Observation error specifications

Gérald Desroziers, with many contributions

Météo-France, CNRS
42 av. G. Coriolis, 31057 Toulouse Cédex, France
gerald.desroziers@meteo.fr

ABSTRACT

The aim of this paper is to propose an overview of the specifications of observation errors in data assimilation schemes. Different ways of diagnosing the statistics of those errors are in particular presented. The evidence of correlations for a given number of observations and the way they can be represented are also discussed.

1 General framework

From a general point of view, the specification of observation errors is related to the Kalman Filter formalism, where one wants to obtain an estimate x^a of the true state, from two pieces of information: a background x^b associated with an error covariance matrix B and observations y^o associated with an error covariance matrix R. The analysis is given by the following equation

$$x^a = x^b + \delta x = Kd = BH^T(HBH^T + R)^{-1}d \tag{1}$$

and is then obtained by adding a correction δx to the background. This correction δx is itself given by the application of the gain matrix K to the innovation vector d containing the differences between observations and their equivalents for the background. The gain matrix is a particular expression, where B, H the linear or linearized observation operator and R appear. The quality of the analysis will then depend on the correct specification of those three ingredients and thus particularly on the correct specification of matrix R.

It is easy to see that such a solution x^a is also the expression that minimizes the following quadratic cost function:

$$J(\delta x) = 1/2\ [\delta x^T B^{-1}\delta x + (H\delta x - d)^T R^{-1}(H\delta x - d)]. \tag{2}$$

In the case where the observation operator H is non-linear, the incremental formulation proposes a way to minimize the original non-quadratic cost function, by minimizing a set of successive quadratic cost-functions (Courtier *et al*, 1994):

$$J(x) = 1/2\ [(x - x^b)^T B^{-1}(x - x^b) + (H(x) - y^o)^T R^{-1}(H(x - y^o)]. \tag{3}$$

Even in such a slightly non-linear problem, one can consider that analysis, background, model and observation errors are linked, at first order, by such a linear relation:

$$\varepsilon^a = (I - KH)\varepsilon^b + K\varepsilon^o, \tag{4}$$

with $\varepsilon^a = x^a - x^t$, $\varepsilon^b = x^b - x^t$ and $\varepsilon^o = y^o - H(x^t)$, where x^t is the unknown true state.

The background error ε^{b^+} for the next assimilation is the result of the application of the tangent-linear model M to the current analysis error ε^a plus an error ε^m intrinsically related to the model:

$$\varepsilon^{b^+} = M\varepsilon^a + \varepsilon^m. \tag{5}$$

It is easy to check that the exact covariance matrix for the analysis error is given by the expression,

$$A^t = (I - KH)B^t(I - KH)^T + KR^tK^T, \tag{6}$$

where B^t and R^t are the exact covariance matrices and K is the possibly sub-optimal or inexact gain matrix.

A correct representation of the background error covariance matrix B^{t^+} for the next analysis is then given by the expression

$$B^{t^+} = MA^tM^T + Q^t, \tag{7}$$

where Q^t is the model error covariance matrix.

From these expressions, it is interesting to note that R^t is rather an input in these equations as Q^t. It is not an output as matrix B^{t^+}. That is, it has to be specified, but cannot be documented by an algorithm such as Ensemble Kalman Filtering.

Actually, what are called observation errors, in assimilation schemes, are the differences between observations and their equivalents for the virtual true model state x^t, obtained by the application of the observation operator H to this true state. These differences can be developed as

$$\begin{aligned} \varepsilon^o &= y^o - H(x^t) \\ &= y^o - y^t + y^t - H(x^t) \\ &= \varepsilon^o_i - \varepsilon^o_H, \end{aligned} \tag{8}$$

where y^t is the true state equivalent of y^o. This expression makes appear the actual instrument observation error ε^o_i and an error ε^o_H, which is a complex function of

- the type of observation (and in particular, whether they are in situ or integrated observations, like satellite radiances),
- the resolution of the model state, which makes appear what is called the representativeness error,
- and also the precision of the observation operator, especially for satellite observations.

2 Methods for estimating observation error statistics

Observation errors are not explicitely known. However, an information on the statistics of those errors can be found in the innovations, that is in the departures between observations and background. A first method to extract an information on the observation error variance from the innovations is the so-called Hollingsworth and Lönnberg method (Hollingsworth and Lönnberg, 1986). The principle is to calculate an histogram of innovation covariances, stratified against separation. If one assumes that observation errors are spatially uncorrelated and that observation and background errors are uncorrelated, one can fit a model of background covariance to this histogram. The intercept of this covariance model at zero separation is then the background variance and the observation variance is simply the difference between the variance of the innovations at zero separation and the background error variance.

A second diagnostic that was proposed is a diagnostic based on the so-called J_{min} diagnostics, that is the expected value of the above-mentionned cost function at its minimum, for the analysis. It can be shown

that the statistical expection for the total cost function J, including the J^b and J^o term, should simply be equal to p, the total number of observations in an optimal assimilation (Bennett, 1993). More precisely, it can be shown (Talagrand, 1997) that the statistical expectation of any sub-part J_i^o of the J^o term of the cost-function is given by the expresssion

$$E[J_i^o(x^a)] = p_i - Tr(H_i A H_i^T {R_i}^{-1}), \tag{9}$$

where p_i is the number of observations associated with this sub-part. H_i and R_i respectiveley are the corresponding observation operator and error covariance matrix. A is the analysis covariance matrix for the resulting estimation.

If the actual mean value of a sub-part of J^o deviates from the previous optimal value, it can be expected that this may be partly due to a bad specification of observation errors for the corresponding observations. Then, a tuning procedure of the observation error variance is to determine a normalization coefficient as the ratio between the observed value of $J_i^o(x^a)$ and its expected value:

$$\begin{aligned} s_i^{o2} &= J_i^o(x^a)/E[J_i^o(x^a)] \\ &= J_i^o(x^a)/[p_i - Tr(H_i A H_i^T {R_i}^{-1})]. \end{aligned} \tag{10}$$

This is possible, because it happens that this complicated expression can be computed, even in a variational scheme, by a randomization procedure based on a perturbation of observations (Desroziers and Ivanov, 2001; Chapnik *et al*, 2004).

By the way, the expression $Tr(H_i A H_i^T {R_i}^{-1})$ is nothing else that the mean sensitivity of the analysis to the particular subset of observations i.

It happens then that an ensemble variational assimilation, where observations are perturbed provides, as a by-product, a way to tune the error variance of the observations and also a way to measure their mean impact in the analysis (Desroziers *et al*, 2009).

It can also be shown that such a tuning procedure for the variance of observation errors is equivalent to a maximum likelihood approach (Dee, 1998). Assuming that the innovation vector d is a Gaussian random vector with mean 0 and covariance matrix D and that D is a function of the parameter vector s, then the conditional pdf of a certain realization of d knowing s is given by the likelihood function

$$f(d|s) = \frac{1}{\sqrt{(2p)\det(D(s))}} \exp(-1/2\, d^T D(s) d). \tag{11}$$

The maximum likelihood estimate of the true coefficient vector s^t is the one that mimimizes the Log-likelihood function

$$L(s) = -\log(f(d|s)). \tag{12}$$

Other diagnostics are possible. In particular, it can be shown that the statistical expectation of the cross-products between the departures "observations - analysis" d^{oa} and the departures "observations - background" d should be equal to the observation error covariance matrix R in a consistent analysis.

A simple geometrical interpretation of these relations can found in Figure 1, which symbolizes the simplified case of the construction of the analysis x^a of a single parameter x, using an observation y^o distant of ε^o from the truth x^t and a background x^b distant of ε^b from x^t. If the scalar product of two vectors of errors is the statistical expectation, then observation error and background error are orthogonal at x^t, if one assumes that they are statistically uncorrelated. The triangle (x^b, x^t, y^o) is then a right triangle at x^t.

By construction, the analysis is a linear combination of x^b and y^o and is then on the line defined by x^b and y^o. On the other hand, the analysis error ε^a must be minimal and then orthogonal to $y^o - x^b$ that is

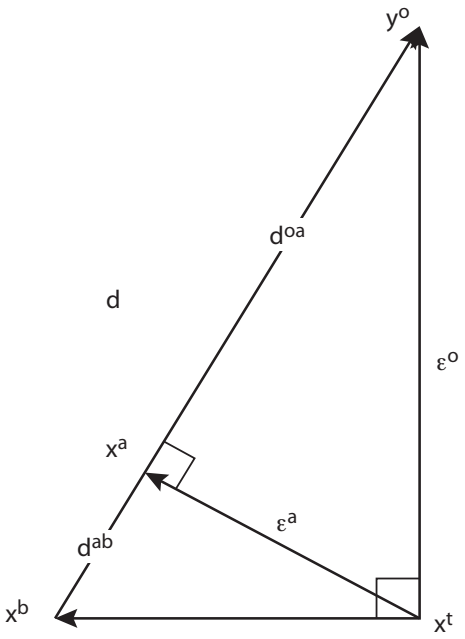

Figure 1: Geometrical representation of the analysis.

to the innovation vector. Then, the application of the Euclide theorem $d^{oa}d = \varepsilon^{o2}$ leads, in particular, to this relation for R.

A main interest of this diagnostic is that it is very simple to implement in any data assimilation scheme. The only thing to do is to compute a posteriori cross-products between the departures "observations - analysis" and "observations - background". Moreover, one can also use such computations to diagnose cross-covariances between different observations. This has been applied, in particular, at some centres to diagnose the correlations between satellite channels or spatial correlations.

However, it has to be kept in mind that all these diagnostics are not exact procedures. They all have limitations and it is important to be aware of these limitations. In particular, it appears that those diagnostics have to rely on implicit assumptions.

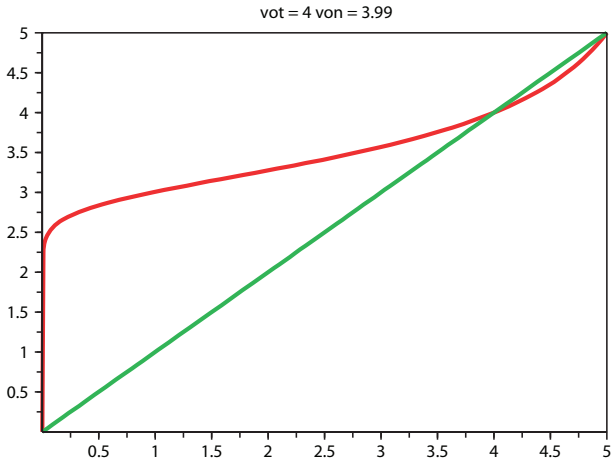

Figure 2: Function $v^o_{diag}(v^o_{spec})$, with L^b = 300 km and L^o = 0 km, and true value of observation error variance v^o = 4.

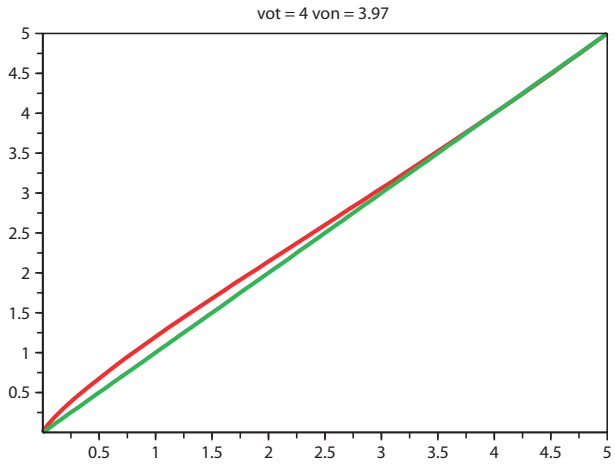

Figure 3: Same as Fig. 2, but with L^b = 300 km and L^o = 200 km.

A first assumption that is made and that has to be true to make such diagnostics work is that there is a spatial correlation in background errors and no spatial correlation or a different spatial correlation in observation errors. Figure 2 shows the ability of the observation space diagnostic to recover the right variance of observation in a toy analysis on a circle. It shows that if the two lengthscales are very different, then starting from an incorrect variance of observation v^o_{spec} equal to 1, as the exact variance is equal to 4, then the diagnosed value v^o_{diag} is already equal to 3 after one analysis with the incorrect value. The diagnosed value after a new analysis with the improved value will still improve the estimation of the variance. Then, it appears that such an algorithm will very quickly converge towards the exact value in this case and that the value obtained after a single application of the diagnostic is already very close to the exact one.

The behaviour of the convergence is completely different if the lengthscales of background and observation errors become closer. In this case (Figure 3), it can be seen that the value obtained after a single iteration is not very different from the mis-specified value and that the process will very slowly converge towards the exact value.

3 Diagnostic of observation error variances

Figure 4 shows an example of the application of the diagnostics based on the J_{min} diagnostics for the French 4D-Var assimilation scheme. The figure shows the diagnosed reduction factor that should be applied to the different observation errors. Indeed, one can see that the σ^o values are rather over-estimated in the assimilation scheme. This is especially the case for the SATWIND or the radiance observations where the diagnostics indicate that the sigmao are overestimated by a factor 2. Of course, it is known that the overestimation of the errors in the assimilation is quite often done on purpose in order to compensate the lack of correlations in the specified matrix R.

Figure 5 shows another example, provided by Niels Bormann et al (2011), of the application of different kinds of diagnostics of the σ^o values for AMSUA channels. This figure shows that all diagnostics give similar values and that these values are much lower than the specified values in the ECMWF system. So, it also appears that the σ^o values were rather overestimated in the ECMWF 4D-Var.

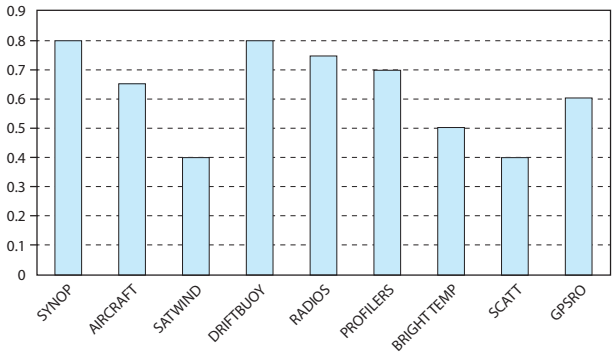

Figure 4: Normalization coefficients of the σ^o values in the French Arpege 4D-Var.

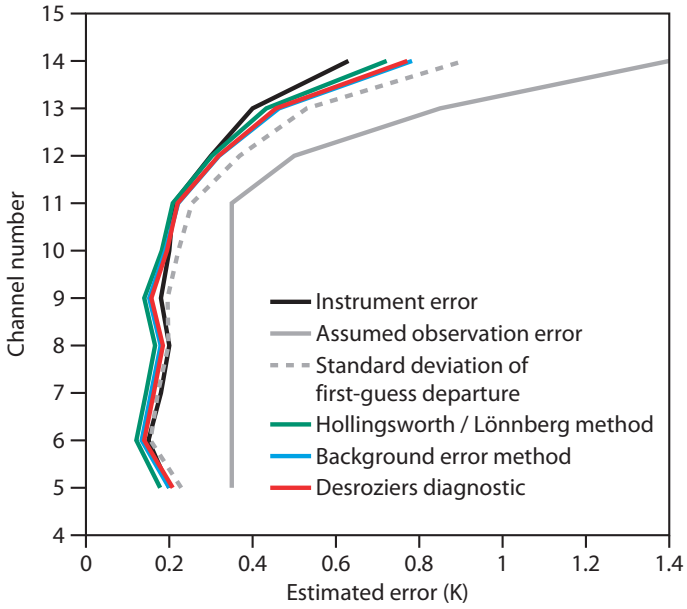

Figure 5: Estimated σ^o values for AMSU-A channels (from Bormann et al, ECMWF, 2011).

4 Diagnostic of observation error correlations

Bormann and Bauer at ECMWF (2010) have also produced different diagnostics to measure the inter-channel correlations for AMSUA errors (Figure 6). The three diagnostics applied actually show that there is no inter-channel correlations for this specific instrument. They also showed that there is no spatial correlations in those observations and then that their variances could be reduced according to the previous diagnostic on the variances. Indeed, ECMWF obtained a very nice positive impact on the forecast skills by just giving more confidence to the AMSUA data.

A different situation is found for IASI channels (Bormann *et al*, 2010), since in this case channels sensitive to water vapor or with strong surface contributions show considerable inter-channel correlations (Figure 7).

The same kind of results were obtained by Stewart (2009) at the University of Reading, since she also found strong correlations for specific IASI channels (Figure 8).

Similarly, this was also obtained by Garand *et al* (2007), at Environment Canada, for AIRS channels, who used a Hollingsworth and Lonnberg like method to diagnose these inter-channel correlations (Figure

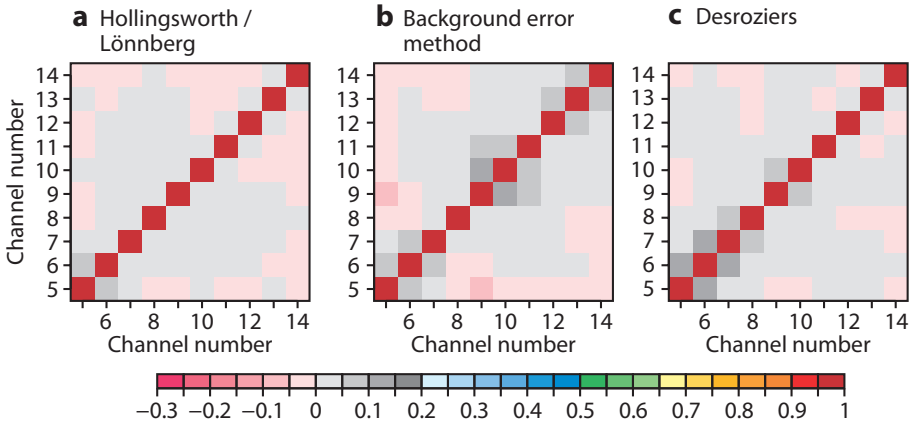

Figure 6: AMSU-A inter-channel error correlations (from Bormann et al, ECMWF, 2010).

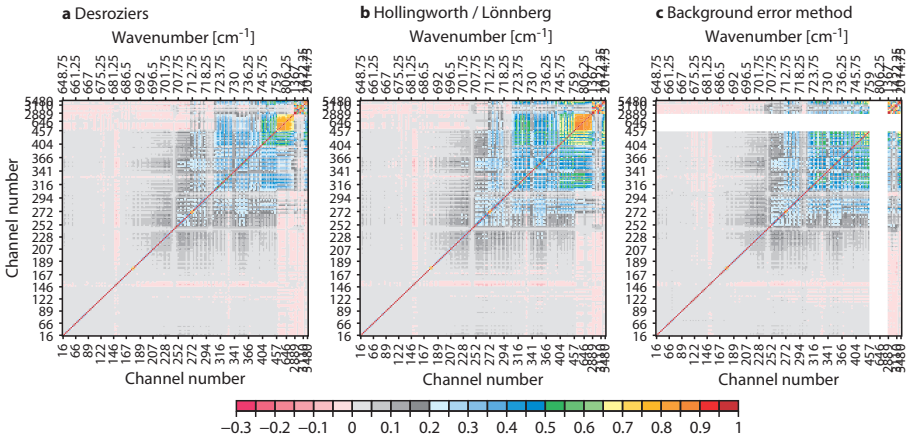

Figure 7: IASI inter-channel error correlations (from Bormann et al, ECMWF, 2010).

9).

Then, there are evidences of inter-channel correlations for satellite observations, but there are also evidences of spatial correlations for some observations. In particular, Bormann *et al* (2003) showed that there are large spatial correlations in the SATWIND observations (Figure 10).

As they found no spatial correlations for the AMSUA data, Bormann *et al* (2011) diagnosed, on the contrary, significant spatial correlations in microwave imager radiances, especially for cloudy observations (Figure 11).

Similarly, spatial observation error correlations can be found in Doppler radar winds as shown by Xu *et al* (2007), at the NOAA (Figure 12).

5 Observation error correlation specification in the assimilation

Other observation errors show temporal correlations as the SYNOP ground stations. Since they especially need to be taken into account in a 4D-Var scheme, Järvinen *et al* (1999), at ECMWF, proposed to represent them by a simple exponential correlation function:

$$c(t_1,t_2) = a\exp(-((t_1-t_2)/b)^2), \tag{13}$$

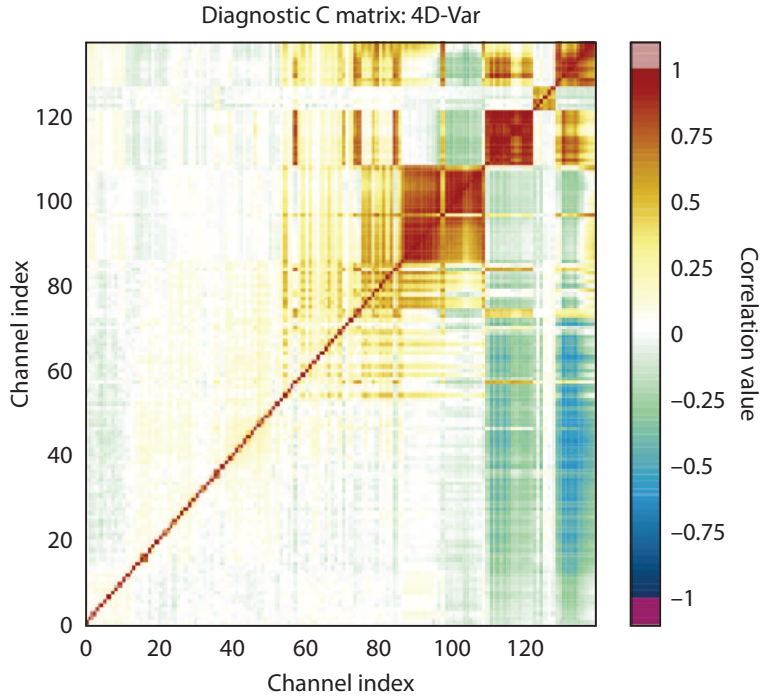

Figure 8: IASI inter-channel error correlations (from Stewart, University of Reading, 2010).

with $b = 6h$.

In the 4D-Var minimization used at ECMWF and at Météo-France, the current departures between observations and the state are normalized by the observation error standard-deviations:

$$z_i = S_i^{-1}(y_i^o - H_i(x^b) - H_i\delta x), \tag{14}$$

where S_i is the diagonal matrix containing the error variances for the observation subset i. Then, the J^o term for a subset of observations is written as a simple scalar product of these normalized departures. The way time-correlated observations are taken into account is then to compute what is called effective departures z_i^{eff} by solving a linear system, $z_i^{eff} C = z_i$ implying the above temporal correlation, and to use these effective departures as for the uncorrelated observations.

Inter-channel correlations can also be relatively easily taken into account in an analysis scheme at least in a 1D-Var scheme. This has been done, for instance, by Garand *et al* (2007) at Environment Canada. Figure 13 shows the mean temperature and humidity increments for an ensemble of 1D-Var analyses and proves that taking into account those correlations really makes a difference. The right pannel shows those increments in observation space and it logically appears that increments are smaller when inter-channel correlations are represented for the water vapor channels affected by those correlations.

The representation of spatial observation error correlations is more complicated, especially in a variational formulation. Fisher and Radnoti (2006) proposed and implemented an elegant representation of such spatial correlations. It relies on a construction of a square-root correlation model of correlations:

$$R_i = \Sigma_i^{-1} C_i \Sigma_i^{-T}, \tag{15}$$

and

$$C_i = U_i U_i^T. \tag{16}$$

The square-root U_i of the correlation matrix C_i is constructed as a sequence of operators

$$U_i = T_i S_i^{-1} G_i^{1/2}, \tag{17}$$

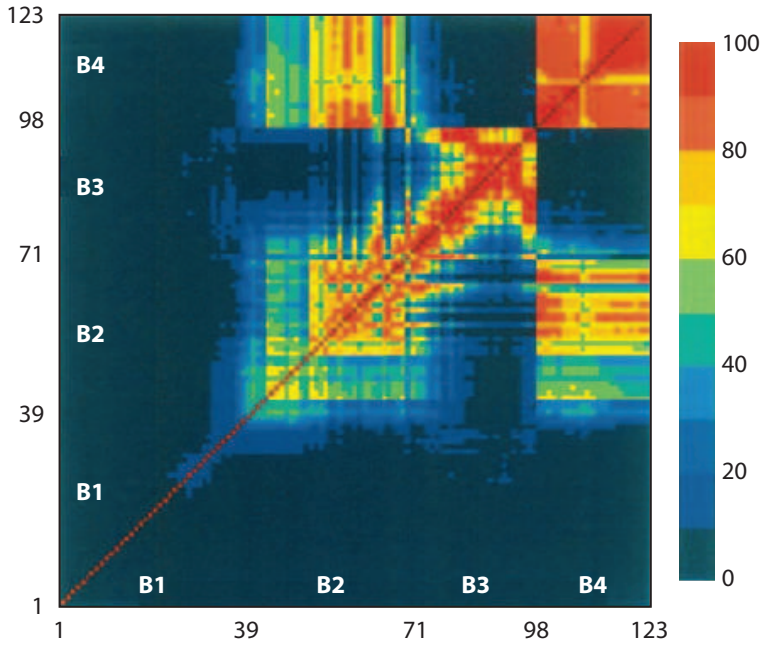

Figure 9: AIRS inter-channel error correlations (from Garand et al, Environment Canada, 2007).

where G_i is the spectral (Hankel) transformation of the correlation function, S_i^{-1} is the inverse spectral transformation and T_i is an interpolation operator at observation locations.

It has to be pointed out that such a formulation is already very useful at this stage to represent realistic observation perturbations in ensemble assimilation, as for instance for SATWIND observations and even if those correlations are not taken into account in the assimilation scheme itself. The code developed at ECMWF is thus used in ensemble assimilation as it has been implemented at Météo-France (Berre *et al* 2007) and at ECMWF (Isaksen *et al* 2010).

The representation of spatial correlations in R^{-1} is a next step, again proposed by Fisher and Radnoti at ECMWF (2006). It relies on an eigenpair decomposition of matrix C_i, using a Lanczos algorithm, and on the use of a limited number of eigenpairs in the construction of the inverse of C_i:

$$C_i = \Sigma_1^K (1/\lambda_{i,k} - 1) v_{i,k} v_{i,k}^T. \tag{18}$$

As for the time-correlated observations the practical implementation of those spatially correlated observations relies on the computation of so-called effective departures.

It is also interesting to introduce a discussion on the use of observations with correlated errors. Figure 14, produced by Liu and Rabier (2003), shows the impact of observation density on the quality of the retrieved analysis. It shows that, if observation errors are uncorrelated, then the use of denser and denser observations always improves the precision of the analysis. On the contrary, if there is observation correlation and if this correlation is not represented in the analysis then the analysis will be degraded if the observations are too dense and this degradation will start when using observations with an interdistance roughly equal to twice the lengthscale of observation correlation. Another interesting and striking point is that even if observation error correlation is well represented in matrix R, then the precision of the analysis will very quickly saturate when augmenting the density of observations.

Figures 15 and 16 show an illustration of such a behaviour, for a very simple analysis framework, on a circle, and this for two observation densities, 200 km and 50 km. In Figure 15, where there is no

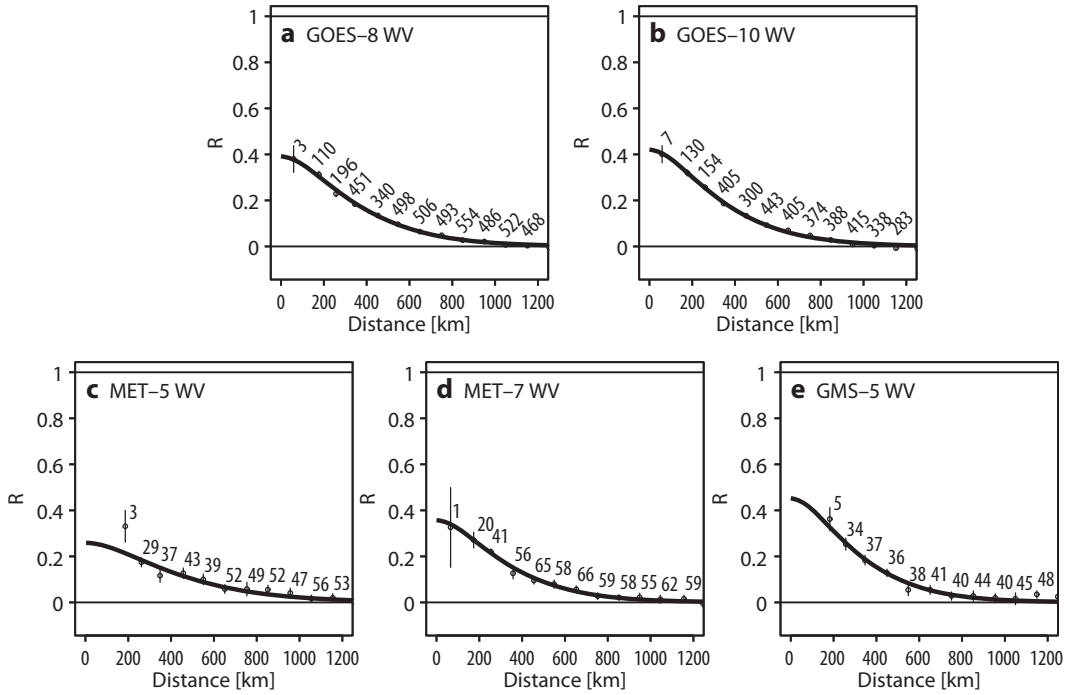

Figure 10: SATWIND spatial error correlations (from Bormann et al, ECMWF, 2003).

correlation in observation errors, it appears that the increase of the density of observations brings the analyis closer to the true state. This is also clear in the large reduction of the root mean square error of the analysis when compared to the truth.

A different behaviour is found when there is correlation in observation errors (Figure 16). On the left pannel, the interdistance between observations is roughly equal to twice the lengthscale of the correlation, which is here equal to 100 km. In this case, there is no difference between the errors of the optimal and of the suboptimal analyses. On the contrary, if the density of observations is increased, then the error in the suboptimal analysis will be much increased, but actually the precision of the optimal analysis will not be improved when compared to the precision of the analysis with 4 times less observations. This confirms the findings of Liu and Rabier (2002), at least in this simple case.

6 Conclusion

Observation errors are not explicitely known. They can be inferred by a comparison with other observations or with the background, using innovations.

There are diagnostics of observation errors (variances and correlations), but relying on explicit or implicit hypotheses.

Correlation of observation errors can be found in many datasets: SYNOP time-correlations, AIRS, IASI inter-channel correlations, SATWIND, SSM/I, radar spatial correlations. Those correlations are often neglected, but with an empirical thinning and/or an inflation of error variance. Correlations can be more or less easily taken into account. A relevant formulation for spatial error correlation has been proposed and implemented in a real size system at ECMWF.

In any case, one has to keep in mind that correlated observations are less informative than uncorrelated observations, even if R is well specified. It may thus appear inefficient to add too many correlated

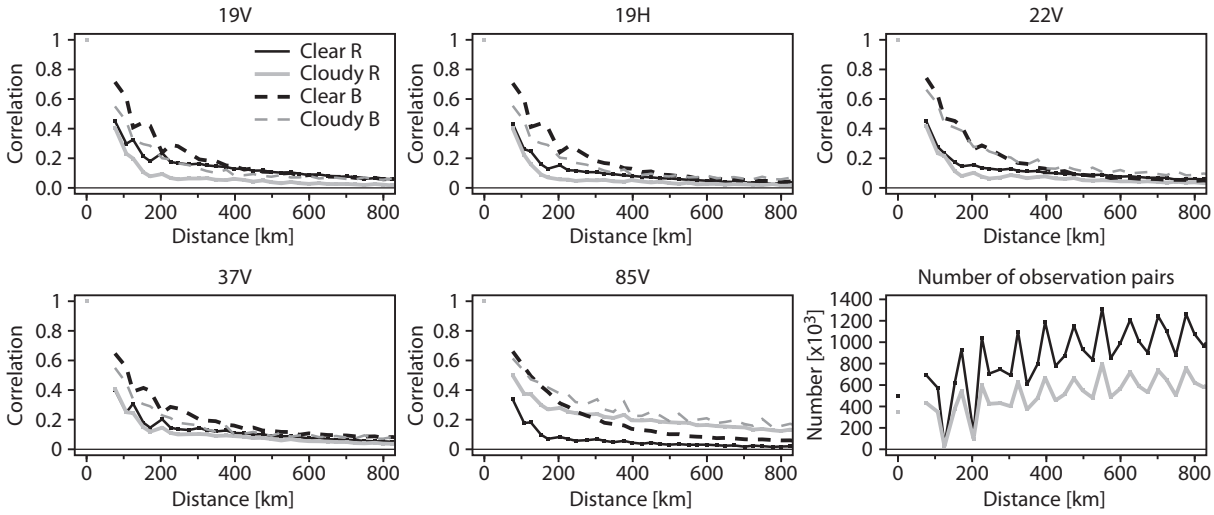

Figure 11: SSM/I spatial error correlations (from Bormann et al, ECMWF, 2011).

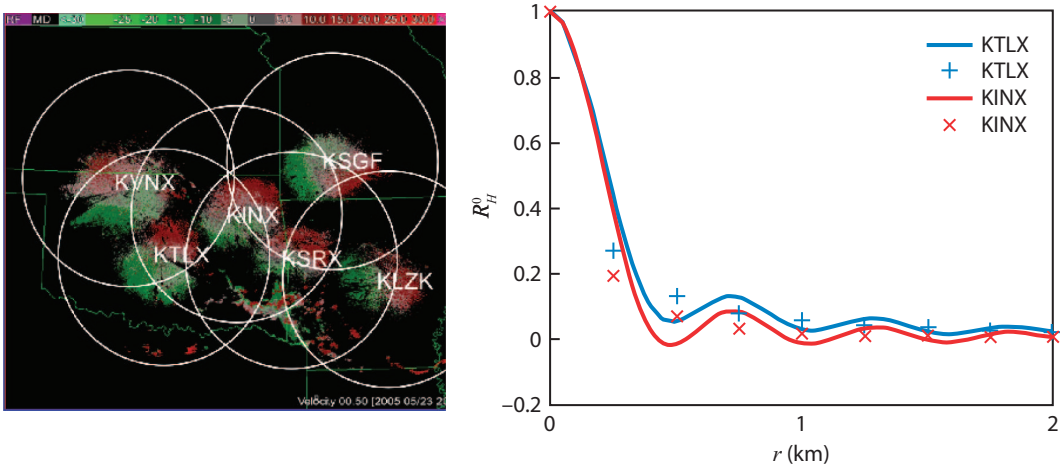

Figure 12: Doppler radar wind spatial error correlations (from Xu et al, NOAA, 2007).

observations.

Finally, the tuning of R must be consistent with the tuning of B, in order to avoid inconsistencies in assimilation schemes.

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] L. Berre, O. Pannekoucke, G. Desroziers, S. E. Stefanescu, B. Chapnik, and L. Raynaud. A variational assimilation ensemble and the spatial filtering of its error covariances: increase of sample size by local spatial averaging. In *Proceedings of the ECMWF Workshop on Recent Developments in Data Assimilation for Atmosphere and Ocean, Reading, 11-13 June 2007*, pages 151–168, 2007.

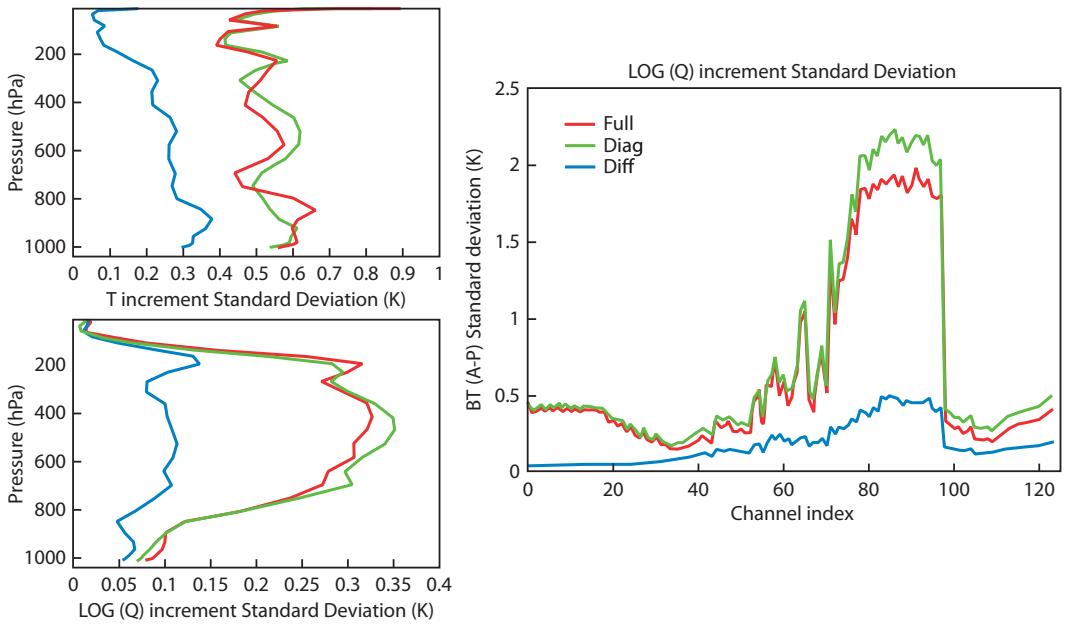

Figure 13: Representation of inter-channel error correlations in R (from Garand et al, Environment Canada, 2007).

[3] N. Bormann and P. Bauer. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to atovs data. *Quart. J. Roy. Meteorol. Soc.*, 136:1036–1050, 2011.

[4] N. Bormann, A. Collard, and P. Bauer. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II: Application to airs and iasi data. *Quart. J. Roy. Meteorol. Soc.*, 136:1051–1063, 2011.

[5] N. Bormann, A. Collard, and P. Bauer. Observation errors and their correlations for satellite radiances. *ECMWF Newsletter*, 128, 2011.

[6] N. Bormann, A.J. Geer, and P. Bauer. Estimates of observation-error characteristics in clear and cloudy regions for microwave imager radiances from numerical weather prediction. *Quart. J. Roy. Meteorol. Soc.*, 2011.

[7] N. Bormann, S. Saarinen, G. Kelly, and J.-N. Thépaut. The spatial structure of observation errors in atmospheric motion vector from geostationnary satellite data. *Mon. Wea. Rev.*, 131:706–718, 2003.

[8] B. Chapnik, G. Desroziers, F. Rabier, and O. Talagrand. Properties and first application of an error statistics tuning method in variational assimilation. *Q. J. R. Meteor. Soc.*, 130:2253–2275, 2004.

[9] 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.

[10] D. P. Dee and A. da Silva. Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Mon. Wea. Rev.*, 127:1822–1834, 1999.

[11] G. Desroziers, L. Berre, V. Chabot, and B. Chapnik. A posteriori diagnostics in an ensemble of perturbed analyses. *Mon. Wea. Rev.*, 138:3693–3720, 2010.

[12] 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.

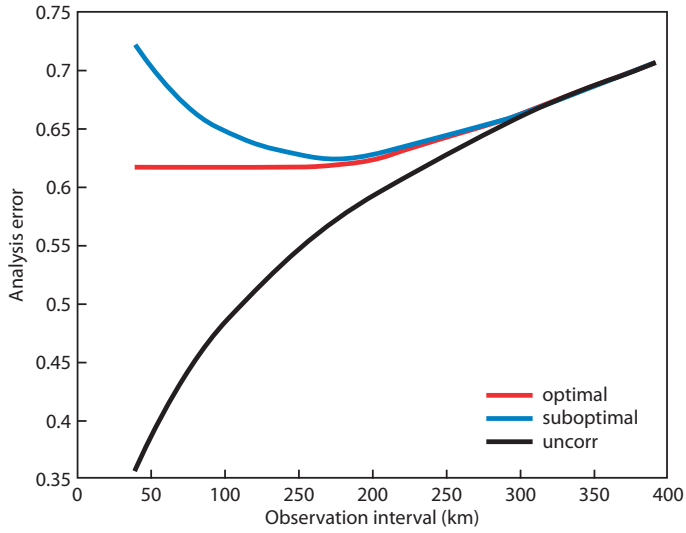

Figure 14: Representation of spatial error correlations in R (from Liu and Rabier, Météo-France, 2003).

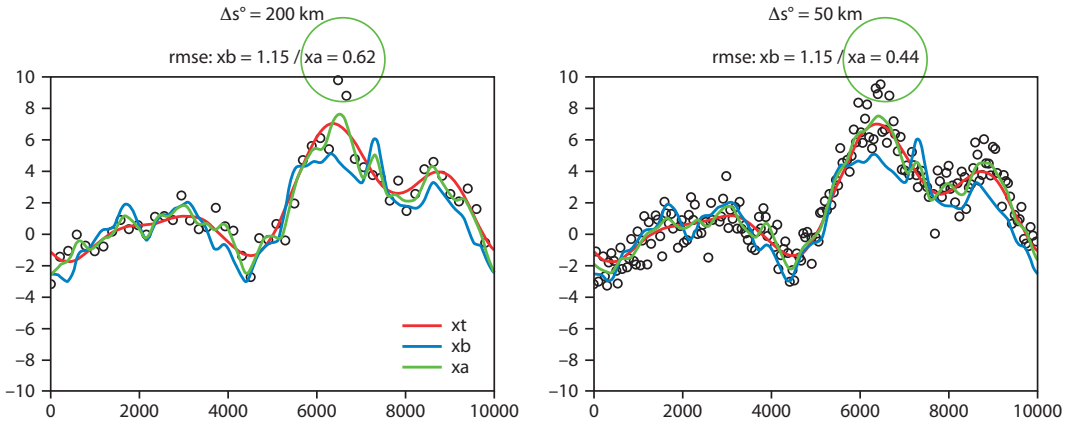

Figure 15: Case with no spatial correlation in observation errors. Red curve: x^t; blue curve: x^b; black circles: y^o; green curve: x^a.

[13] M. Fisher. Accounting for correlated observation error in a variational analysis system, 2006. 7th Workshop on Adjoint Applications in Dynamic Meteorology, 8-13 October, Obergurgl, Austria.

[14] L. Garand, S. Heilliette, and M. Buehner. Interchannel error correlation associated with airs radiance observations: Inference and impact in data assimilation. *J. Appl. Meteor.*, 46:714–725, 2007.

[15] A. Hollingsworth and P. Lönnberg. The statistical structure of short-range forecast errors as determined from radiosonde data. part I: The wind field. *Tellus*, 38A:111–136, 1986.

[16] L. Isaksen, J. Haseler, R. Buizza, and M. Leutbecher. The new ensemble of data assimilations. 2010. ECMWF Newsletter No 123, 17-21.

[17] H. Jarvinen, E. Andersson, and F. Bouttier. Variational assimilation of time sequences of surface observations with serially correlated errors. Technical Report 266, ECMWF Reading, 1999. 28 pp.

[18] Z.-Q. Liu and F. Rabier. The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Q. J. R. Meteor. Soc.*, 129:3013–3035, 2003.

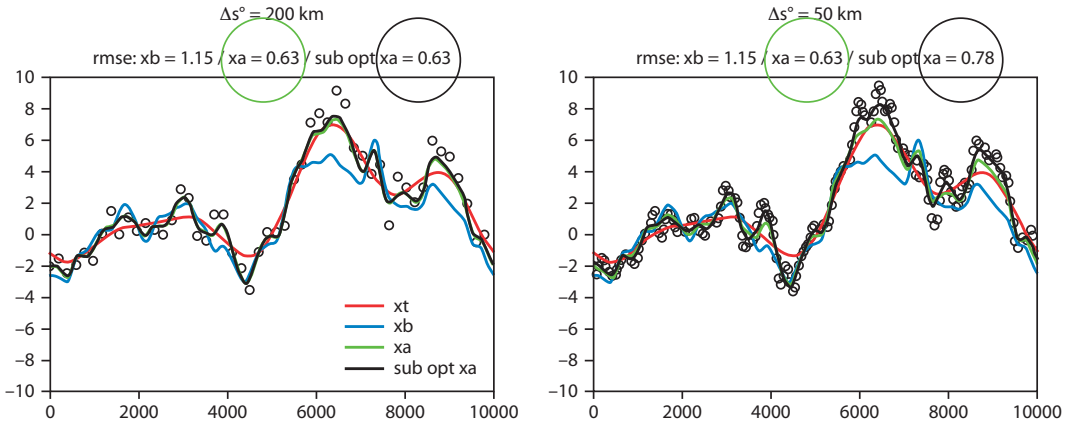

Figure 16: Case with spatial correlation in observation errors. Red curve: x^t; blue curve: x^b; black circles: y^o; green curve: optimal x^a; black curve: sub-optimal x^a.

[19] L. Stewart. Correlated observation errors in data assimilation, 2009. PhD Thesis of the University of Reading.

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

[21] Q. Xu, K. Nai, and L. Wei. An innovation method for estimating radar radial-velocity observation error and background wind error covariances. *Q. J. R. Meteor. Soc.*, 133:407–415, 2007.