Data assimilation.
Advanced Methods (2)

Olivier Talagrand
ESA summer school
Earth System Monitoring & Modelling
ESRIN, Frascati, Italy
2 August 2006

Adjoint Method (continued 2)

$\mathcal{J}(\xi_0) \;=\; (1/2)\,(\boldsymbol{x}_0^{\,b} - \xi_0)^{\mathrm{T}}\,[\boldsymbol{P}_0^{\,b}]^{-1}\,(\boldsymbol{x}_0^{\,b} - \xi_0) + (1/2)\,\Sigma_k[\boldsymbol{y}_k - \boldsymbol{H}_k\xi_k]^{\mathrm{T}}\,\boldsymbol{R}_k^{-1}\,[\boldsymbol{y}_k - \boldsymbol{H}_k\xi_k]$

subject to $\xi_{k+1} = \boldsymbol{M}_k\xi_k$, $\quad k = 0, \ldots, K\text{-}1$

Control variable $\quad \xi_0 = \boldsymbol{u}$

Adjoint equation

$$\boldsymbol{\lambda}_K = \boldsymbol{H}_K^{\mathrm{T}}\boldsymbol{R}_K^{-1}\,[\boldsymbol{H}_K\,\xi_K - \boldsymbol{y}_K]$$

$$\boldsymbol{\lambda}_k = \boldsymbol{M}_k^{\mathrm{T}}\boldsymbol{\lambda}_{k+1} + \boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{R}_k^{-1}\,[\boldsymbol{H}_k\,\xi_k - \boldsymbol{y}_k] \qquad k = K\text{-}1, \ldots, 1$$

$$\boldsymbol{\lambda}_0 = \boldsymbol{M}_0^{\mathrm{T}}\boldsymbol{\lambda}_1 + \boldsymbol{H}_0^{\mathrm{T}}\boldsymbol{R}_0^{-1}\,[\boldsymbol{H}_0\,\xi_0 - \boldsymbol{y}_0] + [\boldsymbol{P}_0^{\,b}]^{-1}\,(\xi_0 - \boldsymbol{x}_0^{\,b})$$

$$\nabla_{\boldsymbol{u}}\mathcal{J} = \boldsymbol{\lambda}_0$$

Result of direct integration (ξ_k), which appears in quadratic terms in expression of objective function, must be kept in memory from direct integration.

Adjoint Method (continued 3)

Nonlinearities ?

$$\mathcal{J}(\xi_0) = (1/2)\,(\boldsymbol{x}_0^{\,b} - \xi_0)^{\mathrm{T}}\,[\boldsymbol{P}_0^{\,b}]^{-1}\,(\boldsymbol{x}_0^{\,b} - \xi_0) + (1/2)\,\Sigma_k[\boldsymbol{y}_k - \boldsymbol{H}_k(\xi_k)]^{\mathrm{T}}\,\boldsymbol{R}_k^{-1}\,[\boldsymbol{y}_k - \boldsymbol{H}_k(\xi_k)]$$

subject to $\xi_{k+1} = \boldsymbol{M}_k(\xi_k)$, $\quad k = 0, \ldots, K\text{-}1$

Control variable $\quad \xi_0 = \boldsymbol{u}$

Adjoint equation

$$\boldsymbol{\lambda}_K = \boldsymbol{H}_K'^{\mathrm{T}}\,\boldsymbol{R}_K^{-1}\,[\boldsymbol{H}_K(\xi_K) - \boldsymbol{y}_K]$$

$$\boldsymbol{\lambda}_k = \boldsymbol{M}_k'^{\mathrm{T}}\boldsymbol{\lambda}_{k+1} + \boldsymbol{H}_k'^{\mathrm{T}}\,\boldsymbol{R}_k^{-1}\,[\boldsymbol{H}_k(\xi_k) - \boldsymbol{y}_k] \qquad k = K\text{-}1, \ldots, 1$$

$$\boldsymbol{\lambda}_0 = \boldsymbol{M}_0'^{\mathrm{T}}\boldsymbol{\lambda}_1 + \boldsymbol{H}_0'^{\mathrm{T}}\,\boldsymbol{R}_0^{-1}\,[\boldsymbol{H}_0(\xi_0) - \boldsymbol{y}_0] + [\boldsymbol{P}_0^{\,b}]^{-1}\,(\boldsymbol{x}_0^{\,b} - \xi_0)$$

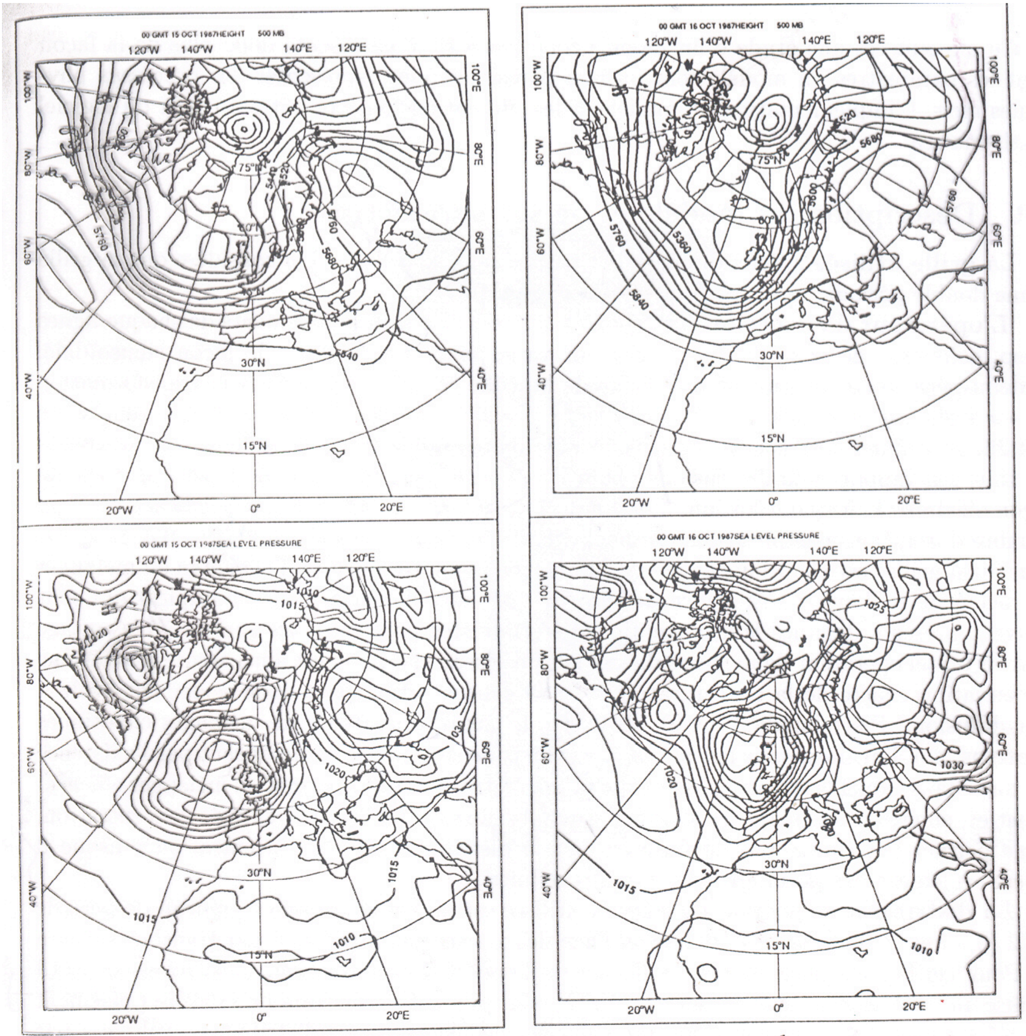
$$\nabla_{\boldsymbol{u}}\mathcal{J} = \boldsymbol{\lambda}_0$$

Not heuristic (it gives the exact gradient $\nabla_{\boldsymbol{u}}\mathcal{J}$), and really used as described here.

Adjoint Method (continued 4)

It works (Le Dimet, Courtier *et al*.) !

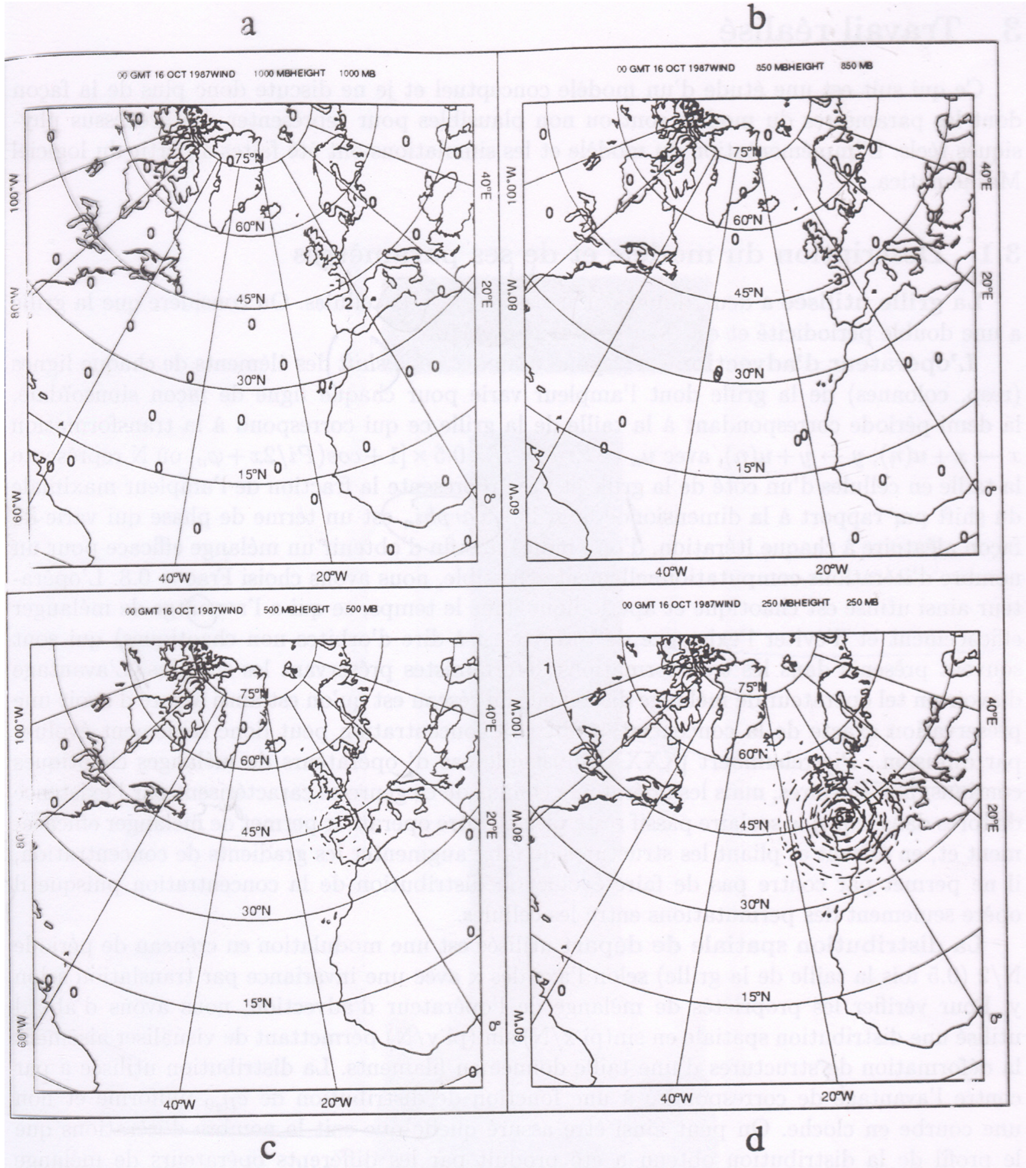
'4*D-Var*'



b d

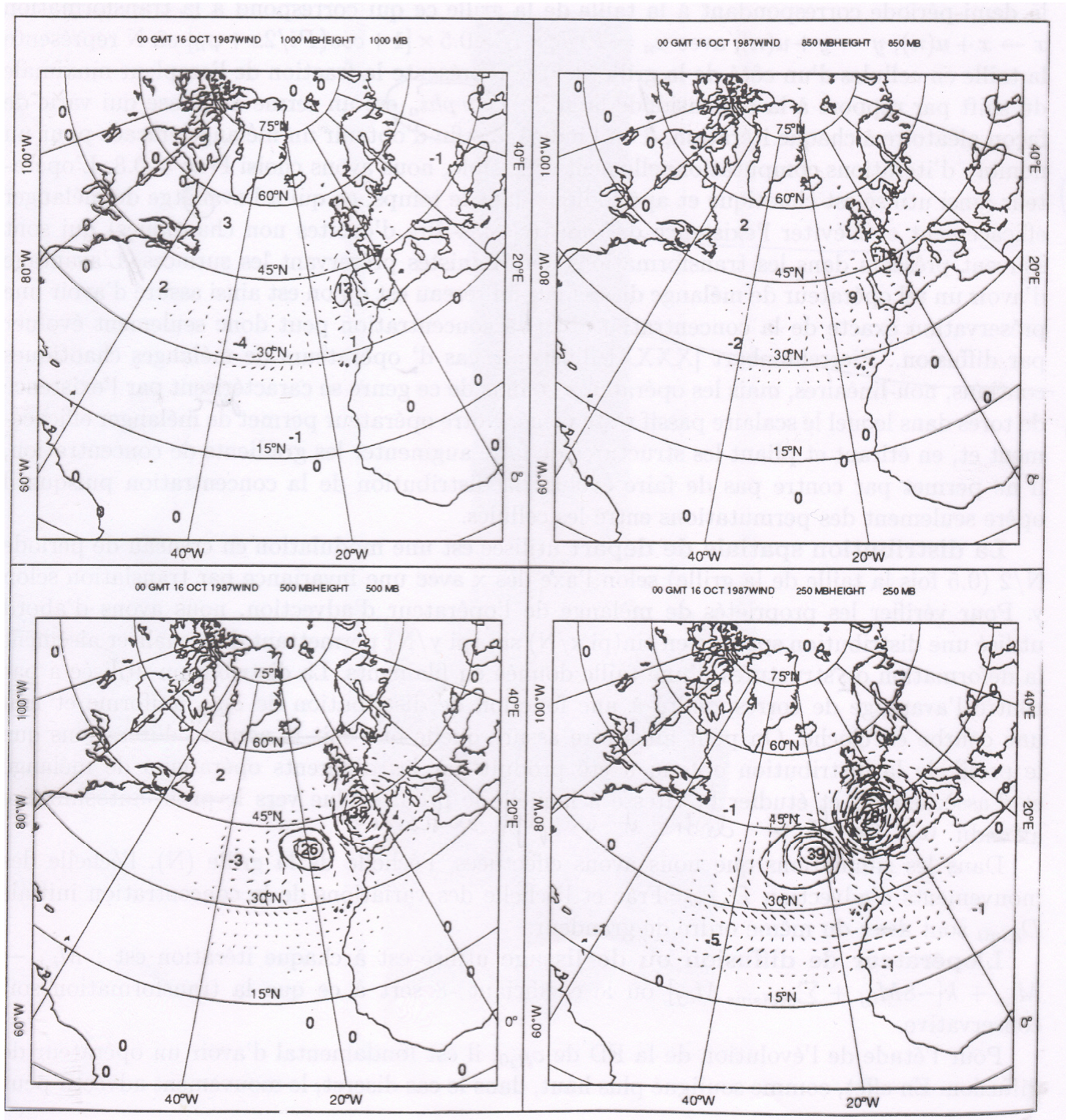
FIG. 1. Background fields for 0000 UTC 15 October–0000 UTC 16 October 1987. Shown here are the Northern Hemisphere (a) 500-hPa geopotential height and (b) mean sea level pressure for 15 October and the (c) 500-hPa geopotential height and (d) mean sea level pressure for 16 October. The fields for 15 October are from the initial estimate of the initial conditions for the 4DVAR minimization. The fields for 16 October are from the 24-h T63 adiabatic model forecast from the initial conditions. Contour intervals are 80 m and 5 hPa.

Thépaut *et al.*, 1993, *Mon. Wea. Rev.*, **121**, 3393-3414



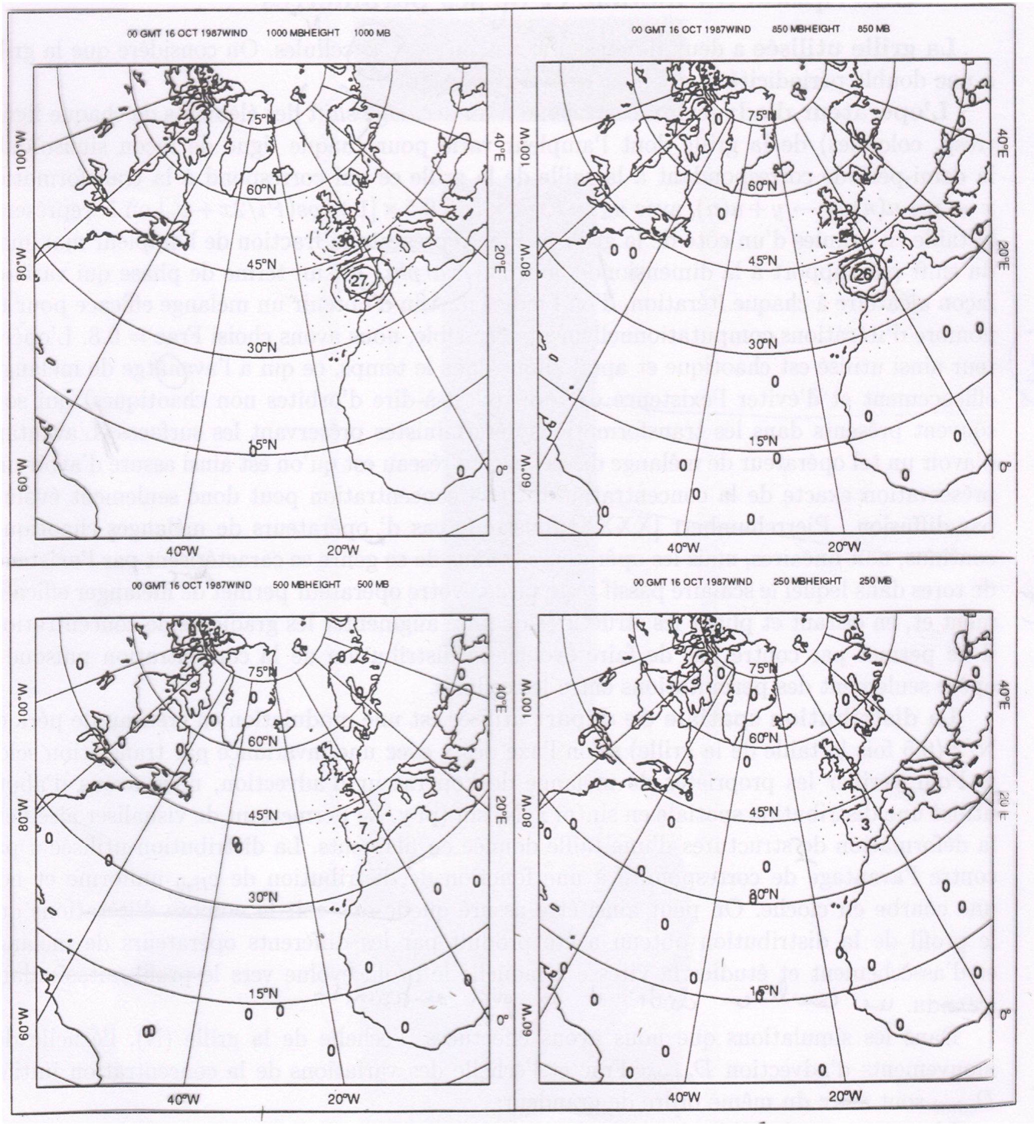
Analysis increments in a 3D-Var corresponding to a height observation at the 250-hPa pressure level (no temporal evolution of background error covariance matrix)

Thépaut *et al*., 1993, *Mon. Wea. Rev.*, **121**, 3393-3414



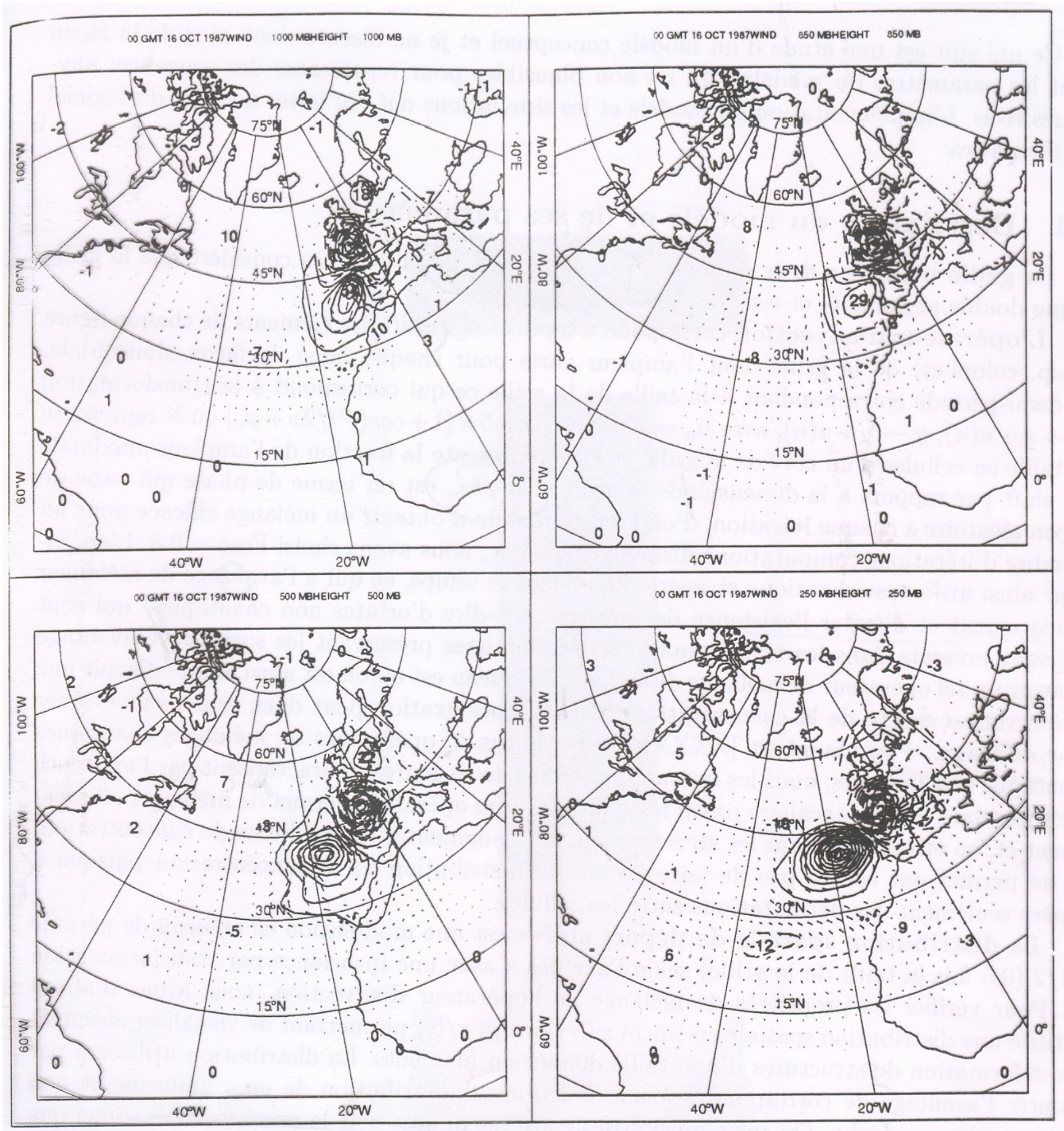
Same as before, but at the end of a 24-hr 4D-Var

Thépaut *et al.*, 1993, *Mon. Wea. Rev.*, **121**, 3393-3414



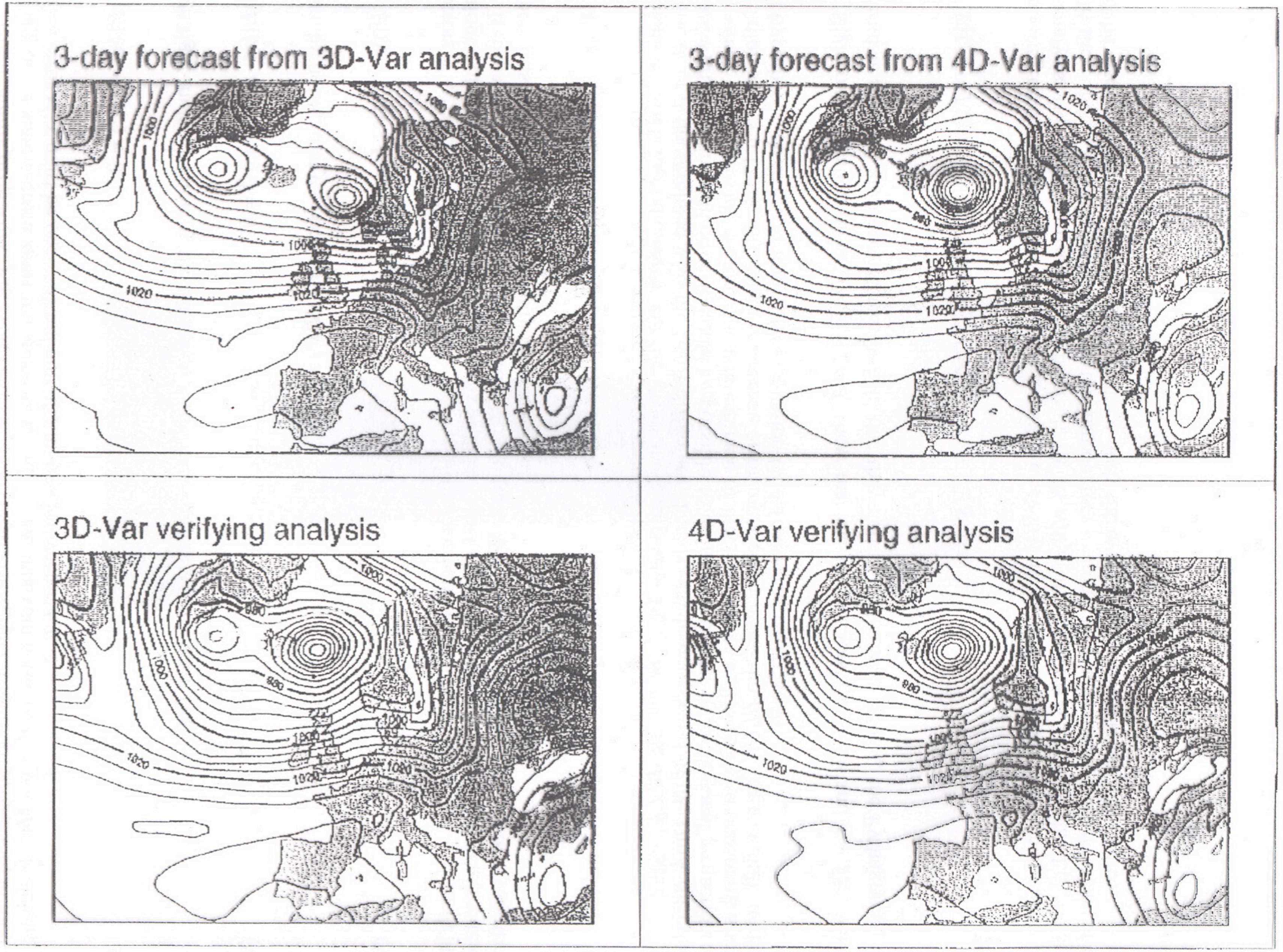
Analysis increments in a 3D-Var corresponding to a *u*-component wind observation at the 1000-hPa pressure level (no temporal evolution of background error covariance matrix)

Thépaut *et al.*, 1993, *Mon. Wea. Rev.*, **121**, 3393-3414



Same as before, but at the end of a 24-hr 4D-Var

Thépaut *et al*., 1993, *Mon. Wea. Rev.*, **121**, 3393-3414



ECMWF, Results on one FASTEX case (1997)

4D-Var is now used operationally at ECMWF, Météo-France, Meteorological Office (UK), Canadian Meteorological Service (together with an ensemble assimilation system), Japan Meteorological Agency

Model error is ignored

Strong Constraint Variational Assimilation

Weak constraint variational assimilation allows for errors in the assimilating model

Data

- Background estimate at time 0

$$x_0^b = x_0 + \zeta_0^b \qquad E(\zeta_0^b \zeta_0^{bT}) = P_0^b$$

- Observations at times $k = 0, \ldots, K$

$$y_k = H_k x_k + \varepsilon_k \qquad E(\varepsilon_k \varepsilon_k^T) = R_k$$

- Model

$$x_{k+1} = M_k x_k + \eta_k \qquad E(\eta_k \eta_k^T) = Q_k \qquad k = 0, \ldots, K\text{-}1$$

Errors assumed to be unbiased and uncorrelated in time, H_k and M_k linear

Then objective function

$$(\xi_0, \xi_1, \ldots, \xi_K) \to$$

$$\mathcal{J}(\xi_0, \xi_1, \ldots, \xi_K)$$

$$= (1/2)\ (x_0{}^b - \xi_0)^{\mathrm{T}}\ [P_0{}^b]^{-1}\ (x_0{}^b - \xi_0)$$

$$+ (1/2)\ \Sigma_{k=0,\ldots,K}[y_k - H_k\xi_k]^{\mathrm{T}}\ R_k{}^{-1}\ [y_k - H_k\xi_k]$$

$$+ (1/2)\ \Sigma_{k=0,\ldots,K-1}[\xi_{k+1} - M_k\xi_k]^{\mathrm{T}}\ Q_k{}^{-1}\ [\xi_{k+1} - M_k\xi_k]$$

Can include nonlinear M_k and/or H_k.

Dual Algorithm for Variational Assimilation (aka *Physical Space Analysis System*, *PSAS,* pronounced *'peezaz',* developed at Data Assimilation Office, NASA, Greenbelt)

$$x^a = x^b + P^b H^T [HP^bH^T + R]^{-1} (y - Hx^b)$$

$$x^a = x^b + P^b H^T \Lambda^{-1} d = x^b + P^b H^T m$$

where $\Lambda \equiv HP^bH^T + R$, $d \equiv y - Hx^b$ and $m \equiv \Lambda^{-1} d$ maximises

$$\mu \rightarrow \mathcal{K}(\mu) = -(1/2)\ \mu^T \Lambda \mu + d^T\mu$$

Maximisation is performed in (dual of) observation space.

Dual Algorithm for Variational Assimilation (continuation 2)

Extends to time dimension, and to weak-constraint case, by defining state vector as

$$x \equiv (x_0^{\mathrm{T}}, x_1^{\mathrm{T}}, \ldots, x_K^{\mathrm{T}})^{\mathrm{T}}$$

or, equivalently, but more conveniently, as

$$x \equiv (x_0^{\mathrm{T}}, \eta_0^{\mathrm{T}}, \ldots, \eta_{K-1}^{\mathrm{T}})^{\mathrm{T}}$$

where, as before

$$\eta_k = x_{k+1} - M_k x_k \quad , \qquad\qquad k = 0, \ldots, K\text{-}1$$

The background for x_0 is x_0^b, the background for η_k is 0. Complete background is

$$x^b = (x_0^{b\mathrm{T}}, 0^{\mathrm{T}}, \ldots, 0^{\mathrm{T}})^{\mathrm{T}}$$

It is associated with error covariance matrix

$$P^b = \mathrm{diag}(P_0^b, Q_0, \ldots, Q_{K\text{-}1})$$

Dual Algorithm for Variational Assimilation (continuation 3)

For any state vector $\xi = (\xi_0^{\mathrm{T}}, \upsilon_0^{\mathrm{T}}, \ldots, \upsilon_{K\text{-}1}^{\mathrm{T}})^{\mathrm{T}}$, the observation operator H

$$\xi \rightarrow H\xi = (u_0^{\mathrm{T}}, \ldots, u_K^{\mathrm{T}})^{\mathrm{T}}$$

is defined by the sequence of operations

$$u_0 = H_0 \xi_0$$

then for $k = 0, \ldots, K\text{-}1$

$$\xi_{k+1} = M_k \xi_k + \upsilon_k$$
$$u_{k+1} = H_{k+1}\, \xi_{k+1}$$

The observation error covariance matrix is equal to

$$R = \mathrm{diag}(R_0, \ldots, R_K)$$

Dual Algorithm for Variational Assimilation (continuation 4)

Maximization of dual objective function

$$\mu \rightarrow \mathcal{K}(\mu) = -(1/2)\ \mu^{\mathrm{T}} \Lambda\, \mu + d^{\mathrm{T}}\mu$$

requires explicit repeated computations of its gradient

$$\nabla_{\mu}\mathcal{K} = -\,\Lambda\mu + d = -\,(HP^{b}H^{\mathrm{T}} + R)\mu + d$$

Starting from $\mu = (\mu_0^{\mathrm{T}}, \ldots, \mu_K^{\mathrm{T}})^{\mathrm{T}}$ belonging to (dual) of observation space, this requires 5 successive steps

- Step 1. Multiplication by H^{T}. This is done by applying the transpose of the process defined above, *viz.*,

Set $\chi_K = 0$

Then, for $k = K\text{-}1, \ldots, 0$

$$\nu_k = \chi_{k+1} + H_{k+1}^{\mathrm{T}}\, \mu_{k+1}$$
$$\chi_k = M_k^{\mathrm{T}}\, \nu_k$$

Finally $\lambda_0 = \chi_0 + H_0^{\mathrm{T}}\, \mu_0$

The output of this step, which includes a backward integration of the adjoint model, is the vector $(\lambda_0^{\mathrm{T}}, \nu_0^{\mathrm{T}}, \ldots, \nu_{K\text{-}1}^{\mathrm{T}})^{\mathrm{T}}$

Dual Algorithm for Variational Assimilation (continuation 5)

- Step 2. Multiplication by P^b. This reduces to

$$\xi_0 = P_0^{\,b}\,\lambda_0$$
$$\upsilon_k = Q_k \nu_k \ , \quad k = 0, \ldots, K\text{-}1$$

- Step 3. Multiplication by H. Apply process defined above to vector $(\xi_0^{\mathrm{T}}, \upsilon_0^{\mathrm{T}}, \ldots, \upsilon_{K-1}^{\mathrm{T}})^{\mathrm{T}}$, thereby producing vector $(u_0^{\mathrm{T}}, \ldots, u_K^{\mathrm{T}})^{\mathrm{T}}$.

- Step 4. Add vector $R\mu$, *i. e.* compute

$$\varphi_0 = \xi_0 + R_0\,\mu_0$$
$$\varphi_k = \upsilon_{k\text{-}1} + R_k\,\mu_k \qquad , \quad k = 1, \ldots, K$$

- Step 5. Change sign of vector $\varphi = (\varphi_0^{\mathrm{T}}, \ldots, \varphi_K^{\mathrm{T}})^{\mathrm{T}}$, and add vector $d = y - Hx^b$,

Dual Algorithm for Variational Assimilation (continuation 6)

The model error covariance matrix Q_k is present in the algorithm only in its direct (not inverse form). Dual algorithm remains regular in the limit of vanishing model error. Can be used for both strong- and weak-constraint assimilation.

No significant increase of computing cost in comparison with standard strong-constraint variational assimilation (Louvel)

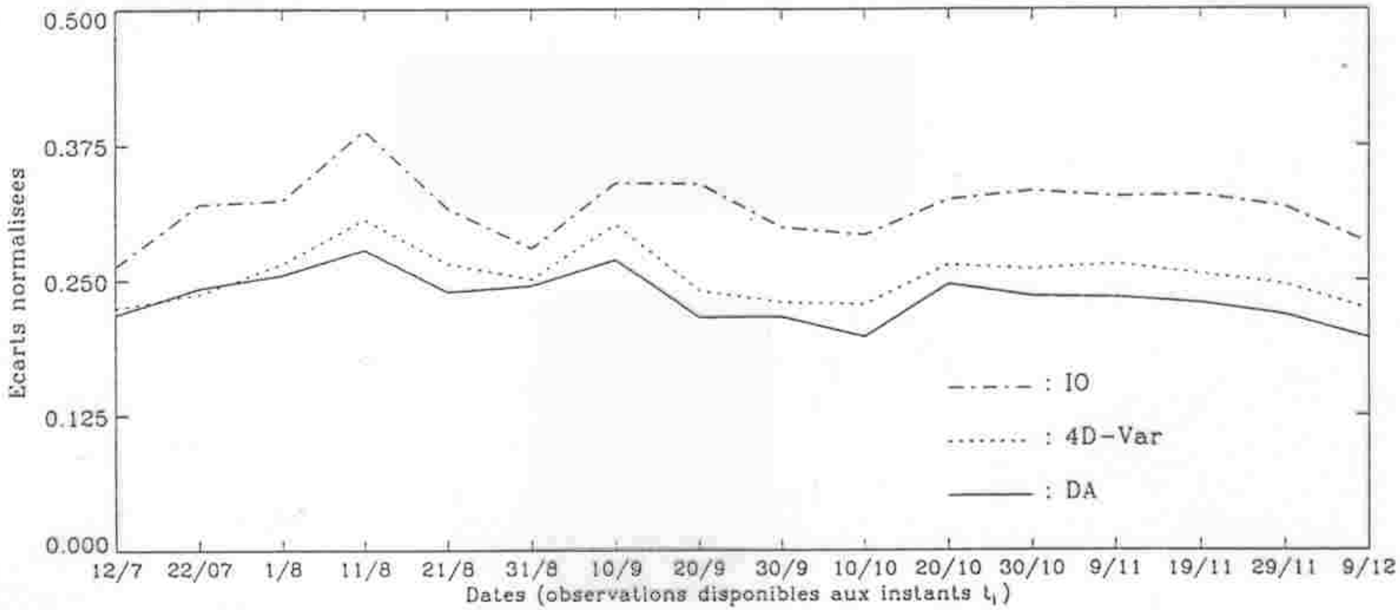


FIG. 9.11 – *Ecarts normalisés prévision/observations sur l'ensemble de la période étudiée*

Louvel, Doctoral Dissertation, Université Paul-Sabatier, Toulouse, 1999 nous permet d'observer les performances des différentes techniques d'assimilation.

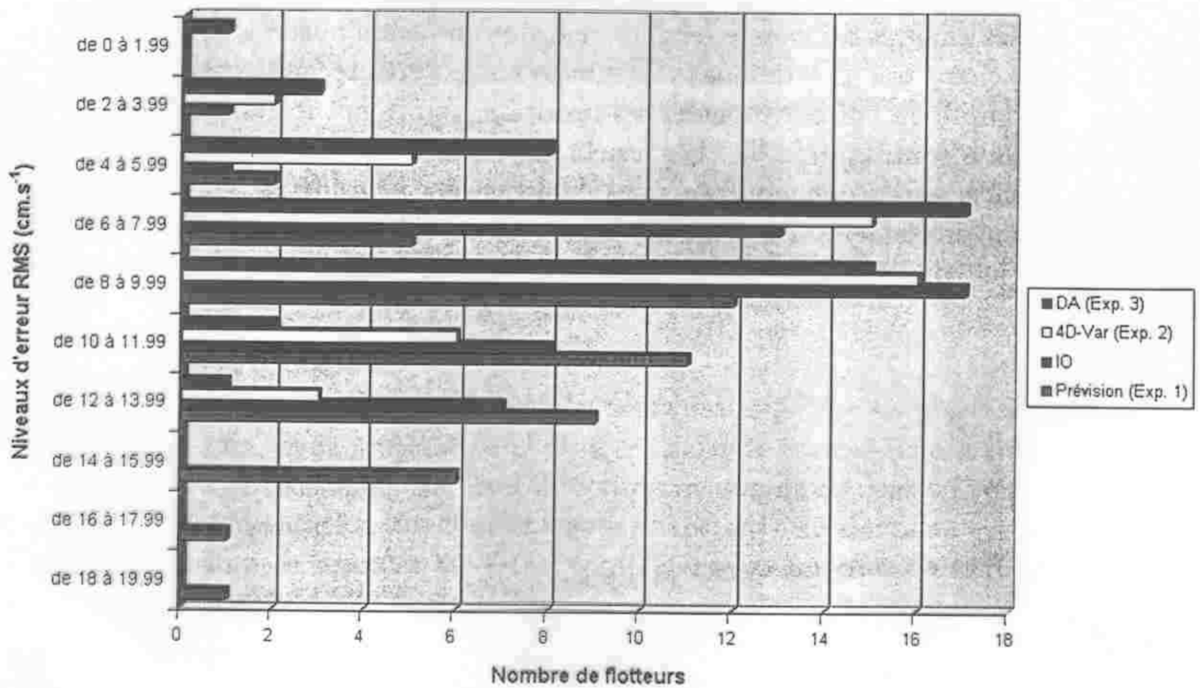


FIG. 9.15 – *Description des écarts flotteurs/modèle en terme de vitesse (à 150 m de profondeur) pour les différents algorithmes d'assimilation*

Louvel, Doctoral Dissertation, Université Paul-Sabatier, Toulouse, 1999

Dual Algorithm for Variational Assimilation (continuation)

Requires

- Explicit background (not much of a problem)
- Exact linearity (much more of a problem). Definition of iterative nonlinear procedures is being studied (Auroux, …)

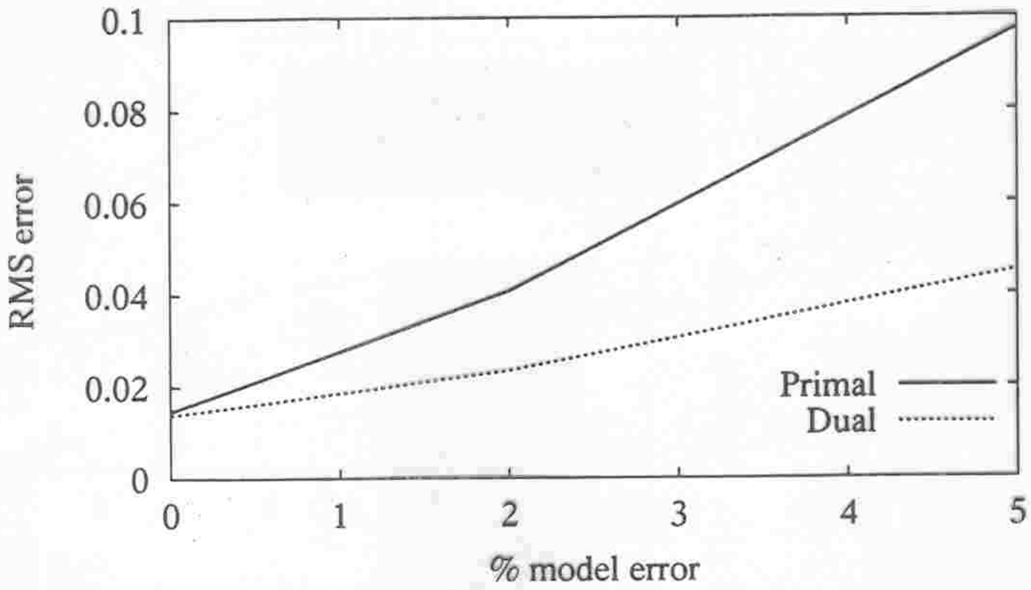


FIG. 6.13 – Normes RMS des erreurs d'assimilation obtenues pour les deux méthodes en fonction de l'erreur introduite dans le modèle au cours de la période d'assimilation.

Auroux, Doctoral Dissertation, Université de Nice-Sophia Antipolis, Nice, 2003

Sequential Assimilation

Pros

'Natural', and well adapted to many practical situations

Provides explicit estimate of estimation error

Cons

Carries information only forward in time (of no importance if one is interested only in doing forecast; and smoother exists in principle)

Optimality is possible only if errors are independent in time

Cost of computation of temporal evolution of estimation error very high, and often prohibitive

$$P^b_{k+1} = M_k P^a_k M_k^{\mathrm{T}} + Q_k$$

Variational Assimilation

Pros

Carries information both forward and backward in time (important for reassimilation of past data).

Can take into account temporal statistical dependence (Järvinen *et al.*)

Does not require explicit computation of temporal evolution of estimation error

Very well adapted to some specific problems (*e. g.*, identification of tracer sources)

Cons

Does not readily provide estimate of estimation error

Requires development and maintenance of adjoint codes.

- Dual approach seems most promising. But still needs further development for application in non exactly linear caes.

- Is ensemble variational assimilation possible ? Probably yes. But also needs development.

EU-funded project *Assimilation of Envisat data* (ASSET)

2003-2006. Coordinator W. Lahoz (University of Reading)

~ 10 groups from different European countries, using different models of atmospheric chemistry and different assimilation algorithms, have assimilated observations from instruments on board ENVISAT (mostly observations of O_3 and NO_x atmospheric contents).

Several articles are in the publication process.

Assimilation, which originated from the need of defining initial conditions for numerical weather forecasts, has progressively extended to many diverse applications

- Oceanography
- Atmospheric chemistry (both troposphere and stratosphere)
- Oceanic biogeochemistry
- Ground hydrology
- Terrestrial biosphere and vegetation cover
- Glaciolology
- Planetary atmospheres (Mars, …)
- Reassimilation of past observations (mostly for climatological purposes, ECMWF, NCEP/NCAR)
- Identification of source of tracers
- Parameter identification
- *A priori* evaluation of anticipated new instruments
- Definition of observing systems *(Observing Systems Simulation Experiments)*
- Validation of models
- Sensitivity studies (adjoints)
- …

Assimilation is related to

- Estimation theory
- Probability theory
- Atmospheric and oceanic dynamics
- Atmospheric and oceanic predictability
- Instrumental physics
- Optimisation theory
- Control theory
- Algorithmics and computer science
- …

References on assimilation of observations

Books

- *Inverse Problems in general*

- Tarantola, A., 2005, *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia, USA, 342 pp., ISBN 0-89871-572-5 (can be downloaded freely from address http://www.ipgp.jussieu.fr/~tarantola/).

- *Filtering Theory*

- Jazwinski, A. H., 1970, *Stochastic Processes and Filtering Theory*, Academic Press, New York, USA, 376 pp..

- Gelb, A., 1974, *Applied optimal estimation*, Cambridge University Press, Cambridge, UK, 374 pp..

- Anderson, B. D. O., and J. B. Moore, 1979, *Optimal Filtering*, Prentice Hall, Englewood Cliffs, New Jersey, USA, 357 pp..

- *Meteorological and Oceanographical Applications*

- Daley, R., 1991, *Atmospheric Data Analysis*, Cambridge University Press, Cambridge, UK, 457 pp..

- Bennett, A. F., 1992, *Inverse Methods in Physical Oceanography*, Cambridge University Press, Cambridge, UK, 346 pp..

- Wunsch, C., 1996, *The Ocean Circulation as an Inverse Problem*, Cambridge University Press, Cambridge, UK, 442 pp..

- Rodgers, C. D., 2000, *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific Publishing Co. Ltd, London, UK, 238 pp..

- Kalnay, E., 2002, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge, UK, 341 pp..

- (Bennett, A. F., 2002, *Inverse Modeling of the Ocean and Atmosphere*, Cambridge University Press, Cambridge, UK, 265 pp.).

- **Papers in Journals**

- Ghil, M., and P. Malanotte-Rizzoli, 1991, Data assimilation in meteorology and oceanography, *Adv. in Geophys.*, **33**, 141-266.

- Talagrand, O., 1997, Assimilation of Observations, an Introduction, *J. Meteor. Soc. Japan*, **75** (1B, Special Issue *Data Assimilation in Meteorology and Oceanography: Theory and Practice),* 191-209 (and other papers in the same issue).