \mathbf{L}_{t_1\to t_2}^{\mathrm{T}} \ldots \mathbf{L}_{t_{n-1}\to t}^{\mathrm{T}}$$

4d-VAR Procedure

- Choose $\mathbf{x}_0 , \mathbf{x}_0^b$ for example.
- Integrate full (non-linear) model forward in time and calculate $\mathbf{d}$ for each observation.
- Map $\mathbf{d}$ back to t=0 by backward integration of TLM, and sum for all observations to give the gradient of the cost function.
- Move down the gradient to obtain a better initial state (new trajectory “hits” observations more closely)
- Repeat until some STOP criterion is met.

note: not the most efficient algorithm

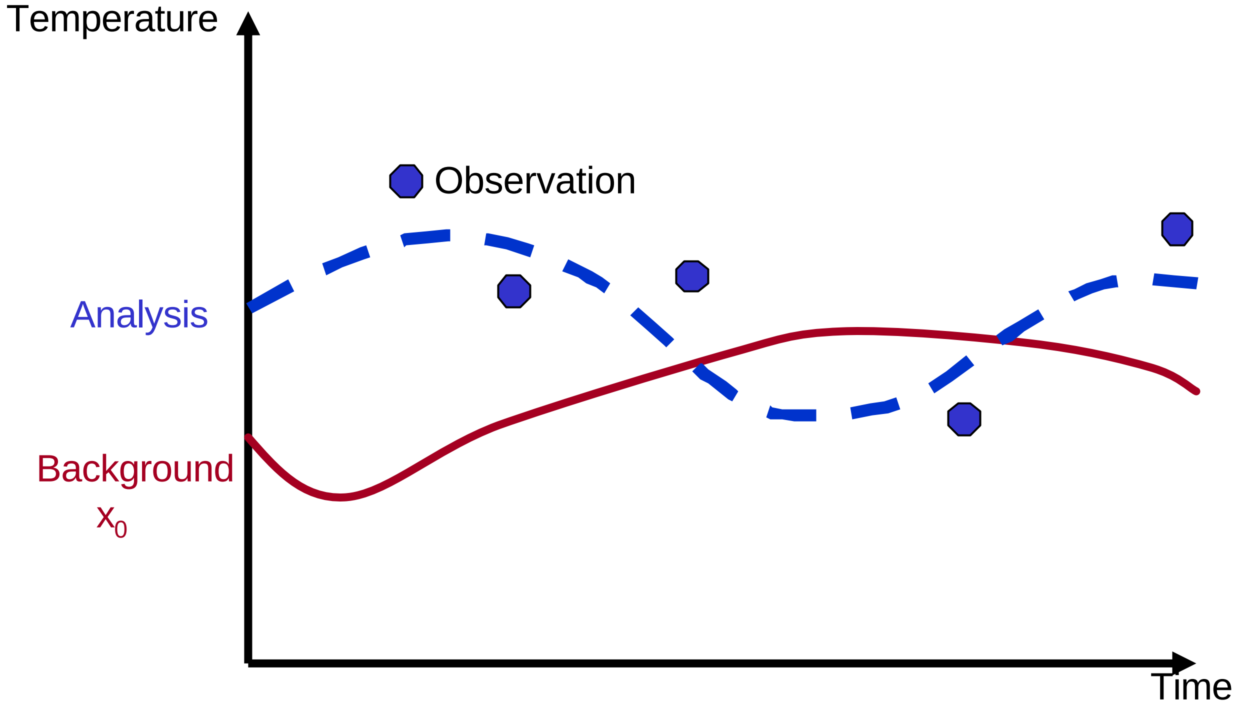
Comments

- 4d-VAR can also be formulated by the method of Lagrange multipliers to treat the model equations as a constraint. The adjoint equations that arise in this approach are the same equations we have derived by using the chain rule of partial differential equations.
- If model is perfect and B_0 is correct, 4d-VAR at final time gives same result as extended Kalman filter (but the covariance of the analysis is not available in 4d-VAR).
- 4d-VAR analysis therefore optimal over its time window, but less expensive than Kalman filter.

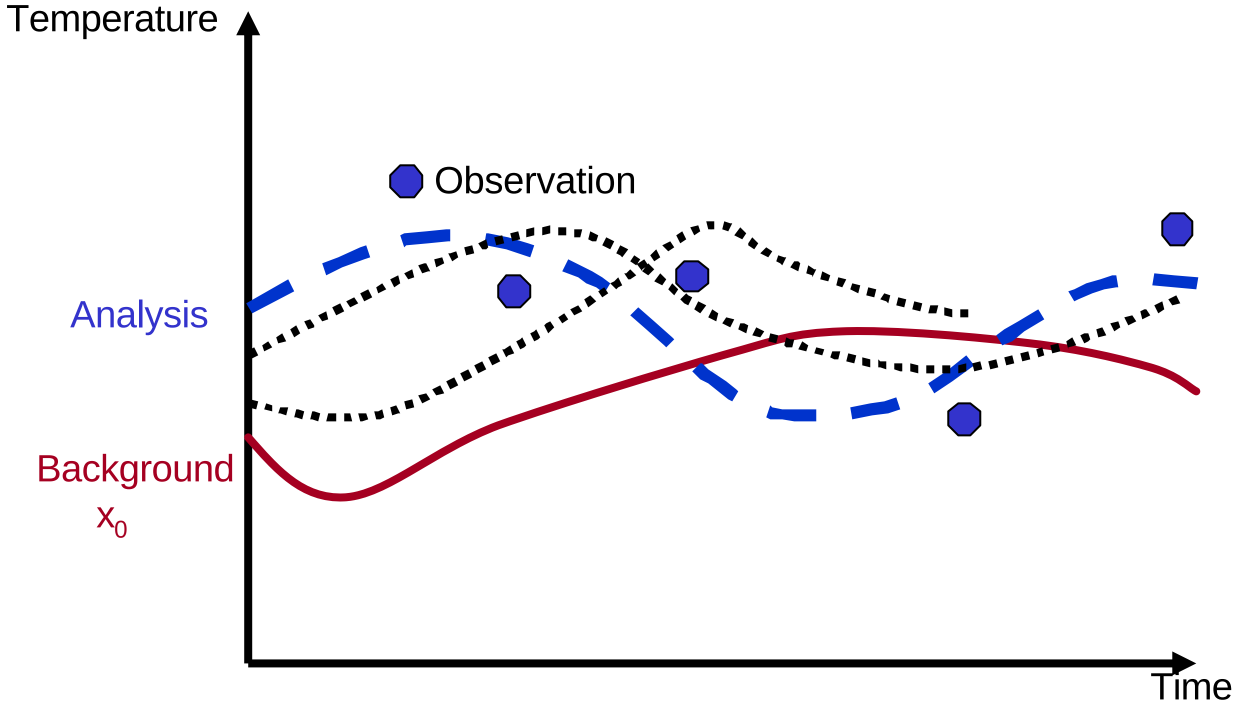
Incremental Form of 4d-VAR

- The 4d-VAR algorithm presented earlier is expensive to implement. It requires repeated forward integrations with the non-linear (forecast) model and backward integrations with the TLM.
- When the initial background (first-guess) state and resulting trajectory are accurate, an incremental method can be made much cheaper to run on a computer.

Full 4D-Var



Incremental 4D-Var



Incremental Form of 4d-VAR

The incremental form of the cost function is defined by

$$J(d\mathbf{x}_0) = \frac{1}{2}(d\mathbf{x}_0)^{\mathrm{T}}\mathbf{B}_0^{-1}(d\mathbf{x}_0)$$

where $d\mathbf{x}_0 = \mathbf{x}(t_0) - \mathbf{x}^b(t_0)$

$$+\frac{1}{2}\sum_{i=0}^{N}[\mathbf{y}_i - \underbrace{H(\mathbf{x}^f(t_i)) - \mathbf{H}_i\mathbf{L}(t_0,t_i)d\mathbf{x}_0}]^{\mathrm{T}}\mathbf{R}_i^{-1}[\mathbf{y}_i - \underbrace{H(\mathbf{x}^f(t_i)) - \mathbf{H}_i\mathbf{L}(t_0,t_i)d\mathbf{x}_0}]$$

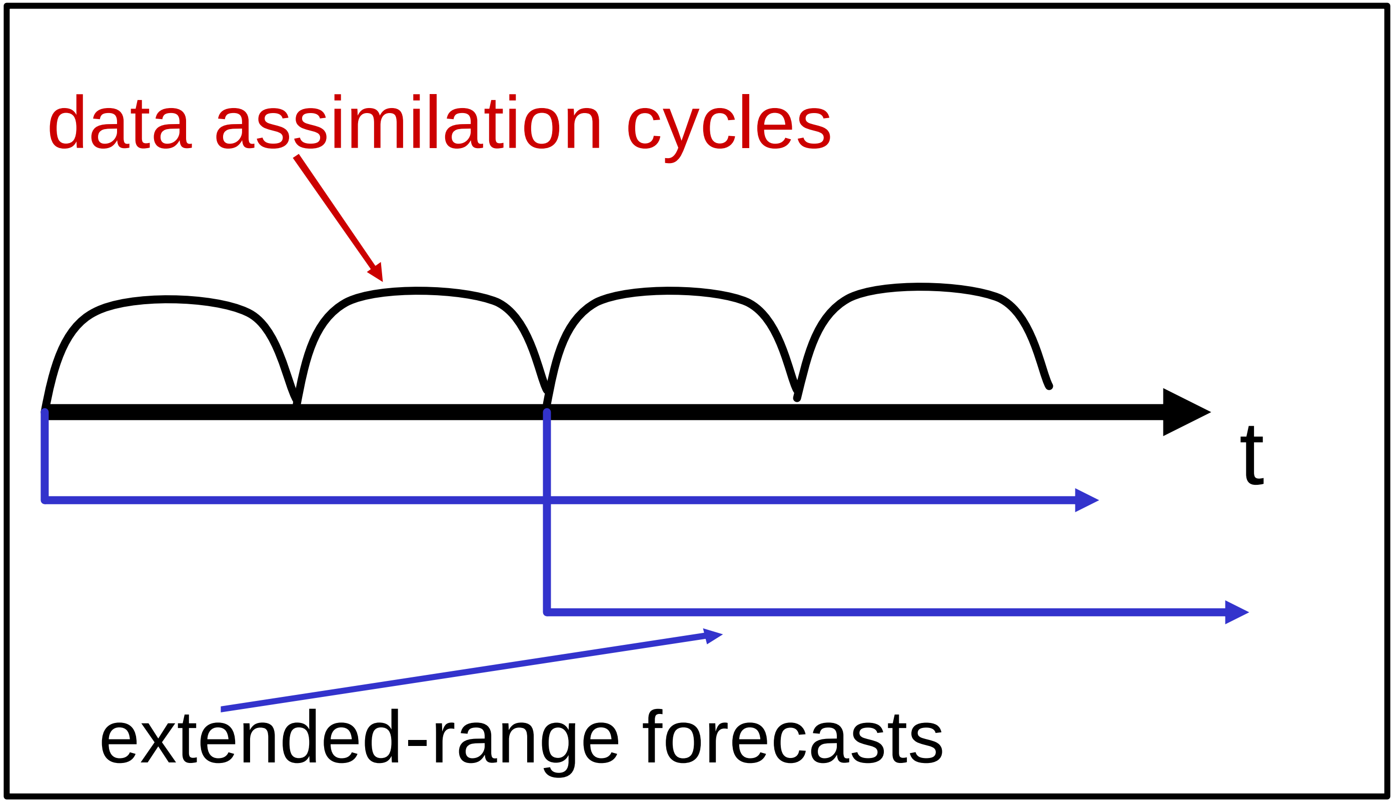
Taylor series expansion
about first-guess trajectory
$\mathbf{x}^f(t_i)$

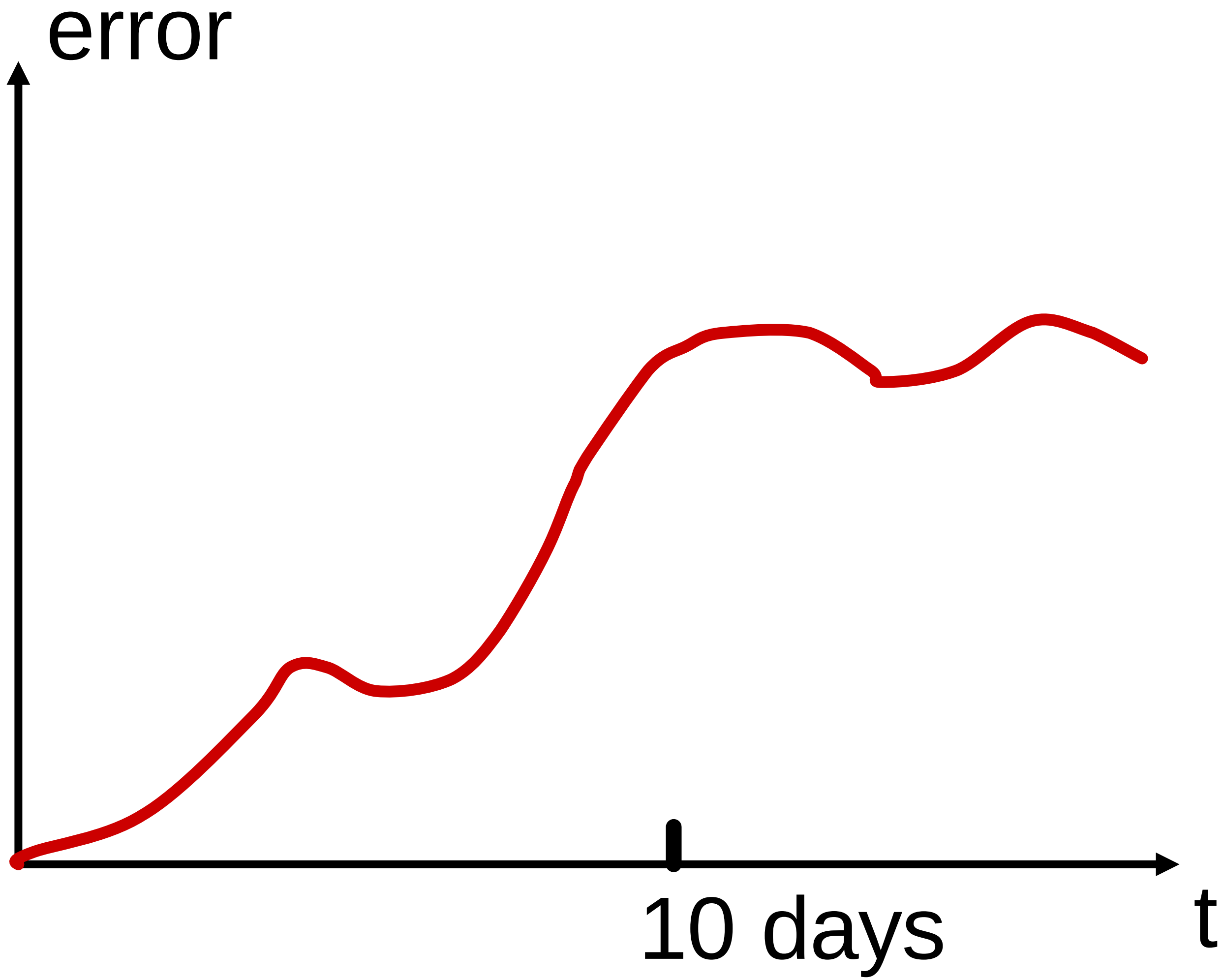
Minimization can be done in lower dimensional space

4D Variational Data Assimilation

- Advantages
  - consistent with the governing eqs.
  - implicit links between variables
- Disadvantages
  - very expensive
  - model is strong constraint

Testing Earth System Models




error
10 days
t

END