SOME EXPERIMENTS WITH OBSERVATION ERROR MODELLING FOR AMSU-A OBSERVATIONS

Harald Schyberg, Frank Thomas Tveter and Ole Vignes

Norwegian Meteorological Institute
Oslo, Norway

ABSTRACT

The degree of positive impact of ATOVS observations in data assimilation depends on the realism of the observation error model used. At the Norwegian Meteorological Institute (DNMI) two new approaches to handling the observation error properties of AMSU-A have been attempted. We describe preliminary results from these attempts. The work is a part of the HIRLAM (High Resolution Limited-Area Model) cooperation, where we will utilize locally received, AAPP processed radiances in a 3D-Var scheme.

Firstly, as a method to implicitly take care of data control to deal with cloud and precipitation contaminated observations, a control variable transformation was applied in the observation cost term. This was applied in a 1D-Var scheme to account for the skew non-gaussian error distributions of the channel errors.

Secondly, the effect of correlated channels is incorporated in 1D-Var and 3D-Var by first projecting the channels onto the eigenvectors of an estimate of the observation channel error covariance matrix. These projections can then be treated as independent observations.

1. A NON-GAUSSIAN APPROACH TO 1D- AND 3D-VAR

The 1D-Var or 3D-Var least squares estimator expressed in terms of quadratic cost terms is only optimal if the observations (and also the parameters that are adjusted) all have Gaussian error characteristics. However, this is usually not the case for the error characteristics of the ATOVS channels. For AMSU-A (and other sensors), the effect of clouds and precipitation not accounted for in the forward radiative transfer model (RTM) contributes to inflating the tail in the error probability distribution.

To illustrate this, the Figure 1 shows the histogram for the departures of AMSU-A from RTTOV-5 applied on HIRLAM analyses.

The implicit quality control usually applied in Var schemes assumes a symmetric distribution of gross errors where positive and negative departures are equally probable, so that approach is not appropriate here. As an alternative to a pre-processing threshold based data control, we attempted a different approach to be presented here: A control variable transform to bring the observation error distribution closer to a Gaussian.

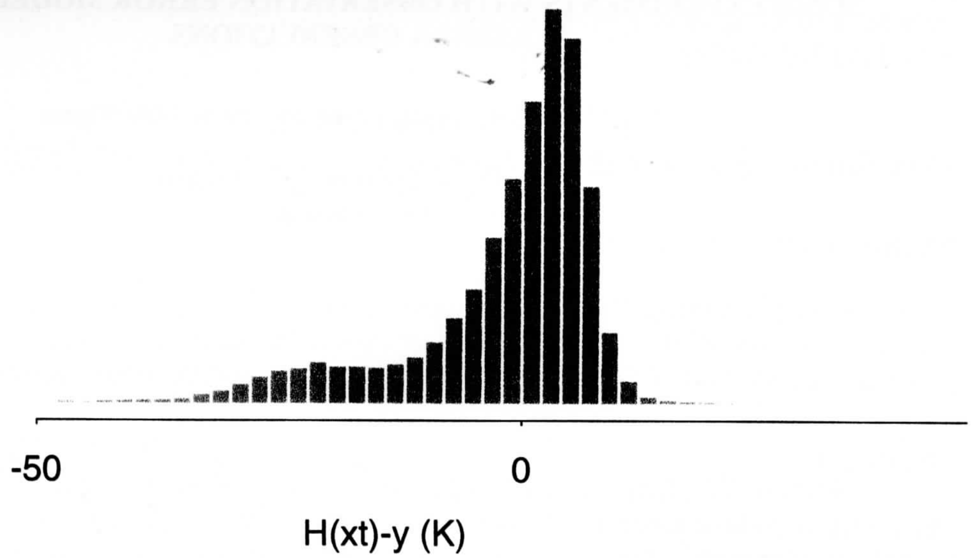

Figure 1: Typical distribution of the error statistics for AMSU-A channel 2.

1.1 One-dimensional control variable transform

From Lorenc (1986), we realize that the observation cost function in variational assimilation schemes is determined by the observation error probability density function (PDF). The quadratic form of the cost function terms arise from the fact that the corresponding probability density functions are Gaussian.

For a one-dimensional PDF, a variable with *any* PDF can be transformed into one with standard normal distribution by a variable transform, as shown in Fig.2.

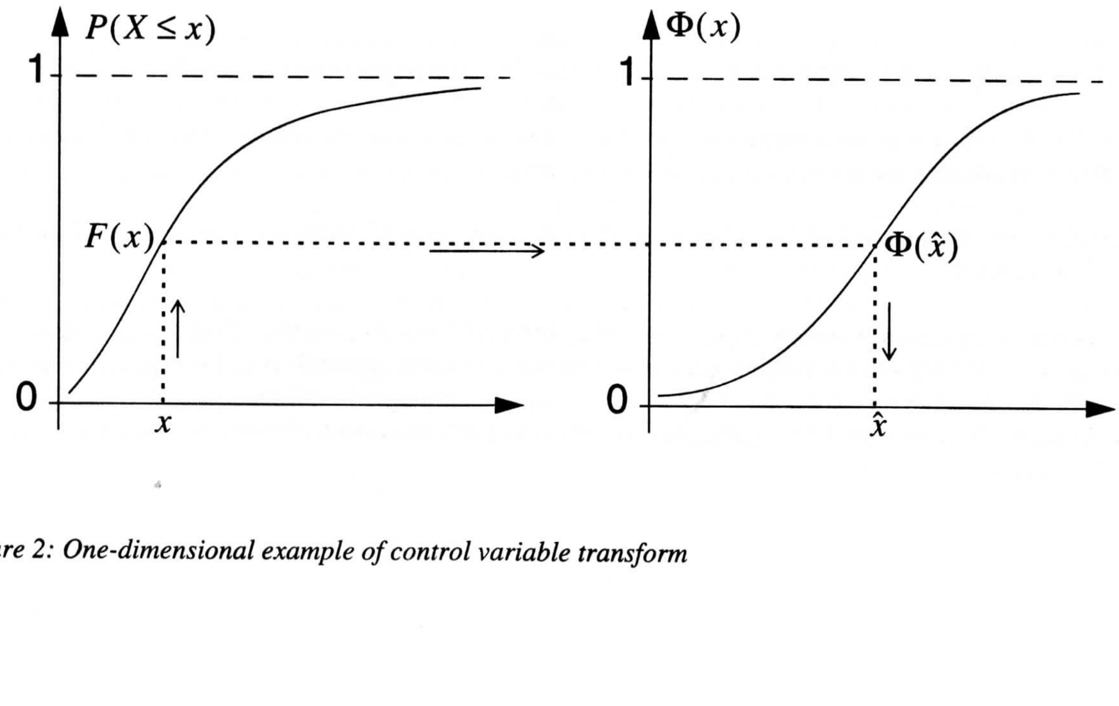

Figure 2: One-dimensional example of control variable transform

Such a transformation takes place using the cumulative distribution $F(x)$ of a variable x to transform to the variable $\hat{x}$ with standard normal distribution with cumulative distribution $\Phi(\hat{x})$. The transformation $\hat{x} = f(x)$ is defined by

$$F(x) = \Phi(\hat{x}),$$

which gives the expression for the transformed variable $\hat{x}$, $\hat{x} = f(x) = \Phi^{-1}F(x)$.

1.2 Generalization to more dimensions

In the case of AMSU-observations, we have a set of channels, so we would like to generalize the case of a scalar variable to an observation vector with higher dimension. Following the one-dimensional case we can try to generalize this result to a N-dimensional distribution using the N-dimensional cumulative distributions

$$F(\boldsymbol{x}) \equiv P(X_1 \le x_1, \ldots, X_N \le x_N), \quad \Phi^{(N)}(\boldsymbol{x}) \equiv \frac{1}{(2\pi)^{N/2}} \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_N} e^{-\frac{1}{2}\sum \xi_k^2} d\xi_1 \ldots d\xi_N$$

to define a transform $\hat{\boldsymbol{x}} = f(\boldsymbol{x})$.

This does not define a unique transform however, since it is only the scalar probability distribution that needs to be defined. In addition, it is difficult to estimate the probability distribution $P(\boldsymbol{y}|\boldsymbol{x}=\boldsymbol{x}_t) = f(\boldsymbol{y} - \boldsymbol{H}(\boldsymbol{x}))$ from empirical data. In the case of 10 AMSU-A channels this leads to a 10-dimensional PDF. By dividing each y-component (channel) axis into K intervals, we need to split the dataset into K^{10} bins. We therefore chose another approach.

Instead of performing the transform using full N-dimensional distribution, we transform each component separately as if the elements of $\delta\boldsymbol{y} \equiv \boldsymbol{y} - \boldsymbol{H}(\boldsymbol{x}_t)$ were independent,

$$P(\delta Y_1 \le \delta y_1, \ldots, \delta Y_N \le \delta y_N) = P_1(\delta Y_1 \le \delta y_1) \cdot \ldots \cdot P_N(\delta Y_N \le \delta y_N)$$

Each variable δy_k are then transformed to $\delta\hat{y}_k$ according to

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\delta\hat{y}_k} e^{-\frac{\xi^2}{2}} d\xi = P_k(\delta Y_k \le \delta y_k)$$

This defines a component-wise transform $\delta\hat{y}_k = g_k(\delta y_k)$ or in vector form $\delta\hat{\boldsymbol{y}} = \boldsymbol{g}(\delta\boldsymbol{y})$.

Generally $\delta\hat{\boldsymbol{y}}$ will not have a normal distribution, since we have only guaranteed that the integral of the probability cloud orthogonal to each component axis is Gaussian, not that the joint N-dimensional distribution is multinormal. Nevertheless, we assume the components of new transformed vector $\delta\hat{\boldsymbol{y}}$ are multi-normally distributed and estimate the covariance matrix for this new vector to enter into the observation cost function.We then hope to obtain variables which are "more Gaussian" even if observation errors are not uncorrelated between channels as assumed. This transformation is also easy to find from empirical data because we use only distributions along each δy_k channel axis, not full joint PDF. It is estimated by using empirical datasets and applying some smoothing. The cost function then corresponds to tails of PDF which are inflated on one side of the PDF as observed, and we expect this to give a more realistic implicit quality control for the cloud contamination. The method can be implemented in 1D, 3D- or 4D-Var schemes without too much complication (details not presented here).

1.3 Comments

The method was implemented in a 1D-Var scheme using AMSU-A observations and HIRLAM forecast first guess fields. This lead to a slightly slower convergence of the scheme, as could be expected. To estimate the PDFs we need to use a reference dataset whose errors are difficult to separate out (in our case we used HIRLAM analyses for estimating $\boldsymbol{H}(\boldsymbol{x}_t)$). The scheme as implemented in 1D-Var was compared both to 1D-Var without any quality control and to a traditional 1D-Var with a pre-check with a threshold of 3K on channel 8. AMSU-A observations and HIRLAM forecast background profiles were co-located with radiosondes, and improvements obtained with 1D-Var over the background fields versus radiosondes were found.

It became clear that a non-linear variable transformation is a very powerful way of controlling and differentiating the weight given to an observation in the data assimilation. The effects of non-linearity in the forward operator also became very visible. For the new scheme we found that the scheme performed clearly better than not doing any quality control.

On datasets where the samples rejected by the pre-check were removed, the two schemes performed comparably. It seems that the gross error check made the observation error statistics almost Gaussian. In other words, the gross error check worked quite well on its own, and the transformation method had problems making good use of the residual non-Gaussian information that was available in the error statistics. We therefore concluded not to insert the new method into the HIRLAM 3D-Var for AMSU-A data.

The approach is however still interesting for treating HIRS data, where a cloud contamination pre-check is not as easy as for microwave, but this has not been attempted so far.

2. CORRELATED CHANNEL ERRORS

Even if AMSU-A observation errors are difficult to estimate, it is reasonable to assume that there may be error correlations between the channels. Estimating a covariance matrix using departures from a reference model analysis shows channel correlations, but those are not necessarily identical to the true observation error correlations.

Usually the various channels have been assumed uncorrelated in the assimilation, since many Var assimilation schemes are coded under the assumption of uncorrelated observations. Since HIRLAM 3D-Var is among these schemes, we can not insert directly a non-diagonal covariance matrix into the scheme. It can however be easily implemented by projecting the channels onto the eigenvectors of an estimate of the observation error covariance matrix and use these projections as independent observations.

2.1 Estimating error correlations

The starting point for the AMSU-A observation error covariance matrix was statistics from the deviation of AMSU-A observations from RTTOV simulated observations using a HIRLAM analysis where only conventional observations were used. The matrix elements of this estimate is expected to be larger than the most optimal AMSU-A error covariance matrix, since it contains a contribution from the analysis error and since the conventional observation errors are independent of AMSU-A observation errors. The reader should note that if we had assimilated AMSU-A observations in the HIRLAM analysis, the resulting error covariance matrix would have been smaller than the most optimal (a priori) AMSU-A error covariance matrix. The statistics found is presented in Fig. 3, which shows standard deviations for departures of AMSU-A from RTTOV-5 applied on (conventional) HIRLAM analyses, for a period of one month (January 2000) and in Fig. 4 which shows channel correlations for these departures.

Because of the difficulty of separating out the part of this covariance matrix coming from the reference dataset (analysis) error correlation, we chose to downscale the matrix elements to account for error contributions from the analysis error. The actual covariance matrix used in 3D-Var has been scaled down by a factor 0.25, as tuning tests showed this to give the best forecast verification results towards radiosondes.

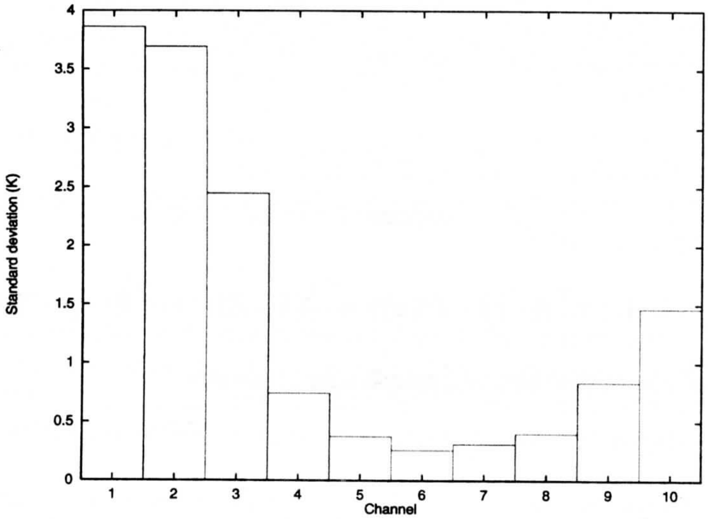

Figure 3: RMS of channel departures from RTTOV-5 applied on HIRLAM analyses

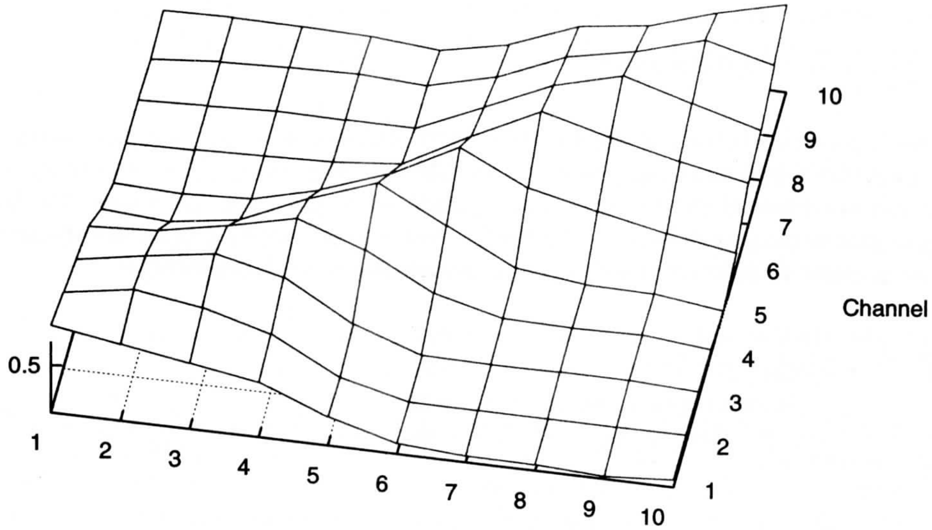

Figure 4: Channel covariance matrix for channel departures from RTTOV-5 applied on HIRLAM analyses

2.2 Method

With a covariance matrix estimated as above, the basic idea is to find the matrix of eigenvectors Z, associated with the observation error covariance matrix $\boldsymbol{R}$. In this case we may write that

$$\boldsymbol{Z}\boldsymbol{R}\boldsymbol{Z}^T = \hat{\boldsymbol{R}},$$

where $\hat{\boldsymbol{R}}$ is the matrix with the eigenvalues on the diagonal and $\boldsymbol{Z}$ is the matrix with rows of eigenvectors, which are found with standard mathematical subroutines. Given an observation vector y, the corresponding "rotated" observation $\hat{y}$ is given by

$$\hat{y} = \boldsymbol{Z}y.$$

The components of this vector now become independent, and the observation cost term is written

$$J_o = \frac{1}{2}(y - H(x))^T \boldsymbol{R}^{-1}(y - H(x)) = \frac{1}{2}(\hat{y} - \boldsymbol{Z}H(x))^T \hat{\boldsymbol{R}}^{-1}(\hat{y} - \boldsymbol{Z}H(x)),$$

where the right hand term can be rewritten as a simple sum of squares.

2.3 Comments

The general feeling at DNMI is that channel rotation probably does not introduce any fundamentally new information to the variational assimilation scheme. However, this approach may make it easier to balance the weight of the different channels against one another. It is of interest to study further the difference in quality from using a tuned diagonal AMSU-A observation covariance matrix, i.e. assuming uncorrelated observations. The observation error for the correlated channels should in this case be appropriately increased, thus reducing the weight of these channels in the data assimilation.

3. Conclusions, further work

It is possible to precondition variational schemes to better deal with the non-Gaussian observation error statistics of ATOVS by using a simple control variable transform. For AMSU-A information it turned out that a simple window channel pre-check is a very good tool in detecting cloud contamination, so there is not anything to gain in using such an approach for these data. It would however be interesting to try the approach for the use of infrared information, where cloud detection is much more difficult.

The HIRLAM 3D-Var was extended to be able to treat cross-channel AMSU error covariances by a preprocessing projecting the observation vector onto the observation covariance matrix eigenvectors. Further work is needed to obtain an optimal estimate of the observation error covariance matrix, particularly work to study more in detail the effect of the non-diagonal elements, and how to estimate the reference dataset error contribution.

ACKNOWLEDGEMENTS

The work undertaken was supported financially by the Norwegian Space Centre.

REFERENCES

Lorenc, A.C., 1986: Analysis methods for numerical weather prediction. *Quart. J. R. Met. Soc.* **112**, pp. 1177-1194.

TECHNICAL PROCEEDINGS OF THE ELEVENTH INTERNATIONAL ATOVS STUDY CONFERENCE

Budapest Hungary

20-26 September, 2000

Edited by

J.F. Le Marshall and J.D. Jasper

Bureau of Meteorology Research Centre, Melbourne, Australia

Published by

Bureau of Meteorology Research Centre

PO BOX 1289K, GPO Melbourne, Vic., 3001, Australia

June 2001