Improved Use of AIRS Data at ECMWF

A.D. Collard, A.P. McNally

European Centre for Medium-Range Weather Forecasts, Reading, U.K.

W.W. Wolf

QSS Group, Inc., NOAA Science Center, 5200 Auth Road, Camp Springs MD, U.S.A.

Abstract

Two possibilities for the improved use of AIRS data in the ECMWF assimilation scheme are explored. Improved spatial sampling has been achieved through the use of the "Warmest Field of View" dataset where a higher number of clear fields of view are assimilated. Improved spectral sampling has been investigated through the use of "reconstructed radiances" where each reconstructed channel is a linear combination of the whole AIRS spectrum. Forecast improvements from both of these approaches have been encouraging.

Introduction

The initial implementation of the assimilation of AIRS radiances at ECMWF was necessarily conservative and resulted in a small but significant positive impact on forecast scores. Experiments have been run investigating more aggressive use of AIRS data. Two such possibilities are the use of more data spatially through the use of the "Warmest Field of View" dataset and the use of reconstructed radiances — a proxy for greater spectral use of the data.

Use of the Warmest Field of View Dataset

More data may be used spatially in the lower troposphere by more intelligent pre-thinning of the AIRS data. In the current operational system, only the central of the nine AIRS fields of view (FOVs) in each AMSU-A FOV is provided in near real time by NOAA/NESDIS. However, NOAA/NESDIS also provide a warmest FOV (used as a proxy for the clearest observation) dataset which is routinely transferred to ECMWF via FTP. Two experiments have been run with the alternative dataset (plus control runs).

The first experiment used the ECMWF assimilation setup (CY28R4) in operational use between October 2004 and April 2005. The experiment ran between 21st December 2004

and 2nd February 2005 (although 19 days in total suffered from problems with the transfer of the AIRS data between NESDIS and ECMWF and are excluded from this analysis).

Figure 1 shows the change in used observation numbers (i.e., after quality control, cloud detection and thinning) for the first ten days of this experiment. The number of used observations for the highest peaking channels is virtually unchanged, while those channels that are significantly affected by the surface are increased by a factor of two.

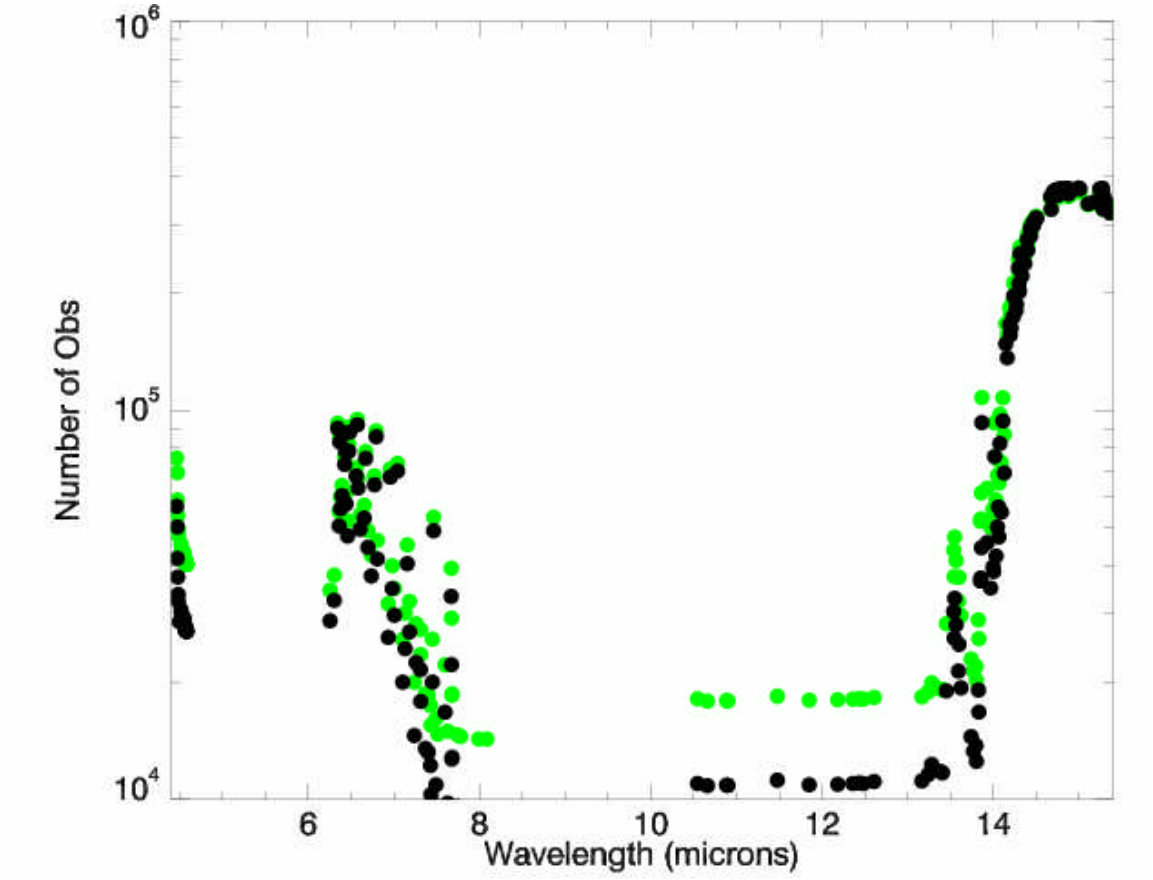


Fig. 1: Change in the number of used observations globally for the first ten days of the experiment. The black dots are for the central field of view dataset and the green dots are for the warmest FOV dataset. Note the logarithmic scale.

In general the observation statistics of other instruments are unchanged. The AIRS statistics are also very similar to those when the central field of view is used except for a 0.2-0.3K increase in the observation–background biases in the AIRS water vapour band in the tropics (Figure 2). This is a feature of the cloud detection scheme in the water vapour band that is exaggerated by the different sampling in the warmest FOV case. The effect of this change in bias is a decrease of less than 2% in the mean water vapour abundances.

Figure 3 gives an example of the distribution of clear fields of view (fields of view marked as clear in AIRS channel 787 which is a window channel at 10.9μm) compared to the cloud fraction from the MODIS cloud mask.

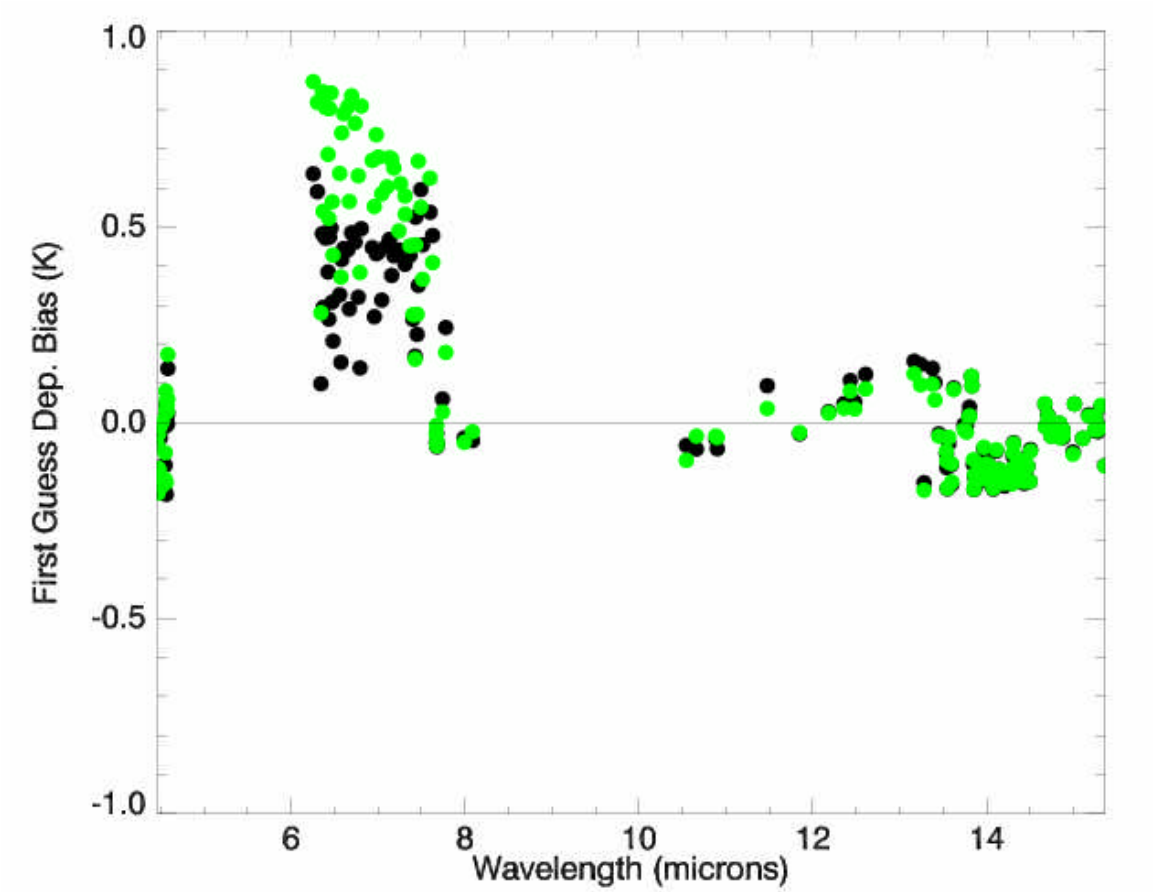


Fig. 2: Tropical observation–background bias for the first ten days of the experiment. The black dots are for the central field of view dataset and the green dots are for the warmest FOV dataset. The only significant difference seen in the bias is in the tropical water vapour band where the bias increases by around 0.3K.

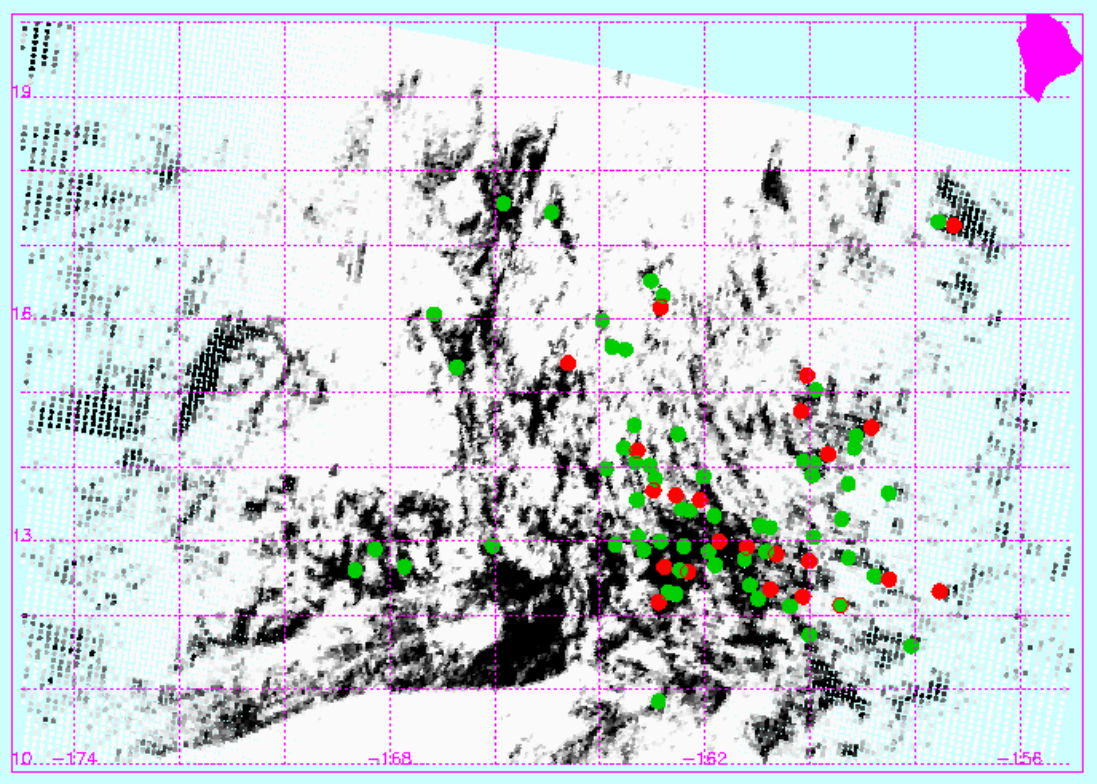
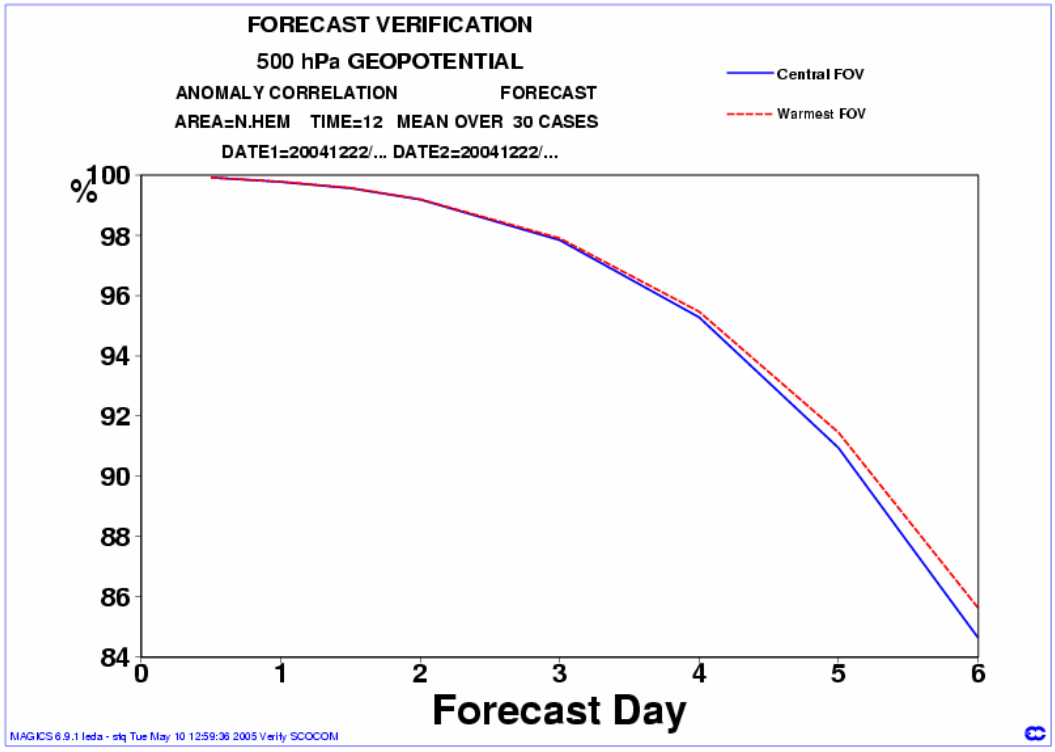


Fig.3. Example distribution of clear fields of view for AIRS channel 787 at 10.9μm (as determined by the ECMWF AIRS cloud detection algorithm) for the central FOV (red) and warmest FOV (green) datasets. The background image is the cloud fraction as determined by MODIS with black indicating clear and white overcast. There are 23 fields of view from the central FOV dataset and 59 from the warmest.

The verification scores based on a total of 30 days are shown in Figures 4–7. Figures 4 and 5 show positive impact from the use of the Warmest FOV dataset in the extra-tropical 500hPa anomaly correlation forecast scores. Figures 6 and 7 show the tropical vector wind scores for 200hPa and 850hPa. The impact is neutral and indicates that the extra bias seen in the tropical water vapour brightness temperature differences are not having an adverse effect.

The second experiment was based ifs cycle CY29R2 which went operational in June 2005 and included some recent modifications to the AIRS processing including a change to thinning based on clearest observations instead of random thinning. This experiment ran for 61 days between 1st March 2005 and 30th April 2005.

For this experiment, the change to the thinning algorithm results in a smaller increase in the number of used observations (Fig. 8), but the differences in biases are now much smaller (the biases in the control run are now similar to the warmest FOV run in the first experiment). The forecast verification scores (Fig. 9) show neutral impact in the Northern Hemisphere, but a robust positive impact in the Southern Hemisphere.



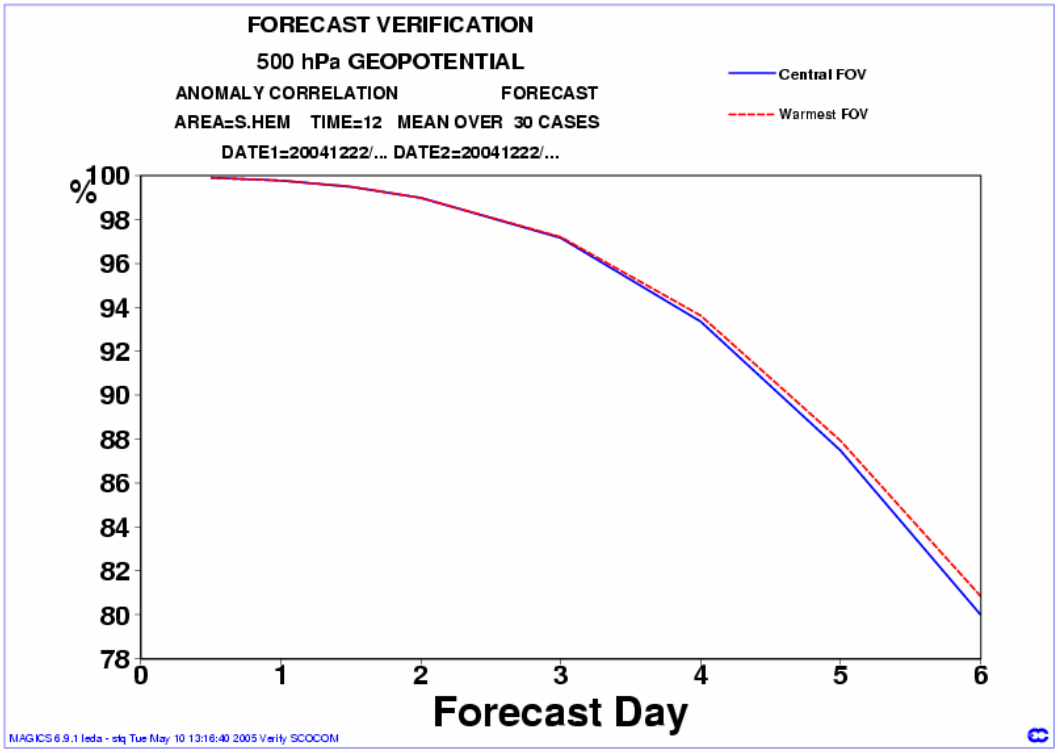


Fig. 4: Northern and southern hemisphere extra-tropical 500hPa anomaly correlation forecast scores

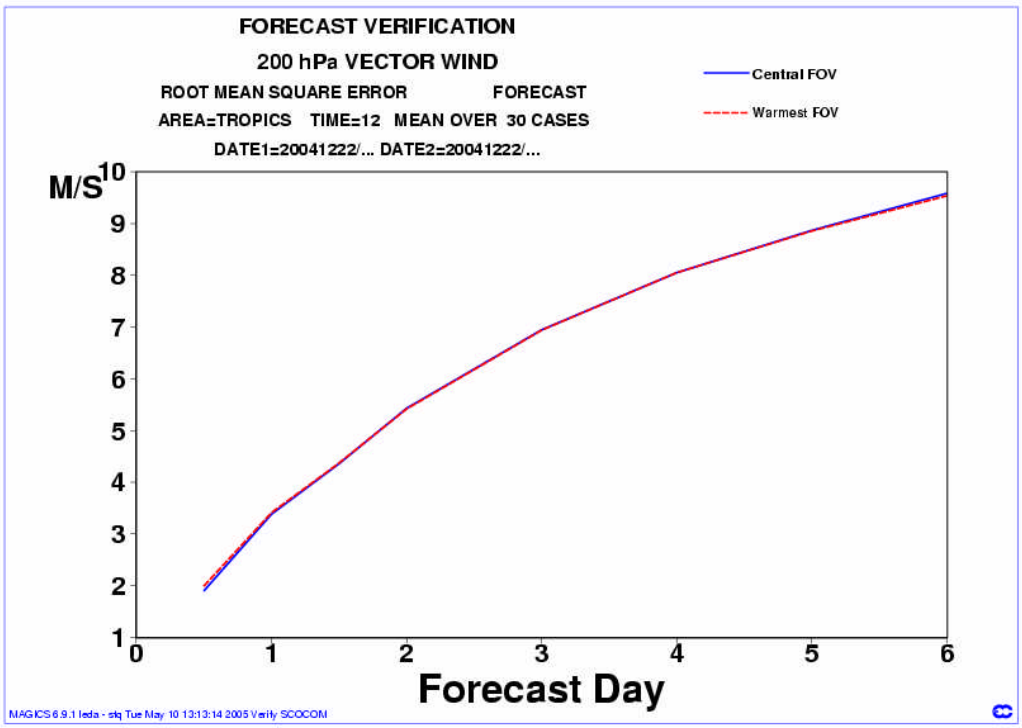


Fig. 6: Tropical 200hPa vector winds RMS forecast errors.

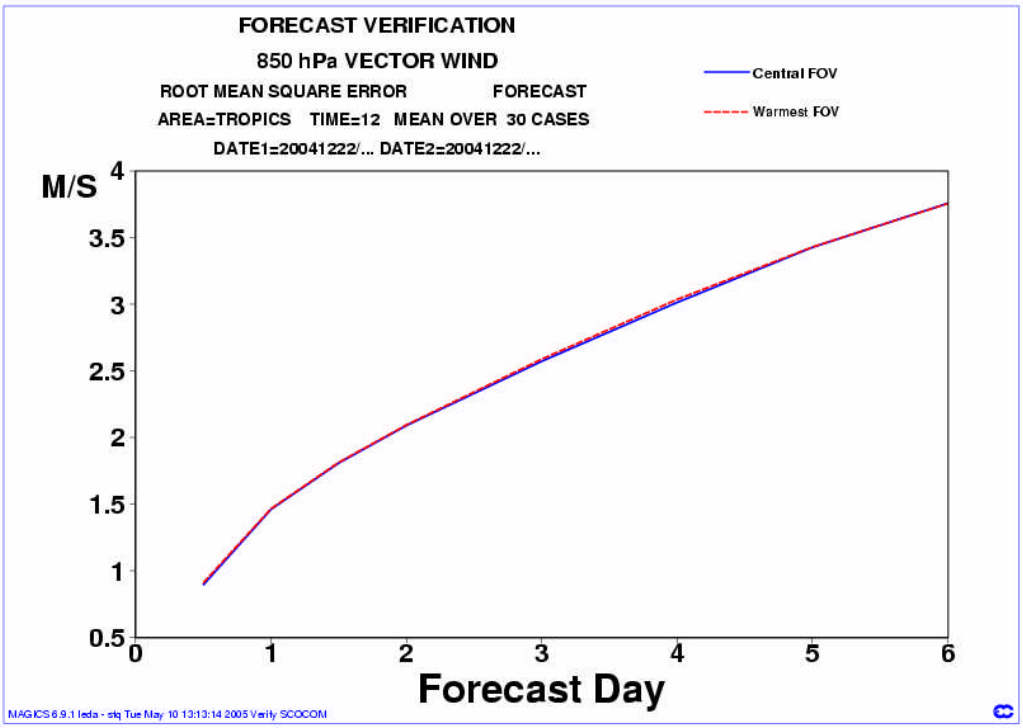


Fig. 7: Tropical 850hPa vector winds RMS forecast errors

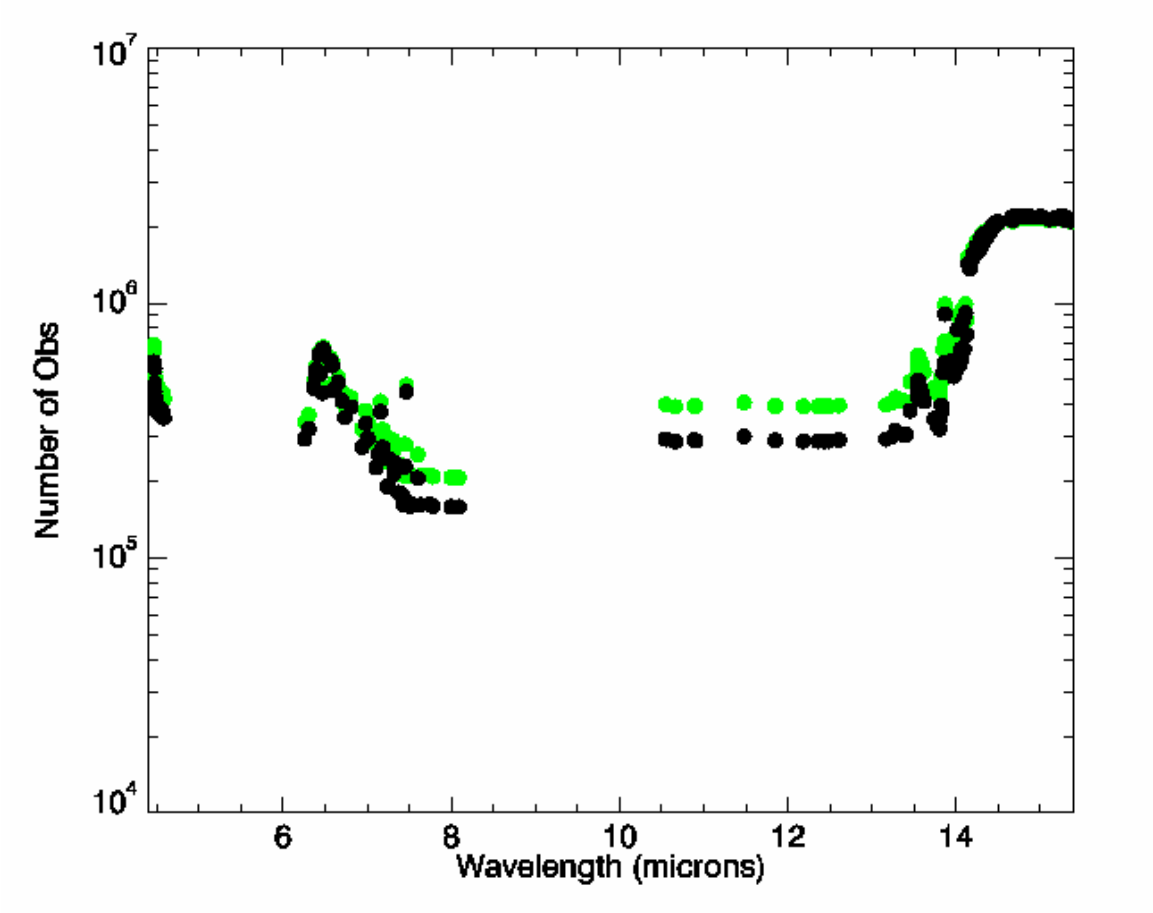


Fig. 8: Change in the number of used observations globally for the second experiment. The black dots are for the central field of view dataset and the green dots are for the warmest FOV dataset. Note the logarithmic scale.

Use of Reconstructed Radiances

Reconstructed radiances (Antonelli *et al.*, 2004) use prior knowledge of the variability of the observations to optimally smooth the AIRS spectrum so that the information from the entire spectrum can be best represented by a subset of channels. Thus one might hope that more of the information from the full 2378 channel AIRS spectrum can be represented in the 157 channels used for assimilation.

Reconstructed radiances are formed through the evaluation of the amplitudes, **p**, of the principal components, **L,** of the observed spectrum. Here **L** is the set of N_p leading eigenvectors of the covariance matrix of a set of thousands of observations. **p** is related to **y** (the noise normalized observed radiances with the mean observation subtracted) through.

$$\mathbf{p} = \mathbf{L}^T\mathbf{y}$$

The reconstructed radiance $\tilde{\mathbf{y}}$ is then calculated from

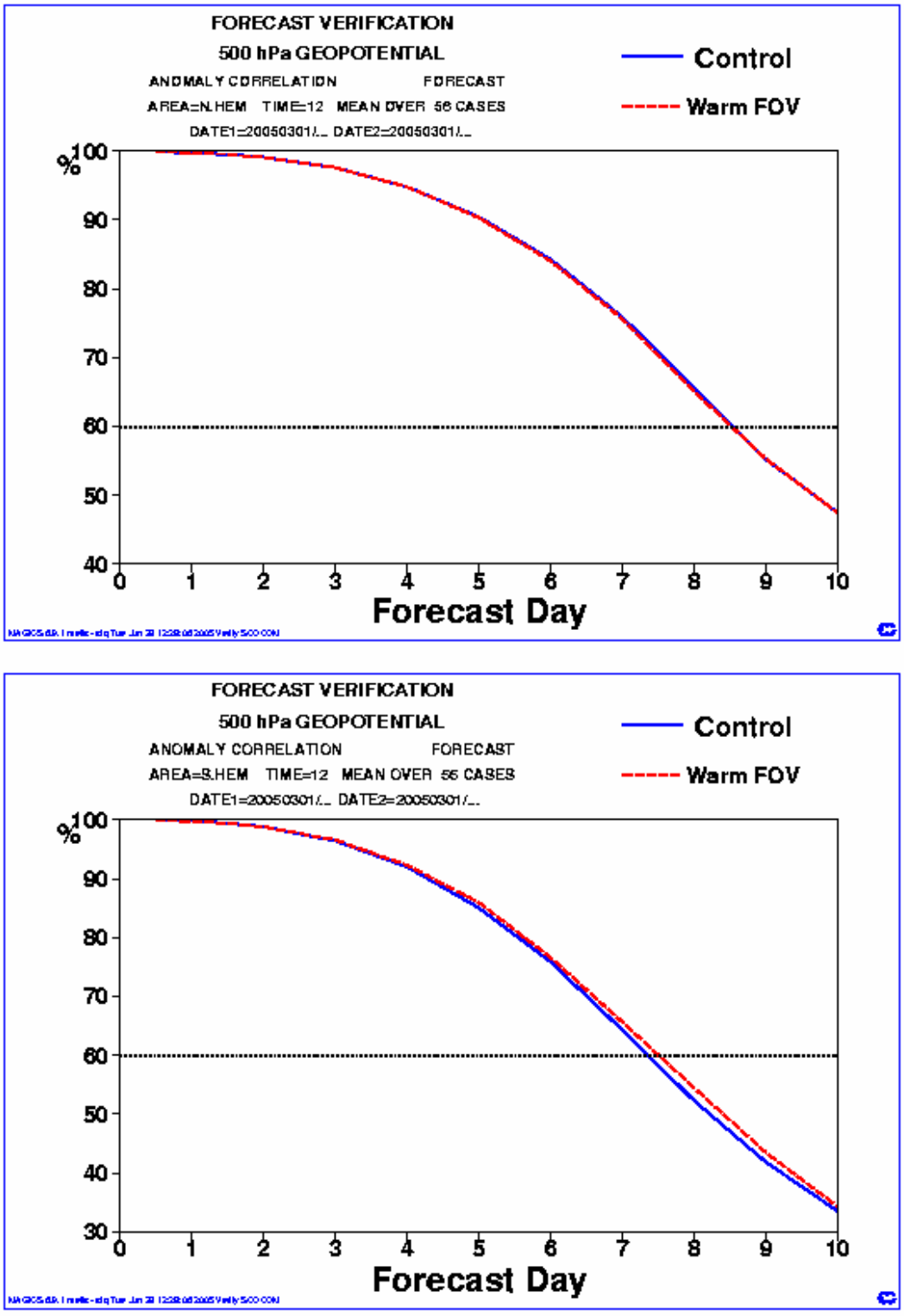


Fig. 9: Northern and southern hemisphere extra-tropical 500hPa anomaly correlation forecast scores for the second experiment.

$$\tilde{\mathbf{y}} = \mathbf{Lp} = \mathbf{LL}^T\mathbf{y}$$

If we restrict $\tilde{\mathbf{y}}$ to a subset of channels, N_R, by replacing the first $\mathbf{L}$ above with $\mathbf{L}_{NR}$, those channels will contain all of the information present in the N_p principal components provided $\mathbf{L}_{NR}$ has $\geq N_p$ positive singular values. The minimum criterion for this is that $N_R \geq N_p$ and in practice this criterion is usually sufficient.

Reconstructed radiances are produced routinely in near real time at NOAA/NESDIS from the 200 leading principal components of the full AIRS spectrum. The impact on forecast scores of using these reconstructed radiances, while keeping all other parameters fixed, has been found to be close to neutral (although observed minus calculated radiance standard deviations are greatly reduced).

The observation error for the reconstructed radiances can be derived from the raw observational error through

$$\varepsilon_{\tilde{\mathbf{y}}} = \mathbf{LL}^T \varepsilon_{\mathbf{y}}$$

Therefore

$$\mathrm{Cov}(\varepsilon_{\tilde{\mathbf{y}}}) = E\{\varepsilon_{\tilde{\mathbf{y}}} \varepsilon_{\tilde{\mathbf{y}}}^T\} = \mathbf{LL}^T \mathbf{O}\, \mathbf{L}^T\mathbf{L} = \mathbf{LL}^T\mathbf{L}^T\mathbf{L}$$

as the radiances are noise normalized and therefore $\mathbf{O}=\mathbf{I}$.

Using the 200 PCs distributed in near-real-time by NOAA/NESDIS, the reconstructed radiances' correlation matrix for the 157 channels used for assimilation at ECMWF is shown in Figure 10. The standard deviations of the reconstructed radiances are reduced, relative to the normal radiances, by a factor of around 4, but these errors are now correlated between channels.

Figure 11 shows an example linear analysis of retrieval errors illustrating the effect of assuming the wrong observational error covariance matrix for the 157 channels used at ECMWF. The background error is shown by the red line, while the retrieval using the normal (unreconstructed) radiances is shown in black. The cyan curve show the expected retrieval error if the assumed errors are kept the same but reconstructed radiances are used — not a large difference — while the green curve shows that a much larger impact is obtained when the correct, correlated errors are assumed for the reconstructed radiances. Finally, if just the diagonals of the reconstructed radiances' covariance matrix are assumed, the result can even be a degradation (blue curve). Therefore, for the maximum impact of the reconstructed radiances to be obtained one must fully account for the correlations in the error covariance matrix.

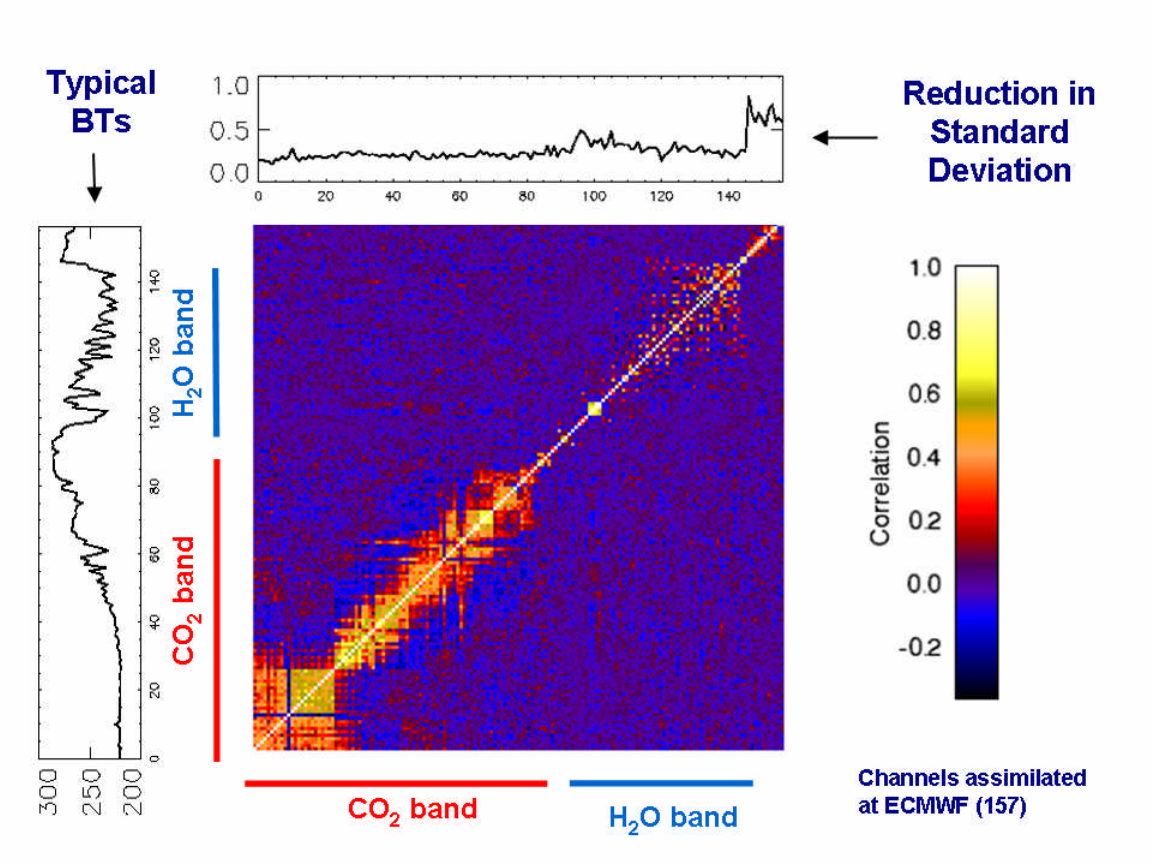


Fig. 10: The correlation matrix of the observational errors for the NOAA/NESDIS reconstructed radiances product for the 157 channels used in assimilation. Also shown are the reduction in standard deviation of the observational noise relative to the unreconstructed radiances and a sample spectrum to show the channels being used.

The correlated observation error is not the only source of correlated error. In particular forward model error is highly correlated. To date, a diagonal error covariance matrix has been assumed when assimilating AIRS (and all other satellite) radiances with error inflation to partly allow for correlations.

Some initial assimilation experiments have been run with reconstructed radiances and a correlated error covariance matrix. The most interesting results occurred when the original, conservative diagonal errors were used but the reconstructed radiances' correlation matrix was employed. While the differences between background, analysis and observations were largely unchanged, there were a number of cases in the analysis increments where they changed markedly. An example is shown in Figure 12.

The forecast errors that result from the use of the correlated errors are improved relative to the use of reconstructed radiances without correlated errors. While the above configuration is far from optimal it illustrates the fact that the introduction of correlation can improve the assimilation of AIRS radiances. The next step is to include the correlated forward model error.

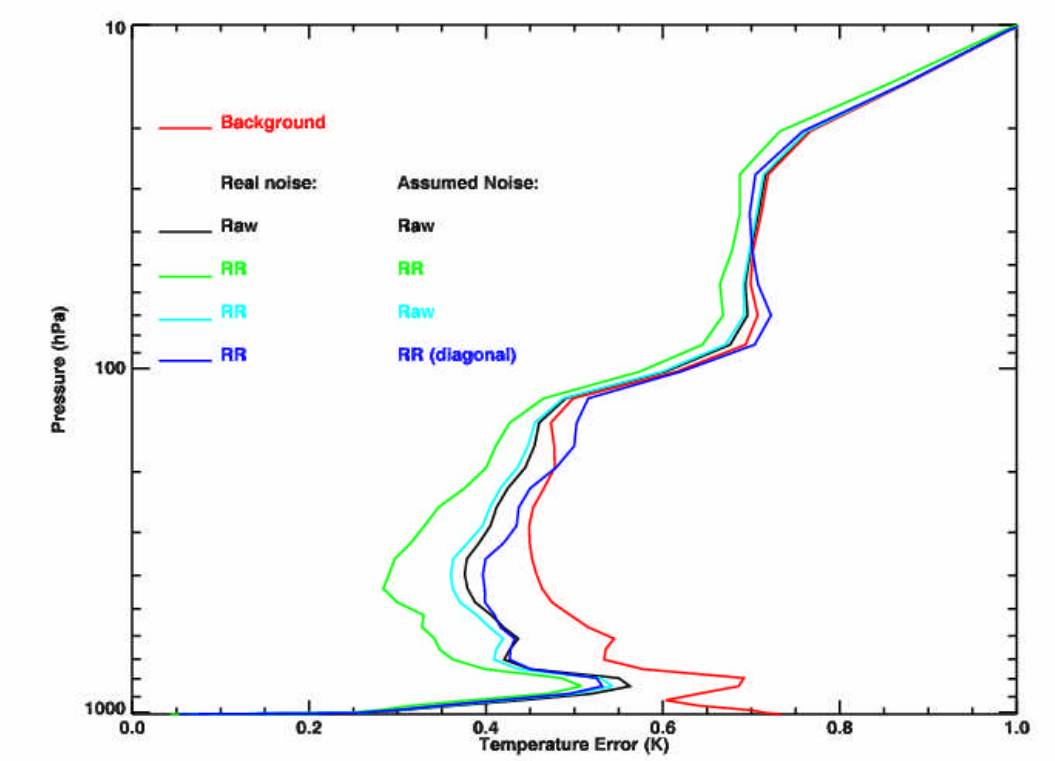


Fig. 11: A comparison of the expected temperature retrieval errors on using normal and reconstructed radiances and various assumptions for the observational error covariance. See text for a full explanation.

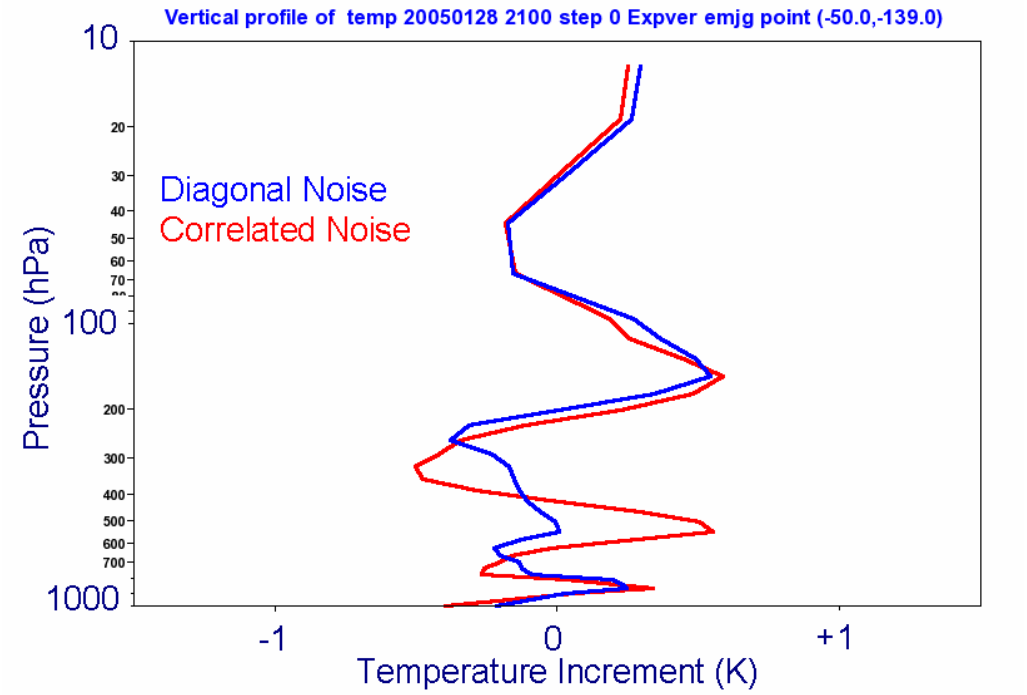
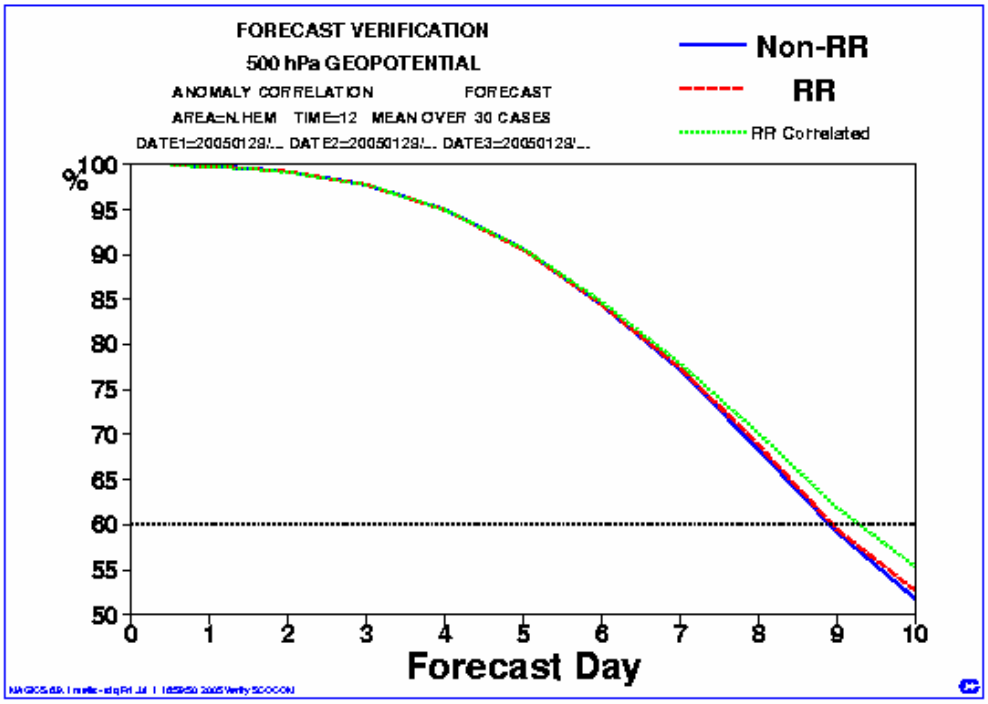


Fig. 12: An example change in analysis increment when assimilating reconstructed radiances on changing from a diagonal noise assumption to correlated noise.



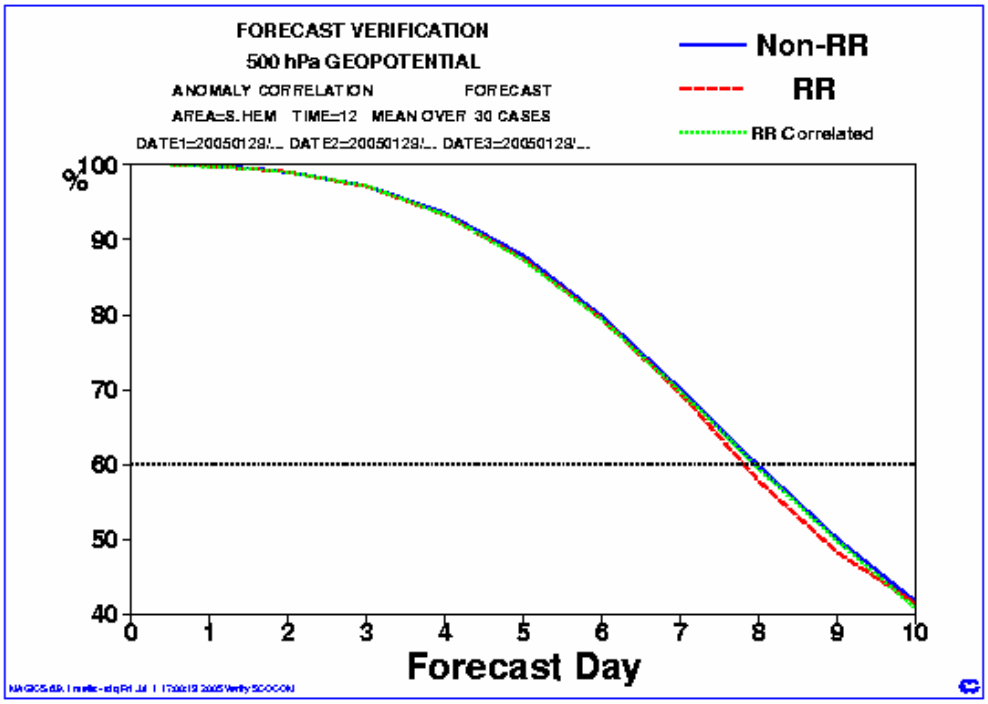


Fig. 13: Northern and southern hemisphere extra-tropical 500hPa anomaly correlation forecast scores on assimilating reconstructed radiances with and without assumed correlated observational errors. The introduction of correlation seems to be resulting in positive impact.

Summary

The use of the Warmest FOV AIRS dataset greatly increases the number of used observations in the lower troposphere. A marked improvement in forecast skill is seen as a result. It is expected that this dataset should become the operational stream distributed to European NWP centres around August 2005.

The use of correlated observational error in the context of reconstructed radiances has been investigated with promising results.

IASI data will be provided in near real time at full spatial and spectral resolution to European NWP centres. This will allow the spatial thinning and the calculation of reconstructed radiances to be performed at the NWP centres themselves, if desired. Optimal strategies for the communication, use and storage of IASI data will be developed partially based on AIRS experience.

Acknowledgments

The authors acknowledge Sean Healy and Jean-Noel Thépaut for useful discussions.

Reference

Antonelli, P., H. E. Revercomb, L. A. Sromovsky, W. L. Smith, R. O. Knuteson, D. C. Tobin, R. K. Garcia, H. B. Howell, H.-L. Huang, and F. A. Best (2004), A principal component noise filter for high spectral resolution infrared measurements, *J. Geophys. Res.*, **109**, D23102-23124.