Accounting for correlated observation error in the assimilation of high resolution sounders

Peter Weston, James Cameron and Laura Stewart

1) Introduction

Currently data from high resolution sounders, such as AIRS (Atmospheric Infrared Sounder on the NASA satellite Aqua) and IASI (Infrared Atmospheric Sounding Interferometer on the EUMETSAT satellite Metop-A), are used with diagonal observation error covariance matrices (**R**) within the Met Office 4D-Var assimilation scheme, assuming no correlation between channels. This is inadequate due to the presence of errors of representativeness, forward model error and errors associated with the pre-processing of the data. Previous work both at the Met Office (Stewart et al. 2009) and ECMWF (Bormann et al. 2010) has demonstrated that correlations exist in IASI data particularly for channels sensitive to water vapour. It is likely that a better description of the error correlations in 4D-Var will allow for improved use of the water vapour channels. This poster shows the results of performing a diagnostic technique described by Desroziers [2005] on AIRS and IASI data to estimate the true structure of the **R** matrices. Initial tests using the full matrices resulted in the 4D-Var minimisation becoming unstable leading to non-convergence and increased computational cost. To counter this, the raw matrices have been reconditioned. Initial results from trialling these matrices in the Met Office assimilation scheme are also shown.

2) Desroziers Diagnostic

To estimate the structure of the full **R** matrix I have used the diagnostic procedure introduced by Desroziers et al. (2005). This uses observation minus background (O-B) and observation minus analysis (O-A) statistics to produce observation error variances and covariances. The formula is:

$$\mathbf{R} = E((y - H(x_a))(y - H(x_b))^T)$$

A key assumption which is used in the derivation of the above formula is that the **R** and **B** matrices used in the assimilation to produce the O-A and O-B stats are exactly correct. However in this project we know that the **R** matrix is not correct initially. Therefore the results should not be entirely trusted. However it has been shown, in very simple examples, that iterating the Desroziers diagnostic after starting with incorrect errors can lead to convergence to the true errors which is encouraging.

3) IASI channels used

Window channels

Temperature sounding channels
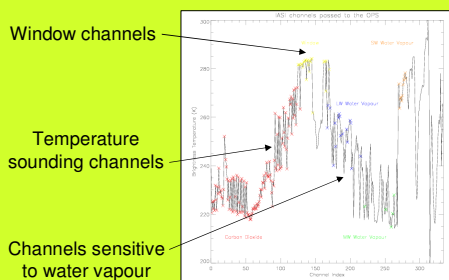
Channels sensitive to water vapour

Figure 1 – Sample brightness temperatures measured by each of the IASI channels assimilated

5) Timings

After symmetrising the diagnostic matrices and making sure they were positive definite the full matrices were used in a 4D-Var assimilation cycle. The results are as follows:

Full **R** used for	Overall time (s)	Time per iteration	Number of iterations
No obtypes	3605	91.16	39
IASI only	25.21%	0.29%	25.64%
AIRS only	101.83%	0.54%	102.56%
AIRS & IASI	127.02%	0.42%	128.21%

Figure 3 – Table showing the increase in time, processing time and iterations when using the full **R** matrices for IASI and AIRS in 4D-Var

4) 4D-Var Results

Weak correlations   Moderate correlations   Very strong correlations

The Desroziers diagnostic was run on output from the Met Office's 4D-Var assimilation scheme (VAR). The results from this are shown in figure 2.

There are weak error correlations between temperature sounding channels, slightly stronger error correlations between window channels and very strong error correlations between water vapour sensitive channels.

The diagnosed standard deviations are much smaller than the currently used operational ones. The water vapour sensitive channels have the largest standard deviations. This is due to larger representativeness errors for these channels.

Inflated operational standard deviations

Figure 2 – The raw diagnostic **R** matrices on the left, the correlation (**C**) matrices in the middle and the diagnosed and operational standard deviations (square root of the diagonal of **R**) on the right, for AIRS (top) and IASI (bottom) data

6) Conditioning

A suggested reason for the increased number of iterations was that the matrices had high condition numbers. The condition number of a matrix is defined as the ratio of the largest eigenvalue to the smallest eigenvalue. It is a measure of how easily invertible the matrix is, so the smaller the condition number the easier it is to invert. Figure 4 shows the condition numbers of the raw diagnostic matrices.

Instrument	Processed channels	Assimilated channels	Full matrix condition number	Sub matrix condition number
IASI	314	138	26668.3	1956.6
AIRS	324	142	12337.9	12287.5

Figure 4 – Table showing the condition numbers of the diagnostic full and sub matrices for IASI and AIRS

Matrix used	Sub matrix condition number (IASI, AIRS)	Iterations	Cost function value (IASI, AIRS)
Diagonal	64.0, 256.0	39	39252, 68895
Full	1956.6, 12187.5	89	242020, 269904
Full	67,66.9	60	128448, 133220

Figure 5 – Table showing the results of using full matrices with different condition numbers in the assimilation

Figure 5 shows that reducing the condition numbers of the matrices (by manipulating the eigenvalues) results in a decrease in the number of iterations. However, even with very low condition numbers there are still more iterations than when using the diagonal matrices.

Increase in standard deviations

Operationally at the Met Office a conjugate gradient based minimisation is now used where a fixed number of 60 iterations are performed. Putting the full matrices into this scheme leads to mixed results. The minimisation fails when using the raw matrices, but using the reconditioned matrices results in a successful convergence of the minimisation. Figure 6 compares the standard deviations from the raw matrices and the reconditioned ones.

Figure 6 – Comparison between the raw diagnostic and reconditioned standard deviations (matrix condition number 67) for IASI.

7) Initial Trial Results
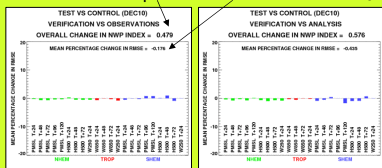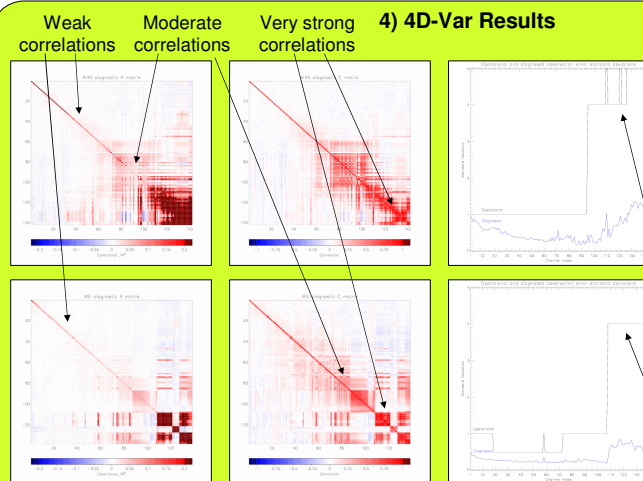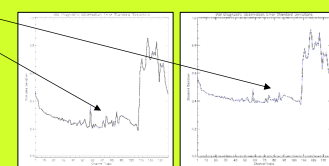
Positive impact   RMSE reduction

Figure 7 – Impact of using the full IASI **R** matrix in the Met Office forecast system

A month long trial using just the full **R** matrix for IASI is currently running. A full **R** matrix is used as part of the quality control procedure in the OPS and results in less observation rejections due to non convergence or bad retrieved brightness temperatures. A different full **R** matrix is also used in the 4D-Var assimilation scheme and results in more weight being given to the IASI observations.

Figure 7 shows results from the first 17 days of this trial and that these changes result in a positive impact and an overall reduction in RMS error in the Met Office forecast system.

8) Conclusions

The Desroziers diagnostic is not perfect but gives a good general idea of what the correlation structure is. Therefore the raw diagnostic matrices need to be edited to make sure they are symmetric and positive definite when being tested in the assimilation.

The inter-channel error correlations are largest in channels sensitive to water vapour and so, as predicted, the biggest positive impact of modelling the correlations should be in these channels.

Using the full **R** matrices in the assimilation scheme results in a very small increase in processing time but does result in the minimisation needing many more iterations to converge. Because of this the raw matrices need to be reconditioned before use operationally.

Initial results from trialling the use of these reconditioned full **R** matrices show positive impact. Future work will include further testing and trialling and hopefully operational implementation.

References

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

Stewart L, Cameron J, Dance S, English S, Eyre J, Nichols N. 2009. Observation error correlations in IASI radiance data. Mathematics Report Series 1/2009, 26pp. Univ. Reading: Reading, UK.

Bormann N, Bauer P. 2010. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. II: Application to AIRS and IASI data. Q. J. R. Meteorol. Soc. 136: 1036–1063.