Objective validation of data assimilation systems: diagnosing sub-optimality

Gérald Desroziers, Loïk Berre and Bernard Chapnik

Météo-France, CNRM-GAME
Toulouse, France

1 Introduction

Most operational assimilation schemes rely on the theory of least-variance linear statistical estimation (Talagrand 1997). Within this framework, analysis systems are in particular dependent on statistics for observation and background. Those statistics are not perfectly known or specified, which can lead to a sub-optimality of the analysis step. It is shown in this paper that a posteriori diagnostics can help to check the consistency of an analysis scheme. The general framework of such diagnostics is presented in section 2. Section 3 focuses on diagnostics of the differences between analysis and background and observation information. Observation space diagnostics are presented in section 4 and their ability to allow a tuning of observation or background covariances is investigated. Finally, different measures of observation impacts on analyses and forecasts are discussed in section 5.

2 General framework

In an assimilation cycle, the background x^b is given by the evolution of the previous analysis x^{a-} by the forecast model M. The subsequent analysed state x^a is obtained as an optimal combination of the background and the observations y^o. The two forecast and analysis steps write

$$x^b = M(x^{a-})$$
$$x^a = A(x^b, y^o),$$

where A stands for the possibly nonlinear analysis operator.

The estimate x^a can be classically obtained as the solution of the minimization of the following cost-function :

$$J(x) = J^b(x) + J^o(x) = 1/2[(x^b - x)^T B^{-1}(x^b - x) + (y^o - H(x))^T R^{-1}(y^o - H(x))],$$

where $J^b(x)$ and $J^o(x)$ respectively are the background and observation terms. Matrices B and R respectively stand for the background and observation error covariance matrices, and H is the possibly nonlinear observation operator including model integration in the 4D-Var formalism. Such a cost-function can be solved by the minimization of a series of quadratic cost-function with observation operators linearized around successive trajectories (Courtier et al 1994).

It can be shown that the following two equations stand for the evolution of the different errors involved in a forecast / analysis scheme, even in a slightly non-linear analysis scheme such as 4D-Var:

$$\varepsilon^b = M\varepsilon^{a-} + \varepsilon^m$$

$$\varepsilon^a = (I - KH)\varepsilon^b + K\varepsilon^o,$$

where M and H respectively are linearized versions of the model M and the observation operator H. Matrix K is defined by

$$K = BH^T(HBH^T + R)^{-1},$$

vectors ε^b, ε^a, ε^m, ε^o respectively contain background, analysis, model and observation errors, and ε^{a-} is the analysis error vector at the previous analysis step.

It is easy to check that, if the gain matrix K is consistent with the true covariances for background and observation errors, innovations d and analysis errors ε^a should be de-correlated from a statistical point of view:

$$E[\varepsilon^a d^T] = 0. \tag{1}$$

A direct consequence of the previous important property is that lagged innovations should also be de-correlated in time (Daley 1992). Equation (1) also translates into two other properties: lagged increments should be de-correlated (see Chapnik 2006 for an investigation of this diagnostic), and the differences between un-assimilated observations and the analysis should be orthogonal to the innovation vector d (Talagrand 2004).

3 "Jmin" diagnostics

3.1 Analysis consistency diagnostics

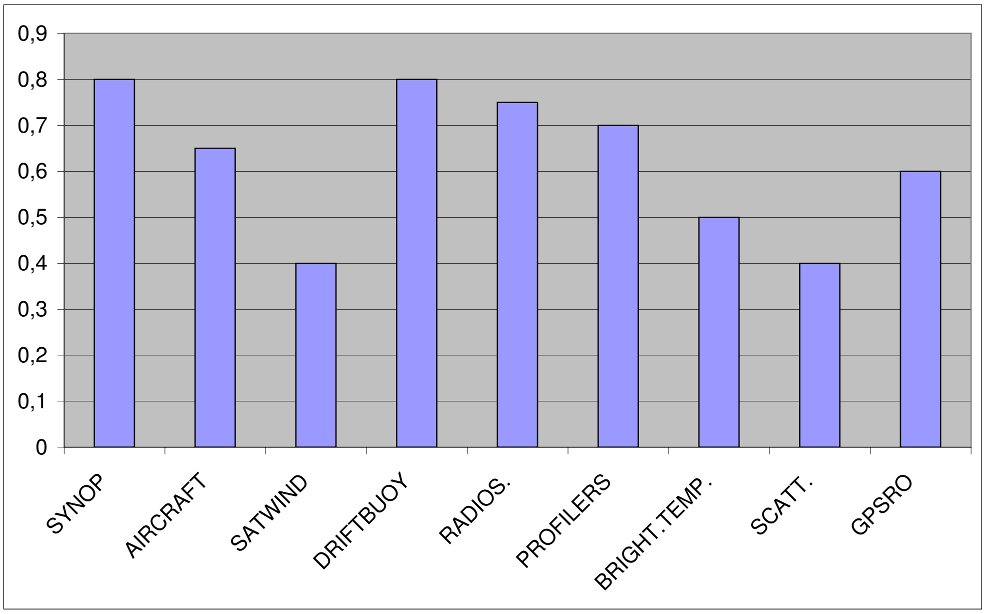

Figure 1: Tuning coefficients of observation errors in the French ARPEGE 4D-Var.

As in Talagrand (1997) and Desroziers and Ivanov (2001), it is possible to introduce an extended vector of observations z, combining the proper observations y^o, with dimension p, and the background vector x^b, with the same dimension n as the unknown true state x^t. This writes $z = \{(x^b)^T(y^o)^T\}^T$ with

$$x^b = x^t + \varepsilon^b,$$

and

$$y^o = H(x^t) + \varepsilon^o.$$

Then, it is possible to write z under the form

$$z = \Gamma(x^t) + \varepsilon,$$

where Γ is an extended observation operator and ε is the vector of background and observation errors with dimension $n+p$.

An important property pointed out by Talagrand (1999) is that if J_i stands for a sub-term of J with p_i elements, then, the statistical expectation of $J_i(x^a)$ should be

$$E[J_i(x^a)] = p_i - Tr(\Gamma_i A \Gamma_i^T S_i{}^{-1}), \tag{2}$$

where Γ_j and S_i respectively stand for the observation operator and the error covariance matrix associated with those p_i elements. Matrix A is the analysis error covariance matrix and is given by

$$A = (B^{-1} + H^T R^{-1} H)^{-1}.$$

In particular, for $\Gamma_i = I_n$ and $S_i = B$, and knowing that $K = AH^T R^{-1}$, it follows that

$$E[J^b(x^a)] = Tr(HK),$$

and for $\Gamma_i = H$ and $S_i = R$,

$$E[J^o(x^a)] = p - Tr(HK),$$

and then

$$E\{J(x^a)\} = E[J^b(x^a)] + E[J^o(x^a)] = p.$$

This means that, if the background and error statistics are correctly specified, then the expectation of the global cost function at its minimum should be equal to p. As pointed out by Bennett et al (1993) and Talagrand (1999), this is a simple *a posteriori* consistency criterion of the analysis scheme.

Desroziers and Ivanov (2001) have shown that the quantities $Tr(\Gamma_i A \Gamma_i^T S_i{}^{-1})$, appearing in expression (2), could be evaluated even if matrix K is not explicitly known, as in a variational formulation, by a Monte-Carlo procedure with

$$Tr(\Gamma_i A \Gamma_i^T S_i{}^{-1}) \simeq 1/L \sum_l \delta_{l\,i}^{o\,T} R_i{}^{-1} H_i \delta_l^a, \tag{3}$$

where δ_l^a are L perturbations on the analysis, obtained with perturbations δ_l^o on the whole set of observations, and $\delta_{l\,i}^o$ is the vector of perturbations on the subset of observations i only.

It is shown in Desroziers and Ivanov (2001) that the previous evaluations of $E[J_i(x^a)]$ can be used to tune a weighting factor s_i^{o2} of observation error variances such as

$$s_i^{o2} = J_i^o(x^a)/E[J_i^o(x^a)].$$

Fig. 1 shows the tuning coefficient s_i^o obtained in the operational ARPEGE 4D-Var for the different subset of observations that indicate that observation errors are rather overestimated in the analysis scheme.

Desroziers et al (2009) have pointed out that the quantities $E[J_i^o(x^a)]$ are direct by-products of an ensemble of perturbed assimilations, as it is implemented operationally at Météo-France.

4 Observation space diagnostics

Other diagnostics of the consistency of an analysis scheme are available. It can be simply shown (Desroziers et al 2005) that the covariance between the d_b^a *analysis-minus-background* differences in observation space and the innovations d should be equal to

$$E[d_b^a d^T] = HBH^T, \tag{4}$$

if matrix $HK = HBH^T(HBH^T + R)^{-1}$ is in agreement with the true covariances for background and observation errors.

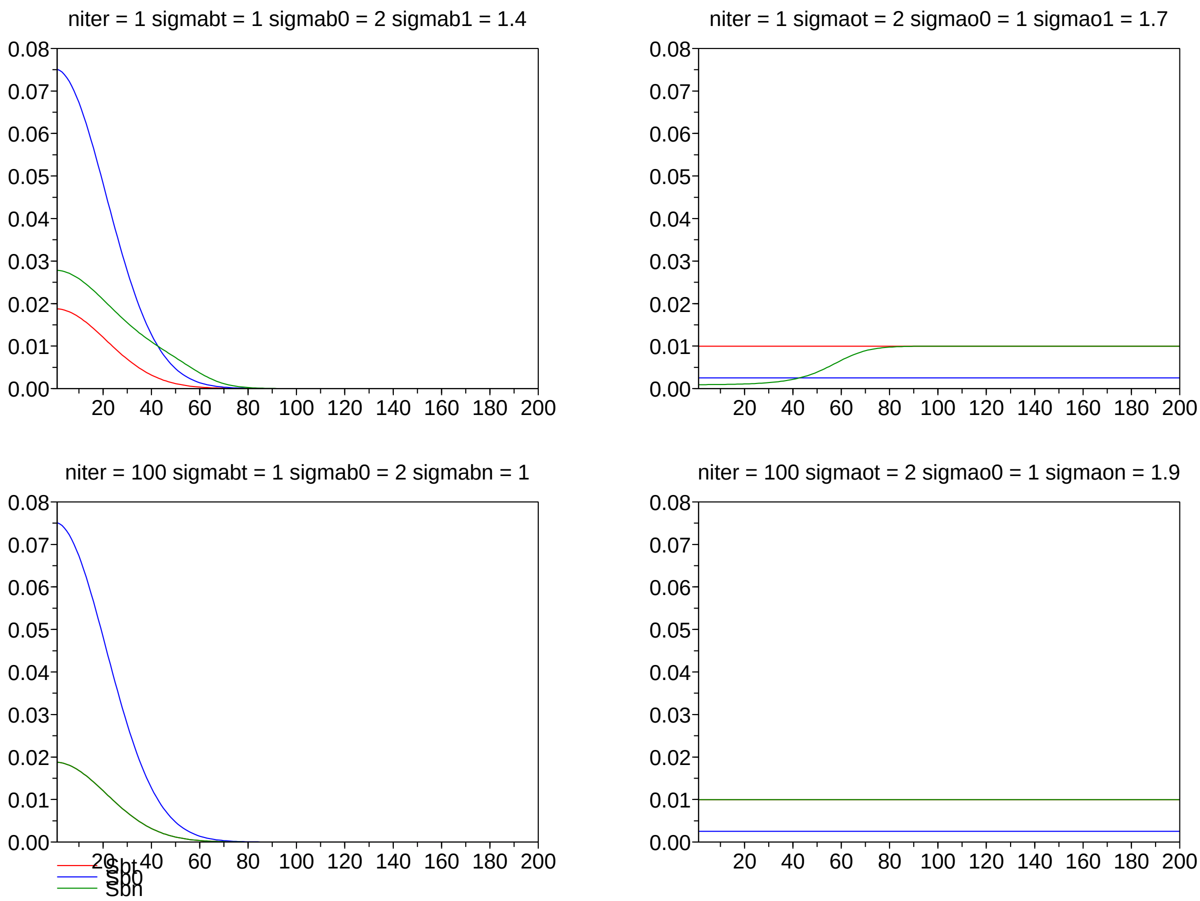

Figure 2: Λ spectra for background (left panels) and observation errors (right panels). Lengthscales for background and observation errors are L^b = 300 km and L^o = 0 km respectively. (Red curves) exact spectra, (Blue curves) (erroneous) spectra specified in the analysis, (Green curves) retrieved spectra after 1 iteration (top panels) or convergence (bottom panels). Exact values of error standard-deviations are sigmabt = 1 and sigmaot = 2. Originally specified values in the analysis are sigmab0 = 2 and sigmao0 = 1. Retrieved values sigmabn/sigmaon after 1 iteration or convergence are displayed in the titles of the figures.

This is a first additional diagnostic to the diagnostic on innovations. It provides a separate consistency check on background error covariances in observation space.

Similarly, the covariance between the d^o_a *observation-minus-analysis* differences and the innovation d should correspond to

$$E[d^o_a d^T] = R. \tag{5}$$

Finally, the cross-product between the d^a_b *analysis-minus-background* differences in observation space and the d^o_a *observation-minus-analysis* differences can also be derived:

$$E[d^a_b d^{oT}_a] = HAH^T. \tag{6}$$

Expressions (4) and (5) can be used in turn, as in Desroziers and Ivanov (2001), to tune background or observation error variances but also correlations. It has been shown in Desroziers et al (2005), relying on a toy analysis problem on a circular domain, that the following fixed-point iteration converges towards the exact values of background and observation error variances v^{ot} and v^{bt} (assumed to be homogeneous over the domain), under the condition that background and observation correlations are well specified and sufficiently different:

$$\left\{ \begin{array}{lll} v^b & = & Tr(E[d^a_b d^T] \\ v^o & = & Tr(E[d^o_a d^T]. \end{array} \right.$$

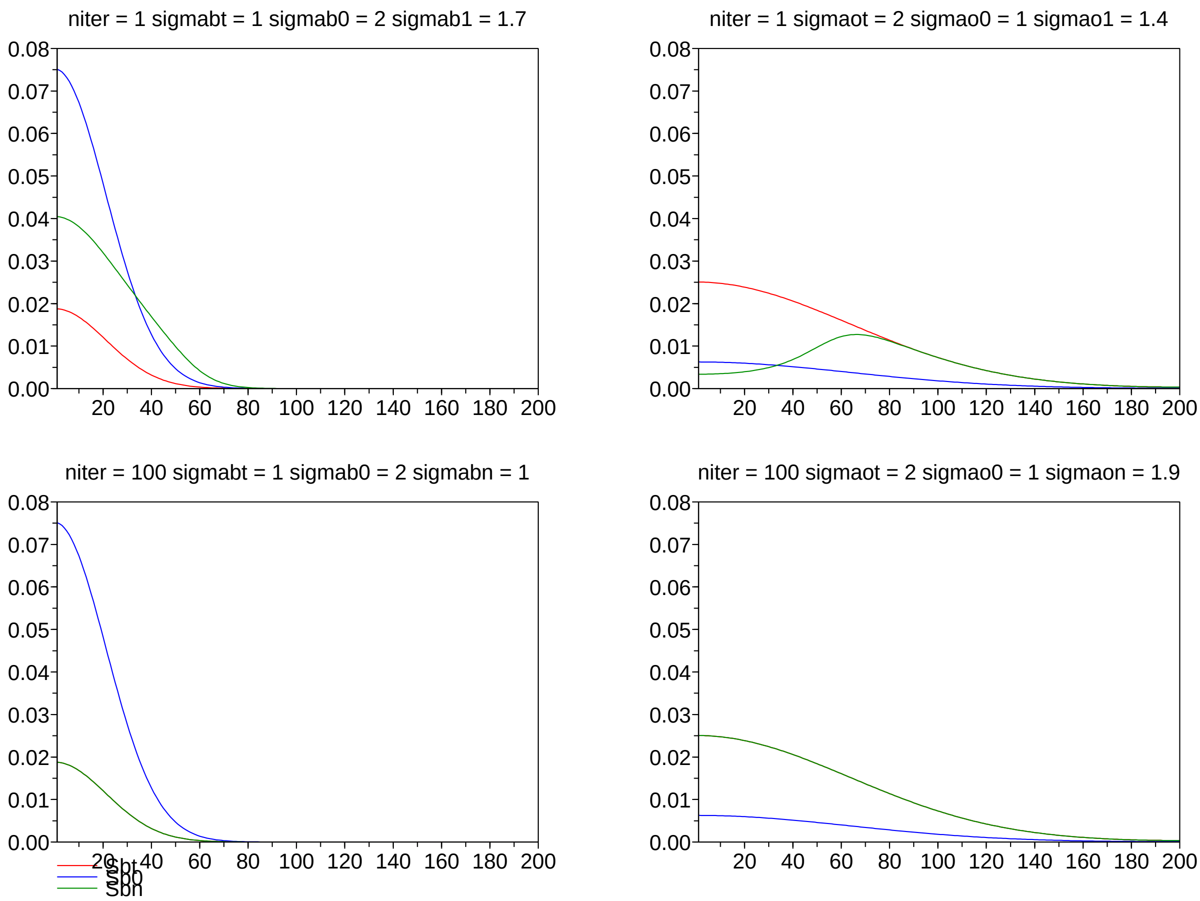

Figure 3: Same as Fig. 2, but with L^b = 300 km and L^o = 100 km.

As in Desroziers et al (2005), and Menard et al (2009), if a uniform data density is assumed over the domain, with one observation at each grid point, a spectral version of the previous iteration can be written:

$$\left\{\begin{array}{lcccl} \Lambda^b & = & \nu^b\lambda^b & & \\ \Lambda^o & = & \nu^o\lambda^o & & \\ \Lambda^b & = & \Lambda^b(\Lambda^{bt}+\Lambda^{ot})./(\Lambda^b+\Lambda^o) & & \\ \Lambda^o & = & \Lambda^o(\Lambda^{bt}+\Lambda^{ot})./(\Lambda^b+\Lambda^o) & & \\ \nu^b & = & \sum_{k=1,p}\Lambda^b_k & = & F(\nu^b) \\ \nu^o & = & \sum_{k=1,p}\Lambda^o_k & = & G(\nu^o), \end{array}\right.$$

where Λ^b, Λ^o respectively stand for the eigenvalues of matrices B and R, λ^b, λ^o their counterparts for the corresponding correlation matrices and p the number of eigenvalues. It is easy to check that Λ^b, Λ^o are such as $\Lambda^b + \Lambda^o = \Lambda^{bt} + \Lambda^{ot}$, where Λ^{bt}, Λ^{ot} are the exact eigenvalues of B and R. Observation and background error variances are also linked by $\nu^b + \nu^o = F(\nu^b) + G(\nu^o) = \nu^{bt} + \nu^{ot}$.

Fig. 2 shows the convergence of the iteration on ν^b and ν^o (and accordingly of Λ^b, Λ^o), in the toy analysis problem treated in Desroziers et al (2005), with a background error lengthscale L^b = 300 km and no correlation in observation errors (L^o = 0 km). In this case, the fixed point iteration converges towards the right values. The plot of $G(\nu^o)$ explains why the convergence towards the exact values is so fast for this case. Note that there are two undesirable fixed points when applying the iteration, which also are the boundary values of the possible interval for ν^o: ν^o = 1 and ν^o = $\nu^{bt} + \nu^{ot}$ and which can be easily eliminated.

The convergence of the algorithm is still guaranteed but slower if the lengthscales of background and observation errors become closer (Fig. 5) (see also Chapnik 2009 for a discussion of the convergence of the algorithm). It could even fail if the two correlation lengthscales are too close and in this case the sum of background and

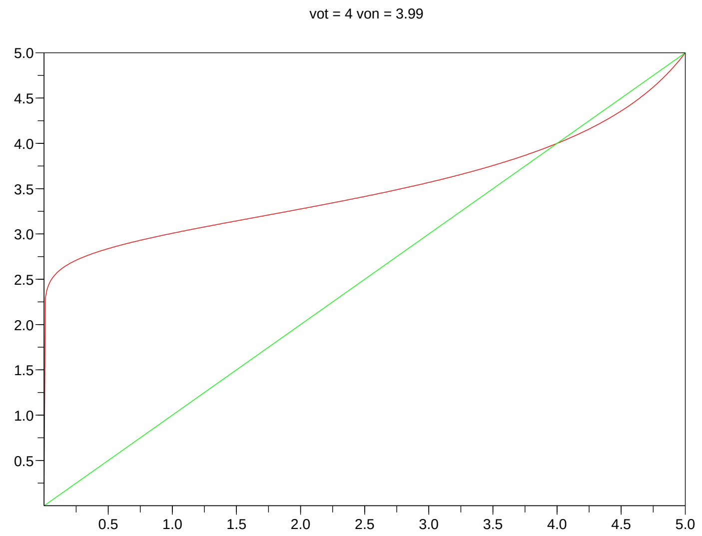

Figure 4: Function $G(v^o)$, with L^b = 300 km and L^o = 0 km, and true value of observation error variance v^o = 4.

observation variances is equal to the right innovation variance ($v^{bt} + v^{ot}$), but the ratio between v^b and v^o will stay equal to the ratio specified at the beginning of the iteration. This case is equivalent to the scalar case mentioned in Menard et al (2009), where no scale separation allows to distinguish background error variance from observation error variance.

Fig. 6 shows that even if the lengthscale of observation error is not perfectly represented in matrix R (exact value L^{ot} = 100 km, but specified value L^o = 0 km), the algorithm converges towards a reasonable value of v^o. However, it has to be noted that the retrieved value overestimates the exact value, as a larger value of v^o would be, on the contrary, required to compensate for the lack of representation of error correlation in R.

As shown in Desroziers et al (2005), there is also a scope for using such diagnostics for the estimation of correlation between observation errors. Nevertheless, it is clear that the application of the diagnostics has still to be better understood from theoretical and practical points of view.

5 Observation impact and optimality

5.1 Degree of Freedom for Signal

The analysis sensitivity to a particular subset i of observations can be given by its Degree of Freedom for Signal (DFS) introduced by Rodgers (2000). This quantity is defined by

$$DFS_i = Tr(\frac{\partial H_i(x^a)}{\partial y_i^o}). \quad (7)$$

It measures the sensitivity of the analysis to a perturbation of a particular subset of observations and then the weights of these observations in the analysis. Cardinal et al (2004) have used the DFS to measure the analysis sensitivity to observations in a real size assimilation system. Rabier et al (2002) have also shown how to use such a diagnostic to select IASI channels in order to extract the useful information from the very large amount of data provided by this instrument.

Fisher (2003) has investigated the possibility to compute the total DFS brought by the complete set of observations in a real size data assimilation. He has compared different methods to compute such a quantity. One of

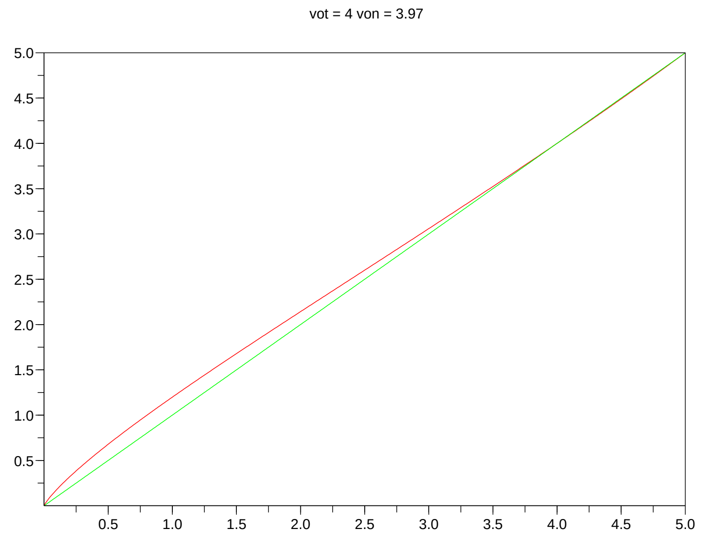

Figure 5: Same as Fig. 4, but with L^b = 300 km and L^o = 200 km.

them is based on the estimation of the so called influence matrix HK by a randomization procedure, as proposed by Girard (1987). Such a randomization procedure has also been used by Wahba et al (1995) to compute the Generalized Cross Validation criterion and inspired the randomized estimation of the $E[J_i^o(x^a)]$ proposed by Desroziers and Ivanov (2001).

Again, since an ensemble of perturbed analyses is based on explicit perturbations of observations and implicit perturbations of the background, it can provide estimations of the DFS_i, with nearly no additional computational cost (Desroziers et al 2009).

Hence, it can be shown (Chapnik et al 2006) that the previous expression of the partial DFS can be re-written

$$DFS_i = Tr(\Gamma_i A \Gamma_i^T S_i^{-1}), \tag{8}$$

where i is a subset of observations.

One can recognize a part of the expression (2) of the expectation of a sub-part of the cost function. Talagrand (1999) has interpreted this expression as a measure of the contribution of the subset of observations i to the overall precision. The DFS_i can thus be computed by the Monte Carlo procedure used in expression (3), corresponding also to the implementation of an ensemble of perturbed assimilations.

Fig. 7 shows the computation of DFS associated with the different sets of observations used in the French ARPEGE 4D-Var. The computation relies on the use of the ensemble assimilation run operationally at Météo-France.

5.2 Other measures of the impact of observations

It is easy to check, that if the gain matrix K is optimal in an analysis system, then the following relation stands:

$$A^{-1} = B^{-1} + \sum_i H_i^T R_i^{-1} H_i, \tag{9}$$

where H and R_i correspond to subsets of observations with independent errors (also independent of background errors). If the inverse of an error covariance matrix is associated with a measure of the precision of the corresponding observations, the previous relation says that the precision of the analysis is equal to the sum of the

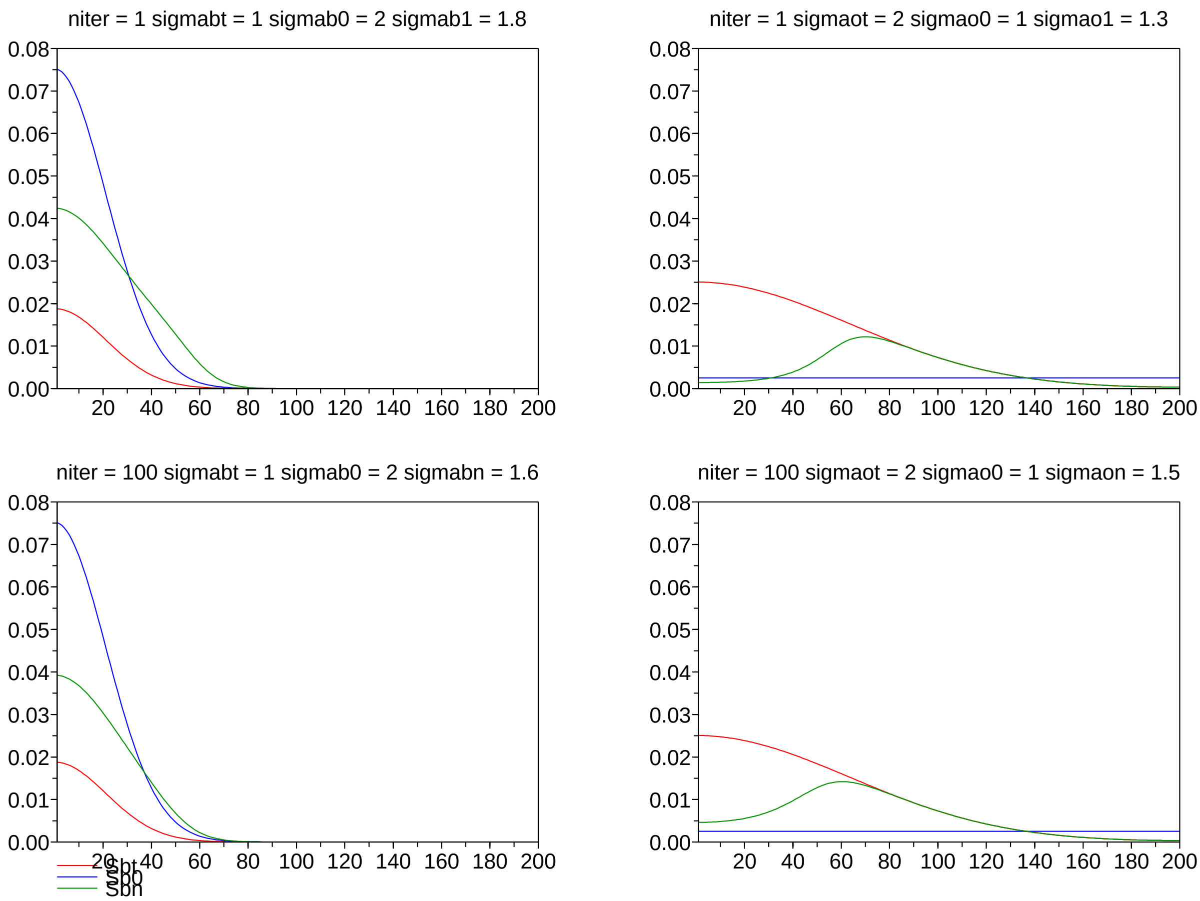

Figure 6: Same as Fig. 3 (L^b = 300 km and L^o = 100 km), but with mis-specified correlation in matrix R (L^o = 0 km in R).

precision of the different sources of independent observations (including background). Multiplying expression (9) by matrix A, it follows that

$$I_n = AB^{-1} + \sum_i AH_i^T R_i^{-1} H_i \tag{10}$$

$$= (I_n - KH) + \sum_i K_i H_i, \tag{11}$$

where K_i is the restriction of K to the independent subset of observations i. The last equation makes appear the weights associated to the background and to the different subsets of observations i. This leads to

$$n = Tr(I_n - KH) + \sum_i Tr(K_i H_i), \tag{12}$$

which expresses, in turn, that the total number of degrees of freedom n for the analysis is given by $Tr(I_n - KH)$, which measures the degrees of freedom for the analysis coming from the background and $Tr(K_i H_i)$, which are the degrees of freedom for the analysis coming from each source of proper observations. Finally, multiplying equation (9) by matrix B, it follows that

$$B = A + \sum_i AH_i^T R_i^{-1} H_i \tag{13}$$

$$= A + \sum_i K_i H_i B. \tag{14}$$

Using linear regression terminology, this last expression says that the covariance matrix of the predicted vector ε_b (the background error vector) is equal to covariance of the residual error vector ε_a plus the sum of the explained covariances by the different subsets of observations i.

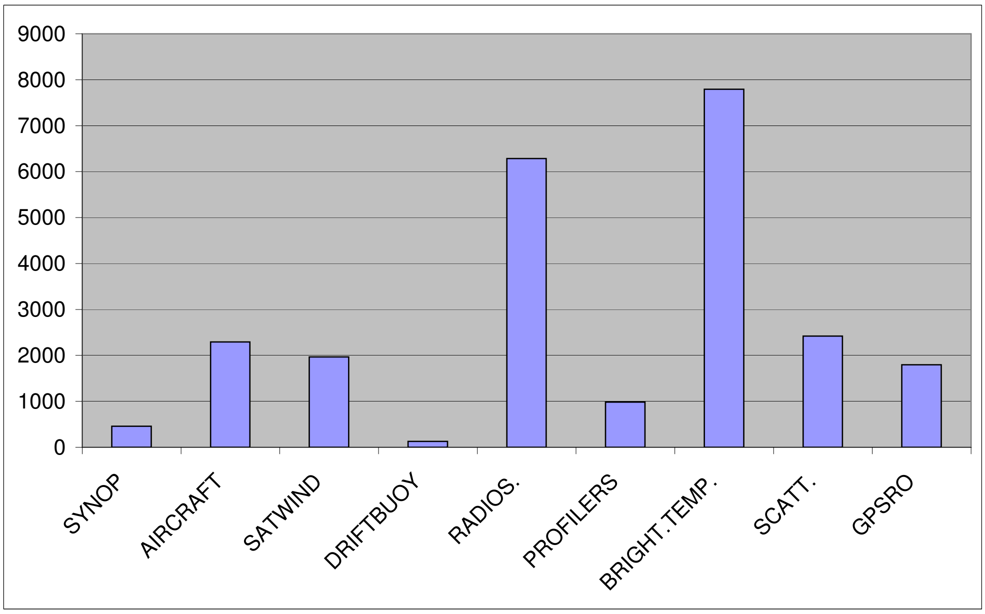

Figure 7: DFS of observations in the French ARPEGE 4D-Var.

5.3 Forecast sensitivity to observations

An increased amount of research has been done recently in numerical weather prediction to assess the observation impact on short-range forecasts. Langland and Baker (2004) have for example proposed a procedure for estimating the impact of observations on a measure of short-range forecast error, using adjoint versions of the forecast model and the data assimilation procedure. A similar approach was followed by Zhu and Gelaro (2008). Desroziers et al (2005) have proposed a randomization procedure for evaluating the error variance reduction brought by observations on analyses and forecasts. Trémolet (2008) has also recently introduced a way of computing the adjoint of the assimilation scheme in an incremental variational formalism.

Following Langland and Baker (2004), the measure of the quality of a forecast $x^f = M(x)$ can be evaluated by the following cost-function:

$$J(x) = (M(x) - x^v)^T C(M(x) - x^v),$$

where C is, for example, the energy norm, and x^v a verifying analysis at final time t^f.

If the error $\varepsilon = x - x^t$ at initial time t^i is not too large, $M(x) - x^v$ can be approximated by the evolution of this initial error ε by the tangent-linear model M, and $J(x)$ can be re-written

$$J(\varepsilon) = (M\varepsilon)^T C(M\varepsilon).$$

The impact of observations on the forecast can be obtained as the difference between $J(\varepsilon^b)$ and $J(\varepsilon^a)$. This difference can be derived by a Taylor expansion at ε^a (Cardinali 2008):

$$\begin{aligned} J(\varepsilon^b) &= J(\varepsilon^a) + (\varepsilon^b - \varepsilon^a)^T J'(\varepsilon^a) + 1/2\,(\varepsilon^b - \varepsilon^a)^T J''(\varepsilon^a)(\varepsilon^b - \varepsilon^a) \\ &= J(\varepsilon^a) + 2\,(\varepsilon^b - \varepsilon^a)^T M^T C M \varepsilon^a + (\varepsilon^b - \varepsilon^a)^T M^T C M(\varepsilon^b - \varepsilon^a) \\ &= J(\varepsilon^a) + 2\,d^T K^T M^T C M \varepsilon^a + d^T K^T M^T C M(\varepsilon^b - \varepsilon^a), \end{aligned}$$

where $J'(\varepsilon^a)$ and $J''(\varepsilon^a)$ respectively are the gradient and the Jacobian matrix of J at ε^a.

The key issue that must be pointed out here is that the first order term should be equal to zero in an optimal analysis because of the orthogonality property between the innovation vector d and the analysis error ε^a.

Alternatively a Taylor expansion at ε^b can also be written:

$$J(\varepsilon^a) = J(\varepsilon^b) + 2\,(\varepsilon^a - \varepsilon^b)^T M^T C M \varepsilon^b + (\varepsilon^a - \varepsilon^b) T M^T C M(\varepsilon^a - \varepsilon^b)$$

$$= J(\varepsilon^b) + 2\, d^T K^T M^T C M \varepsilon^b + d^T K^T M^T C M(\varepsilon^a - \varepsilon^b).$$

It is easy to check that, in this case, the statistical expectation of the first order term is equal to $-2\,Tr(M^T CMKHB)$, which is twice the optimal value of the error reduction by observations.

Thus, it appears that, in both cases, the truncation of the Taylor expansion at first order is not valid (see also Errico 2007). A second order expansion is then needed. Alternatively, the formula used in Langland and Baker leads to the same correct expression and can be interpreted as the application of a trapezoidal rule, as pointed out by Daescu (2008):

$$J(\varepsilon^a) - J(\varepsilon^b) = 1/2\, (\varepsilon^a - \varepsilon^b)^T (J'(\varepsilon^b) + J'(\varepsilon^a)).$$

6 Conclusion

Relying on statistical linear estimation theory, a set of a posteriori diagnostics of the data assimilation system can be defined. Diagnostics of internal consistency of the assumed covariances of background or observation error covariances have in particular been presented. It is clear that those diagnostics do not suffice in determining all these covariances. The tuning they may allow rely on implicit additional assumptions, but they can help greatly in determining a part of unknown statistics such as observation error variances.

Different approaches to measure the impact of observations on analyses or subsequent forecasts have been proposed. The last section showed that they must be implemented and interpreted with care.

References

[1] A.F. Bennett, L.M. Leslie, C.R. Hagelberg, and P.E. Powers. Tropical cyclone prediction using a barotropic model initialized by a generalized inverse method. *Mon. Wea. Rev.*, 121:1714–1729, 1993.

[2] C. Cardinali. Monitoring the observation impact on the short-range forecast. *Q. J. R. Meteor. Soc.*, 135:239–250, 2009.

[3] C. Cardinali, S. Pezzuli, and E. Andersson. Influence matrix diagnostic of a data assimilation system. *Q. J. R. Meteor. Soc.*, 130:2767–2786, 2004.

[4] B. Chapnik. Assessment of the optimality of an operational data assimilation system, 2006. 7th Workshop on Adjoint Applications in Dynamic Meteorology, 8-13 October, Obergurgl, Austria.

[5] B. Chapnik. On-line observation covariance matrix tuning based on optimality diagnostic. 2009. Proceedings of the ECMWF Workshop on Diagnostics of data assimilation system performance, Reading, 15-17 June.

[6] B. Chapnik, G. Desroziers, F. Rabier, and O. Talagrand. Diagnosis and tuning of observational error statistics in a quasi operational data assimilation setting. *Q. J. R. Meteor. Soc.*, 132:543–565, 2006.

[7] P. Courtier, J.-N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteor. Soc.*, 120:1367–1387, 1994.

[8] D. Daescu. On the sensitivity equations of four-dimensional variational (4d-var) data assimilation. *Mon. Wea. Rev.*, 136:3050–3065, 2008.

[9] R. Daley. The lagged innovation covariance: A performance diagnostic for atmospheric data assimilation. *Mon. Wea. Rev.*, 120:178–196, 1992.

[10] G. Desroziers, L. Berre, V. Chabot, and B. Chapnik. A posteriori diagnostics in an ensemble of perturbed analyses. *Mon. Wea. Rev.*, 2009. In press.

[11] G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteor. Soc.*, 131:3385–3396, 2005.

[12] G. Desroziers, P. Brousseau, and B. Chapnik. Use of randomization to diagnose the impact of observations on analyses and forecasts. *Q. J. R. Meteor. Soc.*, 131:2821–2837, 2005.

[13] G. Desroziers and S. Ivanov. Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Q. J. R. Meteor. Soc.*, 127:1433–1452, 2001.

[14] R. Errico. Interpretation of an adjoint-derived observational impact measure. *Tellus*, 59A:273–276, 2007.

[15] M. Fisher. Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. Technical Report 397, ECMWF, 2003.

[16] D. Girard. A fast Monte Carlo cross-validation procedure for large least squares problems with noisy data. Technical Report 687-M, IMAG, Grenoble, France, 1987.

[17] R.H. Langland and N.L. Baker. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, 56A:189–201, 2004.

[18] R. Ménard, Y. Yang, and Y. Rochon. Convergence and stability of estimated error variances derived from assimilation residuals in observation space. 2009. Proceedings of the ECMWF Workshop on Diagnostics of data assimilation system performance, Reading, 15-17 June.

[19] F. Rabier, N. Fourrié, D. Chafaï, and P. Prunet. Channel selection method for infrared atmospheric sounding interferometer radiances. *Q. J. R. Meteor. Soc.*, 128:1011–1032, 2002.

[20] C.D. Rodgers. *Inverse Methods for Atmospheric Sounding Theory and Practice*. World Scientific Publishing, London, 2000. 256 pp.

[21] O. Talagrand. Assimilation of observations, an introduction. *J. Met. Soc. Japan*, 75:191–209, 1997.

[22] O. Talagrand. A posteriori verification of analysis and assimilation algorithms. In *Proceedings of the ECMWF Workshop on Diagnosis of Data Assimilation Systems*, pages 17–28, Reading, 1999. 2-4 November 1998.

[23] O. Talagrand. Objective validation and evaluation of data assimilation. In *Proceedings of the ECMWF Workshop on Diagnosis of Data Assimilation Systems*, pages 287–299, Reading, 2004. September 2004.

[24] Y. Trémolet. Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus*, 60A:964–978, 2008.

[25] G. Wahba, D.R. Johnson, F. Gao, and J. Gong. Adaptative tuning of numerical weather prediction models: randomized GCV in three- and four-dimensional data assimilation. *Mon. Wea. Rev.*, 123:3358–3369, 1995.

[26] Y. Zhu and R. Gelaro. Observation sensitivity calculations using the adjoint of the gridpoint statistical interpolation (GSI) analysis system. *Mon. Wea. Rev.*, 136:335–351, 2008.