O. Talagrand (LMD)

Objective validation and evaluation of data assimilation

Purpose of assimilation

Use all available information in order to determine as accurately as possible the state of the atmospheric or oceanic flow.

Available information consists of

- Observations proper, highly variable in nature, spatial and temporal distribution and accuracy.

- Physical laws governing the flow, available in practice in ~~under~~ the form of a numerical dynamical model.

- statistical or asymptotic properties of the flow (ex: existence of approximat geostrophic balance)

$$\begin{cases} Z_1 = x + \xi_1 \\ Z_2 = x + \xi_2 \end{cases}$$

$$\xi_1 = \mathcal{N}(0, \sigma_1^2) \qquad \sim \exp\left(-\frac{1}{2}\frac{\xi_1^2}{\sigma_1^2}\right)$$

$$\xi_2 = \mathcal{N}(0, \sigma_2^2) \qquad \sim \exp\left(-\frac{1}{2}\frac{\xi_2^2}{\sigma_2^2}\right)$$

Mutually independent!

Probability that $x = \xi$?

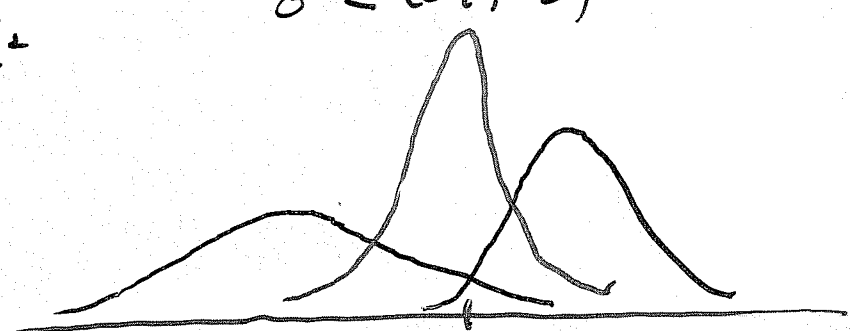
$$x = \xi \iff \xi_1 = z_1 - \xi \text{ and } \xi_2 = z_2 - \xi$$

$$P(x = \xi | z_1, z_2) \sim \exp\left[-\frac{1}{2}\frac{(z_1 - \xi)^2}{\sigma_1^2}\right] \exp\left[-\frac{1}{2}\frac{(z_2 - \xi)^2}{\sigma_2^2}\right]$$

$$\sim \exp\left[-\frac{1}{2}\frac{(\xi - x^a)^2}{\sigma^2}\right]$$

$$P(x = \xi | z_1, z_2) = \mathcal{N}(x^a, \sigma^2)$$

$$\begin{cases} x^a = \sigma^2\left[\frac{z_1}{\sigma_1^2} + \frac{z_2}{\sigma_2^2}\right] \\ \frac{1}{\sigma^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \end{cases} \qquad \sigma^2 < (\sigma_1^2, \sigma_2^2)$$



State vector x, belonging to *state space* $\mathcal{S}$ ($\dim\mathcal{S} = n$), to be estimated.

Data vector z, belonging to *data space* $\mathcal{D}$ ($\dim\mathcal{D} = m$), available.

$$F(z, x, \zeta) = 0 \qquad (1)$$

where ζ is a random element representing the uncertainty on the data (or, more precisely, on the link between the data and the unknown state vector).

For example

$$z = \Gamma x + \zeta$$

Bayesian estimation

Probability that $x = \xi$ for given ξ ?

$$x = \xi \quad \Leftrightarrow \quad F(z, \xi, \zeta) = 0$$

$$P(x = \xi \mid z) = P[F(z, \xi, \zeta) = 0] \;/\; \int_{\xi'} P[F(z, \xi', \zeta) = 0]$$

Unambiguously defined iff, for any ζ, there is at most one x such that (1) is verified.

$\Leftrightarrow$ data contain information, either directly or indirectly, on any component of x. *Determinacy* condition.

More generally

$$z = f(x, \zeta)$$

where ζ is 'error' with known pdf. Then

$$P(\mathbf{x}=\xi | \mathbf{z}) = \int P[\zeta \text{ such that } f(\xi, \zeta) = z]$$

Makes sense only if, for any ζ, there is at most one ξ such that $f(\xi, \zeta) = z$ (*determinacy* condition)

! Bayesian estimation impossible in practice because

- large numerical dimensions (10^7)
- data error statistics are poorly known (if at all ; model errors).

Need for drastically reduced order description of conditional pdf. Two possibilities

- some estimate of $E(x | z)$ and reduced order estimate of $E(xx^T | z)$
- conditional pdf described by finite ensemble (size up to a few 10^2) ;

$$z = \Gamma x + \zeta$$

Γ known (mxn)-matrix, ζ unknown 'error'

Look for estimated state vector x^a of the form

$$x^a = \alpha + Az$$

subject to

1 invariance in change of origin in state space $\Rightarrow A\Gamma = I_m$

2 $E[(x^a_i - x_i)^2]$ minimum for any component x_i.

$$x^a = (\Gamma^T S^{-1} \Gamma)^{-1} \Gamma^T S^{-1} [z - \mu]$$

$$P^a \equiv E[(x^a - x)(x^a - x)^T] = (\Gamma^T S^{-1} \Gamma)^{-1}$$

where $\mu = E(\zeta)$ (expectation) and $S = E\{[\zeta - \mu][\zeta - \mu]^T\}$ (covariance matrix)

Best Linear Unbiased Estimator (*BLUE*) of x from z.

Requires (at least apparently) *a priori* explicit knowledge of first- and second-order statistical moments of error ζ.

Determinacy condition $\Leftrightarrow$ rank$\Gamma = n$. $\Longrightarrow m \geq n$

In case ζ is gaussian, $\zeta = \mathcal{N}[\mu, S]$, BLUE achieves bayesian estimation in the sense that

$$P(x \mid z) = \mathcal{N}[x^a, P^a]$$

Variational form.

BLUE x^a minimizes following scalar *objective function*, defined on state space $\mathcal{S}$

$$\mathcal{J}(\xi) \equiv (1/2)\,[\Gamma\xi - (z-\mu)]^T S^{-1}\,[\Gamma\xi - (z-\mu)]$$

BLUE is invariant in any invertible linear change of coordinates, either in state or data space.

From now on, unless specified otherwise, data assumed to be unbiased ($\mu = 0$).

If determinacy condition is verified, it is always possible to decompose data into

A '*background*' estimate (*e. g.* forecast from the past), belonging to *state space*, with dimension n

$$x^b = x + \zeta^b$$

An additional set of data (*e. g.* observations), belonging to *observation space*, with dimension $m - n = p$

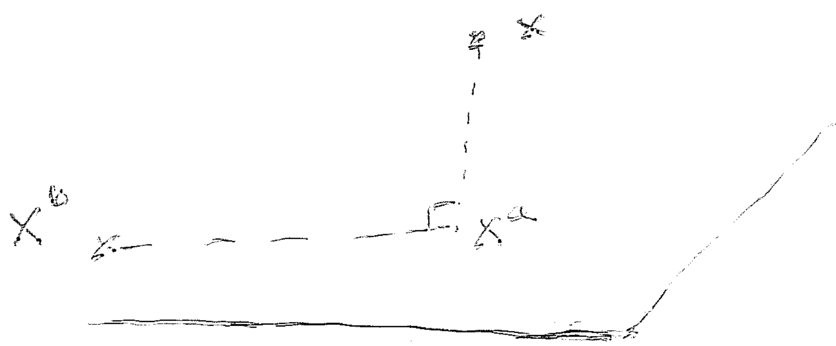
$$y = Hx + \varepsilon$$

Least error variance estimate – BLUE

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

with $P^b \equiv E(\zeta^b \zeta^{bT})$ (also often denoted B), $R \equiv E(\varepsilon\varepsilon^T)$
$E(\varepsilon\zeta^{bT}) = 0$ (not restrictive)

$$P^a \equiv E[(x - x^a)(x - x^a)^T] = P^b - P^b H^T (HP^b H^T + R)^{-1} HP^b$$



$d \equiv y - Hx^b$ is the *innovation vector*

(1) means that the increment $x^a - x^b$ is the orthogonal projection, in the sense of statistical covariance, of the background error $x - x^b$ onto the space spanned by the innovation vector. As a consequence, the estimation error $x^a - x$ is uncorrelated with the innovation vector, *i. e.*

$$E[(x^a - x)\, d^T)] = 0$$

Variational formulation

Analysis x^a minimizes *objective function* defined on state space

$$\xi \rightarrow$$

$$\mathcal{J}(\xi) \equiv (1/2)\,(\xi - x^b)^T [P^b]^{-1} (\xi - x^b) + (1/2)\,(H\xi - y)^T R^{-1} (H\xi - y)$$

Determination of the BLUE requires (at least apparently) the a priori specification of the first- and second-order statistical moments of the errors affecting the data

(general bayesian estimation requires the specification of the entire probability distribution of the errors)

Questions

- Is it possible to objectively evaluate the quality of an assimilation system ?

- Is it possible to objectively evaluate the first- and second-order statistical moments of the data errors, whose specification is (at least apparently) required for determining the *BLUE* ?

- Is it possible to objectively determine if an assimilation system is optimal ?

- More generally, how to make the best of an assimilation system ?

and

- Is assimilation worth all the concern we give to it ?

$$\begin{cases} x^b = x + \zeta^b \\ d \equiv y - Hx^b = \varepsilon - H\zeta^b \end{cases}$$

Innovation vector is the only objective source of information on data error. Implementing assimilation requires knowing, at least to some extent, how the background on the one hand, the observations on the other, contribute to the innovation. That cannot be obtained from the innovation alone.

We will consider assimilation schemes of the form

$$x^a = x^b + K\,d \qquad (2)$$

where K is the *gain matrix* (not necessarily optimal).

(2) $\Leftrightarrow$ if data are exact, then analysis is exact too ($x^a = x$).

After A. Lorenc

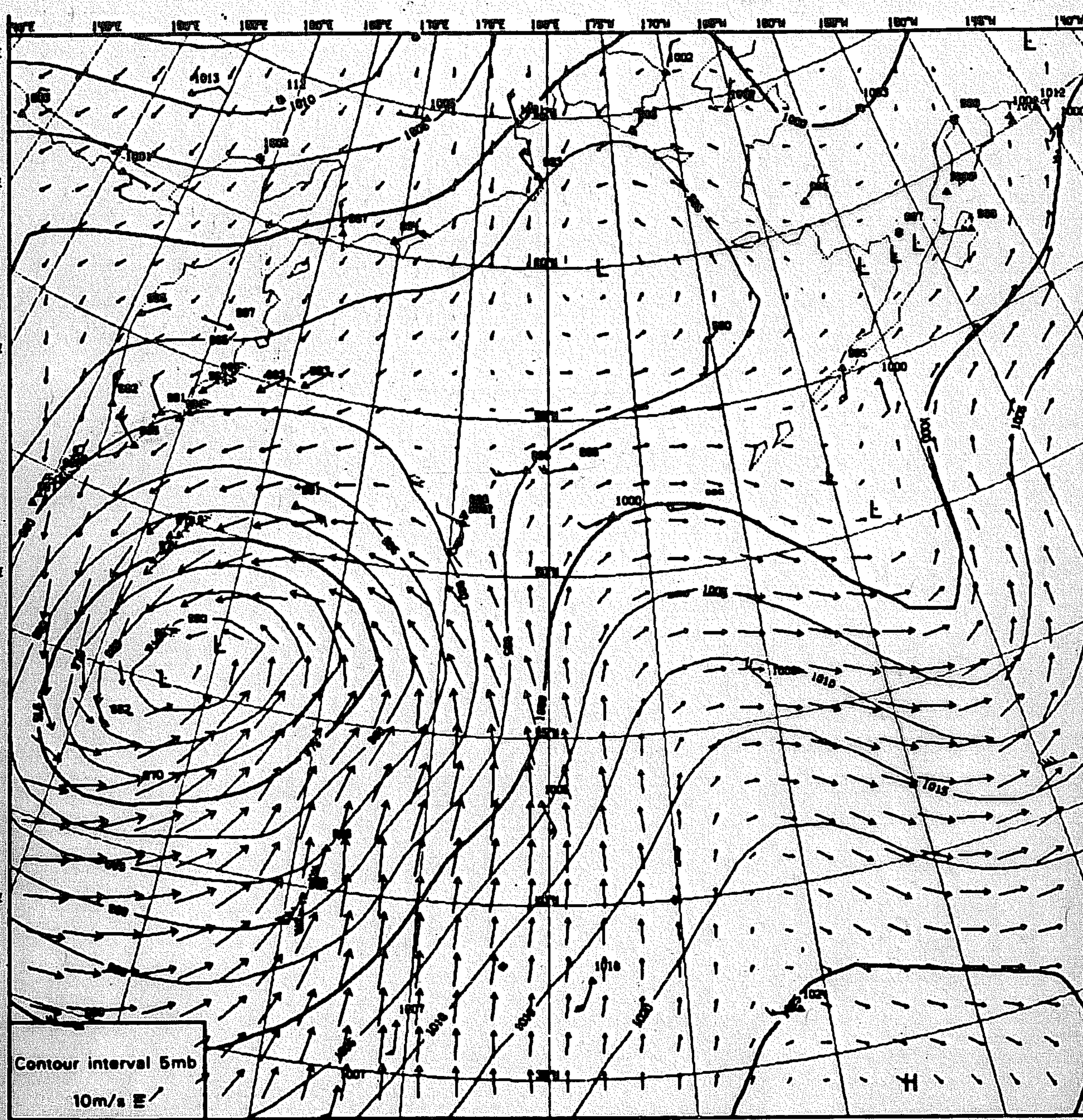


Fig. 4 Sea level pressure and wind forecast corresponding to the central area of Fig. 1, with plotted surface observations of pressure and wind (each fleche = 5 m/s).

After R. Lorenc 15

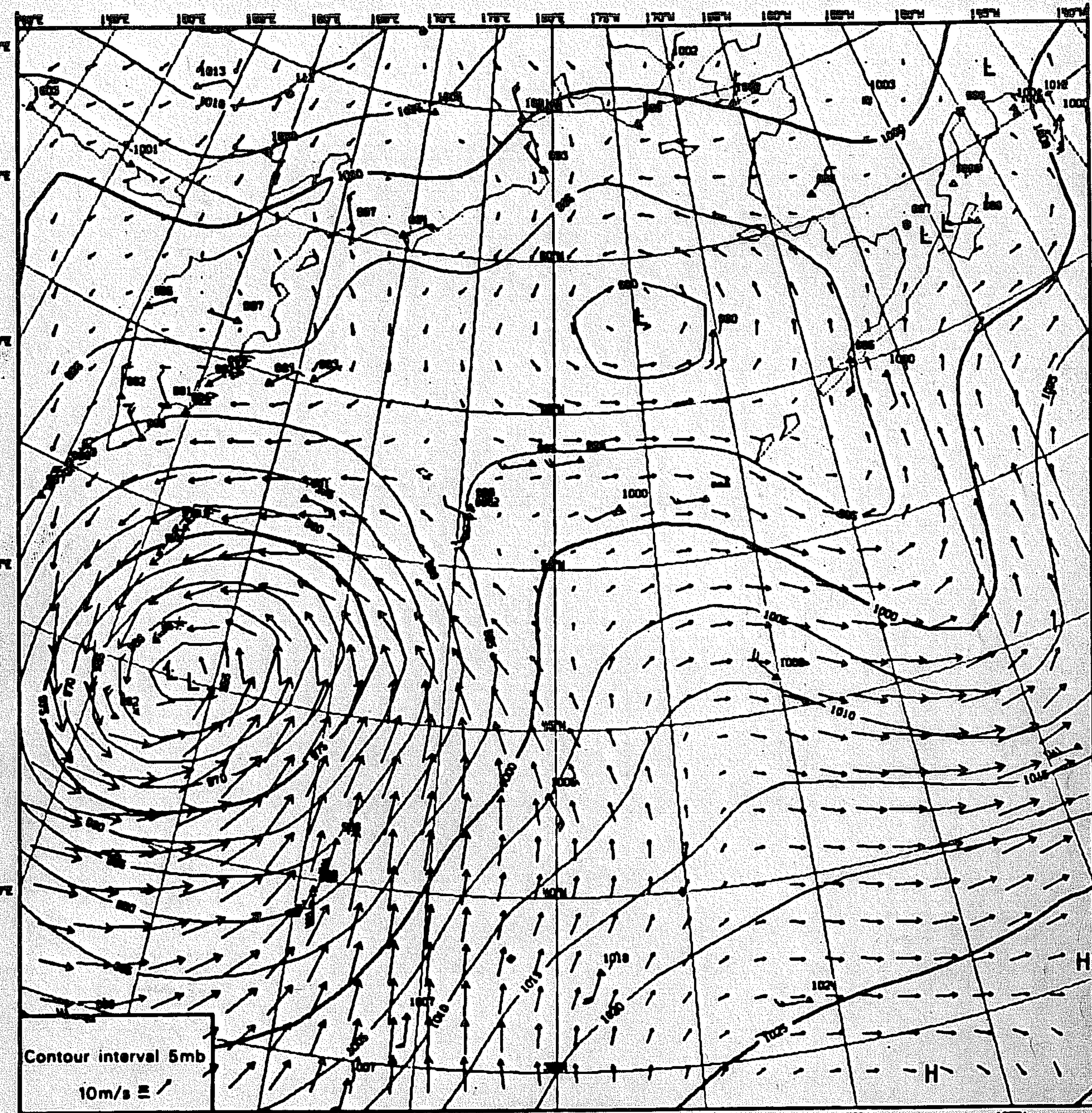


Fig. 5 As Fig. 4 for the analysis in the data assimilation cycle

Difference between data and assimilated fields

$$\delta \equiv \begin{pmatrix} x^b - x^a \\ y - \boldsymbol{H}x^a \end{pmatrix} = z - \Gamma x^a$$

$$\delta = \begin{pmatrix} -\boldsymbol{K}d \\ (I_p - \boldsymbol{HK})d \end{pmatrix}$$

For given gain matrix K, one-to-one transformation between d and δ. Exactly equivalent to perform diagnostics on either innovation or *data-minus-analysis* (*DmA*) difference.

Objective validation of quality of an assimilation system can be made only by comparison with unbiased independent data, i.e. data affected by errors which are statistically independent of errors affecting the data used in the assimilation.

Innovation vector $d = y - Hx^b$ is only source of objective information.

Magnitude: measure of quality of assimilation

~~Info~~ Other diagnostic. Check if statistics of innovation vector are consistent with a priori hypotheses

$E(d) = 0$

$E(dd^T) = HP^bH^T + R$

[Check if d is Gaussian]

If consistency

$$E(\delta) = 0$$

Any systematic bias in δ is the signature of an improperly taken into account bias in the data

$$E(\delta\delta^T) = S - \Gamma(\Gamma^T S^{-1} \Gamma)^{-1} \Gamma^T = S - \underbrace{\Gamma P^a \Gamma^T}_{>0}$$

Assimilated fields must fit data to within assumed accuracy of the latter. System is 'efficient', as defined by Hollingsworth and Lönnberg, 1989)

If they do not, inconsistency between a priori assumed statistics on δ and real statistics.

NCMRWF (K. Bhattacharya)

NORTHERN HEMISPHERE (O-A)RMS/OBS.ERR SSI - JAN 2002

LEVEL(hPa)	RADIOSONDE TD	RAWINSONDE UW	RAWINSONDE VW
50.	1.3	1.4	1.0
70.	0.9	1.1	0.9
100.	0.8	1.0	1.0
150.	0.8	0.9	0.9
200.	0.8	0.8	0.8
250.	0.8	0.8	0.8
300.	0.8	0.8	0.8
400.	0.8	0.8	0.9
500.	0.7	0.9	0.9
700.	0.7	0.9	0.8
850.	0.9	0.9	0.9
1000.	1.3	0.9	0.9

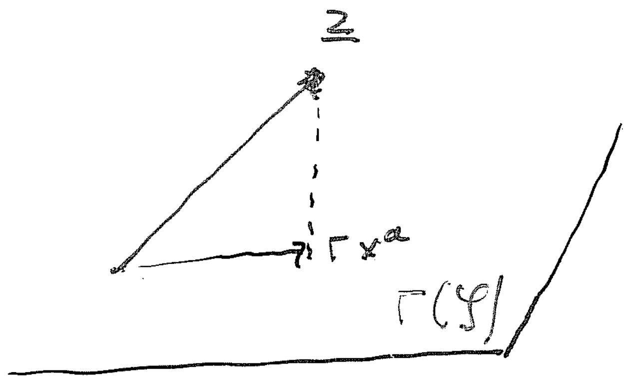
NCMRWF (K. Bhattacharya)

NORTHERN HEMISPHERE (O-A)RMS/OBS.ERR SSI - JUL 2002

LEVEL(hPa)	RADIOSONDE TD	RAWINSONDE UW	RAWINSONDE VW
50.	1.3	1.1	0.8
70.	0.9	0.8	0.8
100.	0.7	0.8	0.8
150.	0.7	0.7	0.7
200.	0.7	0.7	0.7
250.	0.6	0.7	0.7
300.	0.6	0.6	0.7
400.	0.6	0.6	0.6
500.	0.7	0.7	0.7
700.	0.8	0.8	0.8
850.	1.1	0.8	0.8
1000.	1.4	0.8	0.8

$$J(\underline{\xi}) = \frac{1}{2}(\Gamma\underline{\xi} - \underline{z})^T S^{-1}(\Gamma\underline{\xi} - \underline{z})$$

$<\underline{\xi}, \underline{\xi}'> = \underline{\xi}^T S^{-1} \underline{\xi}'$ is a (proper) scalar product in data space $\mathcal{D}$ (Mahalanob

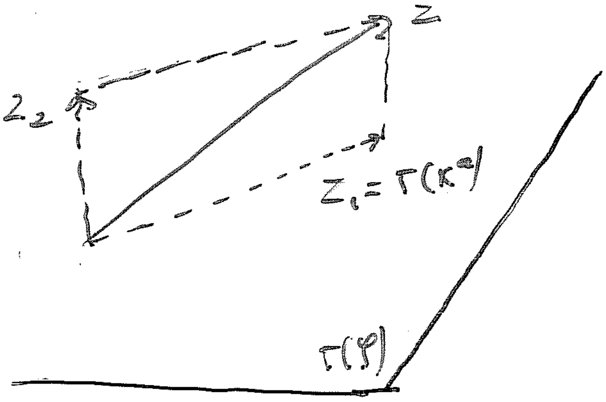


Computing the BLUE amounts to

- Projecting data vector $\underline{z}$ onto image space $\Gamma(\mathcal{S})$ according to the Mahalanobis scalar product
- Taking the inverse of the projection through Γ (unambiguously defined if null space of Γ is void)

The DmA difference $z - \Gamma x^a$ is the component of the data vector that is discarded in analysis.

Decompose data space $\mathcal{D}$ into image space $\Gamma(\mathcal{S})$ and space orthogonal to $\Gamma(\mathcal{S})$ wrt to S-Mahalanobis scalar product.



$$\Gamma = \begin{pmatrix} \Gamma_1 \\ 0 \end{pmatrix}$$

$$z = \begin{pmatrix} z_1 = \Gamma_1 x + \xi_1 \\ z_2 = \xi_2 \end{pmatrix}$$

$$S = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}$$

$$\begin{cases} x^a = (\Gamma_1^T S_1^{-1} \Gamma_1)^{-1} \Gamma_1^T S_1^{-1} z_1 = \Gamma_1^{-1} z_1 \\ P^a = (\Gamma_1^T S_1^{-1} \Gamma_1)^{-1} \end{cases}$$

D m A-difference

~~$z - \Gamma x^a$~~

$$\delta = z - \Gamma x^a = \begin{pmatrix} z_1 - \Gamma_1 x^a = 0 \\ z_2 = \xi_2 \end{pmatrix}$$

$$E(\delta) = E(\xi_2)$$

$$E(\delta' \delta'^T) = E(\xi_2' \xi_2'^T)$$

Inconsistency iff $E(\xi_2) \neq 0$ and/or $E(\xi_2' \xi_2'^T) \neq S$

So what?

$$\begin{cases} z_1 = x + \xi_1 \\ z_2 = x + \xi_2 \end{cases} \qquad n = 1 \quad m = 2$$

We do the analysis under the assumption that

$E(\xi_1) = E(\xi_2) = 0$

$E(\xi_1^2) = E(\xi_2^2) = \sigma^2 \qquad E(\xi_1 \xi_2) = 0$

$\Longrightarrow \quad x^a = \frac{1}{2}(z_1 + z_2) \qquad E[(x^a - x)^2] = \frac{\sigma^2}{2}$

Innovation $\quad d = z_1 - z_2$

We expect $\quad E(d) = 0 \qquad E(dd^T) = 2\sigma^2$

We observe $\quad E(d) = 2b \qquad E(dd^T) = 4b^2 + 2\Delta^2$

Inconsistency iff $\quad b \neq 0$ and/or $\Delta^2 \neq \sigma^2$

$E(d) = 2e$

(A)

$E(d^2) = 4e^2 + 2\Delta^2$

If $E(\xi_1) = -E(\xi_2) = e$

$$E[(\xi_1 - e)^2] = E[(\xi_2 + e)^2] = \sigma'^2$$

$$E[(\xi_1 - e)(\xi_2 + e)] = c\sigma'^2$$

with $\sigma'^2 = \dfrac{\sigma^2 + \Delta^2}{2}$

$$c = \frac{\sigma^2 - \Delta^2}{\sigma^2 + \Delta^2}$$

The quantities x^a and $E[(x^a - x)^2]$ are unchanged, while consistency with (A) is achieved.

In the above example, if it is known that $E(\xi_1 \xi_2) = 0$, then necessarily the weights given to z_1 and z_2 are > 0, and

$$E[(x^a - x)^2] < dE(d^2)$$

$$E[(x^a - x)^2] < E(d^2)$$

~~2 parameters ($E(\xi_1)$ and $E(\xi_2)$) for are available for~~

If either $E(\xi_1)$ or $E(\xi_2)$ is known, then the knowledge of $E(d)$ determines the unknown bias.

Similarly, the knowledge of one or two of the three parameters $E(\xi_1^2)$, $E(\xi_2^2)$ and c determines, together with the knowledge of $E(d^2)$, one or two conditions on the unknown parameters.

In the above example, if it is known, e.g., that $c=0$, one condition is then imposed on the weights given to z_1 and z_2, and on the analysis error. If we want for instance to keep the weights equal, then necessarily $E[(x^a - x)^2] =$ [illegible]

A possible inconsistency between expected or observed statistics of δ (or, equivalently, of innovation d) can always be resolved without changing x^a nor P^a.

Consistency between assumed and observed statistics of innovation is not sufficient for ensuring optimality of assimilation. It is not even necessary.

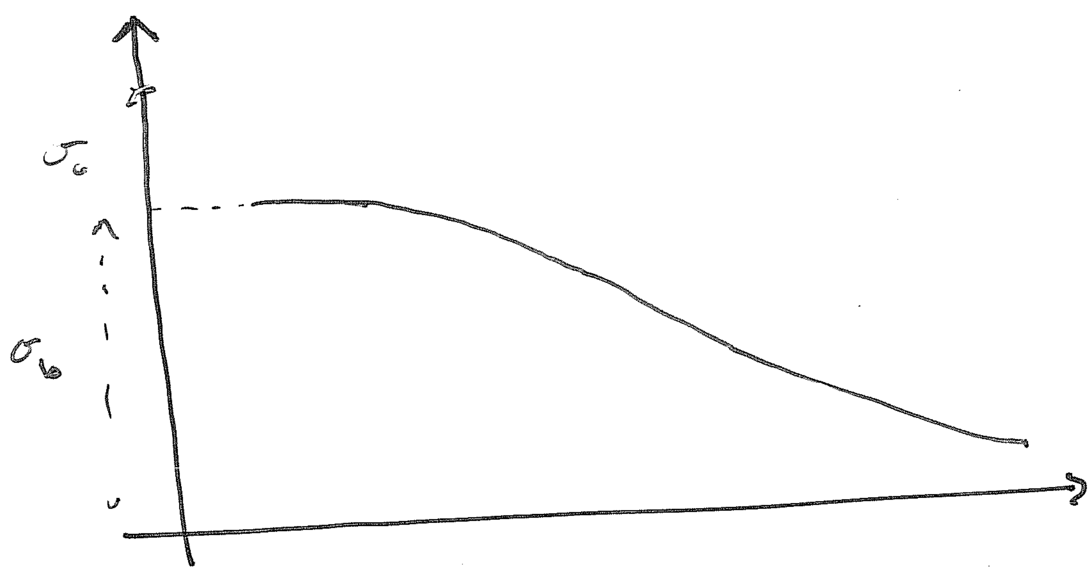
Any inconsistency between the expected and actually observed statistics of the data-minus-analysis difference can always be resolved without modification to the analysis, nor to the associated estimated error.

Independent hypotheses, which cannot be objectively validated (at least on the basis of the data-minus-analysis difference) are necessary.

Problem. Find minimum set of hypotheses, leaving only parameters that can be determined from statistics of the innovation vector.

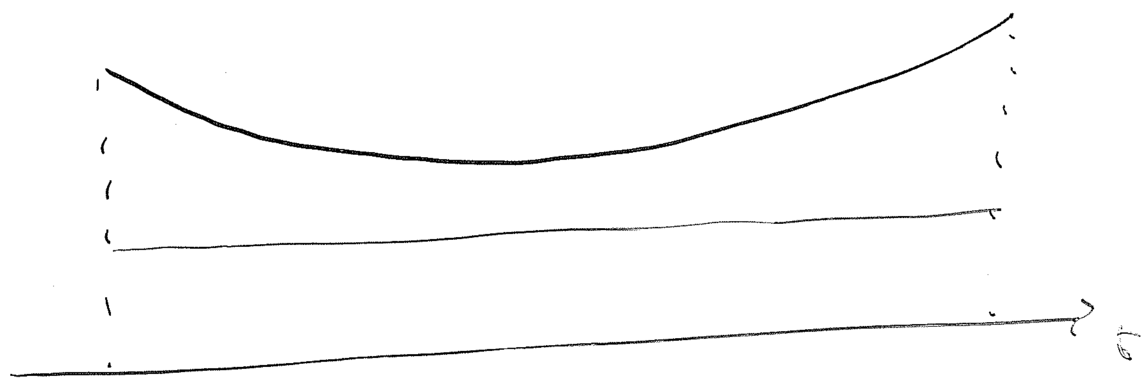
E a priori hypotheses are made, then statistics of the innovation vector (or, equivalently, of the OmA-difference) can be used for determining biases and/or covariances

(R. Daley) Assume observation errors are mutually uncorrelated and uncorrelated with background errors.



Then spatial covariance of innovation is spatial correlation of ~~innovation~~ background error. Variance of background and observation errors can be evaluated by extrapolation to zero distance

Strong-constraint variational assimilation



If observational error known, and uncorrelated with model error, misfit between minimizing solution and observations provides integrated estimate of model error.

Kalman Filter. ~~Model~~ Errors assumed to be uncorrelated in time (necessary for having both sequentiality and optimality) ⇒ adaptive filtering (Hoang et al., ...)

Variational assimilation (either 3-D or 4-D)

$$\mathcal{J}(\xi) \equiv (1/2)\,(\xi - x^b)^T\,[P^b]^{-1}\,(\xi - x^b) \;+\; (1/2)\,(H\xi - y)^T\,R^{-1}\,(H\xi - y)$$

Minimum reached for $\xi = x^a$. For a perfectly consistent system

$$\mathcal{J}(x^a) = (1/2)\,d^T\,[E(dd^T)]^{-1}\,d$$

Minimum of objective function is norm of innovation with respect to its own Mahalanobis scalar product. On expectation

$$E[\mathcal{J}(x^a)] = p/2$$

Often called χ^2-criterion.

3.1.2 Numerical results

The main result is that the two algorithms converge to the true state with identical cost and accuracy. This first experiment show that the practical efficiency of DA in comparison with the usual 4D-Var.

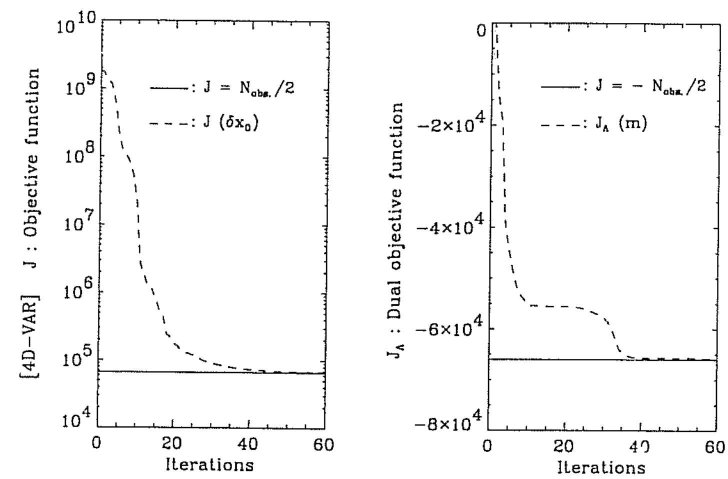


FIG. 2 – *Objective functions*

FIG. 2 gives proof of the property (19) and compare the minimal numerical values of J and J_Λ to the statistical expectation: $J_{\text{stat}} = \text{N}_{\text{obs}}/2$ with $\text{N}_{\text{obs}}=\text{N}_{\text{state}}$. Note that the null initial value of the dual cost function (12) is explained by the use of a null increment at the beginning of the minimization. The decrease of the two objective function is not identical but the primal and the dual algorithms need the same number of iterations. From this experiment, we can assess the equivalence between both methods in term of numerical results at the end of the minimizations.

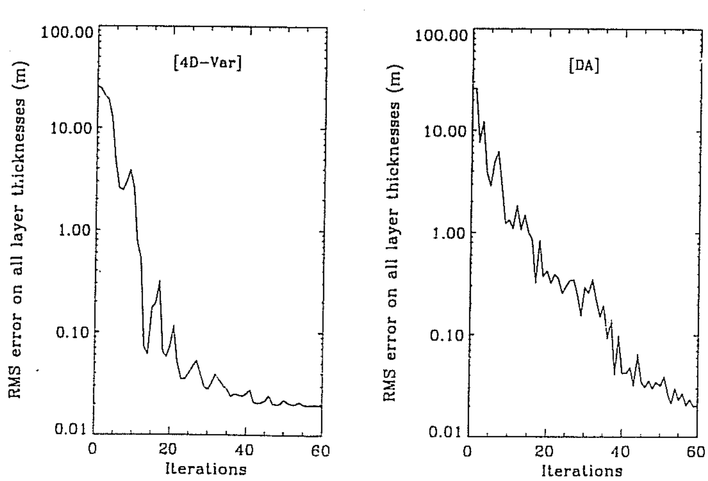


FIG. 3 – *RMS error at the time* t_n

Anomalies operationnelles

nombre d'iterations

```
/scratch/rd/daa/ebmg/logdir/20030101/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1441547          583351.3302514           0.40        0,45
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   583351.330251      60681.8576886        6714.112242
20
/scratch/rd/daa/ebmg/logdir/20030101/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1391118          512489.3745112           0.37        0,41
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   512489.374511      57412.3877855        5234.061281
56
/scratch/rd/daa/ebmg/logdir/20030102/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1460413          560044.4343723           0.38        0,43
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   560044.434372      56877.7970359        5475.952246
10
/scratch/rd/daa/ebmg/logdir/20030102/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1371637          502648.7127915           0.37        0,41
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   502648.712792      53200.0343999        4499.520897
98
/scratch/rd/daa/ebmg/logdir/20030103/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1486660          560261.8512532           0.38
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   560261.851253      56825.0592832        4510.250091
34
/scratch/rd/daa/ebmg/logdir/20030103/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1439923          527864.7124925           0.37
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   527864.712493      56371.5869005        4108.714252
33
/scratch/rd/daa/ebmg/logdir/20030104/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1462278          556644.6455113           0.38
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   556644.645511      56590.8002956        4547.851849
23
/scratch/rd/daa/ebmg/logdir/20030104/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1363128          495154.3086095           0.36
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   495154.308609      52766.7921730        4278.757745
74
/scratch/rd/daa/ebmg/logdir/20030105/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1477371          534654.6665056           0.36
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   534654.666506      54695.4780231        4317.270937
55
/scratch/rd/daa/ebmg/logdir/20030105/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1391729          483022.4228964           0.35
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   483022.422896      51947.4256765        3735.044590
87
/scratch/rd/daa/ebmg/logdir/20030106/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1425827          516035.2023126           0.36
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   516035.202313      53171.8232700        3867.648157
81
/scratch/rd/daa/ebmg/logdir/20030106/main/12/an/4dvar/uptraj_1/ifsmin.1 :    0,38?
            Jo Global :    1376639          466430.9015891           0.34
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   466430.901589      51846.8336970        3822.035276
65
/scratch/rd/daa/ebmg/logdir/20030107/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1476959          526187.0910188           0.36
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   526187.091019      53156.1037852        4074.136909
78
/scratch/rd/daa/ebmg/logdir/20030107/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1458699          524033.4879162           0.36
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   524033.487916      53889.7012286        3936.815589
21
/scratch/rd/daa/ebmg/logdir/20030108/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1500738          554292.5152698           0.37
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   554292.515270      56842.3382909        4313.191710
19
/scratch/rd/daa/ebmg/logdir/20030108/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1478979          519652.0339623           0.35
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   519652.033962      54815.3317792        4088.292663
85
/scratch/rd/daa/ebmg/logdir/20030109/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1501182          536600.4381115           0.36
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   536600.438111      55189.8771365        3954.585491
89
/scratch/rd/daa/ebmg/logdir/20030109/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1458405          521047.2001178           0.36
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   521047.200118      54325.4354694        3858.839886
32
/scratch/rd/daa/ebmg/logdir/20030110/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1479196          551524.9346666           0.37
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   551524.934667      54935.2610658        3966.946995
08
/scratch/rd/daa/ebmg/logdir/20030110/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :    1442781          519009.2804139           0.36        0,40
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   519009.280414      54424.8595290        3934.851931
09
/scratch/rd/daa/ebmg/logdir/20030111/main/00/an/4dvar/uptraj_1/ifsmin.1 :    0,40
            Jo Global :    1498818          545469.1467534           0.36
```

```
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   545469.146753        54571.9950191        3864.932864
00
/scratch/rd/daa/ebmg/logdir/20030111/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1404038              499560.3144969            0.36     0,40
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   499560.314497        52394.3272952        3765.473571
79
/scratch/rd/daa/ebmg/logdir/20030112/main/00/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1452999              531075.9594903            0.37     0,41
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   531075.959490        55941.5710915        4177.026832
15
/scratch/rd/daa/ebmg/logdir/20030112/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1357049              484944.1940747            0.36     0,40
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   484944.194075        52517.7255095        3736.706266
78
/scratch/rd/daa/ebmg/logdir/20030113/main/00/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1337228              536122.2229325            0.40
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   536122.222933        53429.8153097        4053.930255
99
/scratch/rd/daa/ebmg/logdir/20030113/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1311265              478570.4071103            0.36
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   478570.407110        52561.1547953        3904.290460
13
/scratch/rd/daa/ebmg/logdir/20030114/main/00/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1453254              567289.3811072            0.39
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   567289.381107        57475.6966038        4801.895527
70
/scratch/rd/daa/ebmg/logdir/20030114/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1467896              548622.2342518            0.37
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   548622.234252        56298.4672026        4302.401824
15
/scratch/rd/daa/ebmg/logdir/20030115/main/00/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1494508              555721.3641453            0.37
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   555721.364145        57345.3948895        4285.092207
98
/scratch/rd/daa/ebmg/logdir/20030115/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1318768              492015.2817690            0.37
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   492015.281769        53657.0226403        3725.264802
29
/scratch/rd/daa/ebmg/logdir/20030116/main/00/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1202340              502028.6625365            0.42     0,46
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   502028.662537        51815.0327751        4201.291031
35
/scratch/rd/daa/ebmg/logdir/20030116/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1489159              556783.2050098            0.37
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   556783.205010        58472.8524180        4362.181995
21
/scratch/rd/daa/ebmg/logdir/20030117/main/00/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1495427              568478.7591876            0.38
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   568478.759188        57318.6283788        4513.011533
37
/scratch/rd/daa/ebmg/logdir/20030117/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1494243              544629.9026914            0.36
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   544629.902691        56138.0457273        4327.980181
86
/scratch/rd/daa/ebmg/logdir/20030118/main/00/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1430028              540130.5317339            0.38
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   540130.531734        54712.0229546        4332.500891
95
/scratch/rd/daa/ebmg/logdir/20030118/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1437608              508122.5669298            0.35
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   508122.566930        53846.3157778        4024.626659
81
/scratch/rd/daa/ebmg/logdir/20030119/main/00/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1507042              543897.2238212            0.36     0,40
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   543897.223821        54585.7094448        4323.183151
20
/scratch/rd/daa/ebmg/logdir/20030119/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1352213              485534.4666308            0.36     0,40
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   485534.466631        52123.2692774        3939.344145
85
/scratch/rd/daa/ebmg/logdir/20030120/main/00/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1484311              546480.3748742            0.37     0,41
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   546480.374874        54694.5721515        4597.361035
02
/scratch/rd/daa/ebmg/logdir/20030120/main/12/an/4dvar/uptraj_1/ifsmin.1 :
               Jo Global :    1413737              502382.8863849            0.36     0,40
 GREPCOST - ITER,SIM,JO,JB,JC  999 999   502382.886385        53745.6509243        4302.119900
92
```

Anomalie sur utilisation d'observations satellites (G. Kelly)

```
/scratch/rd/daa/ebjb/logdir/20030101/main/00/an/4dvar/uptraj_1/ifsmin.1 :          1,27
            Jo Global :     278737          316472.3489264          1.14
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  316472.348926       34546.5870829       3864.615055
94
/scratch/rd/daa/ebjb/logdir/20030101/main/12/an/4dvar/uptraj_1/ifsmin.1 :          1,34
            Jo Global :     209335          245958.3558854          1.17
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  245958.355885       31350.0436678       3283.661589
09
/scratch/rd/daa/ebjb/logdir/20030102/main/00/an/4dvar/uptraj_1/ifsmin.1 :          1,22
            Jo Global :     270689          293848.3429347          1.09
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  293848.342935       33005.2540858       3137.783602
33
/scratch/rd/daa/ebjb/logdir/20030102/main/12/an/4dvar/uptraj_1/ifsmin.1 :          1,23
            Jo Global :     227041          247046.3459600          1.09
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  247046.345960       29456.7105689       2751.488523
30
/scratch/rd/daa/ebjb/logdir/20030103/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     290767          288908.0527105          0.99
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  288908.052711       32236.7113220       2520.907887
24
/scratch/rd/daa/ebjb/logdir/20030103/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     256155          258081.0190496          1.01
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  258081.019050       31782.0545481       2418.500437
77
/scratch/rd/daa/ebjb/logdir/20030104/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     295499          292890.0818218          0.99
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  292890.081822       32810.1438231       2503.573236
73
/scratch/rd/daa/ebjb/logdir/20030104/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     240572          238527.4531333          0.99
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  238527.453133       30347.0163316       2138.047378
17
/scratch/rd/daa/ebjb/logdir/20030105/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     281473          264091.2223258          0.94
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  264091.222326       31121.5409696       2409.695978
60
/scratch/rd/daa/ebjb/logdir/20030105/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     217305          214307.5882262          0.99
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  214307.588226       28288.2805152       1851.639418
76
/scratch/rd/daa/ebjb/logdir/20030106/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     282812          261219.1117783          0.92
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  261219.111778       30740.8703977       2002.492142
31
/scratch/rd/daa/ebjb/logdir/20030106/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     202966          196973.5616650          0.97
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  196973.561665       27546.4142755       1647.925735
08
/scratch/rd/daa/ebjb/logdir/20030107/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     273642          255595.3144872          0.93
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  255595.314487       30410.5545157       1974.851646
06
/scratch/rd/daa/ebjb/logdir/20030107/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     259690          250369.4266358          0.96
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  250369.426636       31351.3484838       1920.428417
50
/scratch/rd/daa/ebjb/logdir/20030108/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     298968          281955.6530369          0.94
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  281955.653037       32898.8882291       2246.199661
44
/scratch/rd/daa/ebjb/logdir/20030108/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     260220          246674.6748573          0.95
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  246674.674857       30909.5577790       1913.595214
59
/scratch/rd/daa/ebjb/logdir/20030109/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     291484          262459.5897328          0.90          1,02
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  262459.589733       31891.9909930       2062.470751
81
/scratch/rd/daa/ebjb/logdir/20030109/main/12/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     261671          247835.1364361          0.95
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  247835.136436       31778.0656115       1798.778030
94
/scratch/rd/daa/ebjb/logdir/20030110/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     296244          278222.9430185          0.94
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  278222.943018       32205.4800151       2025.182585
97
/scratch/rd/daa/ebjb/logdir/20030110/main/12/an/4dvar/uptraj_1/ifsmin.1 :          1,04
            Jo Global :     259901          237290.8709594          0.91
 GREPCOST - ITER,SIM,JO,JB,JC  999 999  237290.870959       30498.7405193       1911.340193
73
/scratch/rd/daa/ebjb/logdir/20030111/main/00/an/4dvar/uptraj_1/ifsmin.1 :
            Jo Global :     294446          267026.0823887          0.91          1,02
```

```
                                                                                        1996.283467
GREPCOST - ITER,SIM,JO,JB,JC  999 999   267026.082389          30723.7970768
62                                                                                        1,05
/scratch/rd/daa/ebjb/logdir/20030111/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      242028               222600.5295437          0.92
                                                                                        1916.960673
GREPCOST - ITER,SIM,JO,JB,JC  999 999   222600.529544          29443.8972806
57                                                                                        1,09
/scratch/rd/daa/ebjb/logdir/20030112/main/00/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      276392               253085.1235589          0.92
                                                                                        2047.350471
GREPCOST - ITER,SIM,JO,JB,JC  999 999   253085.123559          31601.8268183
92                                                                                        1,14
/scratch/rd/daa/ebjb/logdir/20030112/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      211324               209277.4011967          0.99
                                                                                        1865.544508
GREPCOST - ITER,SIM,JO,JB,JC  999 999   209277.401197          29788.2113005
54
/scratch/rd/daa/ebjb/logdir/20030113/main/00/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      280574               285086.4593491          1.02
                                                                                        2308.529282
GREPCOST - ITER,SIM,JO,JB,JC  999 999   285086.459349          31884.1433896
07
/scratch/rd/daa/ebjb/logdir/20030113/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      212120               222359.1023980          1.05
                                                                                        2042.579575
GREPCOST - ITER,SIM,JO,JB,JC  999 999   222359.102398          29635.3546027
63
/scratch/rd/daa/ebjb/logdir/20030114/main/00/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      284699               296239.7581921          1.04
                                                                                        2686.776592
GREPCOST - ITER,SIM,JO,JB,JC  999 999   296239.758192          34376.8393503
59
/scratch/rd/daa/ebjb/logdir/20030114/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      268225               265914.3565779          0.99
                                                                                        2074.917494
GREPCOST - ITER,SIM,JO,JB,JC  999 999   265914.356578          32292.1346056
36
/scratch/rd/daa/ebjb/logdir/20030115/main/00/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      296343               275720.8803268          0.93
                                                                                        1973.109366
GREPCOST - ITER,SIM,JO,JB,JC  999 999   275720.880327          33038.6553805
94
/scratch/rd/daa/ebjb/logdir/20030115/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      265990               244775.3071775          0.92
                                                                                        1858.572957
GREPCOST - ITER,SIM,JO,JB,JC  999 999   244775.307178          31057.1861847
01
/scratch/rd/daa/ebjb/logdir/20030116/main/00/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      304862               291180.1664443          0.96
                                                                                        2136.046411
GREPCOST - ITER,SIM,JO,JB,JC  999 999   291180.166444          32061.5070201
72
/scratch/rd/daa/ebjb/logdir/20030116/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      265534               266608.6332511          1.00
                                                                                        2154.727906
GREPCOST - ITER,SIM,JO,JB,JC  999 999   266608.633251          33235.3025602
36
/scratch/rd/daa/ebjb/logdir/20030117/main/00/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      305053               285672.5238977          0.94
                                                                                        2297.013787
GREPCOST - ITER,SIM,JO,JB,JC  999 999   285672.523898          33427.5615669
68
/scratch/rd/daa/ebjb/logdir/20030117/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      258824               249297.7748541          0.96
                                                                                        2050.857644
GREPCOST - ITER,SIM,JO,JB,JC  999 999   249297.774854          31236.2493768
44
/scratch/rd/daa/ebjb/logdir/20030118/main/00/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      295150               269780.4915688          0.91
                                                                                        2222.334260
GREPCOST - ITER,SIM,JO,JB,JC  999 999   269780.491569          31959.0114224
81
/scratch/rd/daa/ebjb/logdir/20030118/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      238708               225648.9999502          0.95
                                                                                        2057.144659
GREPCOST - ITER,SIM,JO,JB,JC  999 999   225648.999950          31055.8556146
39                                                                                        1,09
/scratch/rd/daa/ebjb/logdir/20030119/main/00/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      280492               259695.4105145          0.93
                                                                                        2308.011236
GREPCOST - ITER,SIM,JO,JB,JC  999 999   259695.410514          30822.5631781
27                                                                                        1,16
/scratch/rd/daa/ebjb/logdir/20030119/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      211307               213670.7624570          1.01
                                                                                        1903.644222
GREPCOST - ITER,SIM,JO,JB,JC  999 999   213670.762457          29205.2420348
99                                                                                        1,04
/scratch/rd/daa/ebjb/logdir/20030120/main/00/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      282709               260686.0339818          0.92
                                                                                        2297.012622
GREPCOST - ITER,SIM,JO,JB,JC  999 999   260686.033982          30471.3185585
19                                                                                        1,06
/scratch/rd/daa/ebjb/logdir/20030120/main/12/an/4dvar/uptraj_1/ifsmin.1  :
             Jo Global  :      210251               215378.5076491          1.02
                                                                                        2041.798946
GREPCOST - ITER,SIM,JO,JB,JC  999 999   215378.507649          30020.5302435
61
```

Objective function

$$\mathcal{J}(\xi) = \Sigma_k \mathcal{J}_k(\xi)$$

where

$$\mathcal{J}_k(\xi) \equiv (1/2)\,(H_k\xi - y_k)^{\mathrm{T}}\, S_k^{-1}\,(H_k\xi - y_k)$$

with dim$y_k = m_k$

Accuracy of analysis

$$[P^a]^{-1} = \Sigma_k\, H_k^{\mathrm{T}}\, S_k^{-1}\, H_k$$

$$1 = (1/n)\,\Sigma_k\, \mathrm{tr}(P^a H_k^{\mathrm{T}}\, S_k^{-1}\, H_k)$$

$$= (1/n)\,\Sigma_k\, \mathrm{tr}(S_k^{-1/2}\, H_k\, P^a H_k^{\mathrm{T}}\, S_k^{-1/2})$$

Measure of the relative contribution of subset of data y_k to overall accuracy of assimilation.

Invariant in linear change of coordinates in data space ⇒ valid for *any* subset of data.

Can be numerically computed (Wahba, Fisher, Desroziers and Ivanov).

Can be extended to measure of relative contribution of any subset of data to accuracy of any subjet of analysed fields (but practical computation ?).

For a perfectly consistent system

$$E[\mathcal{J}_k(x^a)] = (1/2)\,[m_k - \mathrm{tr}(S_k^{-1/2} H_k P^a H_k^{\mathrm{T}} S_k^{-1/2})]$$

It is possible to compare $E[\mathcal{J}_k(x^a)]$, as determined operationally, and $(1/2)\,[m_k - \mathrm{tr}(S_k^{-1/2} H_k P^a H_k^{\mathrm{T}} S_k^{-1/2})]$, as computed directly, thus providing a check of the consistency of the assimilation system.

Also, $E[\mathcal{J}_k(x^a)]$ must be less than $m_k/2$ $\Leftrightarrow$ every piece of data must fit the analysis to within its assumed accuracy.

In particular

- Relative contribution of background

$(S_k = P^b,\ H_k = I_n)$

$$\frac{1}{n}\, tr\left[P^a (P^b)^{-1}\right]$$

How to compare with data?

$$E[J_k(x^a)] = \frac{1}{2}\left[m_k - tr\left(S_k^{-\frac{1}{2}} H_k P^a H_k^T S_k^{-\frac{1}{2}}\right)\right]$$

$2E[J_k(x^a)]$ must be less than m_k

ECMWF, operations $(p = 1.4 \times 10^6)$

$$\frac{2J_o}{p} = 0.38$$

Relative informative content of observati[ons]

as a whole $(n = 8 \times 10^6)$

$$\frac{p}{n}\left[1 - \frac{2J_o}{p}\right] = \frac{1.4 \times 10^6}{8 \times 10^6}\left[1 - 0.38\right] \simeq 0.11$$

But system globally inconsistent anyway

Assimilation without satellite observations

(globally consistent)

$\frac{2 J_0}{p}$ can be larger than 1

$\frac{2 J_0}{p} \simeq 0.95$

Relative informative content

$$\frac{p}{n}\left[1 - \frac{2 J_0}{p}\right] = \frac{2.5 \times 10^5}{8 \times 10^6}\left[1 - 0.95\right] \simeq 1.5 \times 10^{-3}$$

Seems very small. Observations more accurate than they are ~~sq~~ supposed to be (and background ~~more~~ less accurate)?

Additional work is necessary.

$$x^a = \alpha y + (1-\alpha) x^b$$

$$x^a - y = (\alpha - 1) y + (1-\alpha) x^b$$

$$E[(x^a - y)^2] = (\alpha-1)^2 \underbrace{[\sigma_o^2 + \sigma_b^2]}$$

correcty estimated (innovatio

$(\alpha-1)^2$ too large

α too small

Observations are more accurate than they are supposed to be (and backgroun is less accurate)

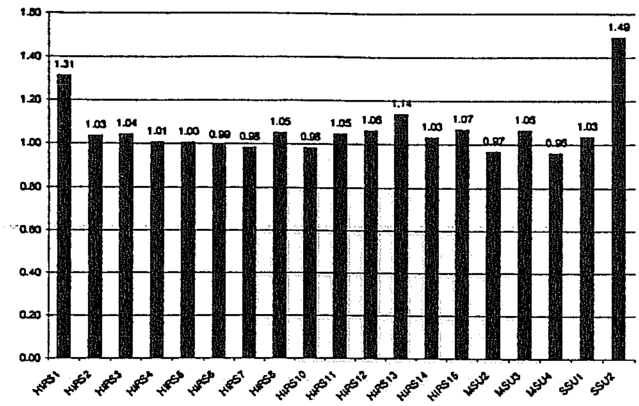


Figure 8. Tuning coefficients for NOAA14 channels after 10 iterations, using Simulated observations.

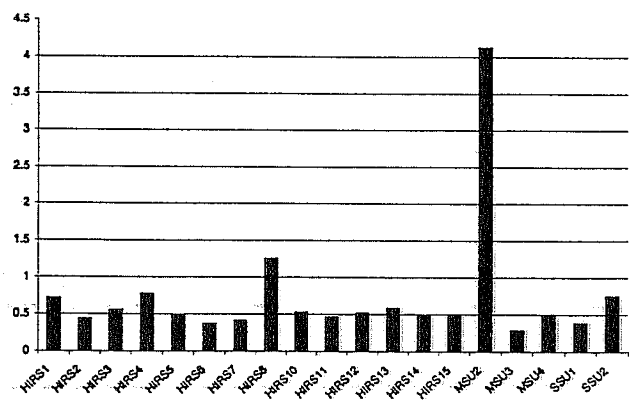


Figure 9. Tuning coefficients for NOAA14 channels after 5 iterations, real data

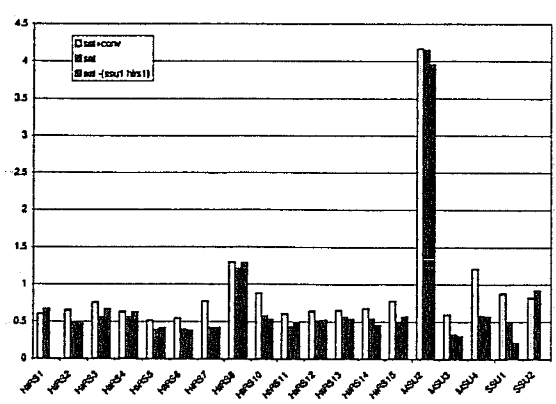


Figure 10. Tuning coefficients for NOAA14 channels computed: - 1 (white) with all other conventional observations analyzed, but not tuned. -2 (black) with NOAA14 channels only -3 (grey) with HIRS1 and SSU2 radiances removed.

Chapnik, 2003

Conclusions

- Only objective source of information on errors affecting the data is innovation vector $d = y - Hx^b$
- One-to-one transform between d and Data-minus-Analysis difference δ. It is exactly equivalent to perform diagnostics on either one of those two vectors
- Any inconsistency between a priori assumed and a posteriori observed statistics can always be mathematically explained out without change to either the analysis or the estimated analysis error. Independent ~~observations~~ hypotheses, which cannot be objectively validated (at least not on the basis of the innovation vector) will always be necessary.
- Once such hypotheses are made, adaptive adjustment of parameters became possible

- It is possible to quantify the relative contribution of any subset ~~to~~ of data to the overall accuracy of the assimilation. This particular diagnostic can be used, among others, for 'tuning' the analysis parameters.

- Do systematically this kind of diagnostics on assimilation system, even (especially?) at early stages of development

- Extension to non linear bayesian estimation?