Impact of RTIASI fast radiative transfer model error on IASI retrieval accuracy

Vanessa Sherlock

February 21, 2001

The Met. Office
NWP Division
Room 344
London Road
Bracknell
RG12 2SZ
United Kingdom

Abstract

This study quantifies the impact of specific forward model and Jacobian error characteristics on retrieval accuracy for RTIASI, a fast radiative transfer operator for IASI. In light of previous studies, improved estimates of forward model error covariance and fast model Jacobian errors have been made for two versions of RTIASI using three-case and single-case water vapour predictor schemes. Forward model errors present strong correlations within spectral bands and within the window regions. Error characteristics are comparable for the two versions with the exception of the H_2O ν_2 band where the errors for the single predictor scheme are larger, and error correlation structures differ. The impact of simplifying assumptions to the structure of the forward model error covariance on retrieval accuracy has been evaluated for two instrumental noise scenarios in two atmospheric sounding scenarios (tropical and sub-arctic atmospheres). In the H_2O ν_2 band instrumental noise and forward model error are comparable in magnitude and forward model errors govern long-range correlation structure. In this case, neglecting correlation structure can degrade retrieval of upper tropospheric temperature and humidity. In other spectral intervals instrumental noise makes the dominant contribution to the observation error covariance and the error incurred by a diagonal approximation to the forward model error covariance is small. In all cases a block diagonal approximation to the forward model error covariance matrix accurately captures all relevant correlation features of the full matrix. Errors in temperature Jacobians have a negligible impact on retrieval accuracy and suggest a target accuracy of 5% for fast model Jacobian calculations. Errors in water vapour Jacobians can degrade the accuracy of both temperature and humidity retrievals significantly. This is particular issue for the three-case water vapour predictor scheme in the tropical atmosphere. All results highlight the sensitivity of retrievals to forward model error characteristics in the H_2O ν_2 band. If they can be generalised to a wide range of atmospheric states, a single water vapour predictor scheme is to be preferred, although further improvements in model performance in the H_2O ν_2 band are desirable even in this case.

1   Introduction

In a previous study[1] referred to hereinafter as VS2000, fast radiative transfer model errors were estimated for two IASI fast models, RTIASI and PFAAST. It was shown that while fast model errors are generally acceptable, both models have specific problems or limitations which must be solved before integration into an operational data assimilation system is feasible. RTIASI, the model currently used in the development of a 1D-Var scheme, has good forward model error characteristics in the CO_2 bands used for temperature sounding, and has the capability to generate analytic Jacobians. However, water vapour absorption is significantly less well modelled: forward model errors in the window regions and the H_2O ν_2 band are larger than those obtained with the PFAAST model, and these errors present a high degree of correlation. Moreover, significant errors were found in modelled water vapour Jacobians.

In this study we seek to quantify the impact of RTIASI-specific forward model and Jacobian error characteristics on retrieval accuracy. Given that water vapour absorption is less accurately modelled in RTIASI (as compared to PFAAST), the performance of two water vapour predictor schemes are considered in detail. These are the RTIASI Version 1 (November 1999) release three-case and single-case water vapour predictor schemes and will be denoted v13 and v11 respectively.

The results of this quantitative impact study are used to identify requirements for future fast model and data assimilation developments. We seek to answer two questions:
- is the RTIASI fast model adequate as it stands ?
- can approximations to the structure of the observation error covariance be made which will simplify computations in a variational retrieval scheme ?

These results also serve as a benchmark for estimating the retrieval errors which could be expected from error sources not considered here, such as spectroscopic uncertainties and representativity errors. These sources of error, and errors due to modelled surface emissivity and solar reflection and climatological variations of modelled fixed gases will be treated in detail in a future study.

Section 2 summarises extensions to forward model and Jacobian error estimates detailed in VS2000. The salient features of full forward error covariance matrix estimates for the v11 and v13 versions of the RTIASI model are described and the treatment of the instrumental error covariance matrix is discussed. The relative contributions of these two sources of error to the observation error covariance matrix are described.

Section 3 presents the results of impact studies. The mathematical framework of these impact studies is described in section 3.1. The effects of different observation error covariance scenarios on retrieval resolution and measurement information content (degrees of freedom for signal) are discussed in section 3.2. We then evaluate the increase in retrieval error due to the introduction of simplifying approximations to the forward model error covariance matrix (section 3.3) and due to errors in fast model Jacobian estimates (section 3.4).

Section 4 summarises the conclusions and recommendations of this study.

This report assumes a certain familiarity with the issues surrounding the use of operational fast radiative transfer models. The reader is referred to VS2000 if a more detailed introduction to this work is required.

2 Observation error covariance and Jacobian error estimates

2.1 Observation error covariance estimates

Under the (reasonable) assumption of independent instrument noise and forward model errors, the observation error covariance matrix O is given by the sum of the the instrumental error covariance matrix E, and the forward model error covariance matrix F, O = E + F. In this section we describe estimates of the instrumental and forward model error covariance matrices for IASI retrievals using the fast forward model RTIASI.

2.1.1 Treatment of the instrumental error covariance matrix

CNES Level 1C instrumental noise specifications are used to specify instrumental variances – the diagonal elements of the instrumental error covariance matrix, E. The CNES Level 1C noise estimates are expressed as a noise equivalent brightness temperature for a scene temperature of 280 K, and will by denoted by NdTb or 'normalised brightness temperature difference' throughout this discussion. Two noise estimates are considered here – original 'typical' Level 1C noise estimates [2], identified by the label O1, and revised Level 1C noise estimates [3], identified by the label O2. These CNES estimates of Level 1C instrumental noise (standard deviation) are illustrated in Figure 1. In the revised CNES Level 1C noise scenario instrumental noise is reduced at all wavenumbers relative to the original noise estimates, with noise levels between 0.1 and 0.2 K throughout the $700 - 2300$ cm^{-1} interval. The reduction in the $1200 - 2000$ cm^{-1} interval encompassing the H$_2$O ν_2 band is of particular note as forward model errors and the revised instrumental noise are comparable in magnitude in this interval.

Instrumental noise may be considered uncorrelated from channel to channel for spectra deduced from unapodised interferograms, but this is not true of the Level 1C radiances: apodisation introduces interchannel noise correlations and the off-diagonal elements of the instrumental error covariance matrix must also be prescribed.

Correlation structure has been estimated following the method of Amato, de Canditiis and Serio [4][1]. Interchannel correlations are essentially homogeneous across the spectrum and may be specified with a single vector of correlation coefficients. Correlation coefficients for the four nearest neighbours are tabulated in Table 1. Note these correlation coefficients are in exact agreement with coefficients published by Barnet and Susskind [5].

Error covariances must be symmetric positive definite matrices. Truncation of the correlation function (only accounting for a limited number of nearest neighbour correlations) can lead to a loss of the property of positive definiteness and a system of inconsistent equations. The question is then raised as to what level of inter-channel error correlation must be specified to ensure E is a consistent symmetric positive definite matrix. A correlation function C is symmetric positive definite if its Fourier transform $\mathcal{F}(C) > 0$ [6]. Truncation of the correlation function will clearly give rise to oscillations in $\mathcal{F}(C)$ (positive and negative values of $\mathcal{F}(C)$) and positive definiteness is no longer guaranteed. If truncation is performed at a point separation where the correlation function tends to zero, then only a limited number of the eigenvalues of the truncated matrix will be negative (strictly, small and negative) and in fact all eigenvalues may be positive. In this case a symmetric positive definite error covariance matrix and its pseudo inverse can be reconstructed from the set of non-negative eigenvalues and eigenvectors. In the approach adopted here we have sought the minimum length correlation vector which gives an instrumental error covariance with strictly non-negative eigenvalues. This condition is satisfied for a correlation specification out to four nearest neighbours i.e. E is a band diagonal matrix with bandwidth 9.

Apodisation will also clearly have an effect on the condition of the instrumental error covariance matrix. In the case of strong apodisation, as used for IASI, the modification of the condition number can be large. In Table 2 we tabulate the spectral norm condition number for E and O for the v13 O2 error covariance scenario for uncorrelated (diagonal) and correlated instrumental noise specifications. The left hand panel relates to calculations using a normalised brightness temperature difference noise specification – the NdTB 'metric', the right hand panel relates to calculations using a radiance space noise specification [W/sr/m^2] – the dR 'metric'. The condition number of E increases by a factor of 1000 when interchannel correlations are taken into account. The fact that the condition number is so large has implications for the numeric stability of matrix inversions: for example, for Cholesky decomposition numerical breakdown occurs when the spectral norm condition number $\kappa_2 \sim 1/\sqrt{u}$, where u is the unit roundoff error. Thus for condition numbers greater than 10^3, real8 precision must be used for all calculations.

The choice of noise metric has a small impact on the condition of E – there is a factor three increase in condition number for the dR noise specification. This is because large instrumental errors in IASI band 3 ($\nu > 2000$ cm^{-1}) tend to increase the condition number of the normalised brightness temperature error covariance specification (because NdTB noise is relatively homogeneous as a function of wavenumber, the minimum/maximum values of instrumental noise for the dR noise specification are governed by the nonlinearity of the Planck function with wavenumber – this more than compensates the increases in NdTB in band 3). The condition of the full error covariance matrix is more sensitive to the choice of metric: there is a factor thirty increase in condition number for the dR noise specification. Forward model error has a small stabilising effect for the NdTB specification, and a destabilising effect for the dR specification.

The Cholesky decomposition L of the symmetric square matrix A: $LL^T = A + dA$ has has well defined bounds for the errors dA [7][8]. These bounds may be used to estimate how errors in the Cholesky decomposition of the innovation and

---

[1]Matrix representations of the discrete fourier transform and apodisation function are used to deduce the correlation stucture introduced on apodisation of a calibrated, bandlimited spectrum/interferogram.

Point spacing	0	1	2	3	4
Correlation	1.000	0.707	0.250	0.044	0.004

Table 1: IASI Level 1C (apodised) instrumental error covariance: correlation coefficients for the four nearest neighbour channels.

Configuration	NdTB			dR		
	λ_{max}	λ_{min}	κ_2	λ_{max}	λ_{min}	κ_2
E corr.	1.0×10^1	3.7×10^{-5}	3×10^5	4.13×10^{-7}	4.7×10^{-13}	1×10^6
E diag.	3.7×10^0	9.0×10^{-3}	4×10^2	1.0×10^{-7}	1.0×10^{-10}	1×10^3
E+F corr.	1.0×10^1	1.0×10^{-4}	1×10^5	1.4×10^{-6}	5.1×10^{-13}	3×10^6
E+F diag.	3.7×10^0	9.0×10^{-3}	4×10^2	1.3×10^{-6}	1.1×10^{-10}	1×10^4

Table 2: Spectral norm condition, v13 O2 n=1057

observation error covariances will affect analysis increments (see Appendix A for details). Of these two sources of error, roundoff error for the Cholesky decomposition of the observation error covariance is the dominant source of error in the evaluation of the Kalman gain matrix (again, this is shown in appendix A), so some consideration should be given to the choice of units for radiance assimilation. Note for real8 precision calculations roundoff errors are acceptably small in all cases considered here.

2.1.2 Forward model error covariance estimates

Estimates of the forward model error covariance matrix follow on from the evaluation of fast model errors described in VS2000. Radiance simulations have been extended to a sample of 117 diverse atmospheric states, as represented by the ECMWF 50-level forecast model [9]. Convolved GENLN2 radiance spectra were provided by Marco Matricardi. These GENLN2 radiative transfer calculations were performed using the RTIASI vertical layering. Thus, layer average temperature differences (absorber density/air density weighting) contribute to observed differences, but representativity errors are not otherwise taken into account.

RTIASI was run to generate radiance spectra for the 117 profiles. RTIASI – GENLN2 radiance differences were evaluated on a channel-by-channel and profile-by-profile basis and processed to generate estimates of the forward model error covariance matrix. Forward model error covariance estimates have been made for RTIASI versions v11 and v13 model, and for the June 1999 RTIASI release with a two-case water vapour predictor scheme (denoted here v02). The forward model error characteristics of the v02 and v13 models do not differ greatly, and only the latter will be discussed here.

The units for the specification of the forward model error covariance are chosen in the processing step. The results presented here are all based on a normalised brightness temperature difference (NdTB) specification. Radiance-space and brightness temperature-space error specifications are also possible. Implications for the numerical stability of matrix inversions and error in the evaluation of the gain matrix have been mentioned briefly in subsection 2.1.1 and are discussed in appendix A.

In Figures 1 and 2 we illustrate the bias, standard deviation and root mean square normalised brightness temperature differences for the v11 and v13 RTIASI models respectively. Also illustrated for reference are two CNES estimates for Level 1C instrumental noise (standard deviation), again expressed as a normalised brightness temperature difference. Original 'typical' Level 1C noise estimates [2] are identified by the label O1. Revised Level 1C noise estimates [3] are identified by the label O2.

The errors show the same features as those revealed by a robust analysis of simulations for a much reduced set of atmospheric states, the AFGL atmospheres [VS2000]: high bias and standard deviation in the H_2O ν_2 band, a structured and relatively high standard deviation in the 8-12 μm atmospheric window, bias and random components to errors in the 9.6 μm O_3 and 4.3 μm CO_2 bands[2]. The v11 and v13 RTIASI model errors differ in the H_2O ν_2 band. In this spectral interval the bias and standard deviation of v11 model errors are greater than those of the v13 model. A more detailed comparison of v11 and v13 errors in the 1300-1700 cm^{-1} interval is given by the scatter plots illustrated in Figure 3. These plots clearly illustrate that the single water vapour predictor scheme (v11) performance is degraded most in channels where the v13 model performance is also poorest. This is true of bias, standard deviation and RMS. Note an approximately 1:1 mapping for biases less than -0.2 K. v11 standard deviations are higher than v13 standard deviations over the whole range, but again, degradation is only really significant for standard deviations greater than 0.2 K.

With the exception of the H_2O ν_2 band the random component of the forward model errors (standard deviations) are significantly less than the instrumental noise. In the H_2O ν_2 band RTIASI version v11 random errors and instrumental noise are comparable for both CNES scenarios, while v13 errors are only comparable with the revised CNES noise estimates

---

[2]The absolute (non-normalised) standard deviations of brightness temperature differences in the 4.3 μm band are of the order of 0.5 K, as found in VS2000. These errors are $\sim$ 6 times smaller than even the revised CNES IASI instrumental noise estimates. Forward model error would not appear to be the limiting factor for IASI retrievals in this spectral interval. The same is not true of AIRS: instrumental noise specifications are for a NEΔT at 280 K of less than 0.1 K at 2400 cm^{-1}.

(O2 scenario). As mentioned above, under the (reasonable) assumption of independent instrument noise and forward model errors, the observation error covariance matrix O is given by the sum of the the instrumental error covariance matrix E, and the forward model error covariance matrix F, $O = E + F$ and the diagonal elements of O are given by $O_{ii} = \sigma^2_{E,i} + \sigma^2_{F,i}$. Thus if $\sigma_F \sim 0.5\sigma_E$, as is often the case here, 'F' makes a relative contribution of 0.25 to the diagonals of O. Figures 1(b) and 2(b) give a useful summary of the relative contributions of instrumental and forward model error to the diagonal elements of the observational error covariance matrix, but this is not the whole picture: interchannel correlation structures are important too.

In the upper panel of Figure 4 we illustrate the forward model error correlation matrix for the v13 RTIASI model. To extract the correlation structures represented in this figure note that yellow indicates high correlation, blue indicates moderate anticorrelation, and deep orange/red indicates low correlation/approximate independance. Thus, there are correlations between the 15 and 4.3 μm CO_2 bands ($\sim$ 650 and 2400 cm^{-1} respectively), and indeed (although poorly represented in this figure) between Q branches in the 15 μm band. Similarly, high correlation exists between window channels (800–1000 cm^{-1}, 1100–1200 cm^{-1}, $\sim$ 2100 cm^{-1} and 2500 cm^{-1} $\rightarrow$ spectral intervals). There is a strong correlation structure within the 9.6 μm O_3 band, and moderate anticorrelation between these errors and errors in the CO_2 bands and the H_2O ν_2 band. Finally, there is significant correlation structure within the ν_2 band. The observed 'star' structure is related to distinct forward model error characteristics in the Q, P and R-branches of the ν_2 band. Specifically, this correlation structure indicates that errors in the Q-branch centre have lower correlation with the regions of low variance in the vicinity of the P and R band heads than with errors in the P and R band wings (see Figure 1(b)).

The picture for the v11 RTIASI model is very similar. The only real difference occurs in the H_2O ν_2 error correlation structure. Errors are more highly correlated throughout the P and R bands, and the Q-branch is actually a local minimum in forward model error. Thus the v13 'star' structure is lost; there is higher, more uniform correlation across the ν_2 band, although the Q-branch remains distinct. These features can be seen in the 'ν_2 block' of the block-F correlation matrix which is illustrated in the lower panel of Figure 4 and which is discussed now.

The spectral structure of forward model error correlations clearly suggests that a block specification of forward model error covariance would capture most of the covariance structure. By block specification we mean that correlation structures within limited spectral intervals (e.g. spectral bands and window regions) are accounted for, outside these intervals correlations are neglected. A suggested block specification is illustrated for the RTIASI v11 model in the lower panel of Figure 4. With the exception of the anticorrelation of errors in the O_3 band and errors in the CO_2 bands and the H_2O ν_2 band, and correlations between errors in the H_2O ν_2 band and errors in the $700 - 800$ cm^{-1} interval, this block approximation gives a good description of the principal correlation structures. A reordering of channels grouping the two CO_2 bands together and grouping window regions together gives a computationally efficient block diagonal structure to the matrix. The accuracy of this block diagonal approximation will be examined in section 3.3.

As described in subsection 2.1.1 above, apodisation introduces correlation in Level 1C instrumental noise linking channels within a proximity of ~ 1.5 cm^{-1} (fourth-nearest neighbours). Clearly then long range correlation structures will be governed by the forward model error contribution to E+F. The significance of these correlations depends on the relative magnitudes of instrumental and forward model error contributions.

In Figure 5 we illustrate the observation covariance matrix $O = E+F$ for the v13 and v11 forward model error covariance, and the revised CNES instrumental error covariance (scenario O2). A pentadiagonal instrumental error covariance is assumed (instrumental correlations greater than 0.25) for illustration here, although clearly these nearest neighbour correlations cannot be distinguished on the scale of this graph. There are significant off-pentadiagonal elements of the E+F covariance matrix in the interval around 720 cm^{-1}, the 800–1000 cm^{-1} window region, the 1050 cm^{-1} (10 μm) O_3 band, and the H_2O ν_2 band. In these intervals forward model error makes a significant contribution to the observation error covariance matrix (the v11 forward model contribution in the ν_2 band is of particular note). Elsewhere instrumental contributions dominate. As instrumental noise increases (e.g. if the observation error covariance is evaluated using the original CNES noise estimates) then obviously the relative contribution of forward model error to the observation error covariance decreases.

In section 3.2 we quantify changes in information content (strictly, degrees of freedom for signal) corresponding to the v11/v13 and O1/O2 forward model and instrumental error covariance scenarios. Then in section 3.3 we examine the effect of simplifying approximations to the forward model error covariance on retrieval accuracy. In addition to the block diagonal approximation described above, we quantify the error associated with the a diagonal approximation to the forward model error covariance matrix. This approximation has been widely adopted in IASI performance studies to date because realistic estimates of forward model error correlations have not been available and because it simplifies and speeds up computations radically – $O(n^3)$ floating point operations are required to evaluate a full matrix inverse, whereas the calculation of the inverse of a diagonal is trivial and requires $O(n)$ flops. These are clearly important considerations for IASI, where the full observation error covariance is given by a 8641$\times$8641 matrix.

A subset of 1057 channels – every eighth channel – have been selected for impact studies. As explained by Collard [10], this channel selection does not significantly modify the sampling of absorption regimes/features, so retrieval error covariances are essentially unchanged apart from a small reduction in absolute accuracy due to a reduction in the number of degrees of freedom for signal. It does simplify calculations through the reduction of storage requirements and computation time. At the resampled resolution (2 cm^{-1}), instrumental noise may be considered uncorrelated from channel to channel. For this reason, the instrumental error covariance is prescribed by a strict diagonal matrix. This assumption can be relaxed in future studies as required. The current studies examine the impact of current forward model error levels and 'long range' forward model error correlations on retrieval accuracy and measurement information content.

2.2 Jacobian error estimates

Forward model error is taken into account explicitly in the retrieval process – the relative weights of the *a priori* information and observations in determining analysis increments are governed by their respective error characteristics. The variational framework makes the implicit assumption that fast model Jacobians are exact. This hypothesis clearly needs to be validated.

One approach to validating fast model tangent linear Jacobians is to compare them with brute force (finite difference) Jacobians calculated using a line-by-line model[3]. Errors in RTIASI version v13 temperature and water vapour Jacobians were evaluated in this manner in VS2000 for the AFGL tropical and sub-arctic winter atmospheres on three targeted wavenumber subintervals: 645-800 cm^{-1}, 885-915 cm^{-1} and 1300-1450 cm^{-1}. This study showed that modelled temperature Jacobians $\partial T_{Bi}/\partial T_j$ were generally very satisfactory (errors typically $\leq 5\%$), but that water vapour Jacobians $\partial T_{Bi}/\partial \ln(q)_j$ were considerably less well modelled. In the AFGL1 atmosphere errors of the order of 10-40% were found in the magnitude of water vapour Jacobians, although this was much improved in the AFGL5 atmosphere. Larger errors still were found in the v02 release in cases where switching between predictor schemes gave rise to large discontinuities in modelled Jacobians. This problem is reduced, but not completely eliminated in the v13 release (discontinuities still occur for some Jacobians, but are smaller in magnitude).

The corresponding analysis has been made for the RTIASI version v11 Jacobians. The accuracy of modelled temperature Jacobians remains essentially unchanged. A small improvement in temperature Jacobians in the 750-800 cm^{-1} interval is noted for AFGL1. A small degradation is noted in the same interval for AFGL5. Significant improvements in the accuracy of modelled water vapour Jacobians are found in both the 885-915 cm^{-1} and the 1300-1450 cm^{-1} interval for the AFGL1 atmosphere. Again, a degradation in accuracy is found for the AFGL5 atmosphere. This implies that the v11 scheme models water vapour absorption better than the v13 scheme in both the window region and the H_2O ν_2 band in the case of the AFGL1 atmosphere, but worse in the case of the AFGL5 atmosphere. Extrapolation of these results to 'humid' and 'dry' atmospheres is not recommended (!)

Comparisons of the maximum relative errors in RTIASI v11 and v13 water vapour Jacobians for the AFGL1 and AFGL5 atmospheres on the 1300-1450 cm^{-1} interval are illustrated in Figure 6. Strictly, the error represented is the maximum relative error for Jacobian elements whose value is $\geq 50\%$ maximum Jacobian element for the given channel and atmosphere, and the error is signed: positive errors indicate GENLN2 > RTIASI (in terms of the absolute value of the Jacobian elements), conversely negative errors indicate GENLN2 < RTIASI. The Jacobians considered are essentially well behaved, smooth functions i.e. there are none of the gross errors (discontinuities) of the v02 release. Thus the maximum relative error characterises the typical magnitude of errors in the region of maximum sensitivity on a channel by channel basis.

The improvement in the v11 AFGL1 water vapour Jacobians is clear - errors are generally of the order of 10% as opposed to 10-30% for the v13 scheme. However, the degradation in the AFGL5 case is as evident. Errors are of the order of 20% as compared with 5-10% for the v13 scheme. Note however that the AFGL5 atmosphere is particularly dry; sensitivity to perturbations in the analysis variable ln(q) ($d ln(q) = \frac{1}{q} dq$) is therefore significantly lower than that found in the AFGL1 case (i.e. the absolute values of the water vapour Jacobians are smaller).

The impact of errors in modelled temperature and water vapour Jacobians on retrieval accuracy is described in section 3.4.

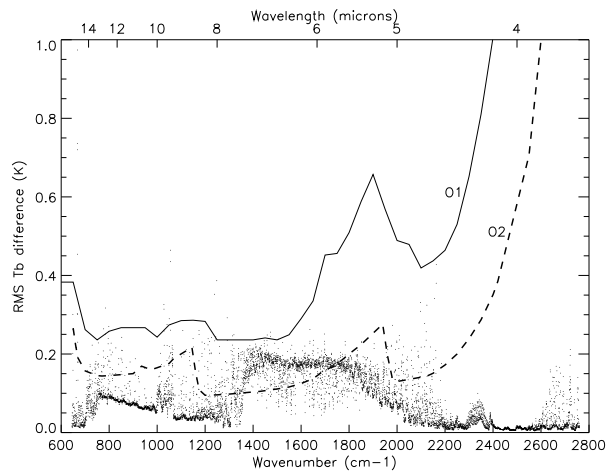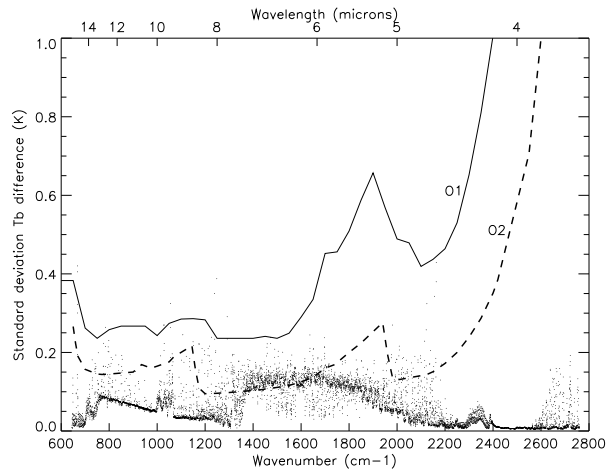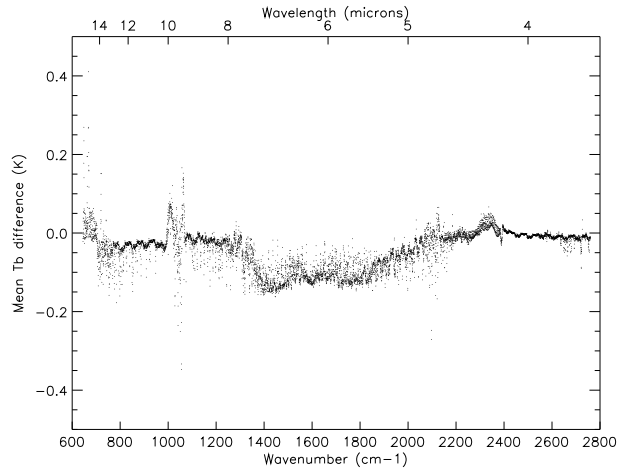[3]Profile perturbations must be small enough to ensure a linear approximation is valid.

Figure 1: Mean, standard deviation and root mean square errors for the RTIASI version 1 release with three water vapour predictor regression schemes. Radiance errors have been converted an equivalent brightness temperature difference for a scene temperature at 280 K in order to compare forward model errors with Level 1C NEDT estimates for IASI instrumental noise: initial CNES noise estimates are traced with the solid curve labelled O1 (typical noise scenario), revised estimates (F. Cayla, August 1999) are traced with the dashed curve labelled O2. Statistics are derived from 117 spectra simulated using atmospheric profiles selected from the 50-level ECMWF profile set.
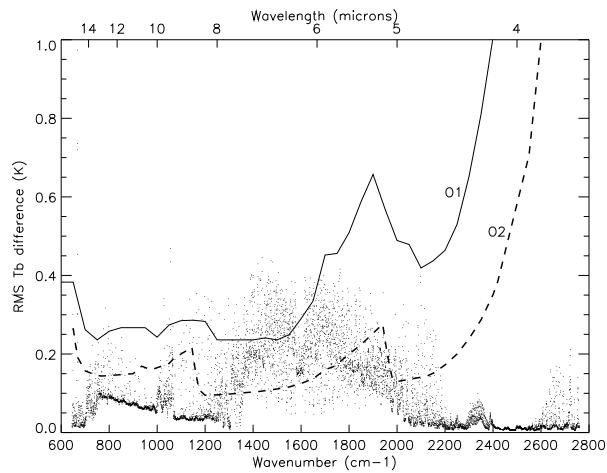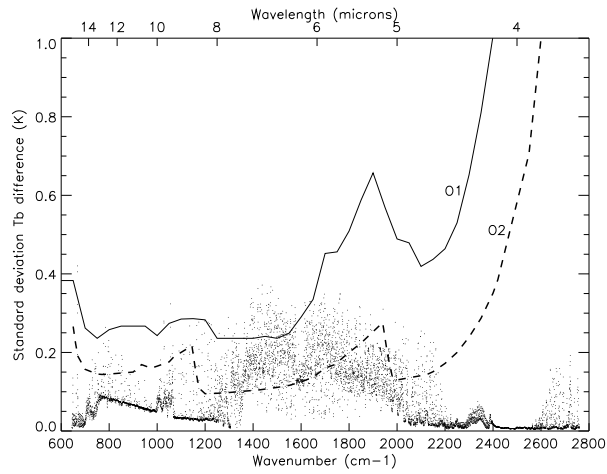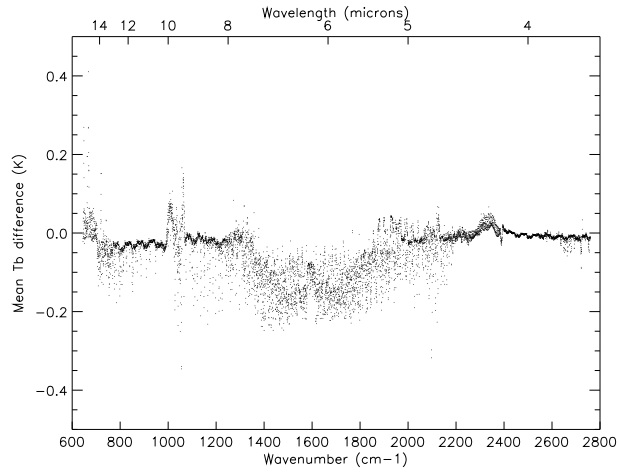
6

Figure 2: Mean, standard deviation and root mean square errors for the initial RTIASI f90 release with one water vapour predictor regression scheme. Radiance errors have been converted an equivalent brightness temperature difference for a scene temperature at 280 K in order to compare forward model errors with Level 1C NEDT estimates for IASI instrumental noise: initial CNES noise estimates are traced with the solid curve labelled O1 (typical noise scenario), revised estimates (F. Cayla, August 1999) are traced with the dashed curve labelled O2. Statistics are derived from 117 spectra simulated using atmospheric profiles selected from the 50-level ECMWF profile set.
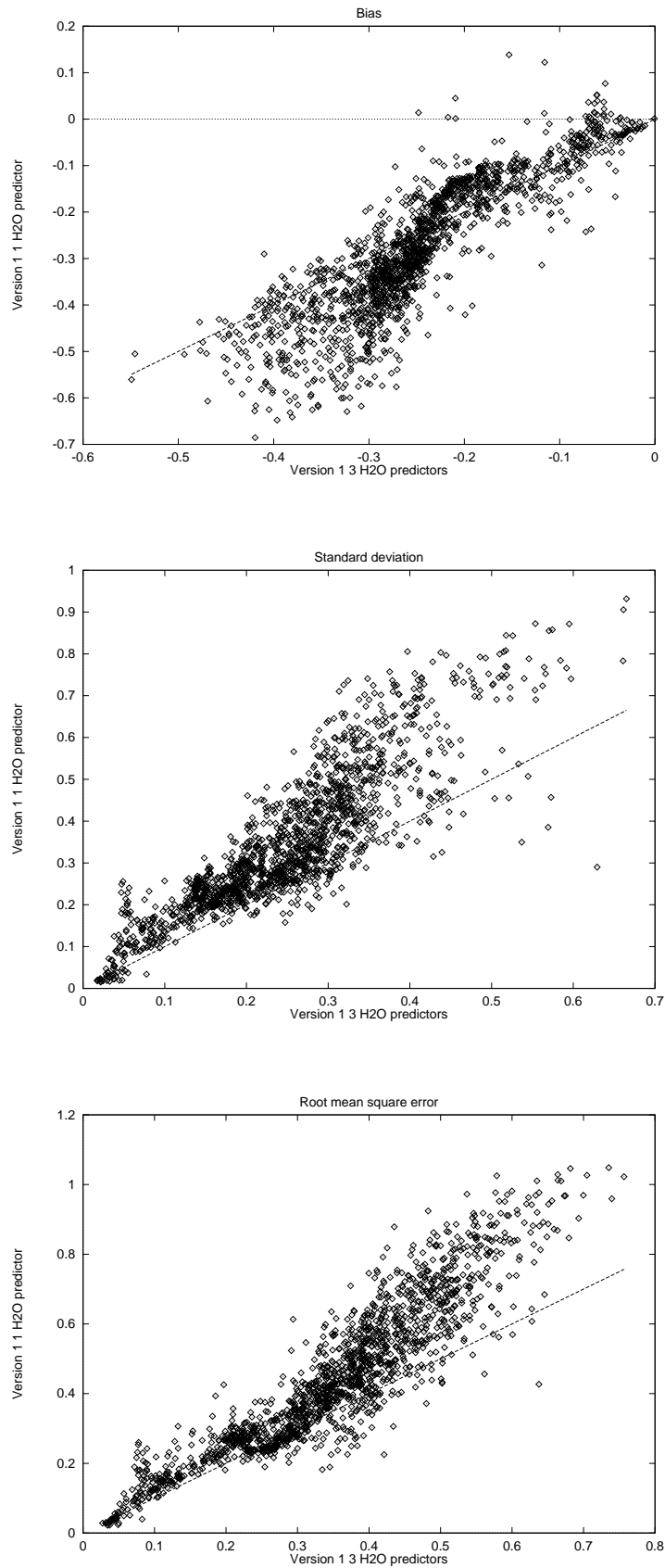
Figure 3: Scatter plots of bias, standard deviation and root mean square brightness temperature errors (K) for RTIASI Version 1 with three or one water vapour regression schemes. Forward model errors on the 1300 to 1700 cm^{-1} interval only are traced here.
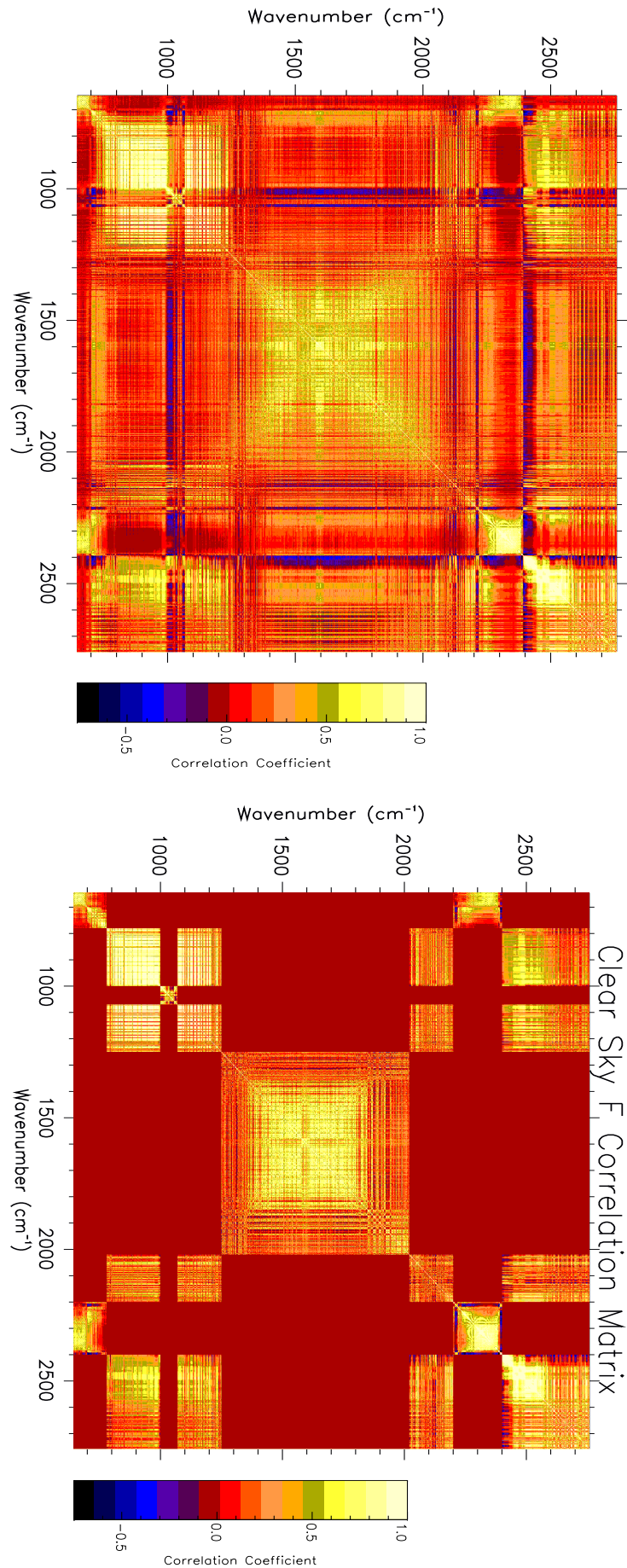
Figure 4: Upper panel: full forward model error correlation matrix for the RTIASI v13 model. Lower panel: block forward model error correlation matrix for the RTIASI v11 model.
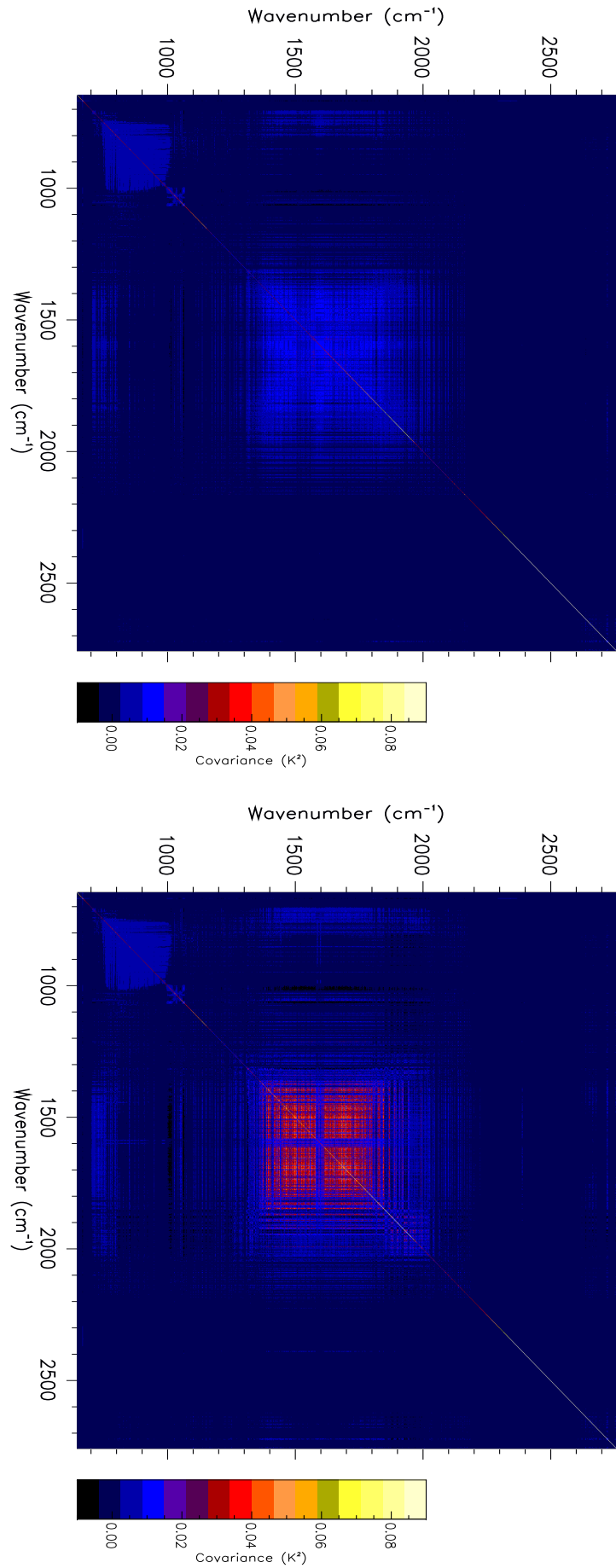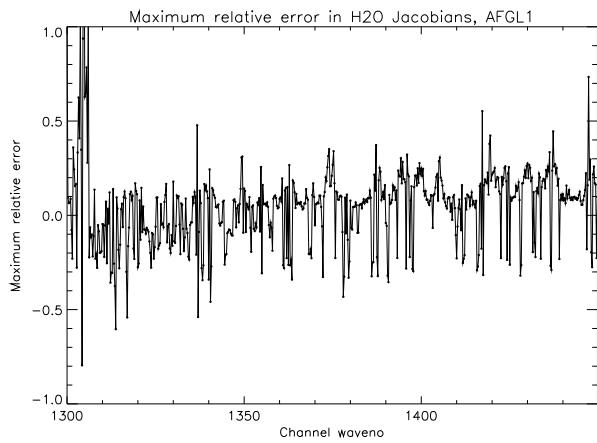
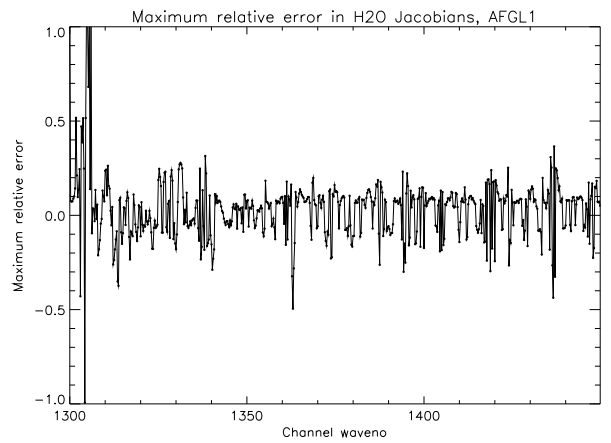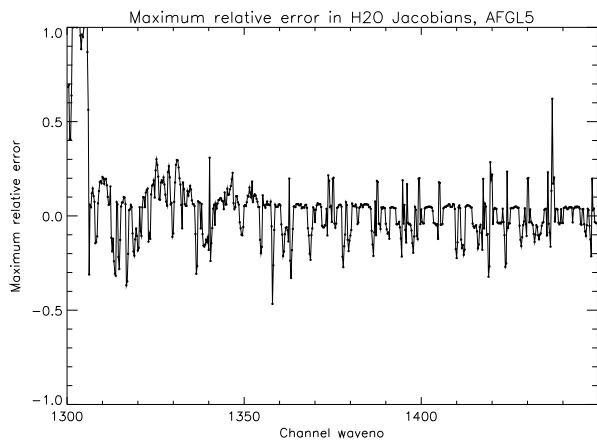Figure 5: Observational error covariance matrices for RTIASI v13 and v11 models (upper an lower panels respectively). In both cases the instrumental noise is given by the revised (O2) CNES Level 1C estimates. Covariance in units of K^2 (normalised brightness temperature differences for a scene temperature of 280 K).
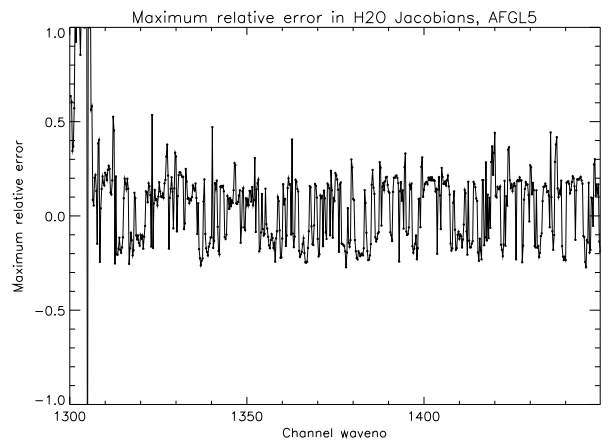
AFGL1 v13

AFGL1 v11

AFGL5 v13

AFGL5 v11

Figure 6: Signed maximum error in the Jacobian central maximum.

3 Impact of fast model errors on retrieval error covariances

3.1 Methodology

The impact of forward model errors on retrieval accuracy is assessed within a linear retrieval framework. The analysis presented here is a direct application of the methodology used by Watts and McNally [11] to assess the sensitivity of a minimum variance retrieval scheme to the values of its principal parameters and by Collard [10] to assess the impact of undetected cloud on IASI retrievals. This methodology is outlined briefly here.

In a minimum variance retrieval scheme the best estimate of atmospheric state $\hat{x}$ is found by minimising the cost function $J(x)$:

$$J(x) = (x - x_0)^T B^{-1}(x - x_0) + (y - \mathcal{H}(x))^T O^{-1}(y - \mathcal{H}(x)), \tag{1}$$

where x_0 is the *a priori* estimate of atmospheric state, with associated error covariance B, the y_i are observations, and $\mathcal{H}(x)$ is the forward operator mapping the state x into observation space. y-$\mathcal{H}(x)$ has an associated error covariance O which is comprised of two terms or error sources: E, the instrumental error covariance and F, the forward model error covariance. These errors are assumed to be independent, thus $O = E + F$.

In the linear or weakly nonlinear case analysis increments $\hat{x} - x_0$ are given by:

$$\hat{x} - x_0 = (B^{-1} + \nabla_x \mathcal{H}^T O^{-1} \nabla_x \mathcal{H})^{-1} \nabla_x \mathcal{H}^T O^{-1}(y - \mathcal{H}(x_0)),$$
$$= W(y - \mathcal{H}(x_0)), \tag{2}$$
$$\text{where} \quad W = (B^{-1} + \nabla_x \mathcal{H}^T O^{-1} \nabla_x \mathcal{H})^{-1} \nabla_x \mathcal{H}^T O^{-1}.$$

W is known as the gain or weight matrix of the analysis. The rows of W describe how the departures y-$\mathcal{H}(x_0)$ are mapped into analysis increments.

If the true atmospheric state is given by x_T, then in the linear case Equation 2 can be expressed as:

$$\hat{x} - x_0 = R(x_T - x_0) + W\epsilon_y, \tag{3}$$

where $R = W\nabla_x \mathcal{H}$ is the averaging kernel or model resolution matrix, and ϵ_y is the realisation of instrumental and forward model noise for the given observations y and mapping $\mathcal{H}(x_0)$. An appropriate analysis of the resolution matrix yields important information on retrieval characteristics and is discussed below.

Equation 3 may be rearranged to give an expression for the difference between the analysis and the true atmospheric state $\hat{x} - x_T$:

$$\hat{x} - x_T = (I - R)(x_0 - x_T) + W\epsilon_y, \tag{4}$$

which may in turn be used to estimate the error covariance for an ensemble of retrievals A (the *a posteriori* error covariance):

$$A = E[(\hat{x} - x_T)(\hat{x} - x_T)^T],$$
$$= (I - W\nabla_x \mathcal{H})B(I - W\nabla_x \mathcal{H})^T + WOW^T, \tag{5}$$

where $E[\ ]$ denotes statistical expectation. The retrieval error covariance is comprised of two terms. The latter, WOW^T describes the propogation of measurement noise and forward model errors into retrieved states $\hat{x}$, while the former describes how errors in the *a priori* estimate of atmospheric state are propogated into the retrieval when (as) the *a priori* information is used to 'fill out' that profile information which cannot be deduced from observations. For this reason this term is described as the null space error by Rodgers [12].

If now we consider an ensemble of retrievals where an observation error covariance O has been assumed in W, but the true observation error covariance is O', then the retrieval error covariance is given by:

$$A' = (I - W\nabla_x \mathcal{H})B(I - W\nabla_x \mathcal{H})^T + WO'W^T,$$
$$A' = A + W(O' - O)W^T. \tag{6}$$

Thus, errors in the assumed observation error covariance matrix are seen to give an additional contribution to the propogated measurement error – the minimum variance solution or optimal retrieval requires $O \equiv O'$.

In a similar manner, if an ensemble of retrievals are performed assuming a Jacobian $\nabla_x \mathcal{H}$, when in fact $\nabla_x \mathcal{H}'$ is the true Jacobian, then the retrieval error covariance is given by:

$$A' = (I - W\nabla_x \mathcal{H}')B(I - W\nabla_x \mathcal{H}')^T + WOW^T. \tag{7}$$

Retrieval errors are increased through the incorrect mapping of *a priori* information: the minimum variance solution requires $\nabla_x \mathcal{H} \equiv \nabla_x \mathcal{H}'$.

3.1.1 Measures for retrieval characterisation

Comparison of *a priori* and *a posteriori* error covariance matrices gives a measure of the benefit of the assimilation of the given observation type. This 'benefit' may be assessed through comparison of the *a priori* and *a posteriori* variances and through measures such as the degrees of freedom for signal and/or measurement information content [13]. In the studies presented here we mainly focus on the modification of the retrieval variance (or standard deviation) – i.e. the diagonal elements of the retrieval error covariance matrix A. In particular, we will quantify the increase in retrieval variance due to simplifying approximations (block diagonal and diagonal approximations) to the full forward model error covariance matrix and due to errors in modelled Jacobians through the relations defined in Equations 6 and 7 above. However, we also evaluate the modification of the degrees of freedom for signal due to approximations to the forward model error covariance and due to errors in modelled Jacobians following the method outlined by Rodgers [13]. Specifically,

$$\text{DFS} = \text{Tr}(I - A^{'}) = \text{Tr}(I - B^{-\frac{1}{2}}A(B^{-\frac{1}{2}})^{T}). \tag{8}$$

Four observation error covariance scenarios have been considered (v11,v13)$\times$(O1,O2) for two relatively extreme atmospheric states (A1$\equiv$AFGL1 tropical and A5$\equiv$AFGL5 sub-arctic winter atmospheres). Two linearisation points have been considered to give some measure of the state dependence of retrieval characteristics. The treatment of nonlinearity errors is beyond the scope of this report (see Eyre [14], Eyre and Collard [15]).

Often results can be better understood in the light of information deduced from an analysis of the corresponding resolution matrix. For this reason, we begin the presentation of results from these impact studies with a discussion of the degrees of freedom for signal and effective vertical resolution of retrievals for the different error covariance and atmospheric state scenarios. The measures derived from the resolution matrix in order to characterise the information content and vertical resolution of the retrieval are:

- Tr(R) and components R_{ii}: if retrievals are performed using the optimal gain matrix W, the trace of the resolution matrix may be interpreted in terms of the number of degrees of freedom for signal in the retrieval [13], or equivalently, as the total effective number of constraints imposed on the retrieval by the observed data [16]. Purser and Huang [16] extend this concept to define a measure of local data density $\rho_i = R_{ii}/\Delta z_i$ and effective vertical resolution ($1/\rho_i$). As this definition of vertical resolution has been used in previous IASI retrieval characterisation studies [17], it is the measure of vertical resolution choosen for illustration here. For an ideal observation R=I, the identity matrix, and Tr(R)=dim(x). For reference, the state vector considered has 75 elements: temperature on 43 pressure levels, surface air and skin temperature, humidity on 28 pressure levels, surface air humidity and surface pressure.

- Eigenvalue/eigenvector decomposition: the rows of the resolution matrix describe how the departures x_T-x_0 are smoothed in the analysis increments $\hat{x}$-x_0, while the columns of R describe how a perturbation at one profile level will be redistributed in analysis increments. An eigenvalue/eigenvector decomposition of the resolution matrix allows (vertically correlated) structures in departures which can be retrieved from observations to be identified [12] providing an alternative method to summarize the structure and smoothing characteristics of the resolution matrix.

A single *a priori* error covariance matrix B is used throughout these studies. This is the ECMWF 40-level background error covariance matrix interpolated onto RTIASI model levels which has been used in previous retrieval characterisation studies [10][17]. The *a priori* error specification has an important role in determining the relative weight of observations in analysis increments, and in determining how increments are smoothed in the vertical: estimates of degrees of freedom for signal and retrieval error covariance depend on the detail of B. We reiterate then that results apply to retrievals/analyses in an operational/NWP framework where the *a priori* estimate of atmospheric state is reasonably well constrained, particularly for the tropospheric temperature field. No evaluation of the sensitivity of results to the specification of B has been undertaken here.

Simulation	DFS full	Simulation	DFS full
v11 O1 A5	10.3	v11 O1 A1	16.3
v13 O1 A5	10.4	v13 O1 A1	16.4
v11 O2 A5	12.7	v11 O2 A1	21.5
v13 O2 A5	12.8	v13 O2 A1	21.6

Table 3: Degrees of freedom for signal, full forward model error covariance specification. For reference, when $F \equiv 0$ degrees of freedom for signal are as follows: O2 A5 DFS=14.1, O2 A1 DFS=23.0.

3.2 A first look at the effects of different observational error scenarios

3.2.1 Features of the resolution matrix for a full F specification

Table 3 summarises the degrees of freedom for signal for the four observational error covariance scenarios (full forward model error covariance) and the two atmospheric states described above. It is clear that for the scenarios considered here – specifically, the relative magnitude of instrumental and forward model error contributions – differences in v11/v13 forward model errors only have a small impact on the information content of the IASI measurements. Instrumental noise does have an impact on information content: depending on atmosphere, two to five degrees of freedom for signal are lost in the O1 noise scenario relative to the O2 scenario (an approximate doubling of instrumental noise at all wavenumbers).

By far the largest variation in degrees of freedom for signal is associated with the variation in atmospheric state. The reasons for this are three-fold. Firstly, the degrees of freedom for signal are in part a measure of the instrumental signal to noise ratios. Signal levels (radiances) are significantly higher for the warm tropical atmosphere. Secondly, water vapour loadings are higher in the AFGL1 atmosphere. Sensitivity to perturbations in ln(q) and temperature in the H_2O ν_2 band is greater as a consequence. Finally, the vertical resolution of a passive infrared measurement is intrinsically linked to the thermal structure (lapse rate) of the atmospheric column sensed: in the limit of an isothermal layer no information can be obtained regarding the vertical distribution of absorbing species. The AFGL5 atmosphere is characterised by an isothermal region extending from ~ 250 to 100 hPa, and weaker vertical temperature gradients in the troposphere than the AFGL1 atmosphere (or indeed any of the other AFGL atmospheres).

The breakdown of degrees of freedom for signal by state vector element is illustrated for the v13 O2 scenario for the two atmospheres in Figure 7 (solid curves). There is a notable increase in the contribution to degrees of freedom for signal for temperature retrieval in the AFGL1 atmosphere below 200 hPa, particularly in the 300 to 500 hPa region. Similarly, a marked increase in the contribution to degrees of freedom for signal for water vapour is observed for the AFGL1 atmosphere below 100 hPa. These results would indicate that water vapour absorption makes a significant contribution to the information content of both temperature and humidity retrievals for the tropical atmosphere. While this result is to be expected, it does suggest that these retrievals will also be most affected by errors in modelled water vapour absorption.

For completeness, plots of data density for the v13 O2 scenario are illustrated in Figure 8. Because the layer geopotential thicknesses do not vary dramatically between the two atmospheres, the differences in data density between the two atmospheric states are as discussed previously. Note the effective vertical resolutions implied by these data density estimates: a temperature sounding resolution of the order of 2 km and humidity sounding resolution of the order of 1 km below ~ 200 hPa in the tropical atmosphere, a temperature sounding resolution of the order of 3 km and a humidity sounding resolution of the order of 2 km below 300 hPa in the sub-arctic winter atmosphere. Note too the marked degradation of resolution in humidity in the isothermal layer above 250 hPa in the subarctic winter atmosphere.

A reduction of two to five degrees of freedom for signal was noted in association with the change from the O2 to the O1 instrumental noise levels. This reduction corresponds to an approximately uniform decrease in the data density for temperature and humidity at all levels illustrated for AFGL5, and to a lesser degree, AFGL1. In the latter case the decreases in data density are slightly greater between 200 and 800 hPa for both temperature and humidity and the absolute reduction in data density is greatest for humidity.

The degrees of freedom for signal for surface (skin) temperature retrieval is essentially equal to unity (≥ 0.997). Corresponding retrieval errors are small $0.03 - 0.06$ K. Even when a diagonal approximation is made to the full F, error amplification only leads to an increase of at most 0.03 K in retrieval error. These estimates are over-optimistic of real retrieval accuracy – uncertainties in modelled surface emissivity, continuum absorption and aerosol attenuation, and cloud clearing errors (residual cloud) [10] will degrade these error estimates. It is none the less interesting to note that for the observation error covariance scenarios considered here (specifically, the relative magnitudes of instrumental noise and forward model error to diagonal and off diagonal elements of the forward model error covariance matrix) retrieval accuracy is not compromised, even with the diagonal error approximation. Surface air temperature is poorly constrained: the corresponding element R_{ii} is close to zero and *a priori* and *a posteriori* variances differ by 0.01 K^2. Surface temperature retrievals will not be discussed further here.

The discussion thus far has focussed on the diagonal elements of the resolution matrix. Off diagonal structure is important as it determines the resolution of the measurement (in the sense of the Backus Gilbert measure of spread [18]). Specifically, the off diagonal elements R_{ij} describe how the departures $x_T - x_0|_j$ are mapped into the analysis increment $\hat{x} - x_0|_i$. A contour plot of the resolution matrix for the AFGL1 v13 O2 case is illustrated in Figure 9. Contributions are maximum along the diagonal axis of the matrix: typically departures within $\pm$ two levels make significant contributions to the analysis

increment for a given level. Averaging kernel widths at half maximum (FWHM) are consistent with the effective vertical resolution estimates given above. Note however how humidity departures contribute to tropospheric temperature increments, reflecting the fundamental ambiguity in interpretation of radiances from the ν_2 H_2O band (absorption by a variable gas).

Structures are similar for the AFGL5 atmosphere. The vertical resolution of the measurements is lower (this may be gauged qualitatively from the density of the contour lines in the resolution plot, and is again consistent with effective vertical resolution estimates) although the contribution of water vapour departures to temperature increments is reduced.

3.2.2 Comparison of the resolution matrices for the three F scenarios

The diagonal elements R_{ii} of the resolution matrices for block diagonal and diagonal forward model error covariance matrix approximations are also illustrated in Figure 7 (dashed and dotted curves respectively). In all cases there is no significant change to the diagonal elements of R on the introduction of the block diagonal approximation. There are some small changes to the diagonal components of R in the case of the diagonal approximation – principally a reduction in the 'resolution' of humidity in the mid troposphere and a small increase in the 'resolution' of upper stratospheric temperature (not shown). The significance of this increase is questionable: eigenvalue/eigenvector truncation tests indicate this effect is associated with/ dependent on the smallest eigenvalue/eigenvector components of the innovation error covariance $(B^{-1} + \nabla\mathcal{H}^T O^{-1} \nabla\mathcal{H})^{-1}$. The off-diagonal structure of the resolution matrices has also been visualised graphically for the three forward model error covariance scenarios (as in Figure 9). Qualitatively the structures are very similar – in particular, the temperature/water vapour mixing in temperature retrievals is a consistent feature for all cases.

A better measure of differences can be gained through comparison of the eigenvectors of the resolution matrices. The leading eigenvectors of the resolution matrix have been examined for the three forward model error covariance scenarios (full, block, diagonal error covariance matrices) for the v13 O2 observational error covariance scenario. Modifications in the leading eigenvectors of R are generally small for the full and block approximations. The most significant differences are associated with temperature structures at pressures less than 100 hPa ($\sim 16 - 17$ km). It is thought that this is due to the fact that anticorrelations between errors in the O_3 band and errors in the CO_2 and H_2O bands are not taken into account in the block diagonal specification of the F matrix (unmodelled correlations between the CO_2 bands and the H_2O ν_2 band may also play a role). Small modifications to the resolution of water vapour between 100 and 300 hPa are also observed. In one case (5th leading eigenvector, AFGL1) there are significant differences in lower tropospheric temperature and humidity in the 400 to 100 hPa region.

The eigenvalues for the resolution matrix for the diagonal F approximation are generally topologically similar to those for the full and block cases. However, the eigenvectors often present significant differences in the magnitude of the 'smoothing' structures. Structures can also be displaced (translation in the vertical). This, and the modification of the diagonal elements of the resolution matrix indicates that the diagonal F approximation can give quite different retrieval resolution and smoothing characteristics. We will return to this point in section 3.4.

The differences in the eigenvectors of R for full, block diagonal and diagonal F matrices are qualitatively similar for the AFGL1 and AFGL5 atmospheres. The fact that there are more leading eigenvectors of the full R matrix which are 'temperature only' modes for the AFGL5 atmosphere is however of note (Eigenvectors 1, 4, 5, 7, (9, 11) are all 'temperature only' modes, as compared to eigenvectors 8, 10, (11, 16, 17, 21) ... for AFGL1). This is consistent with the weaker temperature/water vapour mixing in temperature retrivals noted previously for the AFGL5 atmosphere.

3.2.3 Comparison of the retrieval variance for the three F 'truth' scenarios

The retrieval error covariance matrix has also been compared for the three forward model error covariance 'truth' scenarios. No significant change to the retrieval variance was found: for a given atmosphere and instrumental noise scenario, the retrieval variances were equivalent for all three forward model error scenarios for all practical purposes. Throughout the following sections retrieval variances and standard deviations will be illustrated for various approximations to the forward model error covariance matrix and for different Jacobian error scenarios. Unless stated explicitly, the full forward model error covariance is assumed as truth in all comparisons.
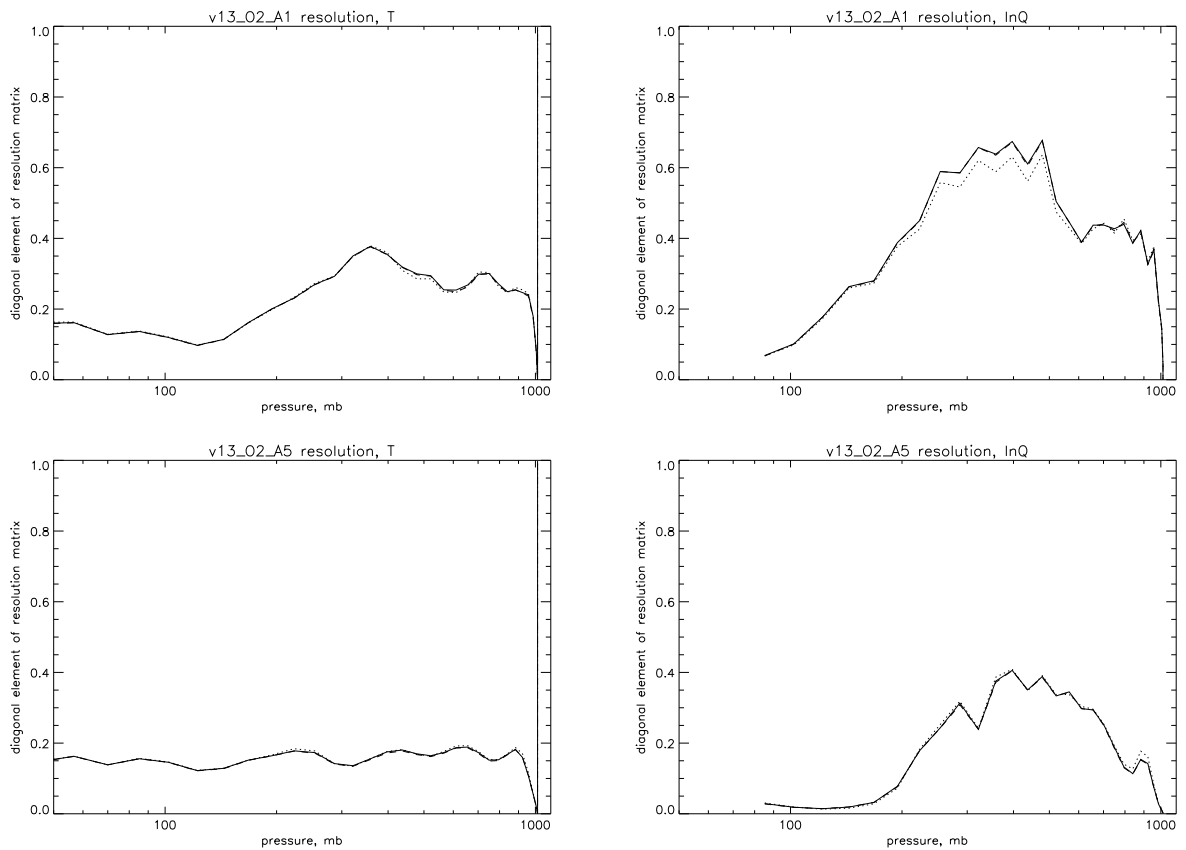
Figure 7: Contributions to degrees of freedom for signal for temperature and humidity for the AFGL1 tropical atmosphere (upper panels) and AFGL5 sub-arctic winter atmosphere (lower panels). Observational error covariance is given by the v13 O2 scenario. The solid curve illustrates the contributions to the degrees of freedom for signal for a full forward model error covariance specification. The dashed curve illustrates the diagonal elements of the resolution matrix for a block diagonal forward model error covariance matrix. The dotted curve illustrates the diagonal elements of the resolution matrix for a diagonal forward model error covariance matrix. Note the marked increases in the degrees of freedom for signal for mid and lower tropospheric temperature and tropospheric humidity in the tropical atmosphere.
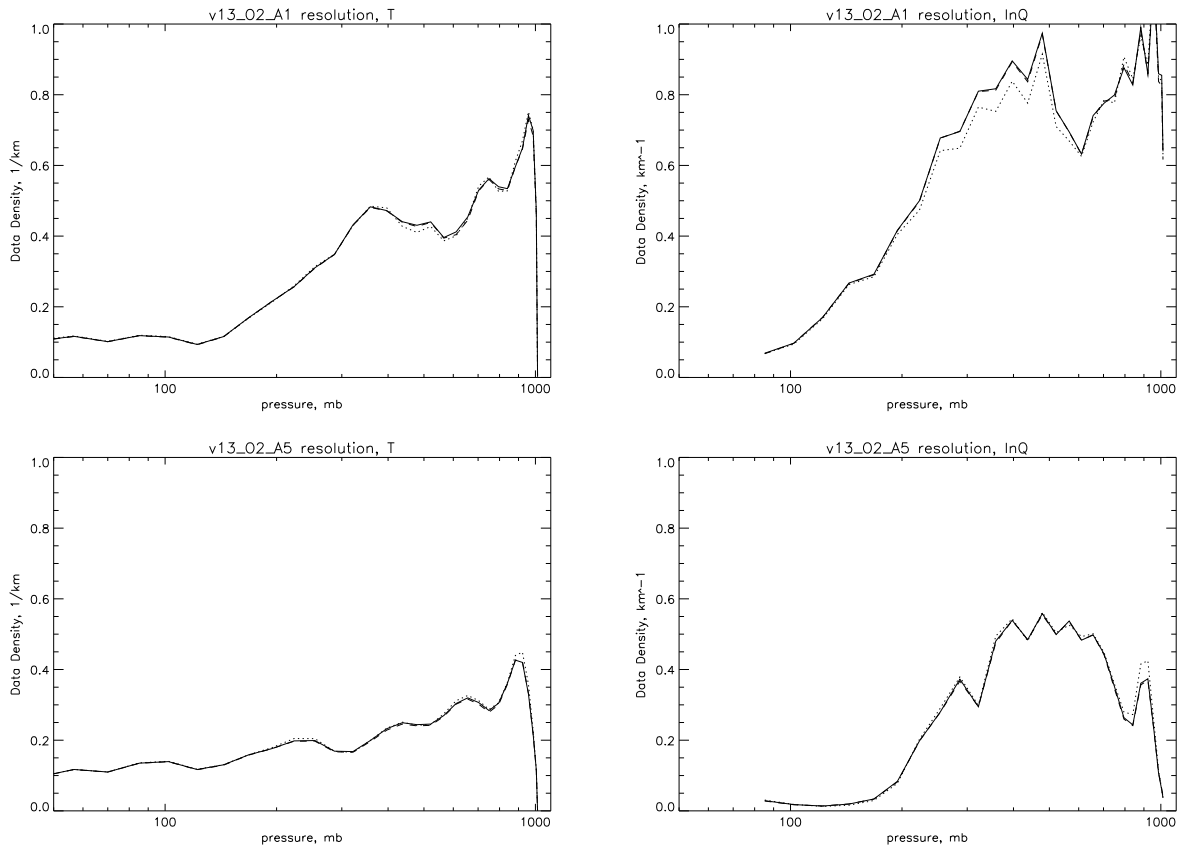
Figure 8: Data density for temperature and humidity for the AFGL1 tropical atmosphere (upper panels) and AFGL5 sub-arctic winter atmosphere (lower panels). Observational error covariance is given by the v13 O2 scenario. The solid curve illustrates the data density for a full forward model error covariance specification. The dashed and dotted curves illustrate $R_{ii}/\Delta z_i$ for block and diagonal forward model error covariance matrices respectively. A measure of the effective vertical resolution of the retrieval (in kilometres) is given by the reciprocal of the data density.
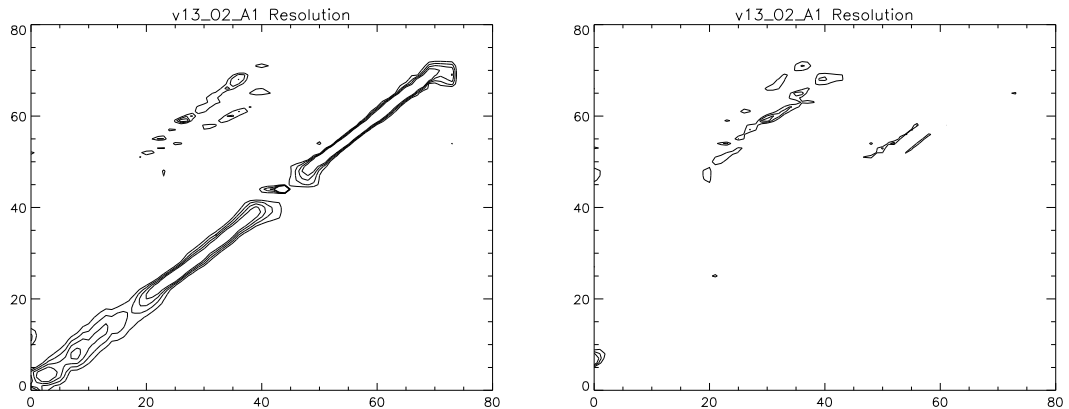

Figure 9: Contour plot representation of the resolution matrix for the AFGL1 v13 O2 scenario. The left panel illustrates contours for the positive elements of R_{ij} for values between 0.05 and 0.2 at contour intervals of 0.05. The right hand panel illustrates the negative elements of R_{ij} for values between -0.05 and -0.2 at 0.05 contour intervals. The rows of R are oriented vertically (strictly, R^T is imaged). Profile elements 0–44 are temperature (from the stratosphere to the surface), profile elements 45–73 are water vapour (from $\sim$ 80 hPa to the surface). Note the contributions from water vapour departures to tropospheric temperature increments.

17

3.3 Impact of simplifying approximations to the forward model error covariance matrix

We now consider the error inherent in two simplifying assumptions to the structure of the full forward model error covariance matrix: block diagonal and diagonal F matrices. As discussed previously, if an observation error covariance O is assumed in retrievals, i.e. assumed in the calculation of the gain matrix W, the retrieval error covariance is increased by a term $W(O - O^{'})W^T$ describing how the incorrect specification of the observation error covariance is mapped into retrievals.

The standard deviations of an ensemble of temperature and humidity retrievals are illustrated for the v13 O2 observation error covariance scenario and the two AFGL atmospheres in Figure 10. Full block and diagonal forward model error scenarios are illustrated as previously with full, dashed and dotted curves. The standard deviations of the *a priori* information are illustrated by the dot-dashed curve.

It is immediately apparant that the block diagonal approximation to the forward model error covariance matrix sucessfully captures all relevant correlation structure. The error introduced by this approximation is negligible for all practical purposes and this is true of all the other observational error covariance scenarios considered in this study. There is a small – arguably tolerable – degradation in retrieval accuracy associated with the diagonal approximation to the forward model error covariance matrix. This error is largest for the AFGL1 mid-troposphere temperature retrievals and the AFGL5 upper tropospheric humidity retrievals.

The breakdown of the null and propagated measurement error contributions for these cases are illustrated in Figure 11. The propagated measurement error contributions to the retrieval variance for full, block and diagonal forward model error scenarios are illustrated as previously with full, dashed and dotted curves. The null space contributions to the retrieval variance are illustrated by the dot-dashed curve. Null and measurement error contributions are comparable below 200 hPa in the AFGL1 case, and the error contribution due to the diagonal approximation is significant for temperature retrievals in the mid troposphere. Null space error dominates the error budget for the AFGL5 atmosphere, although the diagonal approximation makes a significant contribution to humidity measurement errors in the lower and upper troposphere. Note the higher measurement information content for the AFGL1 atmosphere leads to lower null space contributions, and overall, lower retrieval errors.

While the absolute magnitude of the error due to the diagonal F approximation is small, it can be a significant fraction of the reduction in error/uncertainty in the atmospheric state. The reduction in variance on assimilation of (IASI) observations can be usefully measured by the fraction of unexplained variance A_{ii}/B_{ii}. For the AFGL1 atmosphere, the fraction of unexplained variance is of $\sim 20\%$ for both temperature and humidity below 200 hPa. The diagonal forward model error approximation increases the fraction of unexplained variance by 5 to 10%. For comparison, the fraction of unexplained variance in the AFGL5 temperature retrieval is of the order of 60% in the upper troposphere, decreasing to 40% in the lower troposphere. The fraction of unexplained variance in humidity is of the order of 20% between 300 and 700 hPa, but increases rapidly above and below. The diagonal forward model error approximation increases the fraction of unexplained variance by $\leq 5\%$.

Obviously, where the information content of observations allows a significant reduction in variance, then the retrieval (retrieval accuracy) is more sensitive to the detail of the forward model error covariance specification. In the cases illustrated here the degradation is tolerable. Clearly one does not want the situation where the error amplification is such that the assimilation of observations results in a retrieval with greater uncertainty than the *a priori* state.

In Figure 12 we illustrate an equivalent analysis of retrieval standard deviations for the v11 O2 observation error covariance scenario. There is a clear increase in the errors in tropospheric temperature and humidity retrievals associated with diagonal F approximation for both atmospheres. Mid tropospheric temperature retrievals and humidity retrievals in the 100-300 hPa region are most affected in the AFGL1 atmosphere – in these regions propagated measurement errors are greater than or equal to null space errors, and the fraction of unexplained variance increases by ~ 20 to 40% for the diagonal F approximation. In the AFGL5 atmosphere humidity retrievals are significantly degraded in the 200 to 300 and 800 to 1000 hPa regions (amplification of the propagated measurement error structures illustrated in Figure 11(d)). In these intervals the fraction of unexplained variance increases by 20 to 40% for the diagonal F approximation. Increases in the fraction of unexplained variance for temperature are of the order of 5 to 10 % at tropospheric levels.

Finally, in Figure 13 we illustrate retrieval errors for the v11 O1 observation error covariance scenario. The increased instrumental noise gives rise to a slight degradation in retrieval accuracy for temperature and humidity at all levels (as compared to the v11 O2 scenario). With the increased weight of the diagonal elements of the observation error covariance matrix (due to the increase in instrumental noise) the diagonal error assumption is an increasingly reasonable approximation to the full observation error covariance matrix. The impact of the diagonal approximation on retrieval accuracy is small – propagated measurement errors are less than null space errors everywhere (with the exception of the 300 to 400 hPa region for AFGL1 temperature retrievals, where null and measurement errors are equal in magnitude). Errors are smaller still for the v13 O1 scenario (not illustrated) also for this reason.

When retrievals are performed with a suboptimal gain matrix retrieval accuracy is compromised and there is a corresponding reduction in measurement information content or degrees of freedom for signal. To complete the discussion above, Table 4 extends the tabulation of degrees of freedom for signal for the full F matrix given in Table 3 to the degrees of freedom for signal for the block diagonal and diagonal approximations to forward model error covariance matrix. Again, reductions in the degrees of freedom for signal are small – ≤ 0.2 in all cases – for the block diagonal approximation. Reductions in degrees of freedom for signal for the diagonal approximation range from 0.3 to 3.1, and are most marked for the v11 forward model in the O2 instrumental noise scenario.

Simulation	DFS full	DFS block	DFS diag	Simulation	DFS full	DFS block	DFS diag
v11 O1 A5	10.3	10.2	9.5	v11 O1 A1	16.3	16.2	15.5
v13 O1 A5	10.4	10.3	10.1	v13 O1 A1	16.4	16.3	16.1
v11 O2 A5	12.7	12.5	10.0	v11 O2 A1	21.5	21.3	18.2
v13 O2 A5	12.8	12.6	11.4	v13 O2 A1	21.6	21.4	20.1

Table 4: Degrees of freedom for signal for the full forward model error covariance specification and block diagonal and diagonal approximations.

In conclusion then, where the diagonal terms are the dominant elements in the observational error covariance matrix (e.g. $E_{ii}/F_{ii} \sim 4{:}1$) then a diagonal approximation to the forward model error covariance matrix – the unique source of long range error correlations – is adequate. The diagonal approximation might be considered unsatisfactory in cases where forward model and instrumental noise are comparable in magnitude in a spectral interval which is being given significant weight in determining analysis increments - water vapour absorption in the ν_2 band in the v11 O2 A1 scenario, for example. In this situation a block diagonal assumption appears a very good approximation to the full error covariance matrix and this will be true whenever errors present the same type of spectral/spectroscopic correlation structures as those considered here. Given that there may be modifications to the resolution matrix (averaging kernels) on passing from a full/block to a diagonal specification of the forward model error covariance matrix, it is probably well worth performing off-line channel selection studies with full or block F matrices.
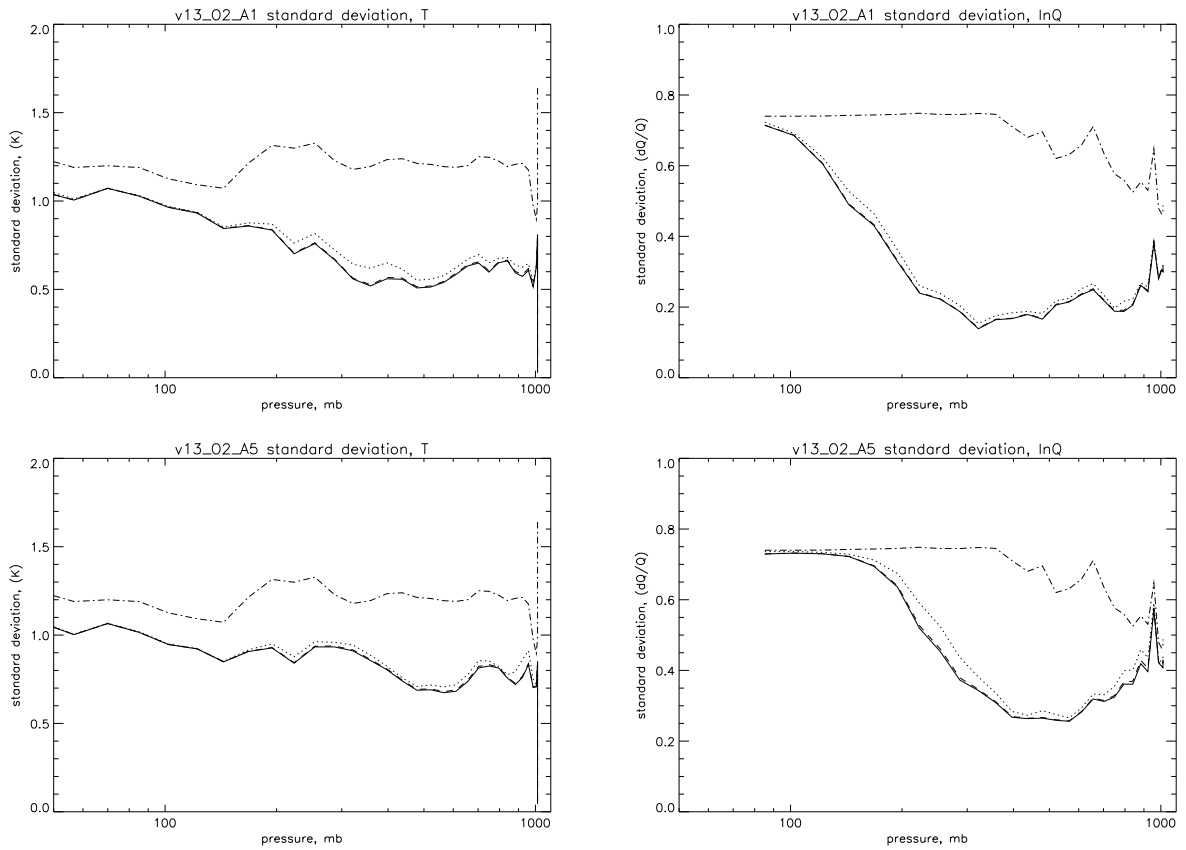
Figure 10: Retrieval standard deviation for temperature and humidity for the AFGL1 tropical atmosphere (upper panels) and AFGL5 sub-arctic winter atmosphere (lower panels). Observational error covariance is given by the v13 O2 scenario. The solid curve illustrates the retrieval error for a full forward model error covariance specification. The dashed curve illustrates the retrieval error for a block diagonal forward model error covariance matrix and the dotted curve illustrates the retrieval error for a diagonal forward model error covariance matrix. The upper dot-dashed curve illustrates the error in the *a priori* profile. The large reduction in skin temperature error and the small modification in surface air temperature uncertainties (P > 1000 hPa in left hand plots) are discussed in section 3.2.1.
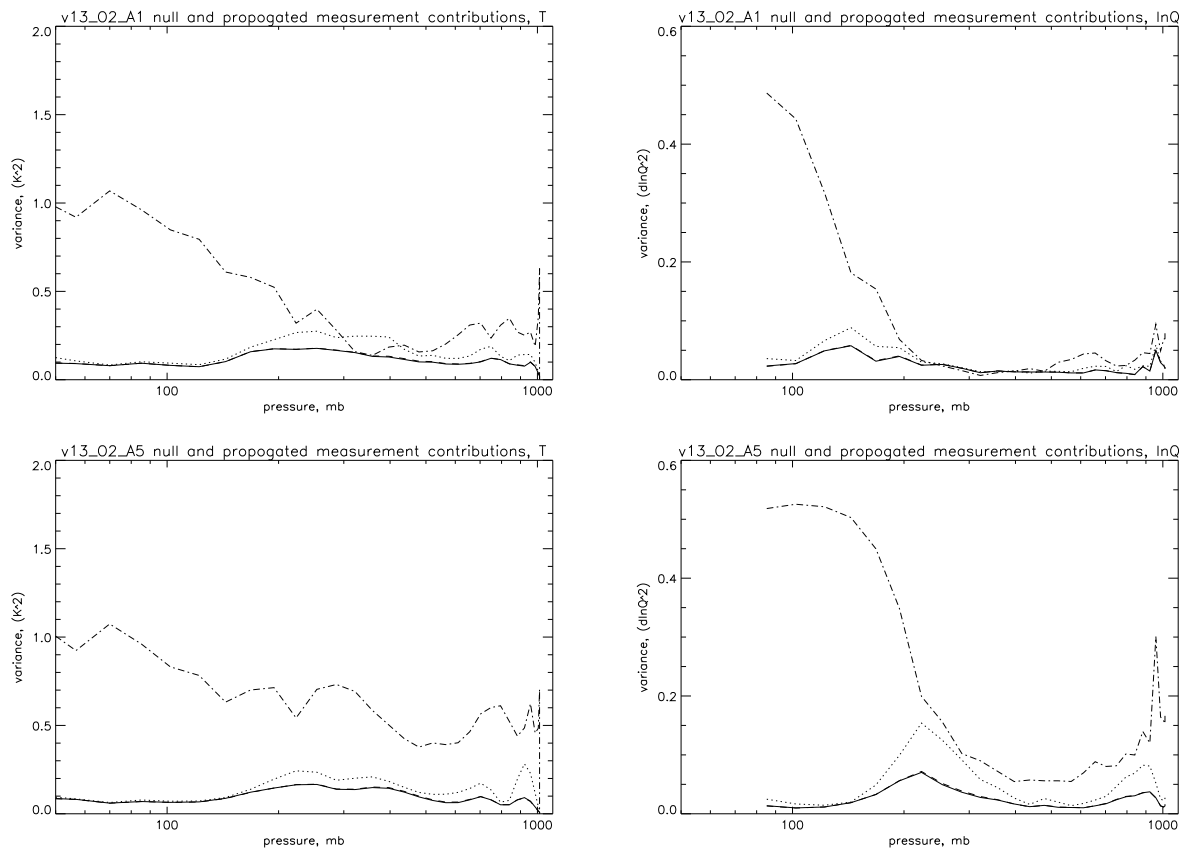
Figure 11: Null and propagated measurement error contributions to the retrieval variance for temperature and humidity for the AFGL1 tropical atmosphere (upper panels) and AFGL5 sub-arctic winter atmosphere (lower panels). Observational error covariance is given by the v13 O2 scenario. The solid curve illustrates the measurement contribution to retrieval variance for a full forward model error covariance specification. The dashed curve illustrates this measurement contribution for a block diagonal forward model error covariance matrix and the dotted curve illustrates this measurement contribution for a diagonal forward model error covariance matrix. The dot-dashed curve illustrates the null space contribution to retrieval variance.
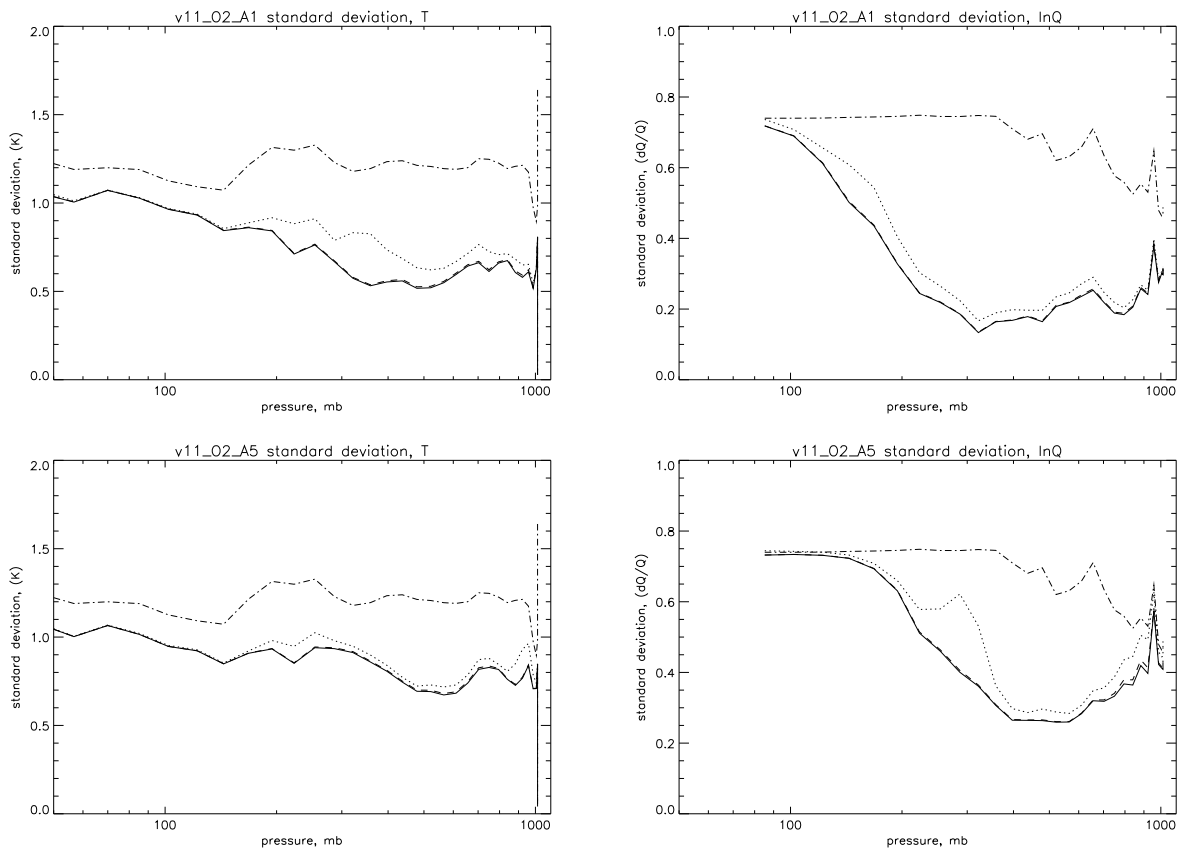
Figure 12: Retrieval standard deviation for temperature and humidity for the AFGL1 tropical atmosphere (upper panels) and AFGL5 sub-arctic winter atmosphere (lower panels). Observational error covariance is given by the v11 O2 scenario. The solid curve illustrates the retrieval error for a full forward model error covariance specification. The dashed curve illustrates the retrieval error for a block diagonal forward model error covariance matrix and the dotted curve illustrates the retrieval error for a diagonal forward model error covariance matrix. The upper dot-dashed curve illustrates the error in the *a priori* profile.
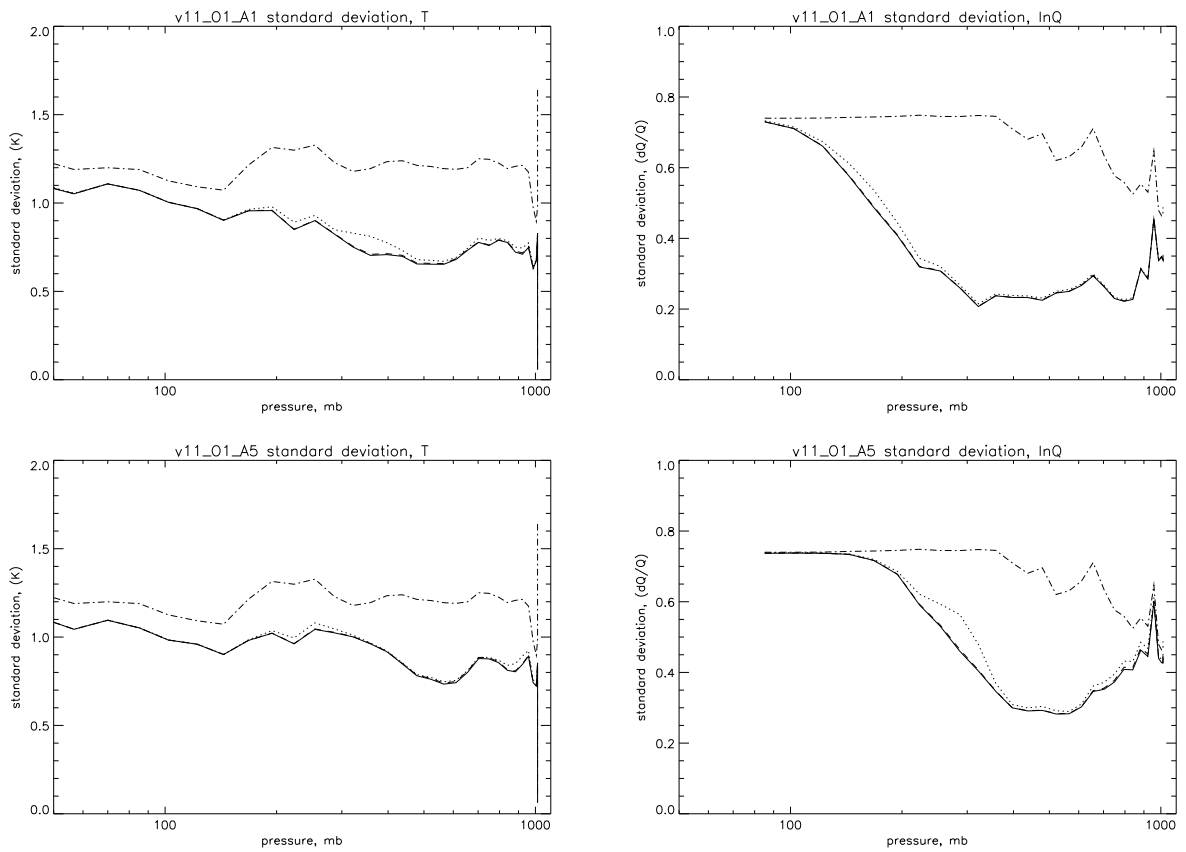
Figure 13: Retrieval standard deviation for temperature and humidity for the AFGL1 tropical atmosphere (upper panels) and AFGL5 sub-arctic winter atmosphere (lower panels). Observational error covariance is given by the v11 O1 scenario. The solid curve illustrates the retrieval error for a full forward model error covariance specification. The dashed curve illustrates the retrieval error for a block diagonal forward model error covariance matrix and the dotted curve illustrates the retrieval error for a diagonal forward model error covariance matrix. The upper dot-dashed curve illustrates the error in the *a priori* profile.

Simulation	DFS R_F	DFS R_J	Simulation	DFS R_F	DFS R_J
v11 O1 A5	10.3	9.1	v11 O1 A1	16.3	12.2
v13 O1 A5	10.4	9.1	v13 O1 A1	16.4	12.5
v11 O2 A5	12.7	11.1	v11 O2 A1	21.5	14.9
v13 O2 A5	12.8	11.0	v13 O2 A1	21.6	15.8

Table 5: Degrees of freedom for signal for the 1057 channel full F matrix simulations R_F, and 163 channel reference (truth) Jacobian simulations R_J=W.H, W≡W(H).

3.4 Impact of errors in modelled Jacobians

In this section we consider the impact of errors in fast model Jacobian on retrieval accuracy. As described previously, Jacobian errors impact on the null space error, giving a non-optimal propagation (an incorrect mapping/interpretation) of *a priori* error in the retrieval.

In order to evaluate the impact of errors in RTIASI Jacobians GENLN2 finite difference Jacobian calculations are taken as truth. To date GENLN2 Jacobians have only been calculated on three targetted subintervals: $645 - 805$ cm^{-1}, $885 - 915$ cm^{-1} and $1300 - 1450$ cm^{-1}. Moreover, water vapour Jacobians have not been evaluated for the first subinterval. The comparisons illustrated here are performed for the 2 cm^{-1} sampled spectra on these three subintervals: 163 channels are analysed in total, reference water vapour Jacobians exist for 86 channels. Only 'well-behaved' water vapour Jacobians are considered, but relative errors of 10-40% in the magnitude of Jacobian elements are still possible with the v13 version of RTIASI (errors are $\leq 10\%$ for the v11 version).

The degrees of freedom for signal for retrievals using the full 1057 channels and retrievals using the 163 channels on the selected spectral subintervals are compared in Table 5. In the latter case, degrees of freedom for signal have been calculated for the truth scenario (R_J=W.H; W≡W(H), H truth). The resolution matrices are denoted R_F and R_J respectively. For the AFGL5 atmosphere the reduction in degrees of freedom for signal in passing from 1057 to the targetted 163 channels is relatively small: only 1.2 to 1.6 degrees of freedom for signal are lost, mostly in humidity. The reductions in degrees of freedom for signal in passing from the 1057 to the targetted 163 channels are greater for the AFGL1 case: $4 - 7$ degrees of freedom for signal are lost. These reductions affect temperature below 200 hPa and humidity below 100 hPa. Reductions in degrees of freedom for humidity are slightly greater than those for temperature.

The question is then raised as to how the results for the subset of channels considered may be generalised to compare with cases with a larger number of channels, or more precisely, a larger number of degrees of freedom for signal. It is reasonable to assume that Jacobian error characteristics will not differ greatly for a larger spectral range: it would appear that errors are associated with distinct absorption features/regimes[4] and hence a result of shortcomings in the fast model formulation. In this case, extension of the spectral interval considered for retrieval is unlikely to improve Jacobian error characteristics (note this is not necessarily true of spectroscopic errors – spectral intervals where line and continuum absorption parameters are more or less well known are likely to exist). With increased degrees of freedom for signal, the null-space contribution to retrieval errors will be smaller. Thus in principle the studies presented here represent a worst case scenario. The results for the AFGL5 atmosphere are presumably closer to the truth (for this linearisation case) as the reduction in degrees of freedom signal is significantly lower than that for the AFGL1 case.

In Figure 14 we illustrate the impact of Jacobian errors on retrieval accuracy for the v13 RTIASI model for the AFGL1 and AFGL5 atmospheres. Three error scenarios are considered: temperature Jacobian errors only (dashed curve), water vapour errors only (dotted curve) and combined temperature and water vapour Jacobians (dot dashed curve). As previously, retrieval errors for the reference simulations are traced with a solid line. The upper triple-dot dashed curve represents the *a priori* error.

In all cases errors in temperature Jacobians have a negligible impact on retrieval accuracy. Errors in water vapour Jacobians have a significant impact on retrieval accuracy for both temperature and humidity in the AFGL1 atmosphere and account for essentially all the degradation in retrieval accuracy in the combined temperature and water vapour Jacobian error case. The observed degree of degradation of the temperature retrieval is of particular concern. In the AFGL5 atmosphere errors in water vapour Jacobians only have a significant impact on the accuracy of humidity retrievals, and even in this case the degradation is small.

The differences in the effects of Jacobian errors on retrieval accuracy in the two atmospheres may be attributed to the marked differences in the magnitude of Jacobian errors in the humid and dry cases (relative errors of $10 - 40\%$ and ≤ 10 % respectively). Resolution effects must also be taken into account however: errors in water vapour Jacobians degrade temperature retrievals because water vapour departures contribute to tropospheric temperature increments, as described previously. R_J and the resolution matrix for the approximation scenario $R_J^\star$ ($R_J^\star$=W.H$'$; W≡W(H), H$'$ truth) are contoured in Figure 15 and Figure 16 for the AFGL1 and AFGL5 atmospheres respectively. In each case the largest modification to the resolution matrices occurs in the off-diagonal contributions from water vapour departures to temperature increments. In absolute terms, modifications are smaller for the AFGL5 atmosphere, consistent with a (much) smaller degradation in the accuracy of temperature retrievals, although lower errors in the v13 AFGL5 water vapour Jacobians are probably the major factor in this instance.

---

[4]Note this implies that Jacobian errors are spectrally correlated.

Simulation	DFS R_J	DFS T	DFS H	DFS T+H	Simulation	DFS R_J	DFS T	DFS H	DFS T+H
v11 O1 A5	9.1	8.9	8.2	8.0	v11 O1 A1	12.2	12.2	10.6	10.5
v13 O1 A5	9.1	9.0	8.8	8.7	v13 O1 A1	12.5	12.4	9.6	9.5
v11 O2 A5	11.1	10.9	9.2	9.0	v11 O2 A1	14.9	14.8	10.9	10.8
v13 O2 A5	11.0	10.9	10.2	10.0	v13 O2 A1	15.8	15.7	7.7	7.5

Table 6: 163 channel Jacobian calculations. Degrees of freedom for signal for for reference calculations ($Tr(R_J)$) and for the Jacobian error cases T: temperature Jacobian errors only, H: water vapour Jacobian errors only, T+H: all Jacobian errors.

As discussed previously (section 3.2), the diagonal approximation to the forward model error covariance matrix can modify the structure (eigenmodes) of the resolution matrix. How does this affect the 'propagation' of Jacobian errors and retrieval accuracy ? In the AFGL1 v13 O2 case, a diagonal approximation to F reduces the error in temperature retrievals (due to errors in modelled water vapour Jacobians) in the upper troposphere by ~ 0.1 K. There is no significant impact in any of the other retrievals illustrated here.

In Figure 17 we illustrate the corresponding Jacobian error analysis for the RTIASI v11 model. There is an overall reduction in errors in the AFGL1 case, consistent with the improvements in the modelled water vapour Jacobians obtained with the v11 RTIASI model for this atmosphere. In the AFGL5 case there is a reduction in retrieval accuracy associated with the increase in errors in modelled v11 water vapour Jacobians for this atmosphere. The degradation in temperature retrievals is however negligible for all practical purposes. Water vapour retrieval accuracies between 500 and 300 hPa are degraded by $\sim5\%$, while water vapour retrieval accuracies for pressures less than 300 hPa are degraded by $\sim10\%$.

To complete the discussion above, the degrees of freedom for signal for retrievals in the presence of Jacobian errors are tabulated in Table 6. Again, temperature and water vapour Jacobian errors are considered separately (columns T and H in Table 6) and conjointly (columns T+H in Table 6) for the eight atmospheric and observation error covariance scenarios.

Errors in temperature Jacobians have a minimal impact on the degress of freedom for signal in all cases – errors in water vapour Jacobians on the other hand can significantly reduce the degrees of freedom for signal. In the AFGL5 atmosphere water vapour Jacbian errors give a reduction of 0.3 to 2.0 degrees of freedom for signal relative to the 163 channel reference calculations. Reductions are greatest for the O2 instrumental scenario (i.e. where more weight is given to observations) and the v11 model. In the AFGL1 atmosphere water vapour Jacobians give reductions of 1.6 to 8.0 degrees of freedom for signal. Again, reductions are greatest for the O2 instrumental noise scenario. In this case, errors in modelled water vapour Jacobians halve the degrees of freedom for signal for the v13 model relative to the 163 channel reference calculations. Even with the improvements in the decription of the water vapour Jacobians in the tropical atmosphere with the v11 model, errors in modelled water vapour Jacobians give a loss of 4 degrees of freedom for signal.

To conclude then, the impact of errors in temperature Jacobians on retrieval accuracy is negligible, suggesting a target accuracy of $\leq5\%$ for relative errors in fastmodel Jacobians. Accurate water vapour Jacobians are critical for upper tropospheric temperature and humidity retrievals, and this is the only instance where the adequacy of the RTIASI fast model is seriously called into question. If the results illustrated here can be generalised to the range of humid and dry atmospheres encountered in reality, then it would appear that the v11 RTIASI model with a full or block diagonal forward model error covariance specification is the most appropriate choice of the two versions of the RTIASI fast model tested here. This hypothesis (generalisation of results to wider range of states) clearly needs to be fully tested, implying the extension of GENLN2 Jacobian calculations to a wider range of atmospheric states on the full IASI spectral interval. If these results indicate the existence of channels with consistently high Jacobian errors (for a wide range of atmospheric states) then the possibility of channel exclusion should also be explored – i.e. the impact of channel exclusion on retrieval accuracy and information content should be assessed. Improvements in modelled water vapour absorption are still necessary, even with the RTIASI v11 model. It would be of interest to evaluate to what extent the dry bias in the set of profiles used to generate the RTIASI transmittance predictor coefficients affects the accuracy of modelled water vapour Jacobians. If the errors due to the regression profile set are not significant, then it would appear that new methods for fast water vapour transmittance calculations must be considered (e.g. OPTRAN) – for in this case only a new methodology could give the required improvements in Jacobian accuracy.

v13_02_A1_0 standard deviation, T

v13_02_A1_0 standard deviation, lnQ

v13_02_A5_0 standard deviation, T

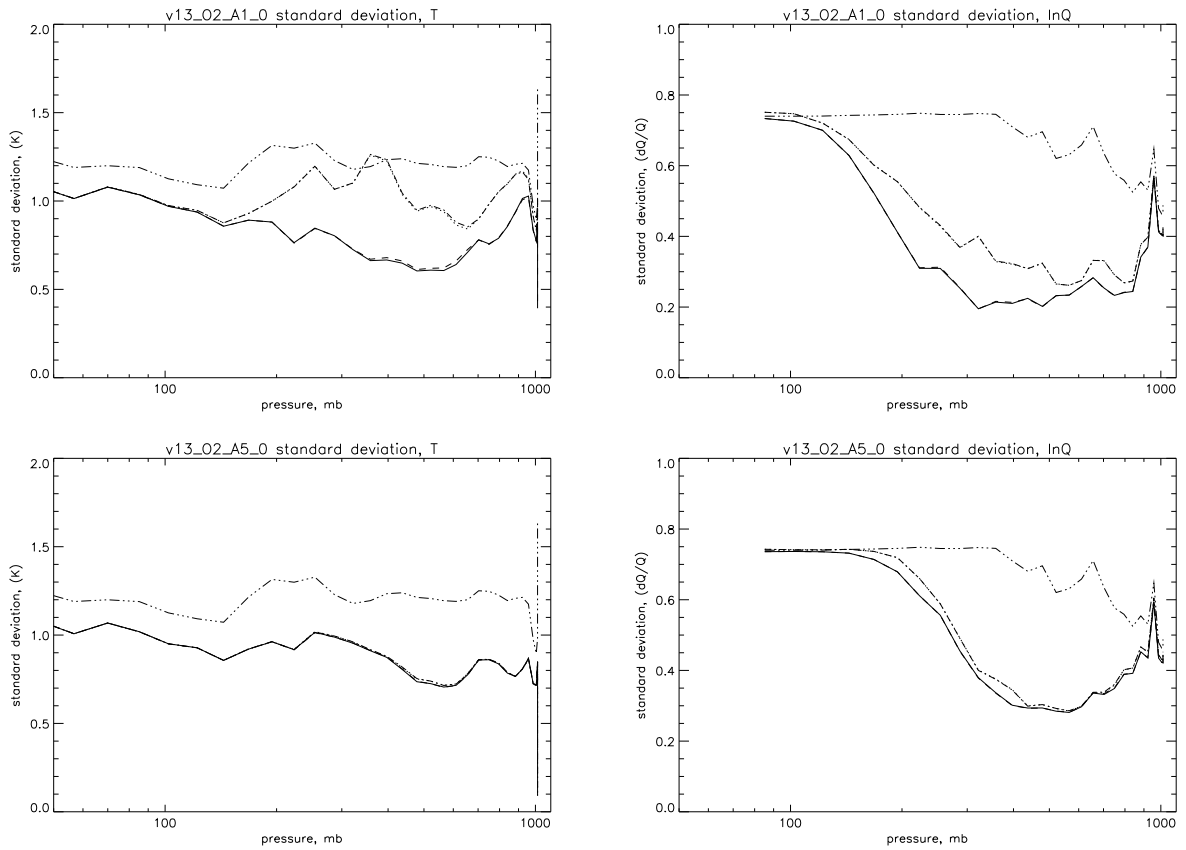v13_02_A5_0 standard deviation, lnQ

Figure 14: Retrieval standard deviations for RTIASI v13 Jacobian error scenarios: dashed curve – temperature jacobian errors only, dotted curve – water vapour jacobian errors only, dot dashed curve – temperature and water vapour jacobian errors. As previously, the solid curve is the reference or truth retrieval error and the upper triple-dot dashed curve is the *a priori error*. Upper panels are for the AFGL1 atmosphere, lower panels are for the AFGL5 atmosphere. The oberservation error covariance is given by the O2 instrumental error scenario and full v13 F matrix. Water vapour Jacobians are the dominant source of error. Thus the water vapour Jacobian error curves (dotted) and total Jacobian error curves (dot dashed) are often quasi-coincident and difficult to distinguish.
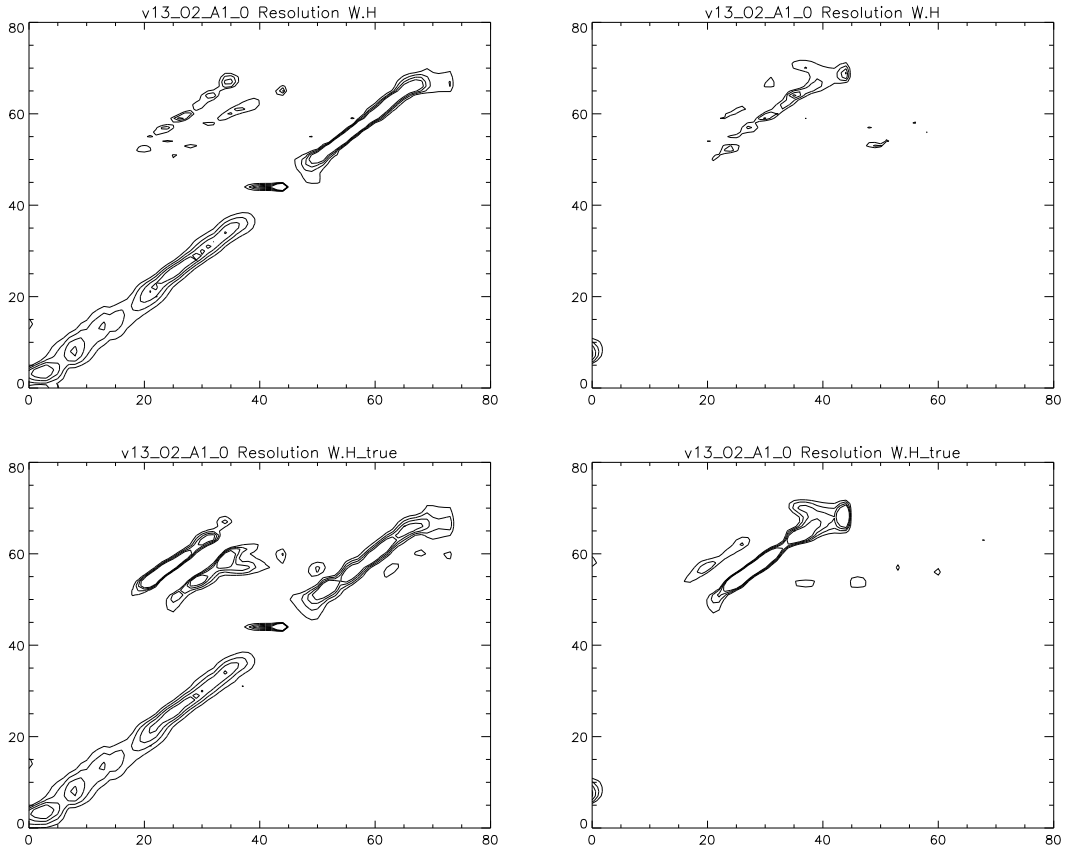
Figure 15: Contour plot representations of the resolution matrix for the AFGL1 v13 O2 scenario. R=W.H is illustrated in the upper panels, R=W.H$_{\mathrm{true}}$ is illustrated in the lower panels. In each case the left panel illustrates contours for the positive elements of R$_{\mathrm{ij}}$ for values between 0.05 and 0.2 at contour intervals of 0.05. The right hand panel illustrates the negative elements of R$_{\mathrm{ij}}$ for values between -0.05 and -0.2 at 0.05 contour intervals. The rows of R are oriented vertically (strictly, R$^{\mathrm{T}}$ is imaged). Profile elements 0–44 are temperature (from the stratosphere to the surface), profile elements 45–73 are water vapour (from $\sim$ 80 hPa to the surface). Note the enhancement of the contributions from water vapour departures to temperature increments in the R=W.H$_{\mathrm{true}}$ scenario.
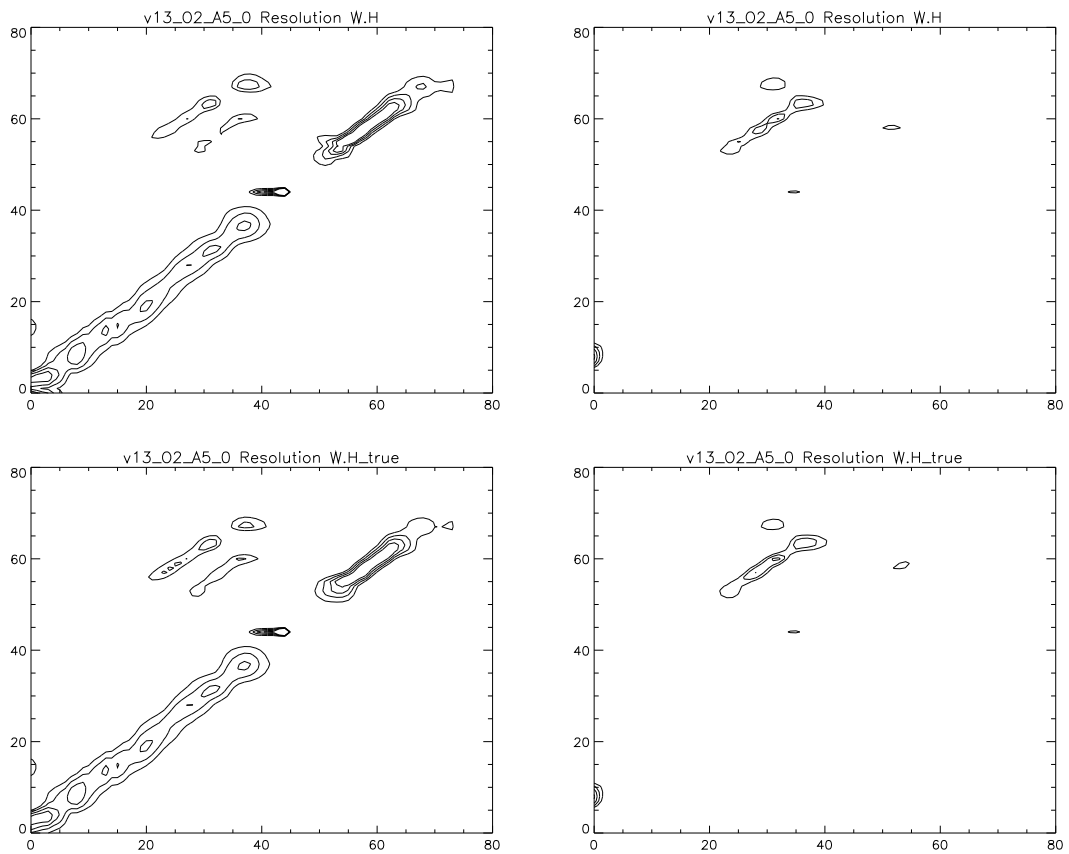
Figure 16: Contour plot representations of the resolution matrix for the AFGL5 v13 O2 scenario. R=W.H is illustrated in the upper panels, R=W.H$_{\mathrm{true}}$ is illustrated in the lower panels. In each case the left panel illustrates contours for the positive elements of R$_{\mathrm{ij}}$ for values between 0.05 and 0.2 at contour intervals of 0.05. The right hand panel illustrates the negative elements of R$_{\mathrm{ij}}$ for values between -0.05 and -0.2 at 0.05 contour intervals. The rows of R are oriented vertically (strictly, R$^{\mathrm{T}}$ is imaged). Profile elements 0–44 are temperature (from the stratosphere to the surface), profile elements 45–73 are water vapour (from $\sim$ 80 hPa to the surface).
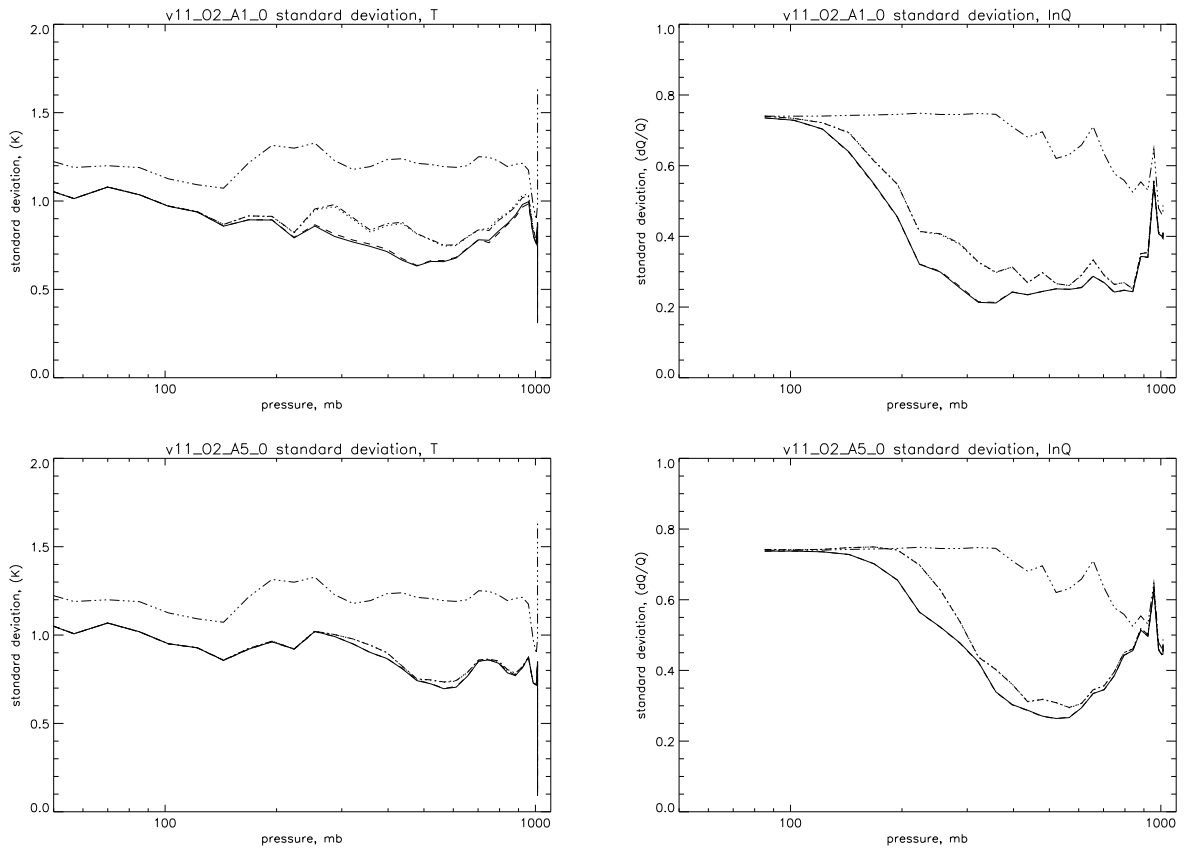
Figure 17: Retrieval standard deviations for RTIASI v11 Jacobian error scenarios: dashed curve – temperature jacobian errors only, dotted curve – water vapour jacobian errors only, dot dashed curve – temperature and water vapour jacobian errors. As previously, the solid curve is the reference or truth retrieval error and the upper triple-dot dashed curve is the *a priori* error. Upper panels are for the AFGL1 atmosphere, lower panels are for the AFGL5 atmosphere. The observation error covariance is given by the O2 instrumental error scenario and full v11 F matrix. Water vapour Jacobians are the dominant source of error. Thus the water vapour Jacobian error curves (dotted) and total Jacobian error curves (dot dashed) are often quasi-coincident and difficult to distinguish.

4 Conclusions and recommendations

Accurate estimates of RTIASI-specific forward model error covariance have been made for two versions of the RTIASI model, v13 and v11, using three-case and single-case water vapour predictor schemes respectively. Errors present a high degree of correlation within spectral bands and within the window regions. The only significant difference between the two schemes occurs in the H_2O ν_2 band, where the v11 scheme forward model errors are higher. However, detailed examination of the errors indicates that the largest degradation occurs for channels which were already relatively poorly modelled in the three-case water vapour predictor scheme. The error correlation structure also differs quite significantly between the two models in the H_2O ν_2 band.

The inter-channel correlation structure of instrumental noise is localised spectrally to fourth nearest neighbours. Thus long-range correlation structure is governed by the forward model error contribution to the observation error covariance matrix. Forward model errors are comparable with revised CNES noise estimates in the $1250 - 1800$ cm^{-1} interval – the H_2O ν_2 band. Accordingly, in this interval forward model error covariance makes a significant contributions to the off-diagonal components of the observation error covariance matrix.

Studies have been undertaken to assess the impact of approximating and/or neglecting forward model error correlations on retrieval accuracy. Where instrumental noise is the dominant term in the observation error covariance, a diagonal (band diagonal) approximation to the forward model error covariance matrix is adequate. The results of the studies presented here suggest that for $E_{ii}/F_{ii} \geq 4{:}1$ a diagonal approximation does not introduce significant error in retrievals. The adequacy of a diagonal approximation for spectral intervals where forward model error and instrumental noise are comparable in magnitude and which are given significant weight in determining analysis increments is less clear. If correlated errors in the H_2O ν_2 band are neglected, temperature and humidity retrievals in the mid and upper troposphere are degraded. In the cases considered here this never leads to retrieval errors which are larger than the uncertainty in the *a priori* estimate of atmospheric state. However, the benefit of the data assimilation process – as measured by the reduction of variance or fraction of unexplained variance – can be significantly reduced, particularly in the case of temperature retrievals. In all cases, the degradation of retrieval accuracy is negligible when a block diagonal approximation to the forward model error covariance matrix is assumed in the calculation of the gain matrix.

Spectroscopic uncertainties and representativity errors are expected to give additional correlated errors in the H_2O ν_2 band – this is also true of the CO_2 and O_3 bands and the window regions (continuum absorption). Although it is not yet clear what the magnitude of these contributions to the observation error covariance (after bias correction) will be, given the current relative magnitudes of forward model error and instrumental noise, it would appear that some treatment of error correlation in the H_2O ν_2 band will probably be required. The validity of a (band) diagonal approximation to the error covariance in other spectral regions will again depend on the magnitude of contributions to the covariance due to spectroscopic uncertainties and representativity errors. Even in the revised CNES scenario instrumental errors are the dominant contribution to O: with the exception of the $750 - 850$ cm^{-1} interval $E_{ii}/F_{ii} \geq 4{:}1$. Thus one can be reasonably optimistic that a (band) diagonal approximation will be appropriate for the CO_2 bands. Treatment of error in the window region will depend on preprocessing – cloud clearing and application of a land/sea mask – and here too a treatment of error correlation may be required.

Discussion has focussed on the specification of the observation error covariance for IASI. If a similar forward model is adopted for AIRS, there is no reason for forward model error characteristics to be substantially modified. The instrumental noise characteristics of the AIRS instrument are quite different – noise is uncorrelated from channel to channel, noise levels are much lower throughout most of the IASI spectral interval[5]. In this case, forward model errors will be the dominant component of the observation error covariance matrix, and the unique source of correlated error. Retrieval error covariance estimates have been made for an AIRS noise scenario. As previously, when error correlations are neglected in the H_2O ν_2 band tropospheric temperature and humidity retrievals are degraded by $0.1 - 0.2$ K, 5% in dq/q for the v13 scenario, and $0.2 - 0.4$ K, 10% in dq/q for the v11 scenario. Stratospheric temperature retrievals are also slightly degraded (≤ 0.1 K) reflecting the influence of error correlation in the CO_2 bands. In the AFGL tropical atmosphere this increase in retrieval error corresponds to an approximate doubling in the fraction of unexplained variance in the mid and upper troposphere – the diagonal approximation is a tolerable but distinctly sub-optimal approximation to the forward model error covariance matrix.

Jacobian error estimates have been made for the v11 model, complementing v13 model Jacobians calculated previously. For the two atmospheres considered, the v11 water vapour Jacobians were more accurate than v13 water vapour Jacobians in the AFGL tropical atmosphere. The inverse is true for the AFGL sub-arctic winter atmosphere. Temperature Jacobians are essentially unmodified in the v11 scheme.

These estimates have been used to quantify the impact of Jacobian errors on retrieval accuracy. The impact of errors in temperature Jacobians on retrieval accuracy is negligible, suggesting a target accuracy of $\leq 5\%$ for relative errors in fast model Jacobians. Accurate water vapour Jacobians are critical for mid and upper tropospheric temperature and humidity retrievals, and this is the only instance where the adequacy of the RTIASI model is called into question. Errors in the v13 model water vapour Jacobians for the AFGL tropical atmosphere ($10 - 40\%$ errors) give significant degradation in retrieval accuracy, and in some situations degrades the *a priori* estimate of temperature. The v11 model AFGL1 water vapour Jacobians are more accurate than the v13 Jacobians, and the impact of Jacobian errors on retrieval accuracy is reduced by about a factor two.

---

[5]AIRS instrumental noise is of the order of 0.05 K (NdTB, 280K scene temperature) throughout the $1000 - 2300$ cm^{-1} interval and less than 0.2 K between 2300 and 2700 cm^{-1}. Between 650 and 750 cm^{-1} AIRS noise levels are comparable with the original CNES IASI noise estimates and comparable or less then CNES revised noise estimates between 750 and 1000 cm^{-1}.

The v11 performance in the sub-arctic winter case is poorer and the contribution to errors in water vapour retrievals from Jacobian errors is approximately twice that of the AFGL5 v13 case. The impact of Jacobian errors on temperature retrieval accuracy is small in all cases for the AFGL5 atmosphere.

If the results presented here can be generalised to the range of humid and dry atmospheres encountered in reality, then it would appear that the v11 RTIASI model with a full or block diagonal forward model error covariance specification is the most appropriate choice of the two RTIASI versions tested. This hypothesis (generalisation of results to a wider range of atmospheric states) needs to be fully tested, and will require the extension of GENLN2 Jacobian calculations to a wider range of states on the full IASI spectral interval. Such an undertaking is computationally demanding, so simulation parameters (e.g. spectral and vertical resolution, magnitudes of parameter perturbations for finite difference Jacobian estimates) should be selected to ensure that these Jacobian calculations may be used as a general reference set.

Improvements to modelled water vapour Jacobians are still necessary, even for the v11 model. Results from VS2000 suggest avenues to be explored. Finite difference water vapour Jacobians for the PFAAST model were comparable in accuracy with RTIASI temperature Jacobians. Now, the PFAAST model is based on the same pressure level optical depth formalism as RTIASI, but differs in the detail of its implementation. Four of these implementation differences are susceptible to account for differences in modelled water vapour: the profile set used to derive predictor regression coefficients, choice of predictors, separation of absorption contributions from water vapour lines and continuum and model vertical resolution.

In the first instance, an assessment of the impact of the dry bias in the set of profiles used to generate the RTIASI transmittance predictor coefficients on the accuracy of modelled water vapour absorption (Jacobian and forward model errors) should be undertaken, as this would appear a logical source for the errors observed. If this is not a major source of error, then the impact of model vertical resolution and the details of the implementation of the predictor scheme should be examined. Should none of these studies give the required improvements, new methods for fast water vapour transmittance calculations (forward and adjoint) will have to be considered.

The results presented here highlight the importance of accurate radiative transfer calculations for the H_2O ν_2 band when undertaking the simultaneous retrieval of temperature and humidity. Based on preliminary intercomparisons of line-by-line radiative transfer calculations and AERI and HIS measurements undertaken in the framework of the ISSWG line-by-line intercomparison study (Peter Rayer, private communication) one might expect quite large errors (of the order of Kelvin) in the H_2O ν_2 band. Whether due to spectroscopic uncertainties or representativity errors, forward model errors of this magnitude imply errors in modelled Jacobians of at least the magnitude considered here. Similarly, the nonlinearity of the radiative transfer in the H_2O ν_2 band (with respect to perturbations of the magnitude of the uncertainty in the *a priori* estimate of water vapour concentrations) will give rise to error in modelled Jacobians and estimates of the gain matrix. The optimal use of high resolution sounder observations will necessitate improvements to water vapour modelling in the largest sense of the problem (improvements to fast models, improved spectrocopic parameter and improved NWP descriptions of the spatio-temporal distribution of water vapour, particularly in the upper troposphere). Indeed, it may be beneficial to make conservative use of IASI data in a day-one assimilation scheme, through the use of CO_2 temperature sounding information only.

A Error analysis for the gain matrix calculation

We wish to evaluate bounds for errors in analysis increments x due to the the propagation of roundoff error when evaluating the gain matrix using Cholesky decomposition. We consider the difference x-x$'$ of the analysis increments x=Wy and x$'$=W$'$y in the spectral norm (denoted $||.||$):

$$||x - x'|| = ||(W - W')y|| \leq ||W - W'||||y||, \tag{9}$$

where the inequality follows from the relation

$$||AB||_p \leq ||A||_p ||B||_p;\ A \in \Re^{m \times n} B \in \Re^{n \times q}, \tag{10}$$

i.e. for a given observation vector y, the spectral norm of x-x$'$ is bounded above by $||W - W'||||y||$.

Denoting the innovation error covariance $(B^{-1} + H^T O^{-1} H)$ by the letter I, the gain matrix W can be expressed as:

$$W = I^{-1} H^T O^{-1}. \tag{11}$$

To first order, W - W$'$ is governed by roundoff errors dI and dO in the evaluation of the Cholesky decomposition of I and O respectively:

$$\begin{aligned} W' &= (I + dI)^{-1} H^T O^{-1}, \\ W' &= I^{-1} H^T (O + dO)^{-1}, \end{aligned} \tag{12}$$

and

$$\begin{aligned} ||W - W'|| &\leq ||I^{-1} - (I + dI)^{-1}||||H^T O^{-1}||, \\ ||W - W'|| &\leq ||I^{-1} H^T||||O^{-1} - (O + dO)^{-1}||. \end{aligned} \tag{13}$$

In order to evaluate $||A^{-1} - (A + dA)^{-1}||$ for a general square matrix A, we apply the following theorem [7]:

If A is nonsingular and r=$||A^{-1}dA|| < 1$ then $||A^{-1} - (A + dA)^{-1}|| \leq ||dA||||A^{-1}||^2/(1 - r)$.

Furthermore, application of equation 10 gives the following relation for r: r=$||A^{-1}dA|| \leq ||A^{-1}||||dA||$. All we require now is a bound for the spectral norm of the error or perturbation dA.

The Cholesky decomposition of a symmetric positive definite matrix A, $LL^T = A + dA$ has well defined bounds for the error in the decomposition dA. Kielbasinski [8] derives error bounds for the spectral norm of dA:

$$||dA|| = 1.06.2^{-t}(n + 1 + 2\ln\sqrt{n})\sqrt{n}||A||, \tag{14}$$

where n is the rank of A and 2^{-t} is the unit roundoff error. For large n and small unit roundoff error, Equation 14 can be expressed as:

$$||dA|| = 2^{-t} n^{\frac{3}{2}} ||A||. \tag{15}$$

Using Kielbasinski's results we find $||A^{-1}||||dA|| \leq 2^{-t} n^{\frac{3}{2}} \kappa_2(A)$ where $\kappa_2(A) = ||A||||A^{-1}||$ is the spectral norm condition number of A. We apply this last relation to determine wether the criterion r < 1 holds for the innovation and observation error covariance matrices. Calculations are performed at real8 precision – unit roundoff error is of the order of 10^{-16}. For the innovation matrix n_x=75 and $\kappa_2(I)$ is of the order of $10^4 - 10^6$. Thus for the innovation matrix $r_x \sim 10^{-7} - 10^{-9} << 1$. For the observation error covariance matrix $n_y \sim 1000 - 10000$ and $\kappa_2(O)$ is of the order of $10^5 - 10^6$. Thus for the observation error covariance matrix $r_y \sim 10^{-3} - 10^{-5}$. As r < 1 in both cases, the theorem above may be applied to determine the bounds for $||W - W'||$:

$$\begin{aligned} ||W - W'|| &\leq 2^{-t} \kappa_2(I) n_x^{\frac{3}{2}} ||I^{-1}||||H^T||||O^{-1}||/(1 - r_x) \\ &\leq 10^{-7} ||I^{-1}||||H^T||||O^{-1}||, \\ ||W - W'|| &\leq 2^{-t} \kappa_2(O) n_y^{\frac{3}{2}} ||I^{-1}||||H^T||||O^{-1}||/(1 - r_y) \\ &\leq 10^{-3} ||I^{-1}||||H^T||||O^{-1}||. \end{aligned} \tag{16}$$

The bound for errors due to errors in the Cholesky decomposition of the observation error covariance matrix are $\leq 10^4$ times greater than the corresponding bound for errors in the innovation matrix. The magnitude of these errors can be reduced by improving the condition of the observation error covariance matrix through choice of metric and by reducing the dimension of the observation error covariance matrix through channel selection. For example for the NdTb metric and the 1057 channel sampling $||W - W'|| \leq 10^{-5} ||I^{-1}||||H^T||||O^{-1}||$. Finally, note that in absolute terms these errors are small, smaller than the corresponding norms for the approximations to the forward model error covariance matrix for example: with real8 precision roundoff errors are a negligible source of error in the calculations presented here.

B   Additional issues relating to choice of noise metric

Forward model errors may be expressed in terms of radiance differences ΔR, brightness temperature differences ΔT_B or normalised brightness temperature differences $N\Delta T_B$. Radiance and brightness temperature differences are related to one another through the derivative of the Planck function with respect to temperature, evaluated at the scene temperature T_B:

$$\Delta T_B = \left(\left. \frac{\partial B}{\partial T} \right|_{T=T_B} \right)^{-1} \Delta R, \tag{17}$$

while the $N\Delta T_B$ metric is a simple rescaling of the radiance differences:

$$N\Delta T_B = \left(\left. \frac{\partial B}{\partial T} \right|_{T=T_{ref}} \right)^{-1} \Delta R, \tag{18}$$

i.e. the derivative of the Planck function with respect to temperature, evaluated at the reference scene temperature is the wavenumber dependent scale factor relating ΔR and $N\Delta T_B$, irrespective of scene temperature. Both the condition of the forward model error covariance matrix and the degree of forward model error correlation may be influenced by the choice of noise metric.

Because the Planck function varies by several orders of magnitude across the IASI spectral interval (see Figure 18 for example), the condition of the O, E and F matrices is improved through transformation of radiance errors to a noise equivalent brightness temperature error or through specification of brightness temperature errors appropriate to the scene temperature directly: errors are more homogeneous in magnitude across the spectral interval for these noise definitions. However, correlations between forward model errors and scene temperature/radiance (for a given channel) must not be overlooked, as they are important in determining the validity (or otherwise) of the $N\Delta T_B$ or ΔT_B noise specification.

Consider forward model errors for a given channel (wavenumber). If forward model brightness temperature errors are independent of scene temperature (ΔT_B uncorrelated with T_B), the magnitude (norm) of the corresponding radiance differences ΔR will increase with increasing scene temperature or equivalently, increasing radiance. This correlation between $|\Delta R|$ and R is introduced through the temperature dependence of the derivative of the Planck function $\partial^2 B/\partial T^2 > 0$ (see also the lower curves of Figure 18), and leads to higher spread in radiance errors. Note that the magnitude of this scene temperature dependent spread will depend on both the channel wavenumber (because $\partial^2 B/\partial T^2$ increases with increasing wavenumber ($\partial^3 B/\partial\nu\partial T^2 > 0$)) and the range of scene temperatures encountered in the channel.

If estimates of variance are made in radiance space without taking scene temperature correlations into account, errors are effectively overestimated for cold scene temperatures and underestimated for warm scene temperatures. Futhermore, in the presence of biased brightness temperature estimates ($E[\Delta T_B]\neq 0$), correlations are introduced between ΔR and R. In this case inter-channel error correlations will be introduced through the scene temperature dependent scaling on conversion to radiance differences because scene temperatures are correlated between channels. These effects are illustrated schematically in the right hand panel of Figure 18. The same is true of $N\Delta T_B$ noise estimates because they are related to the radiance differences ΔR by a simple rescaling. Thus, when estimating the forward model error covariance matrix one should examine forward model bias and forward model error–scene temperature correlation characteristics explicitly and choose a metric accordingly.

Bias is a significant component of the current RTIASI fast model brightness temperature errors in many spectral intervals. In Figure 19 we illustrate the correlation coefficients for ΔR–R and ΔT_B–T_B correlations for each IASI channel, based on the simulations for the 117 diverse profile set. Note that correlations greater than 0.2 or less than -0.2 are significant for this sample size. There are clear correlations between the ΔR and the scene radiance and correlations increase with increasing wavenumber, consistent with the error correlations due to biased brightness temperature estimates described above. Correlations are reduced in the ΔT_B–T_B analysis, however they remain at statistically significant levels throughout the spectral interval illustrated.

In one instance it has been possible to identify the origin of these correlations. Strong anticorrelations temperature in the wing of the CO_2 ν_2 band and in the CO_2 ν_3 band are related to the occurance of two distinct classes of forward model errors: positive radiance errors occur when modelling high latitude profiles (cold scene temperatures) while negative radiance errors tend to occur for other atmosphere types (warmer scene temperatures). State dependent errors may account for the correlations in other spectral regions, but the diagnostics are more complicated for variable gas absorption and have not been pursued in depth here.

These results would indicate that the ΔT_B metric should be used for specification of the RTIASI forward model error covariance matrix. This metric gives improved condition over a ΔR specification, and is expected to reduce the fast model error correlation of the $N\Delta T_B$ specification (through the elimination of correlations due to biased ΔT_B foward model errors). In this case, instrumental noise (ΔR specification) must be mapped to the appropriate scene temperature for each channel on a spectrum by spectrum basis. Results also suggest it may be beneficial to treat the atmospheric state dependence explicitly (specification of F by air mass class) in order to reduce ΔT_B–T_B correlations (c.f. comments on variance estimates in the presence of correlation above).

We do not believe the (nominally incorrect) choice of the $N\Delta T_B$ metric will influence the results and conclusions presented in this Technical Report significantly. Two points are of note regarding this conclusion:

(i) As described above, forward model errors tend to be overestimated at cold scene temperatures and underestimated at warm scene temperature. This implies that slightly more weight should be give to the observations in the AFGL5 case, and slightly less weight should be give to the observations in the AFGL1 case. This in turn would tend to increase errors (small) due to approximations to F or Jacobian errors in the AFGL5 case and decrease them in the AFGL1 case. The AFGL1 cases illustrated can therefore be interpreted in terms of an upper bound.

(ii) The presence relatively strong ΔT_B–T_B correlations observed would tend to suggest $N\Delta T_B$ error covariance estimates are not unrealistic: as noted above, error correlations of this type will give rise to inter-channel forward model error correlations and are undoubtably a significant component of the error covariance estimates for the current RTIASI model (correlation characteristics were not significantly reduced on passage from the ΔR to the ΔT_B metric, illustrated in Figure 19).
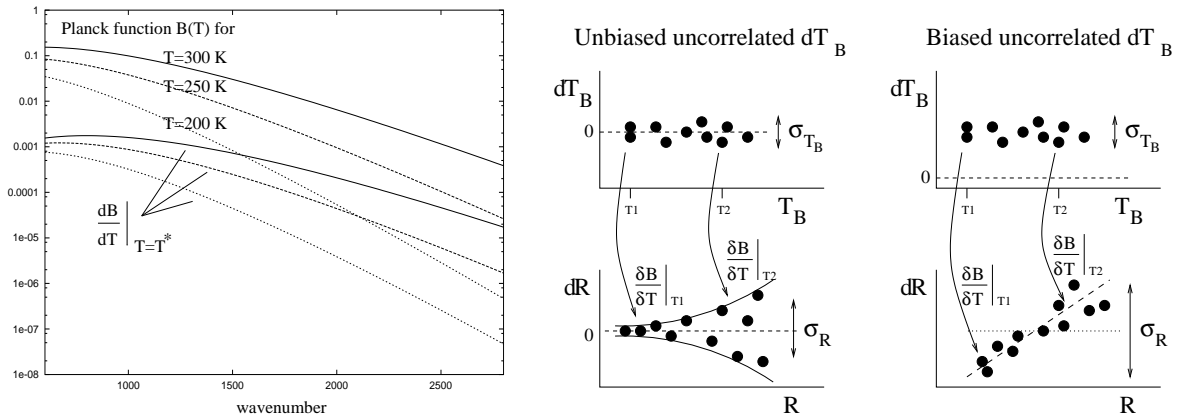
Figure 18: Left hand panel: spectral and temperature dependencies of the Planck function, $B(T)$ [W/sr/m^2/cm^{-1}] and the derivative of the Planck function $\partial B/\partial T$ [W/sr/m^2/cm^{-1}/K] evaluated at the scene temperatures ($T^\star$) for three characteristic atmospheric scene temperatures (300, 250 and 200 K). Note $\partial^2 B/\partial T^2$ increases with increasing wavenumber. Right hand panel: schematic illustration of the introduction of scene-radiance dependent spread in radiance errors and the introduction of correlation with scene radiance due to biased forward model brightness temperature errors through the temperature dependence of the Planck function derivative $\partial B/\partial T$.
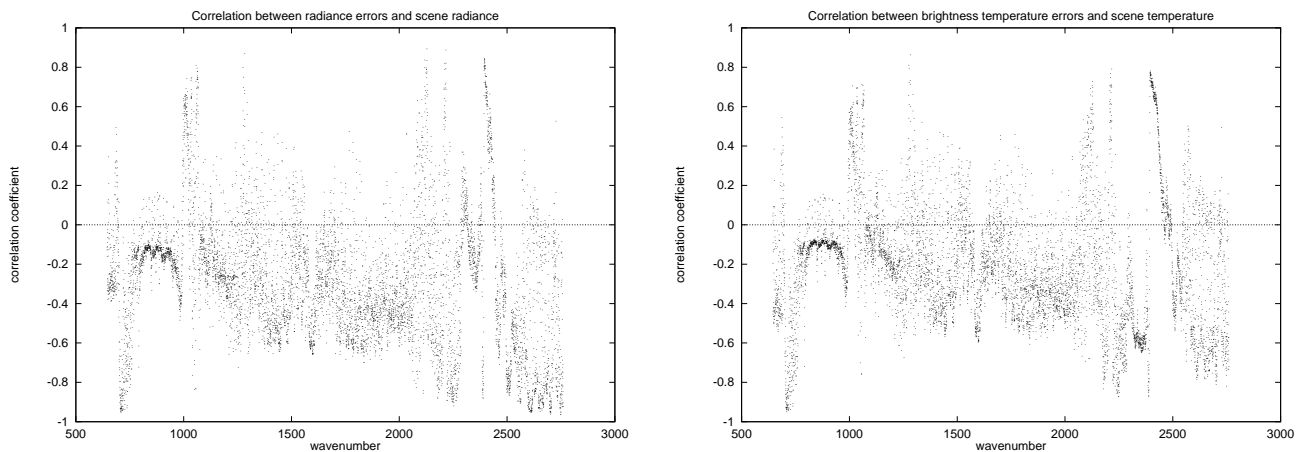

Figure 19: Correlation coefficients for ΔR–R and ΔT_B–T_B correlations for each IASI channel (left and right hand panels respectively), based on the simulations for the 117 diverse profile set. Correlations greater than 0.2 or less than -0.2 are significant for this sample size. Correlations are reduced for the ΔT_B metric, but remain at statistically significant levels for most of the IASI spectral interval.

References

[1] V. J. Sherlock. Results from the first UKMO IASI fast radiative transfer model intercomparison. Technical Report FR-287, The Met. Office, Bracknell, United Kingdom, 2000.

[2] Technical Report IATN-0000-5479-CNE, Centre Nationale d'Etudes Spatiales, Toulouse, France, 1996.

[3] F. Cayla. Communication to ISSWG, Aix-en-Provence, 1999. Report available from P. Schlussel, EUMETSAT.

[4] U. Amato, D. de Canditiis, and C. Serio. Effect of apodisation on the retrieval of geophysical parameters from Fourier Transform spectrometers. *Appl. Optics*, 37(27):6537–6543, 1999.

[5] C. D. Barnet and J. Susskind. Simulation studies of advanced infrared and microwave sounders. In *Proceedings of the Tenth International TOVS Study Conference, Boulder CO*, 1999.

[6] R. O. Weber and P. Talkner. Some remarks on spatial correlation function models. *Monthly Weather Rev.*, 121:2611–2617, 1993.

[7] G. H. Golub and C. F. Van Loan. *Matrix computations*. John Hopkins University Press, 1996.

[8] A. Kielbasinski. A note on the rounding error analysis of Cholesky factorisation. *Linear Algebra and its Applications*, 88/89:487–494, 1987.

[9] F. Chevallier. Sampled databases of atmopsheric profiles from the ECMWF 50-level forecast model. Technical report, European Centre for Medium-range Weather Forecasts, Reading, United Kingdom, 1999.

[10] A. D. Collard. The effect on undetected cloud on IASI retrievals. Technical Report FR-261, The Met. Office, 1999.

[11] P. D. Watts and A. P. McNally. The sensitivity of a minimum variance retrieval scheme to the value of its principal parameters. In *Proceedings of the Fourth International TOVS Study Conference; Igls, Austria; 16-22 March, 1988*. CIMSS, Univ. Wisconsin, Madison, 1988.

[12] C. D. Rodgers. Characterisation and error analysis of profiles retrieved from remote sounding measurements. *J. Geo. Res.*, 95(D5):5587–5595, 1990.

[13] C. D. Rodgers. Information content and optimisation of high spectral resolution measurements. *SPIE* **2830**, *Optical spectroscopic techniques and instrumentation for atmospheric and space research II*, pages 136–147, 1996.

[14] J. R. Eyre. The effects of non-linearity on analysis and retrieval errors. Technical Report FR-252, The Met. Office, 1998.

[15] J. Eyre and A. Collard. The effects of nonlinearity on retrieval errors: implications for the interpretation of advanced IR sounder data. In *Proceedings of the Tenth International TOVS Study Conference, Boulder CO, USA*, pages 191–202, 1999.

[16] H.L. Huang, W. L. Smith, and H. M. Woolf. Vertical resolution and accuracy of atmopsheric infrared sounding interferometers. *J. Appl. Meteorol.*, pages 265–274, 1992.

[17] A. D. Collard. Notes on IASI performance. Technical Report FR-256, The Met. Office, Bracknell, United Kingdom, 1998.

[18] G. E. Backus and J. F. Gilbert. Uniqueness in the inversion of inaccurate gross Earth data. *Philos. Trans. R. Soc. London, Ser. A*, 266:123–192, 1970.