The EUMETSAT Network of Satellite Application Facilities

NWP SAF
Numerical Weather Prediction

Visiting Scientist mission NWP_VS11_02 report

Document NWPSAF-MO-VS-047

Version 1.0

20 July 2012

IASI Channel Selection for Reconstructed Radiances

A. D. Collard

IMSG at NOAA/NCEP/EMC

Met Office · ECMWF · Royal Netherlands Meteorological Institute, Ministry of Transport, Public Works and Water Management · METEO FRANCE Toujours un temps d'avance

IASI Channel Selection for Reconstructed Radiances

A. D. Collard, IMSG at NOAA/NCEP/EMC

This documentation was developed within the context of the EUMETSAT Satellite Application Facility on Numerical Weather Prediction (NWP SAF), under the Cooperation Agreement dated 1 December, 2006, between EUMETSAT and the Met Office, UK. The partners in the NWP SAF are the Met Office, ECMWF, KNMI and Météo France.

Change record			
Version	Date	Author / changed by	Remarks
0.1	Sept 2011	A.D.Collard	Draft version for comment
1.0	20.7.12	A.D.Collard	Final version, incorporating comments

Abstract.

Considerations for channel selection of advanced sounder reconstructed radiances are discussed.

The Rodgers information-content based channel selection algorithm is applied to the reconstructed radiance problem using the approximation of a diagonal observation error covariance matrix. It is found that this over-estimates the performance of the channel selection as in reality the observation error covariance is not diagonal. If one assumes a diagonal error covariance in the analysis the true information content can actually be reduced with additional channels.

If one applies additional constraints on the allowed inter-channel correlations then the sensitivity of the solution to the correlations is reduced.

The effect of these various selection on the condition number of the resulting observation error covariance matrix is also discussed.

Introduction.

The current IASI channel selection for NWP (Collard, 2007) assumes the level 1C IASI noise characteristics. If we wish to assimilate IASI reconstructed radiances rather than raw radiances we should account for the different noise characteristics of the transformed observations in our channel selection.

In this brief report we explore the issues involved with choosing a subset of channels in this context and suggest some possible solutions. The main criterion is that the channel selection should retain as much information as possible while ensuring that the observation error covariance matrix is well conditioned.

One initial rationale for this study was the need to define a set of 300-500 channels to be distributed in near real time if reconstructed radiances are to be distributed instead of raw radiances. However, it has become clear that the choice of channels to be assimilated will potentially be highly dependent on the data assimilation system being used (particularly the assumed background errors and the treatment of clouds).

The most efficient method of the distribution of reconstructed radiances is through distribution of the leading principal component scores and any other method will reduce the users' flexibility.

Reconstructed Radiances Theory.

The theory behind reconstructed radiances has been described in a number of publications (E.g., Antonelli *et al.*, 2004; Collard *et al.*, 2010) and is discussed briefly here.

The basic idea is that the leading eigenvectors, $\mathbf{L}$, of a covariance matrix describing the climatological variation of the instrument's spectrum are used to calculate a number of principal component amplitudes $\mathbf{p}$ which describe the vast majority of the atmospheric information contained in the spectrum (the trailing eigenvectors will contain mostly noise).

These leading eigenvectors can then be used to reconstruct the spectrum, $\tilde{\mathbf{y}}$ so that the noise that was contained in the rejected eigenvectors is filtered out:

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

where $\mathbf{y}$ is the original radiance spectrum. The instrument noise that results from propagating the original noise, $\mathbf{R}$ into reconstructed radiance space is therefore $\mathbf{L}\mathbf{L}^{\mathbf{T}}\mathbf{R}\mathbf{L}\mathbf{L}^{\mathbf{T}}$. In noise-normalised space (where $\mathbf{R} = \mathbf{I}$) the noise error covariance becomes $\mathbf{L}\mathbf{L}^{\mathbf{T}}$ (as $\mathbf{L}^{\mathbf{T}}\mathbf{L} = \mathbf{I}$).

The above transformation of the instrument noise requires the same transformation to be applied to the radiance operator. I.e., if the signals corresponding to the trailing eigenvectors are removed, the same eigenvectors should be removed from the forward model calculation to be consistent. If this is not done, one is effectively considering the reconstructed radiances to be a *retrieval* of the true radiances with an error given by $\mathbf{L}\mathbf{L}^{\mathbf{T}} + \mathbf{F}_{\mathbf{R}}$ where $\mathbf{F}_{\mathbf{R}}$ is the additional reconstruction error due to signals in the true spectrum being removed by the eigenvector truncation. For well-chosen eigenvectors this $\mathbf{F}_{\mathbf{R}}$ term is small, but it may become significant if one tries to extract maximum information content from the reconstructed spectrum in retrieval space. Experience has shown that the correct treatment of this additional $\mathbf{F}_{\mathbf{R}}$ term requires great care in the calculation and interpretation of results and will complicate this study more than is necessary, so throughout this report the radiance operator is transformed into reconstructed radiance space.

Two more additional error terms should be included when these observations are used for assimilation or retrievals. These are due to forward model and representivity errors. Often in studies such as these these errors are assumed to be uncorrelated between channels with a standard deviation of 0.2K. In reality forward model and representivity error will tend to be correlated between channels, but the exact form of this error is hard to estimate accurately. Therefore in this study these error terms will generally not be considered, which also allows us to focus on how the reconstructed radiances transform affects the properties of the instrument error covariance matrix. In any practical system, however, one should bear in mind that it may not be the instrument noise that limits performance.

It is important to note that even though the diagonal of the error covariance matrix is greatly reduced by the reconstructed radiances process (Fig. 1), the resulting error covariance matrix is highly rank-deficient (it can only have a rank as large as the number of eigenvectors used to construct it) and thus there is significant inter-channel correlation (Fig. 2)

Another way of thinking about this is that each channel in the reconstructed spectrum is a linear combination of all the channels of the original spectrum. We are effectively beating down the noise by co-adding channels at the expense of the independence of each channel's measurement. It is basically an intelligent smoothing process.

For this study we have focused on the principal components produced at EUMETSAT by Tim Hultberg for Band 1 of the IASI instrument. Band 1 comprises the first 1997 channels of the instrument (645-944cm^{-1}) and 90 principal components have been found to be adequate to represent atmospheric signal. The first ten principal components are shown in Fig. 3.

One property of reconstructed radiances is that if n reconstructed radiances are calculated using m PCs, $\mathbf{p}$, and $n \geq m$ such that $\tilde{\mathbf{y}}_n = \mathbf{L_n p}$ it is usually the case that $\mathbf{p}$ can be calculated exactly from $\tilde{\mathbf{y}}_n$ (provided $\mathbf{L_n}$ has a generalised inverse and more than m singular vectors, of course). Therefore the task of finding a subset of m or more channels that contain all the available information is effectively trivial - most such combinations do. The problem lies with ensuring that the channel selection chosen has an observation error covariance matrix that is not only theoretically invertible, but is well-conditioned and, ideally, close to diagonal.

As an extreme example, it can be demonstrated experimentally that if one produces reconstructed radiances from the first 200 principal components of the full spectrum (not just Band 1), the complete reconstructed spectrum can be calculated to approaching machine precision from just the first 200 channels and $\mathbf{L_n L^T}$. Therefore the channels around 650cm^{-1} which measure stratospheric temperatures also contain information on the solar contribution to the near-surface channels around 2600cm^{-1}. The amplitudes of these signals are, of course, very small and are contained in subtle correlations between channels that can only be used in retrievals if a precisely-known observation error covariance were used. We would not in practice want to use simply these first 200 channels in our channel selection but would aim to choose a set of channels where the extraction of information can be done in a more robust manner.

Channel Selection Theory.

The IASI channel selection scheme has been detailed by Collard (2007). It is based on a method originally proposed by Rodgers (1996, 2000) where channels are successively chosen based on their ability to improve the information content of the retrieval. In doing so, the Rodgers methodology takes into account the instrument noise for each channel as well as the channels' Jacobians.

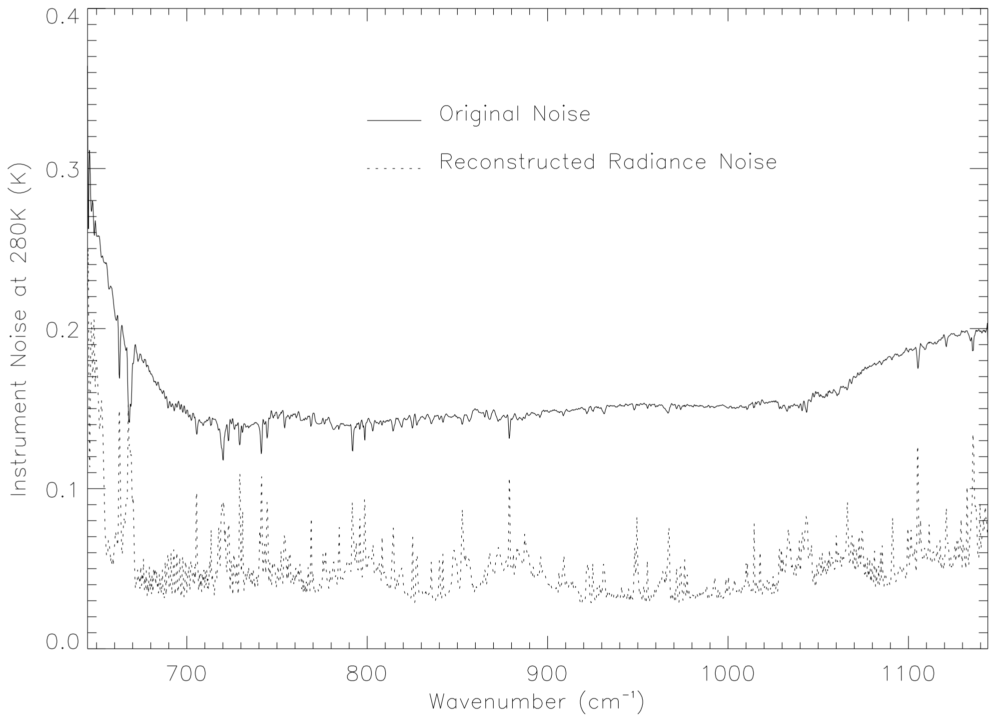

Fig. 1: The diagonal of the level 1c IASI instrument noise error covariance matrix for original and reconstructed radiances. Band 1 (645–944 cm^{-1}) only.

The figure-of-merit used to measure the increase in information content is the degrees of freedom by signal (DFS) which may be calculated from

$$\mathrm{DFS} = Tr(\mathbf{I} - \mathbf{A}\mathbf{B}^{-1})$$

The Rodgers technique uses an efficient method to calculate the DFS by incrementing the previous value but this requires the assumption that the channels' instrument noise is uncorrelated. The method can be modified to account for error correlations explicitly, but this involves multiple matrix inversions and typical run times will be in the days. This option has not been explored as part of this project.

IASI level 1c instrument noise is highly correlated between adjacent channels due to the apodisation function and so during the original channel selection it was forbidden to pick adjacent channels to ensure that the instrument noise covariance matrix of the selected channels was as close to diagonal as possible. It it therefore proposed that one way to pick channels in the reconstructed radiances case is to ensure that inter-channel correlations are kept to a minimum.

The resulting channel selection will still result in an error covariance matrix with correlated errors. If, as seems likely, we might still wish to use a diagonal matrix in assimilation, it would be instructive to investigate the effect of this approximation on retrieval performance.

The error covariance, $\mathbf{A}$, of an optimal retrieval of an atmospheric state $\mathbf{x}$ is given by (Rogers, 1976; Rogers, 2000)

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

where $\mathbf{B}$ is the background error covariance matrix; $\mathbf{R}$ is the observation error covariance matrix and $\mathbf{H} = \nabla_x\mathbf{y}$ is the Jacobian matrix for the observations.

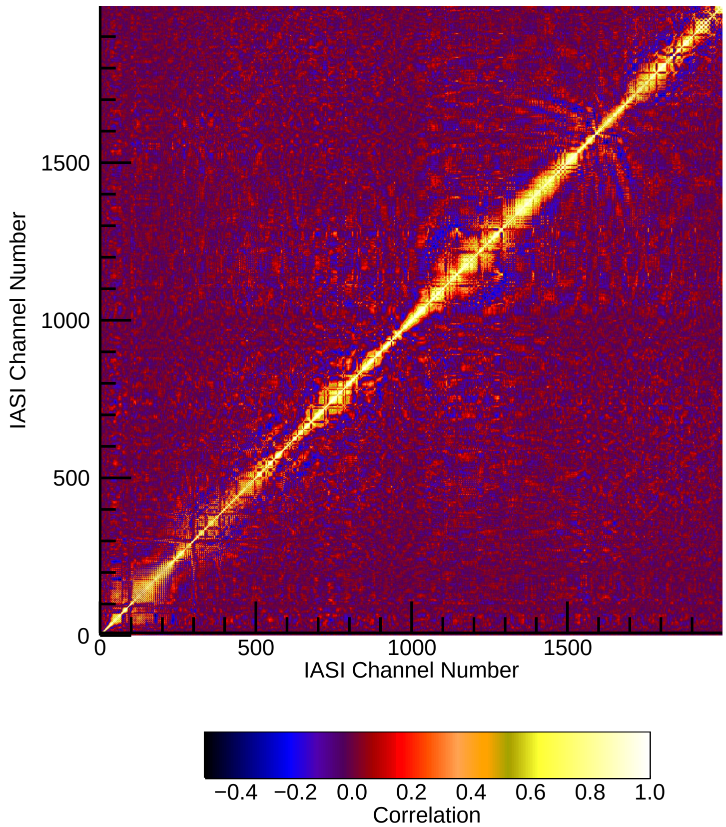

Fig. 2: The noise error correlation matrix for IASI reconstructed radiances. Band 1 (645-944cm^{-1}) only.

It can be shown (Watts and McNally, 1988) that if the wrong $\mathbf{R}$-matrix is assumed, this optimal solution is modified thus:

$$\mathbf{A}_{\text{true}} = \mathbf{A}_{\text{calc}} + \mathbf{W}(\mathbf{R}_{\text{true}} - \mathbf{R}_{\text{assumed}})\mathbf{W}^T$$

where $\mathbf{A}_{\text{calc}}$ is the analysis error covariance matrix calculated with the assumed $\mathbf{R}$-matrix and $\mathbf{W}$ is the Kalman gain matrix, $\mathbf{BH}^T(\mathbf{HBH}^T + \mathbf{R}_{\text{assumed}})^{-1}$.

If we work in noise-normalised radiance space (as us the case with the Rogers algorithm), the Jacobian becomes $\mathbf{H}^* = \mathbf{R}_{\text{assumed}}^{-1/2}\mathbf{H}$ and the instrument noise error covariances become $\mathbf{R}^*_{\text{assumed}} = \mathbf{I}$ and $\mathbf{R}^*_{\text{true}} = \mathbf{R}_{\text{assumed}}^{-1/2}\mathbf{R}_{\text{true}}\mathbf{R}_{\text{assumed}}^{-1/2}$. If $\mathbf{R}_{\text{assumed}}$ is simply the diagonal of $\mathbf{R}_{\text{true}}$ which has the correlation matrix $\mathbf{C}_{\text{true}}$ then

$$\mathbf{R}^*_{\text{true}} = \mathbf{R}_{\text{assumed}}^{-1/2}\mathbf{R}_{\text{assumed}}^{1/2}\mathbf{C}_{\text{true}}\mathbf{R}_{\text{assumed}}^{1/2}\mathbf{R}_{\text{assumed}}^{-1/2} = \mathbf{C}_{\text{true}}$$

.

Then $\mathbf{W}^* = \mathbf{BH}^{*T}(\mathbf{H}^*\mathbf{BH}^{*T} + \mathbf{I})^{-1}$ and

$$\mathbf{A}_{\text{true}} = \mathbf{A}_{\text{calc}} + \mathbf{W}^*(\mathbf{C}_{\text{true}} - \mathbf{I})\mathbf{W}^{*T}$$

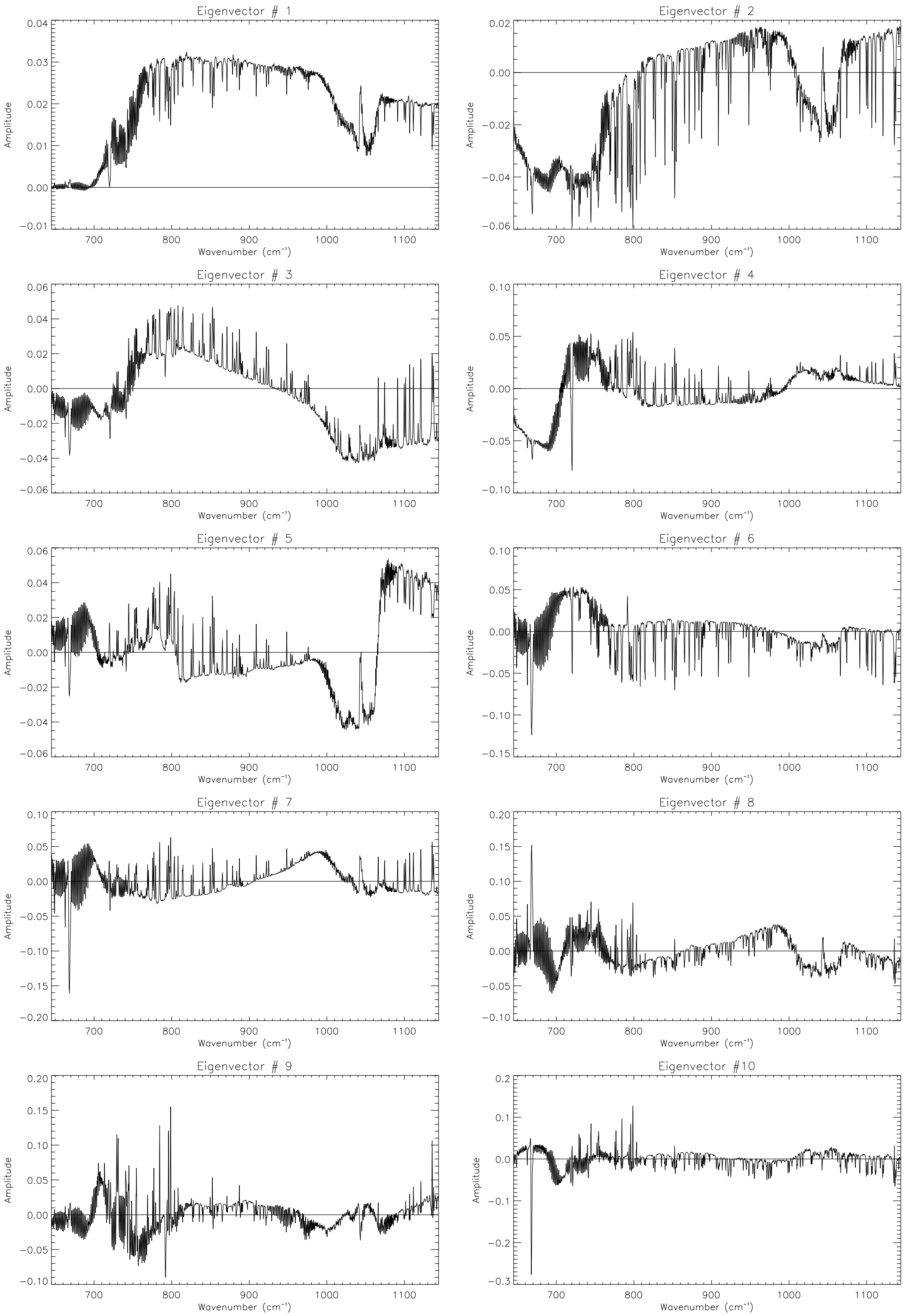

*Fig. 3: The first ten eigenvectors for IASI Band 1 (645-944*cm^{-1}).

Channel Selection Examples.

In the examples in this report, as this is for illustrative purposes, a somewhat simplified scenario is used compared to the full channel selection reported in Collard (2007). Only Band 1 is considered and there is no pre-screening for channels with un-modelled interfering species. Nor is there any attempt to filter out the effect of water vapour on the calculated temperature sounding performance. Also the result of apodisation is ignored resulting in a diagonal error covariance matrix for the original radiances. The B-matrix is the same one used in Collard (2007) and was derived from the ECMWF J_b statistics.

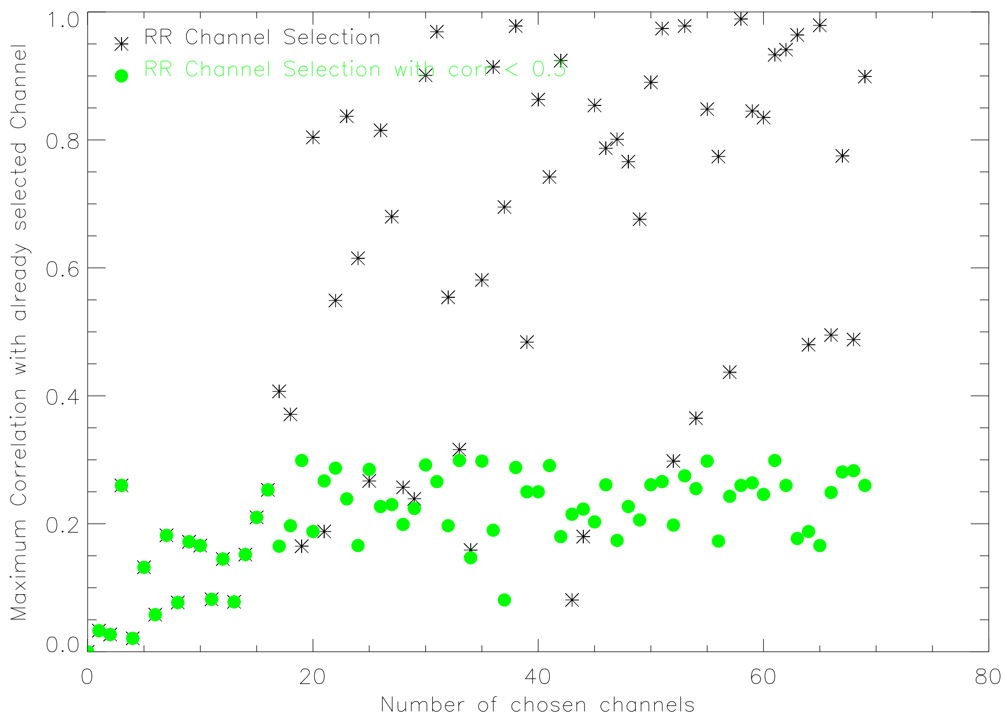

Fig. 4: The evolution of the maximum inter-channel error correlation for the reconstructed radiances channel selection with and without the imposition of a maximum value of 0.3.

As with Collard (2007), the six AFGL standard atmospheres are used and the channel selection is done such that the channel that best improves the DFS for all the profiles simultaneously is chosen.

In each of the cases below, 70 channels are chosen.

Three channel selections were attempted. In each case the errors used in the DFS calculations were assumed diagonal.

- The first just used the original radiance properties, with the IASI level 1-c noise.
- The second channel selection used the noise values for reconstructed radiances but did not account for correlations.

Figure 4 shows for this channel selection the maximum error correlation of each new channel with the channels already chosen. It can be seen that a large number of channels are chosen where this maximum inter-channel correlation is 0.8 or higher. This can be interpreted as the channel selection

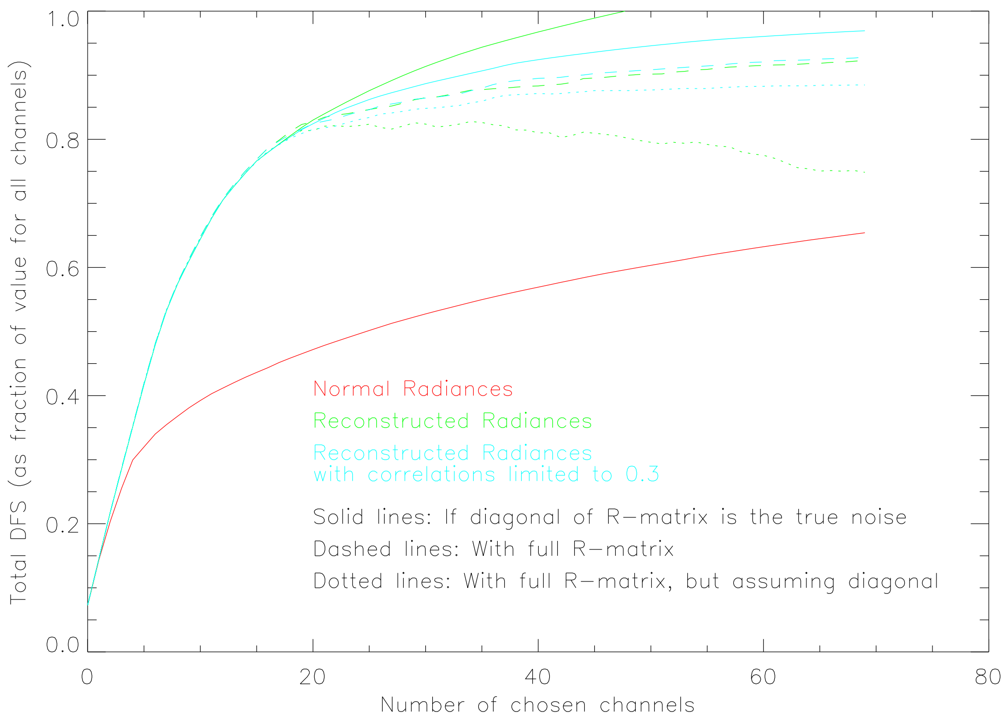

Fig. 5: The evolution of the degrees of freedom for signal for the three scenarios shown in Figure 4. All three solid lines assume the error covariance matrix is diagonal, which is a particularly poor assumption for the green curve where correlations are significant.

algorithm — which is assuming uncorrelated error between channels — choosing similar channels in order to reduce the random error. However, the benefit of doing this is greatly reduced by the fact that these channels are a similar combination of the raw radiances with highly correlated errors.

- In order to mitigate this effect a third selection is tried where the maximum inter-channel correlation is limited to 0.3.

The results of all three channels selections in terms of percentage of maximum DFS are shown in Figure 5. In each case, three calculations are shown. The DFS calculated based on the assumption that the error covariance is diagonal; the DFS calculated using the full error covariance matrix; and the DFS when the real instrument noise error covariance is correlated but a diagonal matrix is assumed. Note that each of the channel selections are performed using the Rodgers method and so the error covariance matrix is *assumed to be diagonal within the channel selection* but the chosen channels are then used to recalculate the DFS in the three ways described above.

It can be seen that the calculations done assuming a diagonal R-matrix are overly-optimistic when there is significant inter-channel error correlations. Indeed after twenty channels have been chosen, the inclusion of additional channels without properly allowing for inter-channels correlations actually reduces the information content. The inclusion of the requirement that inter-channel correlations are less than 0.3 results in a small decrease in the information content (when the correct R-matrix is used: dashed lines in Figure 5) but greatly reduces the degradation when a diagonal R-matrix is

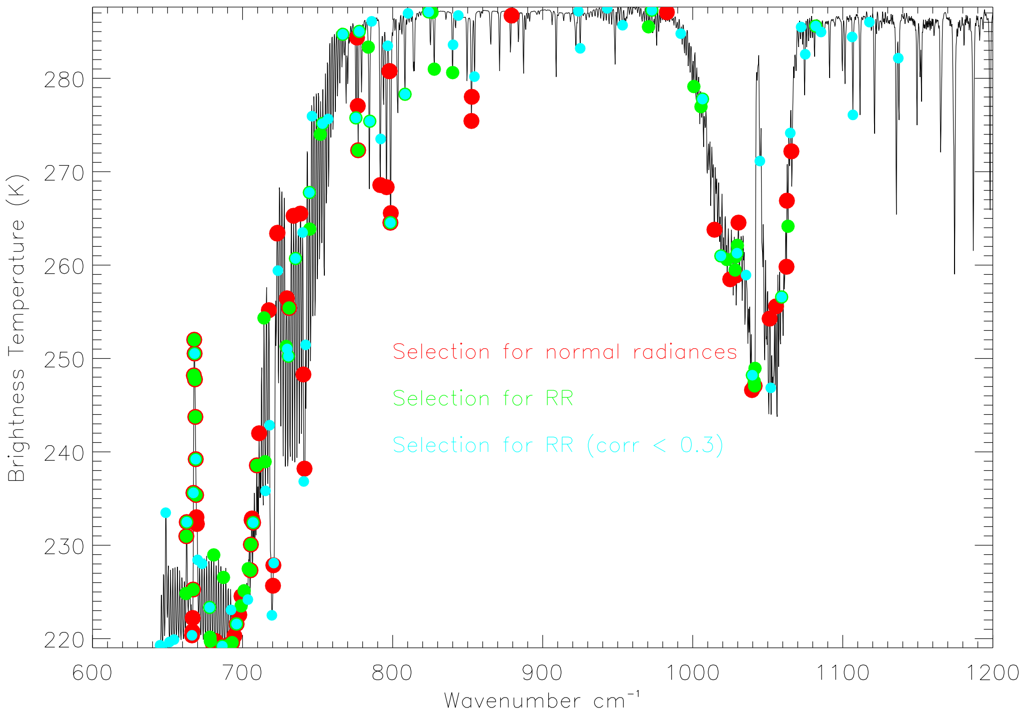

Fig. 6: A typical IASI spectrum upon which are plotted the selected channels for the three scenarios considered here. The red dots are for the normal channel selection, green is for reconstructed radiances and cyan is where the absolute inter-channel correlation is limited to be less than 0.3.

assumed.

In Figure 6, the three channel selections are plotted on a simulated IASI spectrum calculated for a clear-sky US Standard atmosphere. The first two channel selections are broadly the same. With a large number of channels selected between 680cm^{-1} and 740cm^{-1} where there are a number of very similar channels that are sensitive to the mid-to-upper troposphere and lower stratosphere. Because of the correlation restriction, the third selection is prevented from picking more than a few channels in this region and therefore chooses more channels that are sensitive to the lower atmosphere.

Figures 7 and 8 show the resulting observation error correlation matrices from the two reconstructed radiance selection strategies.

It is instructive at this point to investigate the condition numbers of the observation error covariances for these channel selections. These are shown in Figure 9. The **R**-matrix when there is no constraint on inter-channel correlations is poorly conditioned as the number of channels approaches the number of principal components. When the selection is limited to channels with correlations less than 0.3, the condition number is improved but is still around 100 when 70 channels are chosen.

This leads to the question as to whether this condition number can be further reduced. To answer this we start with a selection where we simply choose channels which increase the condition number of the **R**-matrix to the least degree. This will tend to choose channels that are not highly correlated and are therefore more independent.

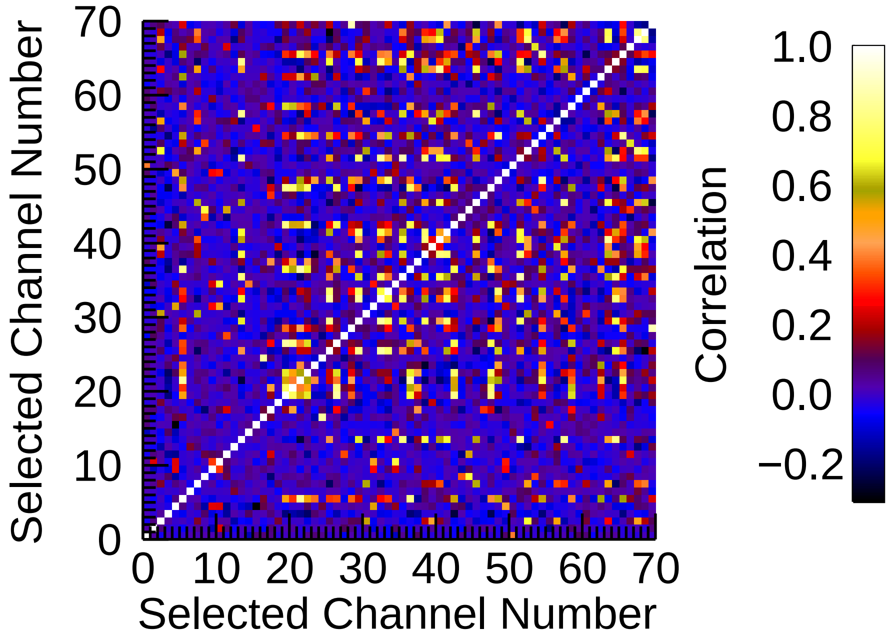

Fig. 7: The observation error correlation matrix for the 70 channels chosen in the reconstructed radiances channel selection. The channels are in the order in which they were chosen.

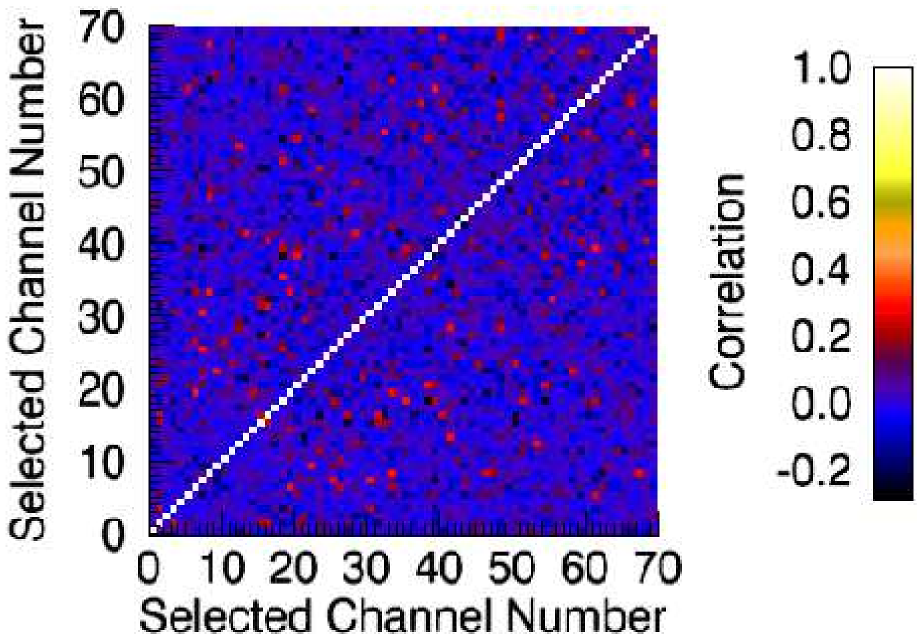

Fig. 8: The observation error correlation matrix for the 70 channels chosen in the reconstructed radiances channel selection when the inter-channel correlation is limited to 0.3.

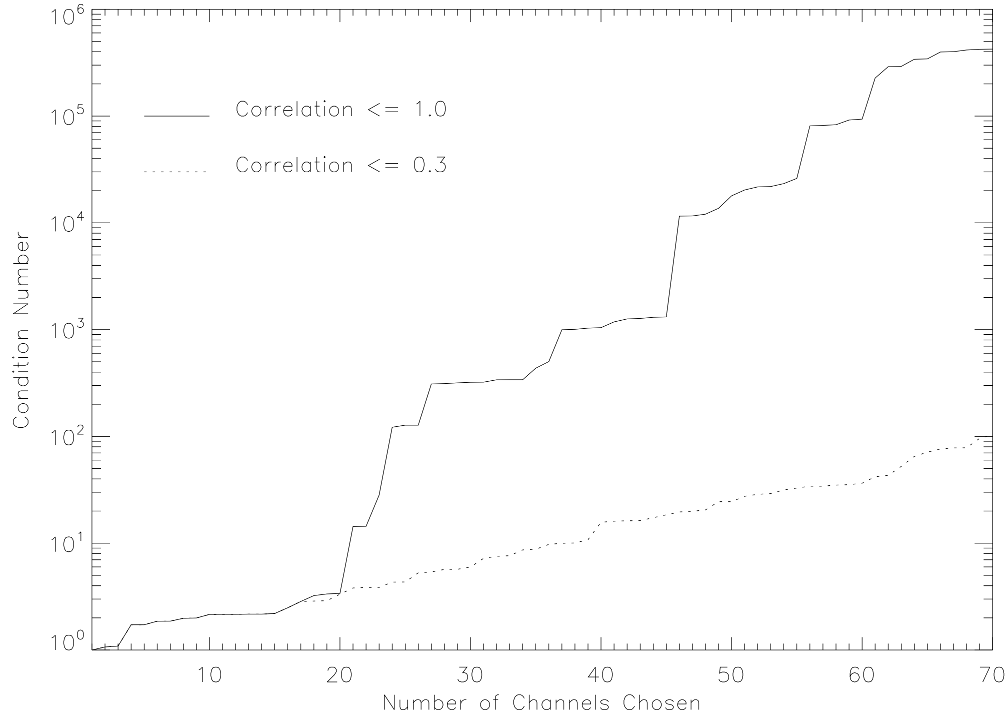

Fig. 9: The condition numbers of the observation error covariance matrix for the reconstructed radiance channel selections in Figure 5. The selection where correlations are limited to less than 0.3 are far better conditioned.

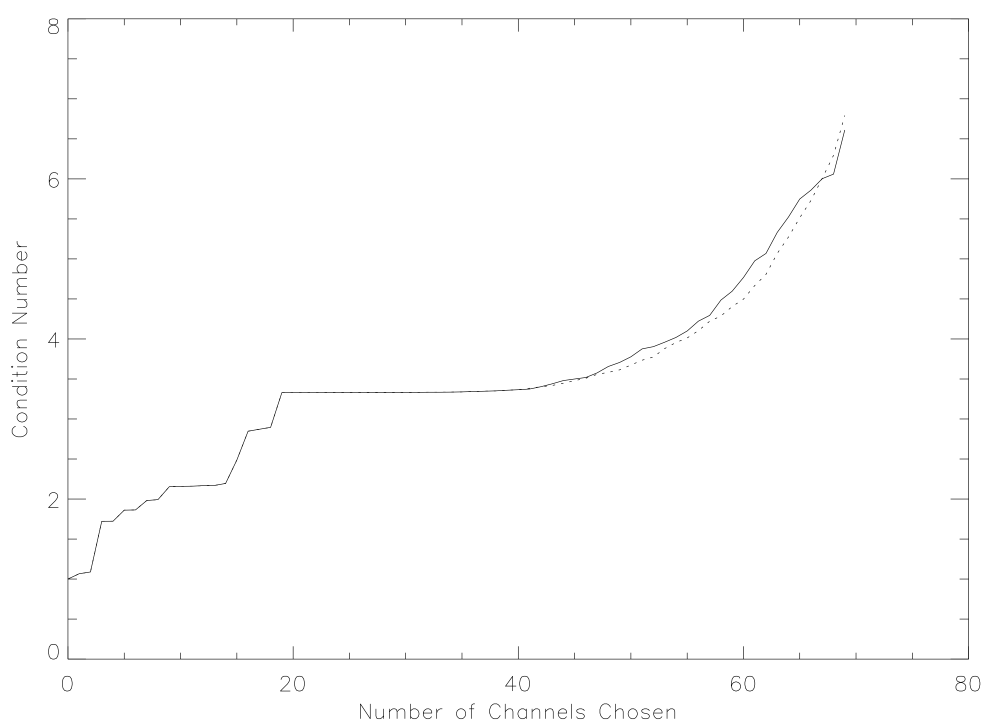

Fig. 10: Evolution of condition number when channels are chosen based on minimising this value. The dotted line includes the additional constraint that the maximum inter-channel correlation is limited to 0.3. The first 20 channels are taken from the Rodgers method selection where a correlation constraint of 0.3 is used.

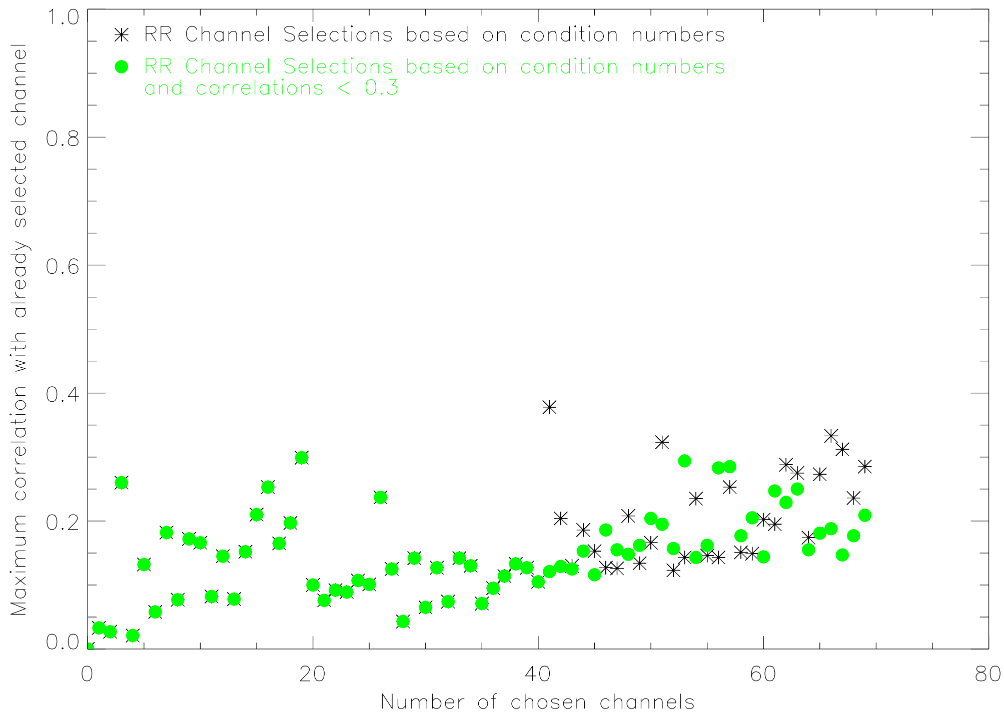

Fig. 11: Maximum correlation for each new channel for the condition-number based channel selection.

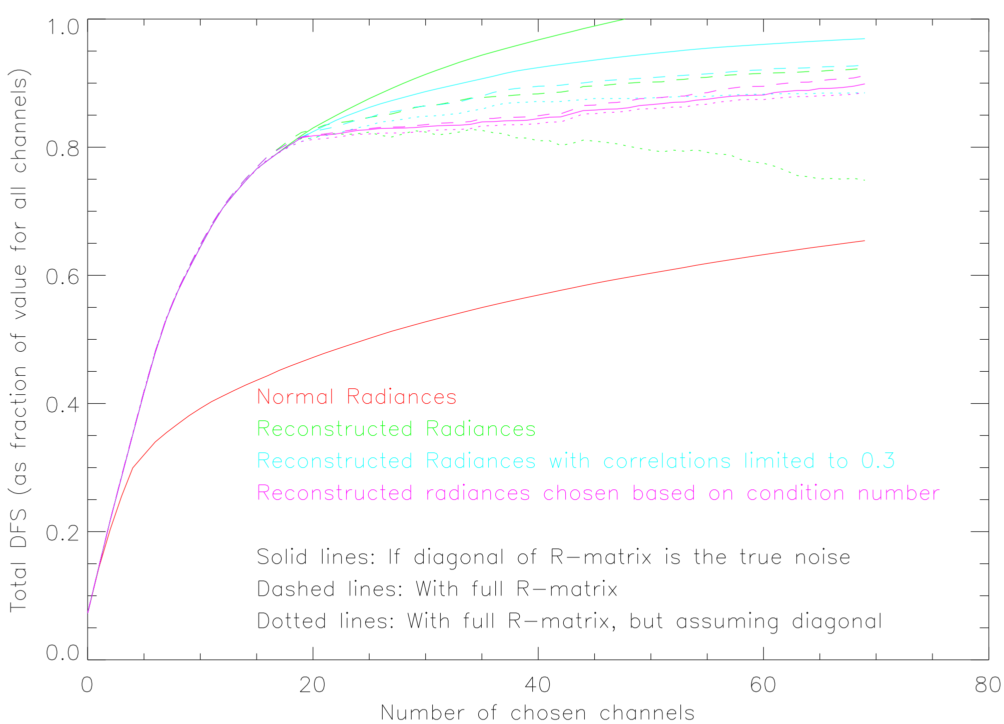

Fig. 12: As Figure 5 but the evolution of the degrees of freedom for signal for the condition number selection has been added. Although, perhaps unsurprisingly, the performance based simply on the condition number is poorer, its sensitivity to the three observation error assumptions is seen to be low.

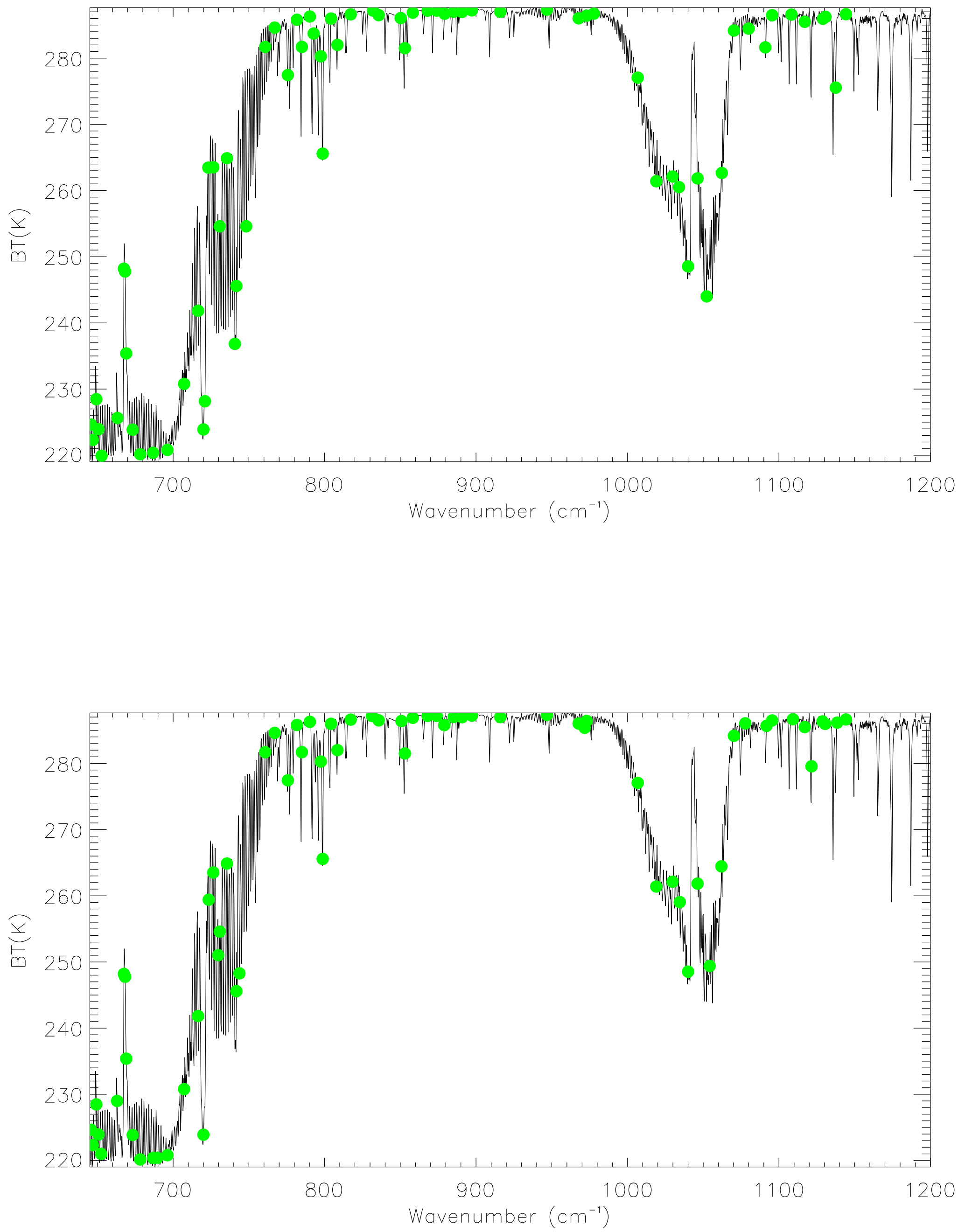

Fig. 13: Selected channels using a condition-number based channel selection. The top plot places no restriction on inter-channel error correlations, while the bottom plot restricts this to less than 0.3.

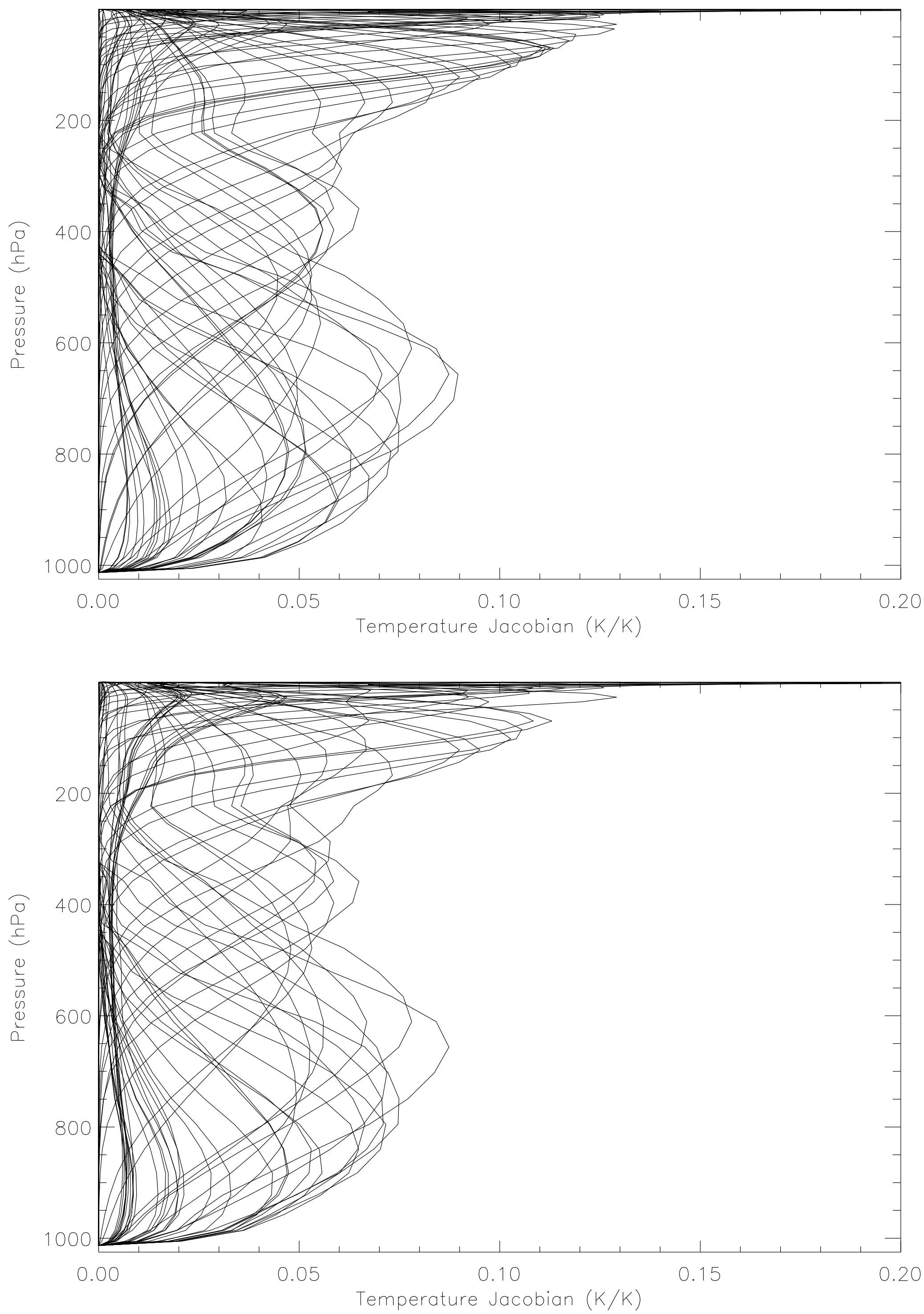

Fig. 14: The temperature Jacobians of the selected channels for the normal radiances (top) and reconstructed radiances (bottom).

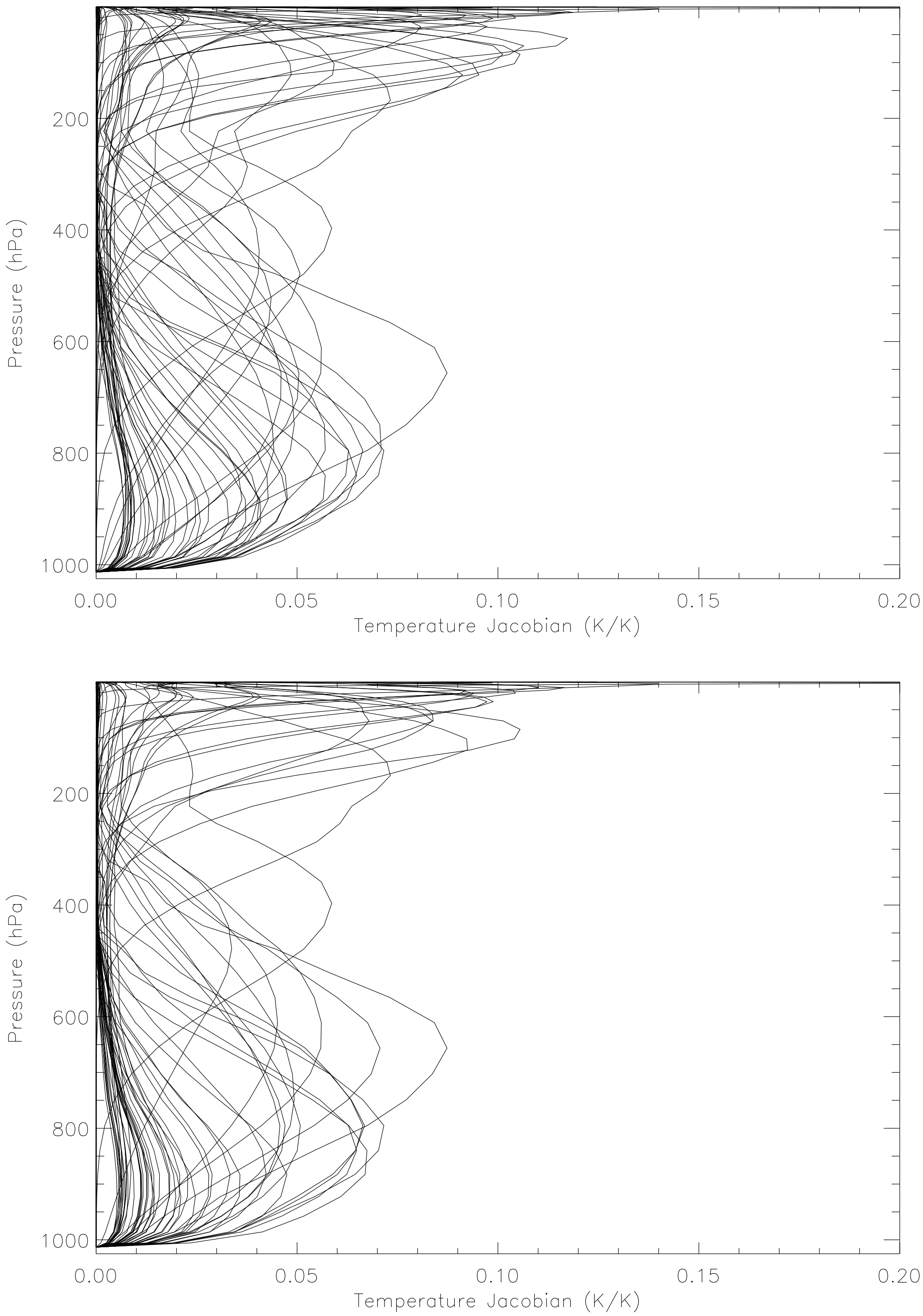

Fig. 15: The temperature Jacobians of the selected channels for the reconstructed radiances with correlations limited to less than 0.3 (top) and when chosen by condition number (bottom).

Figure 10 shows the evolution of the condition number when the channel selection is performed by sequentially choosing the channel that increases the condition number to the least degree. The first 20 channels of the Rodgers method channel selection with the 0.3 correlation limit are used before the correlation criterion is adopted. Also shown in Figure 10 (dotted line) is the evolution of the condition number when the maximum correlation of 0.3 constraint is applied. Interestingly, although the dotted line is initially just above the solid line as expected (at channel 42), the addition of this constraint actually produces better results for much of the rest of the plots. The correlation numbers for these two selections are shown in Figure 11.

When plotting the degrees of freedom for signal for this selection and comparing with the previous selections the condition number only approach does not perform as well. This is probably unsurprising, but it is worth noting that the results when using this selection are particularly insensitive to which of the three observation error scenarios are assumed — indicating that including the condition number in the Rodgers selection process may be worth pursuing.

It should also be noted that if we extend the process to choose 90 channels (the number of PCs) the condition number with this sequential method increases rapidly (to 79.8, not shown). This may indicate that it is not possible to fully represent the full information in the reconstructed radiances spectrum with a limited number of channels and still have a low condition numbers; or it may indicate that the sequential method suggested does not find the optimal solution.

Figure 13 shows the distribution of channels chosen using this approach. While Figures 14 and 15 show the temperature Jacobians corresponding to four of the channel selection approaches explored here.

It can be seen that these approaches tend to preferentially choose channels in the window and around the tropopause. The information in the middle troposphere is then gained through partially through the tails of these Jacobians. Obviously, this can cause problems if there are reasons not to use the window channels (e.g., avoiding low cloud or issues with ill-defined surface radiative properties) and further constraints on the chosen channels may be required.

Finally, Figure 16 shows these results in the context of the standard deviation of the retrieval error. When plotted in this way, the retrieval error is relatively insensitive to the choice of channels. This implies that the differences in the degrees of freedom by signal are in those structures that do not contribute significantly to the bulk RMS values (but which may still be important in the initialisation of forecast model runs) but also that we have the flexibilty to make channel selections based on considerations of numerical robustness without compromising the major component of retrieval performance.

Discussion and Conclusions

This report highlights some of the considerations when choosing a set of channels to assimilate using reconstructed radiances.

Using the channel selection technique of Rogers (2000), a reasonable selection of channels can be achieved (compared to the unreconstructed case) even if diagonal noise is assumed. However, if one assumes diagonal observation error covariance matrices, the effect of adding more channels can be negative once correlations start to be significant (after choosing 20 channels in our case).

Adding a restriction to the allowed correlation improves the robustness of the selection when diagonal

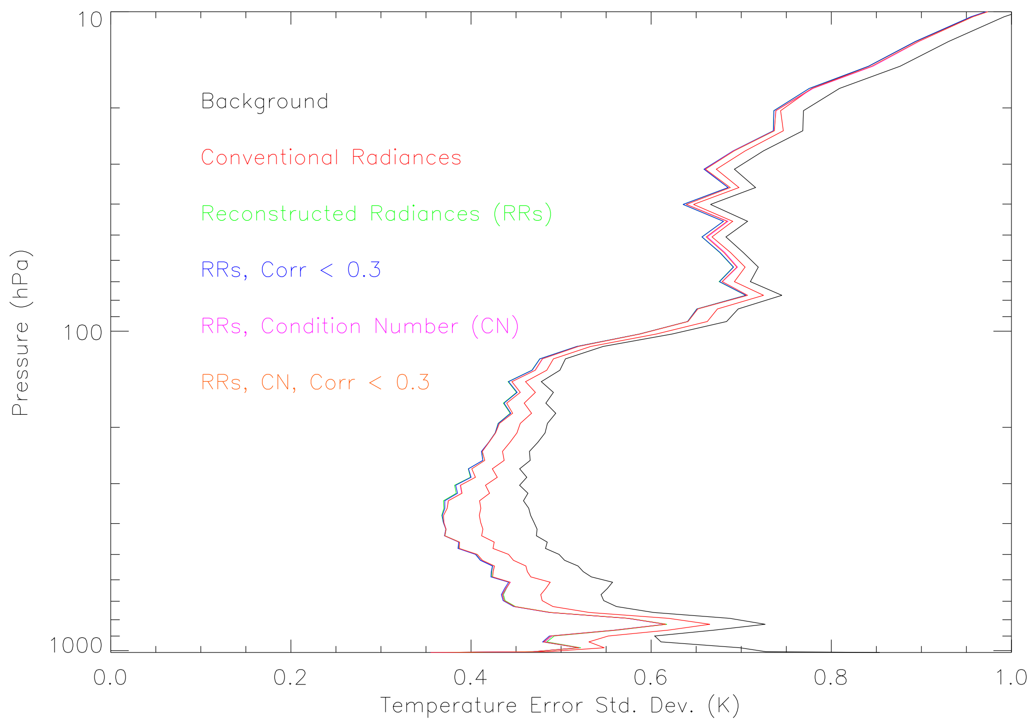

Fig. 16: Temperature retrieval errors for 70 channels for each of the cases discussed in this report. When presented in this manner, the performance of each of the reconstructed radiance channel selections are very similar.

matrices are assumed.

It is noted that for reconstructed radiances a selection of N channels can contain all the information from N eigenvalues. However the condition number of the R-matrix can be very high depending on the choice of this channel subset. Therefore, one might also wish to take into account this condition number in performing the channel selection.

Finally, it should be noted that this discussion only considers one aspect of the total error budget: the instrument's random noise transformed into reconstructed radiance space. In reality other noise sources may well dominate: instrument biases (which may not be completely removed by the chosen bias correction scheme); forward model error (again both random and biases) and representivity error. Some of these errors may be difficult to determine accurately (particularly if we are considering subtle correlations between many channels) and yet they may still affect results. This is one reason why having a well conditioned instrument noise matrix may be advantageous.

In some cases forward model errors might be so large that affected channels should be avoided entirely. Land surfaces and clouds are two such contaminants. In these cases we need to make sure that our channel selection is still adequate if we need to remove these affected channels. This is particularly pertinent if, as in some of the selections above, window channels or channels in the stratosphere are being preferentially chosen. So additional constraints on the height range of the Jacobians might also be required (e.g., Collard (2007) removed all window channels for the initial channel selection).

References

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.

Collard, A.D. (2007). Selection of IASI Channels for use in Numerical Weather Prediction. *Q.J.R. Meteorol. Soc.*, **133**, 1977–1991.

Collard, A.D., A.P. McNally, F.I. Hilton, S.B. Healy and N.C. Atkinson (2010). The use of principal component analysis for the assimilation of high-resolution infrared sounder observations for numerical weather prediction. *Q.J.R. Meteorol. Soc.*, **136**, 2038–2050.

Rodgers, C.D. (1996). Information content and optimisation of high spectral resolution measurements. *SPIE, 2380, Optical spectroscopic techniques and instrumentation for atmospheric and space research II, Paul B. Hays and Jinxue Wang eds.*, pp. 136–147.

Rodgers, C.D. (2000). Inverse methods for atmospheres: Theory and Practice. *World Scientific Publishers, Singapore.*

Watts, P.D., and A.P. McNally (1988). The sensitivity of a minimum variance retrieval scheme to the values of its principal parameters. *Technical Proceedings of the Fourth International TOVS Study Conference.*, 399–407.