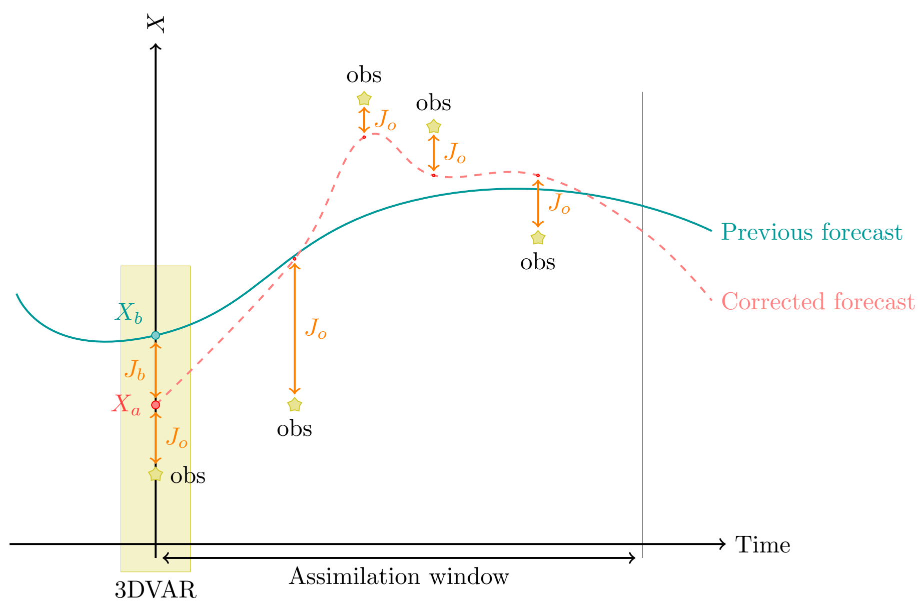
Hybrid Data Assimilation

Mark Asch - CSU/IMU/2023



Outline of the course (I)

Adjoint methods and variational data assimilation (4h)

1. Introduction to data assimilation: setting, history, overview, definitions.
2. Adjoint method.
3. Variational data assimilation methods:
   (a) 3D-Var,
   (b) 4D-Var.
4. EnsVar - Hybrid Ensemble Variational DA

Outline of the course (II)

Statistical estimation, Kalman filters and sequential data assimilation (4h)

1. Introduction to statistical DA.
2. Statistical estimation.
3. The Kalman filter.
4. Nonlinear extensions and ensemble filters.

Reacll: Variational DA - formulation

- In variational data assimilation we describe the state of the system by
  - ⇒ a state variable, $\mathbf{x}(t) \in \mathcal{X}$, a function of space and time that
  - ⇒ represents the physical variables of interest, such as current velocity (in oceanography), temperature, sea-surface height, salinity, biological species concentration, chemical concentration, etc.

- Evolution of the state is described by a system of (in general nonlinear) differential equations in a region Ω,

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathcal{M}(\mathbf{x}) \quad \text{in } \Omega \times [0, T] \,, \\ \mathbf{x}(t=0) = \mathbf{x}_0, \end{cases} \tag{1}$$

where the initial condition is unknown (or poorly known).

- Suppose that we are in possession of observations $\mathbf{y}(t) \in \mathcal{O}$ and an observation operator $\mathcal{H}$ that describes the available observations.

- Then, to characterize the difference between the observations and the state, we define the objective (or cost) function,

$$J(\mathbf{x}_0) = \frac{1}{2}\int_0^T \|\mathbf{y}(t) - \mathcal{H}\left(\mathbf{x}(\mathbf{x}_0, t)\right)\|_{\mathcal{O}}^2 \,\mathrm{d}t + \frac{1}{2}\left\|\mathbf{x}_0 - \mathbf{x}^{\mathrm{b}}\right\|_{\mathcal{X}}^2 \tag{2}$$

where

⇒ $\mathbf{x}^{\mathrm{b}}$ is the background (or first guess)

⇒ and the second term plays the role of a regularization (in the sense of Tikhonov).

⇒ The two norms under the integral, in the finite-dimensional case, will be represented by the error covariance matrices $\mathbf{R}$ and $\mathbf{B}$ respectively, and will be described below.

$\Rightarrow$ Note that for mathematical rigor we have indicated, as subscripts, the relevant functional spaces on which the norms are defined.

- In the continuous context, the data assimilation problem is formulated as follows:

> Find the analyzed state $\mathbf{x}_0^{\mathrm{a}}$ that minimizes J and satisfies
> $$\mathbf{x}_0^{\mathrm{a}} = \operatorname{argmin} J(\mathbf{x}_0).$$

- The necessary condition for the existence of a (local) minimum is (as usual...)

$$\nabla J(\mathbf{x}_0^{\mathrm{a}}) = 0.$$

Variational DA - 3D Var

- Finite-dimensional version of the cost function (2),

$$J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x}-\mathbf{x}^{\mathrm{b}}\right)^{\mathrm{T}}\mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^{\mathrm{b}}\right) \tag{3}$$

$$+\frac{1}{2}\left(\mathbf{H}\mathbf{x}-\mathbf{y}\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{H}\mathbf{x}-\mathbf{y}\right), \tag{4}$$

  where

  ⇒ $\mathbf{x}$, $\mathbf{x}^{\mathrm{b}}$, and $\mathbf{y}$ are the state, the background state, and the measured state respectively;

  ⇒ $\mathbf{H}$ is the observation matrix (a linearization of the observation operator $\mathcal{H}$);

  ⇒ $\mathbf{R}$ and $\mathbf{B}$ are the observation and background error covariance matrices respectively.

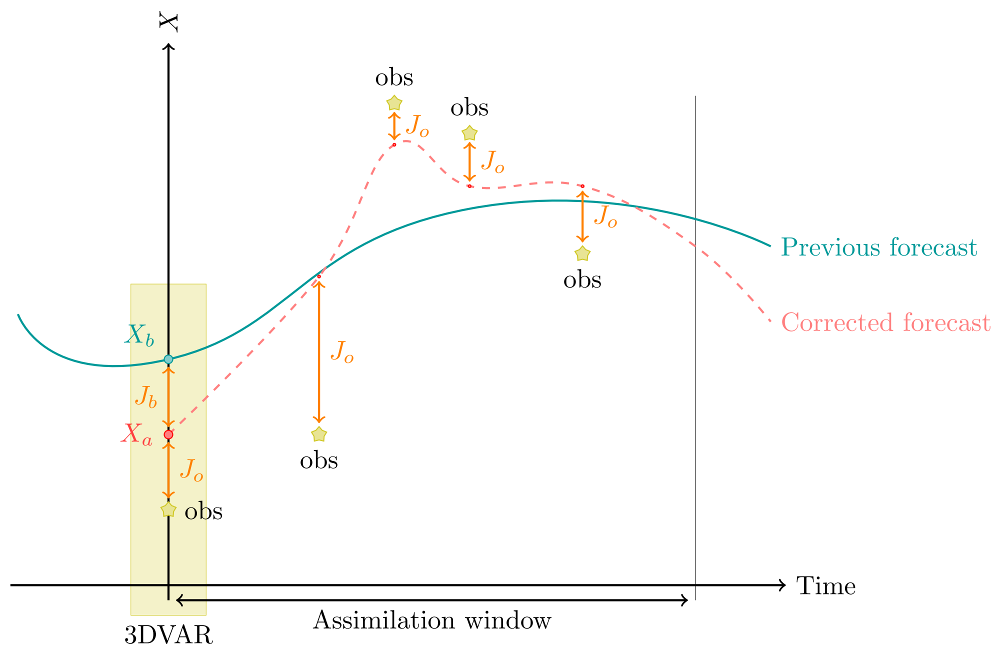
- This quadratic function attempts to strike a balance between some *a priori* knowledge about a background (or historical) state and the actual measured, or observed, state.

- It also assumes that we know and that we can invert the matrices $\mathbf{R}$ and $\mathbf{B}$. This, as we will be pointed out below, is not always obvious.

- Furthermore, it represents the sum of the (weighted) background deviations and the (weighted) observation deviations. The basic methodology is presented in the Algorithm below, which is nothing more than a classical gradient descent algorithm.

Variational DA - 3D Var Algorithm

$j = 0, \;\; x = x_0$
while $\|\nabla J\| > \epsilon$ or $j \le j_{\max}$
    compute J
    compute ∇J
    gradient descent and update of x_{j+1}
    $j = j + 1$
end



Variational DA - 4D Var

- A more realistic, but complicated situation arises when one wants to assimilate observations that are acquired over a time interval, during which the system dynamics (flow, for example) cannot be neglected.

- Suppose that the measurements are available at a succession of instants, $t_k,\ k = 0, 1, \ldots, K$ and are of the form

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \boldsymbol{\epsilon}_k^{\mathrm{o}}, \tag{5}$$

  where

  ⇒ $\mathbf{H}_k$ is a linear observation operator and
  ⇒ $\boldsymbol{\epsilon}_k^{\mathrm{o}}$ is the observation error with covariance matrix $\mathbf{R}_k$,
  ⇒ and suppose that these observation errors are uncorrelated in time.

- Now we add the dynamics described by the state equation,
$$\mathbf{x}_{k+1} = \mathbf{M}_{k+1}\mathbf{x}_k, \tag{6}$$
where we have neglected any model error.[1]

- We suppose also that at time index $k = 0$ we know

  $\Rightarrow$ the background state $\mathbf{x}_0^{\mathrm{b}}$ and
  $\Rightarrow$ its error covariance matrix $\mathbf{P}_0^{\mathrm{b}}$
  $\Rightarrow$ and we suppose that the errors are uncorrelated with the observations in (5).

- Then a given initial condition, $\mathbf{x}_0$, defines a unique model solution $\mathbf{x}_{k+1}$ according to (6).

- We can now generalize the objective function (3),

[1] This will be taken into account below.

which becomes

$$J(\mathbf{x}_0) = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}\right)^{\mathrm{T}} \left(\mathbf{P}_0^{\mathrm{b}}\right)^{-1} \left(\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}\right) \tag{7}$$

$$+ \frac{1}{2}\sum_{k=0}^{K} \left(\mathbf{H}_k \mathbf{x}_k - \mathbf{y}_k\right)^{\mathrm{T}} \mathbf{R}_k^{-1} \left(\mathbf{H}_k \mathbf{x}_k - \mathbf{y}_k\right). \tag{8}$$

- The minimization of $J(\mathbf{x}_0)$ will provide the initial condition of the model that fits the data most closely.

- This analysis is called "strong constraint four-dimensional variational assimilation," abbreviated as *strong constraint 4D-Var*. The term "strong constraint" implies that the model found by the state equation (6) must be exactly satisfied by the sequence of estimated state vectors.

- In the presence of model uncertainty, the state equation becomes

$$\mathbf{x}^{\mathrm{t}}_{k+1} = \mathbf{M}_{k+1}\mathbf{x}^{\mathrm{t}}_{k} + \boldsymbol{\eta}_{k+1}, \tag{9}$$

where

$\Rightarrow$ the model noise $\boldsymbol{\eta}_k$ has covariance matrix $\mathbf{Q}_k$,
$\Rightarrow$ which we suppose to be uncorrelated in time and uncorrelated with the background and observation errors.

- The objective function for the best, linear unbiased estimator (BLUE) for the sequence of states

$$\{\mathbf{x}_k,\ k = 0, 1, \ldots, K\}$$

is of the form

$$
\begin{aligned}
J(\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_K) = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}\right)^{\mathrm{T}} \left(\mathbf{P}_0^{\mathrm{b}}\right)^{-1} \left(\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}\right) \\
+\frac{1}{2}\sum_{k=0}^{K} \left(\mathbf{H}_k\mathbf{x}_k - \mathbf{y}_k\right)^{\mathrm{T}} \mathbf{R}_k^{-1} \left(\mathbf{H}_k\mathbf{x}_k - \mathbf{y}_k\right) \quad (10) \\
+\frac{1}{2}\sum_{k=0}^{K-1} \left(\mathbf{x}_{k+1} - \mathbf{M}_{k+1}\mathbf{x}_k\right)^{\mathrm{T}} \mathbf{Q}_{k+1}^{-1} \left(\mathbf{x}_{k+1} - \mathbf{M}_{k+1}\mathbf{x}_k\right).
\end{aligned}
$$

- This objective function has become a function of the complete sequence of states

$$
\{\mathbf{x}_k,\ k = 0, 1, \ldots, K\},
$$

and its minimization is known as "weak constraint four-dimensional variational assimilation," abbreviated as *weak constraint 4D-Var*.

- Equations (7) and (10), with an appropriate reformulation of the state and observation spaces, are special cases of the BLUE objective function.

EXAMPLES

Codes

Various open-source repositories and codes are available for both academic and operational data assimilation.

1. DARC: `https://research.reading.ac.uk/met-darc/` from Reading, UK.

2. DAPPER: `https://github.com/nansencenter/DAPPER` from Nansen, Norway.

3. DART: `https://dart.ucar.edu/` from NCAR, US, specialized in ensemble DA.

4. OpenDA: `https://www.openda.org/`.

5. Verdandi: `http://verdandi.sourceforge.net/` from INRIA, France.

6. PyDA: `https://github.com/Shady-Ahmed/PyDA`, a Python implementation for academic use.

7. Filterpy: `https://github.com/rlabbe/filterpy`, dedicated to KF variants.

8. EnKF; `https://enkf.nersc.no/`, the original Ensemble KF from Geir Evensen.

References

1. K. Law, A. Stuart, K. Zygalakis. *Data Assimilation. A Mathematical Introduction*. Springer, 2015.

2. G. Evensen. *Data assimilation, The Ensemble Kalman Filter*, 2nd ed., Springer, 2009.

3. A. Tarantola. *Inverse problem theory and methods for model parameter estimation.* SIAM. 2005.

4. O. Talagrand. Assimilation of observations, an introduction. *J. Meteorological Soc. Japan*, **75**, 191–209, 1997.

5. F.X. Le Dimet, O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus,* **38**(2), 97–110, 1986.

6. J.-L. Lions. Exact controllability, stabilization and perturbations for distributed systems. *SIAM Rev.*, **30**(1):1–68, 1988.

7. J. Nocedal, S.J. Wright. *Numerical Optimization*. Springer, 2006.

8. F. Tröltzsch. *Optimal Control of Partial Differential Equations*. AMS, 2010.