Parameterizations in Data Assimilation

(references, summary slides and examples)

ECMWF Training Course
19-21 May 2015

Philippe Lopez

Physical Aspects Section, Research Department, ECMWF
(Room 002)

A few references…

REFERENCES

- **Variational data assimilation:**

Lorenc, A., 1986, *Quarterly Journal of the Royal Meteorological Society*, **112**, 1177-1194.

Courtier, P. *et al.*, 1994, *Quarterly Journal of the Royal Meteorological Society*, **120**, 1367-1387.

Bouttier, F. and P. Courtier, 1999: Data assimilation concepts and methods, ECMWF lecture notes, available at ***http://www.ecmwf.int/newsevents/training/lecture_notes/LN_DA.html***.

Rabier, F. *et al.*, 2000, *Quarterly Journal of the Royal Meteorological Society*, **126**, 1143-1170.

- **The adjoint technique:**

Errico, R.M., 1997, *Bulletin of the American Meteorological Society*, **78**, 2577-2591.

- **Tangent-linear approximation:**

Errico, R.M. *et al.*, 1993, *Tellus*, **45A**, 462-477.

Errico, R.M., and K. Reader, 1999, *Quarterly Journal of the Royal Meteorological Society*, **125**, 169-195.

Janisková, M. *et al.*, 1999, *Monthly Weather Review*, **127**, 26-45..

- **Physical parameterizations for data assimilation:**

Mahfouf, J.-F., 1999, Tellus, **51A**, 147-166.

Janisková, M. *et al.*, 2002, *Quarterly Journal of the Royal Meteorological Society*, **128**, 1505-1527.

Tompkins, A. M., and M. Janisková, 2004, *Quarterly Journal of the Royal Meteorological Society*, **130**, 2495-2518.

Lopez, P., and E. Moreau, 2005, *Quarterly Journal of the Royal Meteorological Society*, **131**, 409-436.

Janisková, M., and P. Lopez, 2012: Linearized physics for data assimilation at ECMWF, *Technical Memorandum 666*, available from ECMWF, Reading, UK.

REFERENCES

- **<u>Microwave Radiative Transfer Model for data assimilation:</u>**

Bauer, P., May 2002, *Satellite Application Facility for Numerical Weather Prediction NWPSAF*-EC-TR-005, 27 pages.

- **<u>Assimilation of observations affected by precipitation and/or clouds:</u>**

Lopez, P. 2007: *Journal Atmospheric Sciences, Special Issue on the Assimilation of Satellite Cloud and Precipitation Observations in NWP Models*, **64**, 3770-3788 (Review paper on the use of moist physical parameterizations in data assimilation).

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, and E. Moreau, 2006a, *Quarterly Journal of the Royal Meteorological Society*, **132**, 2277-2306.

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, S. Saarinen, and M. Bonazzola, 2006b, *Quarterly Journal of the Royal Meteorological Society*, **132**, 2307-2332.

Fillion, L., and R.M. Errico, 1997, *Monthly Weather Review*, **125**, 2917-2942.

Ducrocq, V., *et al.*, 2000, *Quarterly Journal of the Royal Meteorological Society*, **126**, 3041-3065.

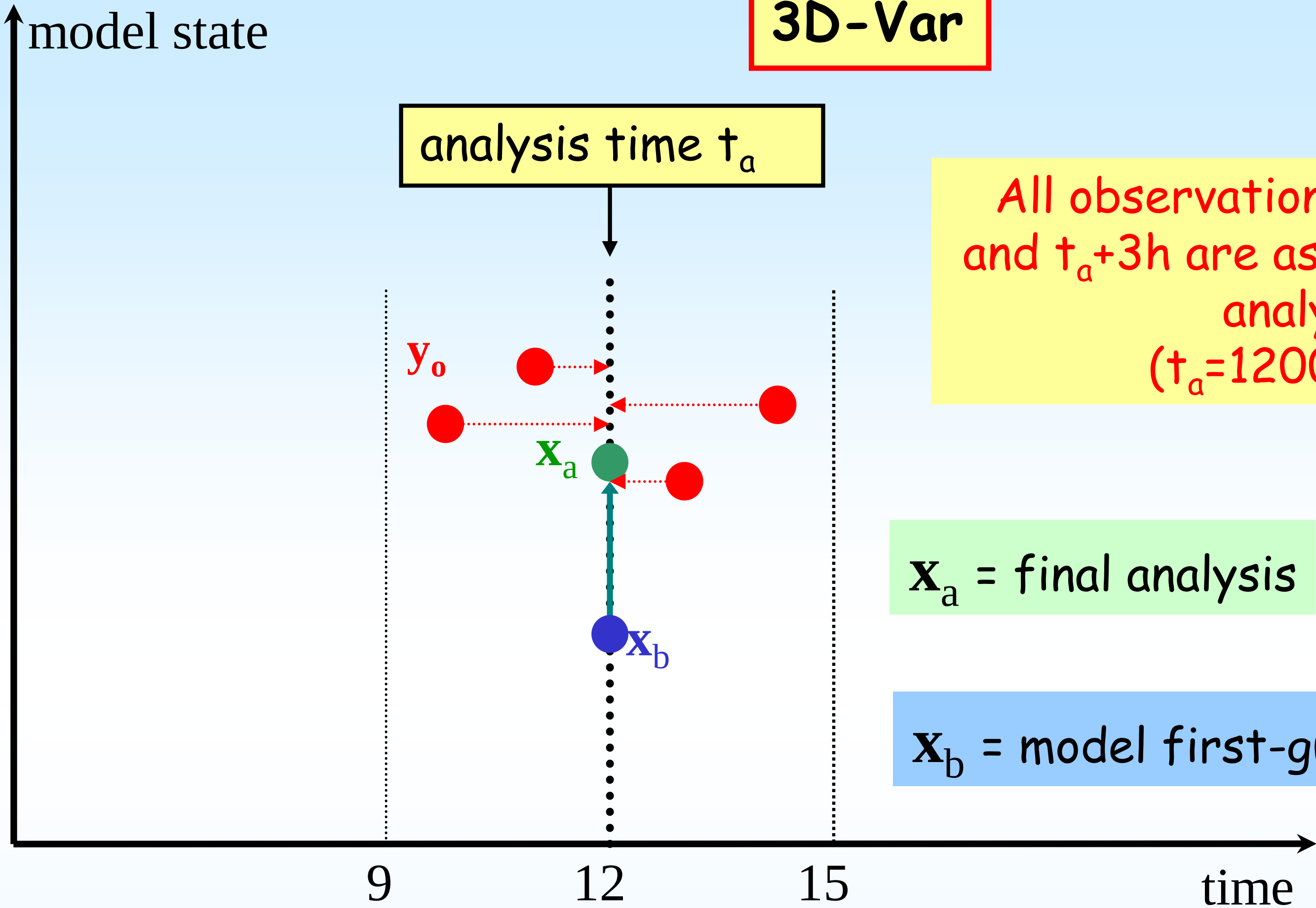
Hou, A. *et al.*, 2000, *Monthly Weather Review*, **128**, 509-537.

Marécal, V., and. J.-F. Mahfouf, 2002, *Monthly Weather Review*, **130**, 43-58.

Marécal, V., and. J.-F. Mahfouf, 2003, *Quarterly Journal of the Royal Meteorological Society*, **129**, 3137-3160.

A few slides of summary…

3D-Var



All observations y_o between t_a-3h and t_a+3h are assumed to be valid at analysis time (t_a=1200 UTC here)

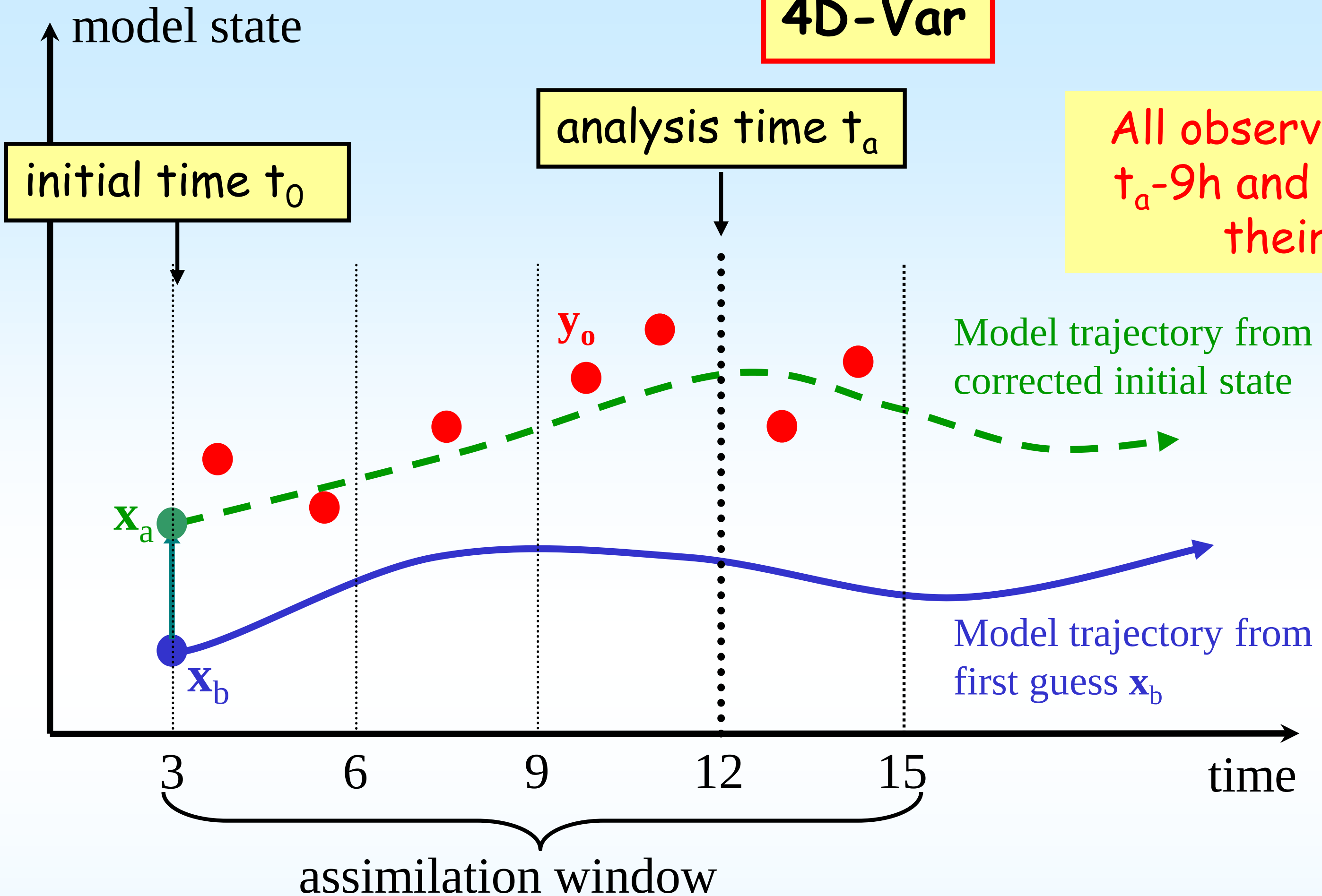
$\mathbf{x}_a$ = final analysis

$\mathbf{x}_b$ = model first-guess

$$\min J = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\left(H(\mathbf{x})-\mathbf{y}_o\right)^T \mathbf{R}^{-1}\left(H(\mathbf{x})-\mathbf{y}_o\right)$$

$$\Leftrightarrow \nabla_{\mathbf{x}} J = \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1}\left(H(\mathbf{x})-\mathbf{y}_o\right) = 0$$

4D-Var



All observations y_o between t_a-9h and t_a+3h are valid at their actual time

Forecast model is involved in minimization

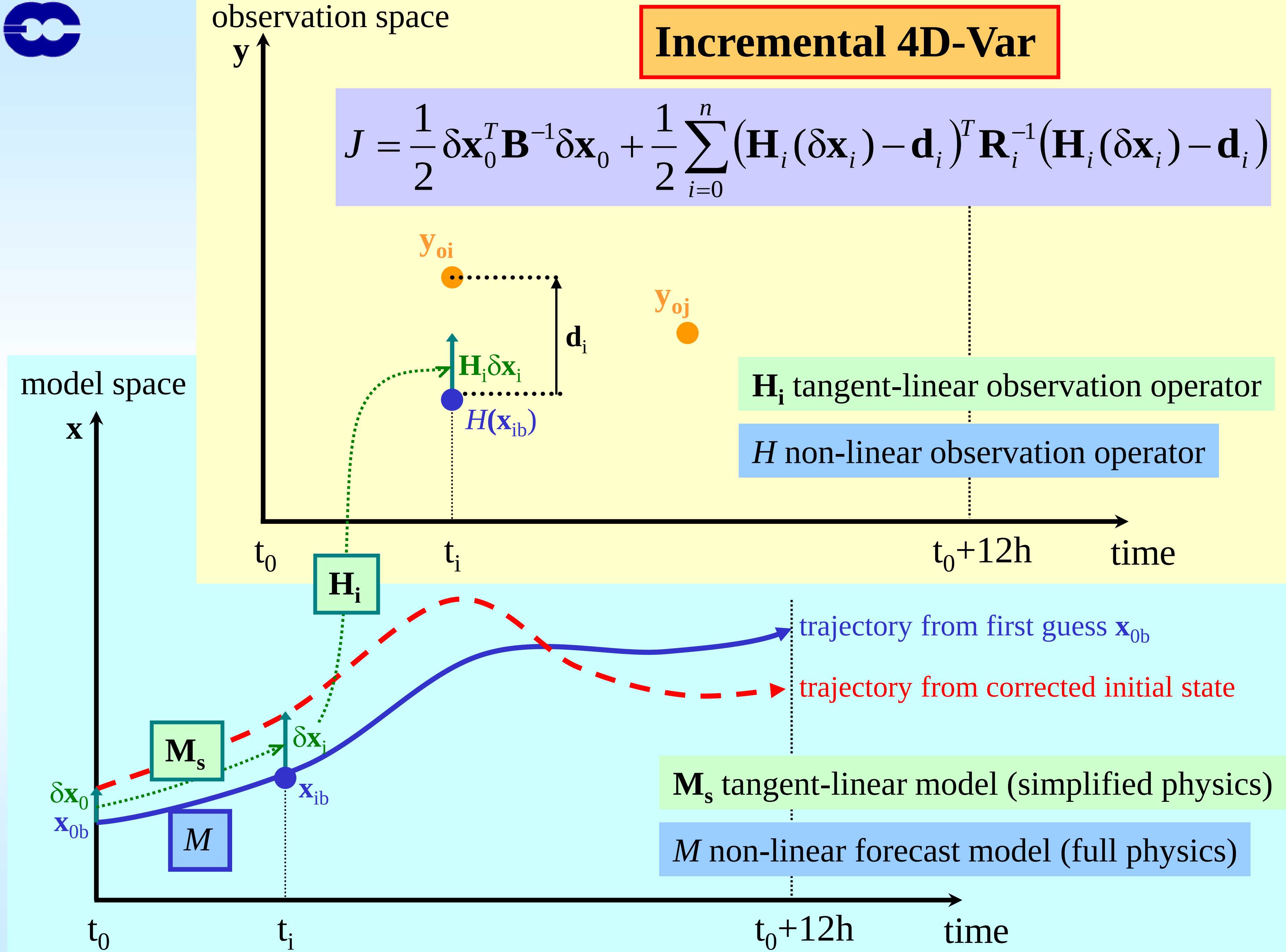
$$\min J = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T \ldots {}_{oi})^T \mathbf{R}_i^{-1}\left(H_i(M[\mathbf{x}_0])\right) - \mathbf{y}_{oi})$$

Adjoint of forecast model
with simplified linearized physics

$$\Leftrightarrow \nabla_{\mathbf{x}} J = \mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \sum_{i=0}^{n} \mathbf{M}^T[t_i, t_0] \mathbf{H}_i^T \mathbf{R}_i^{-1}\left(H_i(M[\mathbf{x}_0])\right) - \mathbf{y}_{oi}) = 0$$

Incremental 4D-Var

$$J = \frac{1}{2}\delta\mathbf{x}_0^T\mathbf{B}^{-1}\delta\mathbf{x}_0 + \frac{1}{2}\sum_{i=0}^{n}\left(\mathbf{H}_i(\delta\mathbf{x}_i) - \mathbf{d}_i\right)^T\mathbf{R}_i^{-1}\left(\mathbf{H}_i(\delta\mathbf{x}_i) - \mathbf{d}_i\right)$$



Summary

- Variational data assimilation relies on some essential assumptions:
  - Gaussian and unbiased model background and observation errors,
  - Quasi-linearity of all operators involved (H, M).

- Given some background fields and a very large set of asynchronous observations available within a certain time window (6 or 12h-long), 4D-Var searches the statistically optimal initial model state $\mathbf{x_0}$ that minimizes the cost function:

$$J(\mathbf{x_0}) = J_b(\mathbf{x_0}) + J_o(HM(\mathbf{x_0}))$$

- The calculation of $\nabla_{\mathbf{x0}} J$ requires the coding of tangent-linear and adjoint versions of the observation operator H and of the full nonlinear forecast model M (including physical parameterizations).

- The tangent-linear and adjoint forecast models, $\mathbf{M}$ and $\mathbf{M}^T$, are usually based on a simplified version of the full nonlinear model, M, to reduce computational cost in the iterative minimization and to avoid nonlinearities.

Summary

- The aim of data assimilation is to produce a statistically optimal model state (the **analysis**) which can be used to initialize a forecast model.

- In variational DA, this is achieved by minimizing a **cost function**, J, that measures the distance to the **model background** and **observations**, weighted by their respective **error statistics.**

  In 3D-Var:

$$J = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\left(H(\mathbf{x})-\mathbf{y}_o\right)^T \mathbf{R}^{-1}\left(H(\mathbf{x})-\mathbf{y}_o\right)$$

- **Parameterizations** are needed during the minimization to convert the model control variables (T,q,u,v,P_s) into observed equivalents (e.g. reflectivities, radiances,…) ("**observation operator**" H).

- Fundamental assumptions:
  - **Background and observation errors** are **Gaussian** and **unbiased**.
  - Observation operator H is **not too non-linear**.

Summary

- The aim of a data assimilation system is to produce a statistically optimal model state (the **analysis**) that can be used to initialize a forecast model.

- In variational DA this is achieved by minimizing iteratively a **cost function** (J) that measures the distance to the **model background** and **observations**, weighted by their respective **error statistics** (**Gaussian** and **unbiased**).

- **Parameterizations** are needed during the minimization to:

  - convert the model variables (T,q,u,v,P_s) into observed equivalents (e.g. reflectivities, radiances,…) (observation operator H),
  - evolve the model state from analysis time to observation time (4D-Var).

- The **tangent-linear** and **adjoint** versions of these usually **simplified** parameterizations must be coded, **tested**, and some **regularization** is usually needed to eliminate **discontinuities/non-linearities**.

- The **adjoint** version of the parameterizations is needed to compute the gradient of the cost function with respect to the initial model state, $\mathbf{x}$:

$$\nabla_{\mathbf{x}_0} J = \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \mathbf{M}[t_i, t_0]^T \mathbf{H}^T \nabla_{\mathbf{y}} J_o \quad \text{with } \nabla_{\mathbf{y}} J_o = \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}_o)$$

A few examples and exercises…

A simple analysis problem

Exercise

- 6-hour forecast of 2m temperature produced by the model:
  $\mathbf{x}_b$ with a standard deviation of forecast error σ_b

- observation of 2m temperature:
  $\mathbf{y}_o$ with a standard deviation of observation error σ_o

- The best estimate of the 2m temperature (analysis) minimizes the departure from the model first-guess and from the observation according to their relative accuracies:

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\left(\frac{\mathbf{x}-\mathbf{x}_b}{\sigma_b}\right)^2 + \frac{1}{2}\left(\frac{\mathbf{x}-\mathbf{y}_o}{\sigma_o}\right)^2$$

Since the analysis $\mathbf{x}_a$ minimizes the cost function, then

$$\frac{\partial \mathcal{J}}{\partial \mathbf{x}}(\mathbf{x}_a) = 0$$

Analysis state can be written as:

$$\mathbf{x}_a = \mathbf{x}_b + \alpha(\mathbf{y}_o - \mathbf{x}_b)$$

A simple analysis problem

Exercise

- **Problem:**
  - Find the coefficient α.
  - Show that the variance of the analysis error is:

$$\frac{1}{{\sigma_a}^2} = \frac{1}{{\sigma_b}^2} + \frac{1}{{\sigma_o}^2}$$

(Note: $\sigma^2 = \overline{(\mathbf{x} - \mathbf{x}_t)^2}$, where $\mathbf{x}_t$ is the unknown true state).

A simple analysis problem
Solution

- Since the analysis $\mathbf{x}_a$ minimizes the cost function, then

$$\frac{\partial \mathcal{J}}{\partial \mathbf{x}}(\mathbf{x}_a) = 0$$

$$\frac{\partial \mathcal{J}}{\partial \mathbf{x}} = \frac{\mathbf{x} - \mathbf{x}_b}{{\sigma_b}^2} + \frac{\mathbf{x} - \mathbf{y}_o}{{\sigma_o}^2} = 0$$

$$\frac{\mathbf{x}{\sigma_o}^2 - \mathbf{x}_b{\sigma_o}^2 + \mathbf{x}{\sigma_b}^2 - \mathbf{y}_o{\sigma_b}^2}{{\sigma_b}^2{\sigma_o}^2} = 0$$

$$(*) \qquad \mathbf{x}({\sigma_o}^2 + {\sigma_b}^2) = \mathbf{x}_b{\sigma_o}^2 + \mathbf{y}_o{\sigma_b}^2$$

$$\mathbf{x} = \frac{\mathbf{x}_b{\sigma_o}^2 + \mathbf{y}_o{\sigma_b}^2 - \mathbf{x}_b{\sigma_b}^2 + \mathbf{x}_b{\sigma_b}^2}{{\sigma_o}^2 + {\sigma_b}^2}$$

$$\mathbf{x} = \frac{\mathbf{x}_b({\sigma_o}^2 + {\sigma_b}^2) + {\sigma_b}^2(\mathbf{y}_o - \mathbf{x}_b)}{{\sigma_o}^2 + {\sigma_b}^2}$$

$$\mathbf{x} = \mathbf{x}_b + \boxed{\frac{{\sigma_b}^2}{{\sigma_o}^2 + {\sigma_b}^2}}(\mathbf{y}_o - \mathbf{x}_b)$$

α

A simple analysis problem

Solution

- Analysis error:
  starting from equation (*) one gets

$$
\begin{aligned}
\mathbf{x}_a &= \mathbf{x}_b \frac{{\sigma_o}^2}{{\sigma_o}^2 + {\sigma_b}^2} + \mathbf{y}_o \frac{{\sigma_b}^2}{{\sigma_o}^2 + {\sigma_b}^2} \\
\mathbf{x}_a - \mathbf{x}_t &= (\mathbf{x}_b - \mathbf{x}_t) \frac{{\sigma_o}^2}{{\sigma_o}^2 + {\sigma_b}^2} + (\mathbf{y}_o - \mathbf{x}_t) \frac{{\sigma_b}^2}{{\sigma_o}^2 + {\sigma_b}^2} \\
\overline{(\mathbf{x}_a - \mathbf{x}_t)^2} &= \overline{(\mathbf{x}_b - \mathbf{x}_t)^2} \frac{{\sigma_o}^4}{({\sigma_o}^2 + {\sigma_b}^2)^2} + \overline{(\mathbf{y}_o - \mathbf{x}_t)^2} \frac{{\sigma_b}^4}{({\sigma_o}^2 + {\sigma_b}^2)^2} \\
&+ \overline{(\mathbf{x}_b - \mathbf{x}_t)(\mathbf{y}_o - \mathbf{x}_t)} \frac{{\sigma_o}^2 {\sigma_b}^2}{({\sigma_o}^2 + {\sigma_b}^2)^2}
\end{aligned}
$$

Since background and observation errors are assumed to be uncorrelated:

$$
Cov(\mathbf{x}_b, \mathbf{y}_o) = \overline{(\mathbf{x}_b - \mathbf{x}_t)(\mathbf{y}_o - \mathbf{x}_t)} = 0
$$

which gives

$$
\begin{aligned}
{\sigma_a}^2 &= \frac{{\sigma_b}^2 {\sigma_o}^4}{({\sigma_o}^2 + {\sigma_b}^2)^2} + \frac{{\sigma_b}^4 {\sigma_o}^2}{({\sigma_o}^2 + {\sigma_b}^2)^2} \\
{\sigma_a}^2 &= \frac{{\sigma_b}^2 {\sigma_o}^2}{{\sigma_o}^2 + {\sigma_b}^2} \iff \boxed{\frac{1}{{\sigma_a}^2} = \frac{1}{{\sigma_b}^2} + \frac{1}{{\sigma_o}^2}}
\end{aligned}
$$

1D-Var assimilation of physical fluxes

Example

- observation operator = physical parametrization
  - example: thermal radiation at the surface *(Brunt, 1934)*

$$R_L = \sigma T^4 (a + b\sqrt{e})$$

    where T is the screen level temperature and e is the water vapour pressure

- model temperature and humidity (T_b, e_b) can be modified to better match an observation of thermal radiation $R_{L\,o}$

- the optimal values of T and e minimize the following cost function:

$$\mathcal{J}(T,e) = \frac{1}{2}\left(\frac{T - T_b}{\sigma_{T\,b}}\right)^2 + \frac{1}{2}\left(\frac{e - e_b}{\sigma_{e\,b}}\right)^2 + \frac{1}{2}\left(\frac{R_L - R_{L\,o}}{\sigma_o}\right)^2$$

gradient of the cost function:

$$\frac{\partial \mathcal{J}}{\partial T} = \frac{T - T_b}{\sigma_{T\,b}^2} + \frac{\partial R_L}{\partial T}\left(\frac{R_L - R_{L\,o}}{\sigma_o^2}\right)$$

$$\frac{\partial \mathcal{J}}{\partial e} = \frac{e - e_b}{\sigma_{e\,b}^2} + \frac{\partial R_L}{\partial e}\left(\frac{R_L - R_{L\,o}}{\sigma_o^2}\right)$$

1D-Var assimilation of physical fluxes

Example

- tangent-linear operator:

$$\delta R_L = \left(\frac{\partial R_L}{\partial T} \quad \frac{\partial R_L}{\partial e} \right) \cdot \begin{pmatrix} \delta T \\ \delta e \end{pmatrix}$$

- adjoint of the tangent-linear operator:

$$\left(\frac{\partial J_o}{\partial T} \quad \frac{\partial J_o}{\partial e} \right) = \begin{pmatrix} \frac{\partial R_L}{\partial T} \\ \frac{\partial R_L}{\partial e} \end{pmatrix} \cdot \left(\frac{\partial J_o}{\partial R_L} \right)$$

with $\frac{\partial R_L}{\partial T} = 4\sigma T^3 (a + b\sqrt{e})$ and $\frac{\partial R_L}{\partial e} = \frac{b\sigma T^4}{2\sqrt{e}}$

EXERCISE 2

- write tangent linear (TL) and adjoint (AD) code of the following non-linear (NL) code (FORTRAN 90)

```
 SUBROUTINE Longwave_Radiation (EA, TA, RL)

! Longwave radiation at the surface (RL in Watts/m2)
! Empirical expression from Brunt (1934) depending upon
! TA = air temperature (K)
! EA = water vapour pressure (hPa)
!
! Non-linear routine
! ------------------

 IMPLICIT NONE
 REAL , INTENT(IN)  :: EA, TA
 REAL , INTENT(OUT) :: RL
 REAL , PARAMETER   :: A=0.75, B=0.003
 REAL , PARAMETER   :: SIGMA=5.67E-8
 REAL               :: ZEMIS

 ZEMIS = A+B*SQRT(EA)
 RL    = ZEMIS*SIGMA*TA**4

 END SUBROUTINE Longwave_Radiation
```

EXERCISE 2 - solution

- tangent linear code

```
SUBROUTINE Longwave_Radiation_TL (EA5, TA5, RL5, EA, TA, RL)

! Longwave radiation at the surface (RL in Watts/m2)
! Empirical expression from Brunt (1934) depending upon
! TA = air temperature (K)
! EA = water vapour pressure (hPa)
!
! Tangent-linear routine
! ----------------------

IMPLICIT NONE
REAL , INTENT(IN)  :: EA5, TA5          ! Trajectory
REAL , INTENT(OUT) :: RL5               ! Trajectory
REAL , INTENT(IN)  :: EA, TA            ! Perturbation
REAL , INTENT(OUT) :: RL                ! Perturbation
REAL , PARAMETER   :: A=0.75, B=0.003
REAL , PARAMETER   :: SIGMA=5.67E-8
REAL               :: ZEMIS5, ZEMIS

ZEMIS5 = A+B*SQRT(EA5)
ZEMIS  = B/(2.*SQRT(EA5))*EA
RL5    = ZEMIS5*SIGMA*TA5**4
RL     = ZEMIS *SIGMA*TA5**4 + 4.*ZEMIS5*SIGMA*TA5**3*TA

END SUBROUTINE Longwave_Radiation_TL
```

EXERCISE 2 - solution

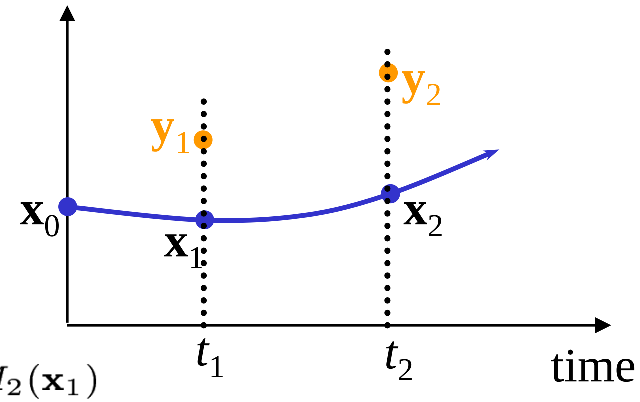
- adjoint code

```
SUBROUTINE Longwave_Radiation_AD (EA5, TA5, RL5, EA, TA, RL)

! Longwave radiation at the surface (RL in Watts/m2)
! Empirical expression from Brunt (1934) depending upon
! TA = air temperature (K)
! EA = water vapour pressure (hPa)
!
! Adjoint routine
! ---------------

IMPLICIT NONE
REAL , INTENT(IN)  :: EA5, TA5        ! Trajectory
REAL , INTENT(OUT) :: RL5             ! Trajectory
REAL , INTENT(IN)  :: EA, TA          ! Perturbation
REAL , INTENT(OUT) :: RL              ! Perturbation
REAL , PARAMETER   :: A=0.75, B=0.003
REAL , PARAMETER   :: SIGMA=5.67E-8
REAL               :: ZEMIS5, ZEMIS

! Trajectory computations

ZEMIS5 = A+B*SQRT(EA5)
RL5    = ZEMIS5*SIGMA*TA5**4

! Initialization of local variables
```

```
ZEMIS = 0.

! Adjoint computation

TA = TA + 4.*ZEMIS5*SIGMA*TA5**3*RL
ZEMIS = ZEMIS + SIGMA*TA5**4*RL
RL = 0.
EA = EA + B/(2.*SQRT(EA5))*ZEMIS
ZEMIS = 0.

END SUBROUTINE Longwave_Radiation_AD
```

A simple 4D-Var analysis problem

Exercise



- Observations $\mathbf{y}_1$ and $\mathbf{y}_2$ at time t_1 and t_2
- Model first guess $\mathbf{x}_1$ and $\mathbf{x}_2$ at time t_1 and t_2
- Time evolution from the initial time t_0:

$$\mathbf{x}_1 = M_1(\mathbf{x}_0) \quad \text{and} \quad \mathbf{x}_2 = M_2(\mathbf{x}_1)$$

- Cost function:

$$\mathcal{J}(\mathbf{x}_0) = \frac{1}{2}\left(\frac{\mathbf{x}_1 - \mathbf{y}_1}{\sigma_1}\right)^2 + \frac{1}{2}\left(\frac{\mathbf{x}_2 - \mathbf{y}_2}{\sigma_2}\right)^2 = \mathcal{J}_1 + \mathcal{J}_2$$

- **Problem:**
  Estimate the gradient of $\mathcal{J}$ with respect to the initial state $\mathbf{x}_0$.

A simple 4D-Var analysis problem

Solution

- At time t_2:

$$\frac{\partial \mathcal{J}_2}{\partial \mathbf{x}_2} = \frac{\mathbf{x}_2 - \mathbf{y}_2}{\sigma_2^2}$$

- At time t_1:

$$\frac{\partial \mathcal{J}_2}{\partial \mathbf{x}_1} = \mathbf{M}_2{}^T \left(\frac{\partial \mathcal{J}_2}{\partial \mathbf{x}_2} \right) = \mathbf{M}_2{}^T \left(\frac{\mathbf{x}_2 - \mathbf{y}_2}{\sigma_2^2} \right)$$

$$\frac{\partial \mathcal{J}_1}{\partial \mathbf{x}_1} = \frac{\mathbf{x}_1 - \mathbf{y}_1}{\sigma_1^2}$$

- At time t_0:

$$\frac{\partial \mathcal{J}_2}{\partial \mathbf{x}_0} = \mathbf{M}_1{}^T \left(\frac{\partial \mathcal{J}_2}{\partial \mathbf{x}_1} \right) = \mathbf{M}_1{}^T \left[\mathbf{M}_2{}^T \left(\frac{\mathbf{x}_2 - \mathbf{y}_2}{\sigma_2^2} \right) \right]$$

$$\frac{\partial \mathcal{J}_1}{\partial \mathbf{x}_0} = \mathbf{M}_1{}^T \left(\frac{\partial \mathcal{J}_1}{\partial \mathbf{x}_1} \right) = \mathbf{M}_1{}^T \left(\frac{\mathbf{x}_1 - \mathbf{y}_1}{\sigma_1^2} \right)$$

- Finally:

$$\frac{\partial \mathcal{J}}{\partial \mathbf{x}_0} = \frac{\partial \mathcal{J}_1}{\partial \mathbf{x}_0} + \frac{\partial \mathcal{J}_2}{\partial \mathbf{x}_0} = \mathbf{M}_1{}^T \left(\frac{\mathbf{x}_1 - \mathbf{y}_1}{\sigma_1^2} \right) + \mathbf{M}_1{}^T \left[\mathbf{M}_2{}^T \left(\frac{\mathbf{x}_2 - \mathbf{y}_2}{\sigma_2^2} \right) \right]$$