Quarterly Journal of the Royal Meteorological Society Q. J. R. Meteorol. Soc. 140: 2111–2118, October 2014 A DOI:10.1002/qj.2280

RMetS
Royal Meteorological Society

# Improving the selection of IASI channels for use in numerical weather prediction

L. Ventress* and A. Dudhia
*AOPP, Department of Physics, Oxford University, UK*

*Correspondence to: L. Ventress, Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford OX1 3PU, UK.
E-mail: ventress@atm.ox.ac.uk

**High-resolution infrared sounders, such as the Infrared Atmospheric Sounding Interferometer (IASI) on the current MetOp series of satellites, produce several orders of magnitude more data per location than previous instruments used in operational retrieval and data assimilation schemes. Using the full spectrum (8641 channels for IASI) is impractical and a common approach is to identify a subset of channels which, ideally, conveys the most information on the target parameters (e.g. atmospheric temperature and water vapour) but using a relatively small number of measurements.**

**Representing the problem as a one-dimensional retrieval, optimal estimation provides an efficient framework for channel selection, and is the basis of several current schemes. However, while modelling the propagation of random (spectrally uncorrelated) errors into the retrieval, the standard algorithm does not allow for spectrally correlated errors, particularly arising from the radiative transfer modelling, which are often the limiting factor in retrieval accuracy. Such errors are either ignored or represented only approximately during the selection.**

**This article describes a modification to the standard algorithm which allows spectrally correlated errors to be properly modelled, and quantified, within the channel selection process. Comparing the results with an established selection scheme shows that significant improvements can be obtained when retrieving temperature regarding water vapour as an error term, but are less dramatic when both are retrieved together. The concept of 'total' information available from an IASI spectrum is also re-assessed.**

*Key Words:* IASI; NWP; channel selection; retrieval; error analysis; optimal estimation

*Received 12 June 2013; Revised 3 October 2013; Accepted 15 October 2013; Published online in Wiley Online Library 3 December 2013*

## 1. Introduction

The Infrared Atmospheric Sounding Interferometer (IASI) is one of eleven instruments flown on the MetOp series of polar-orbiting satellites. It is a nadir-viewing infrared Fourier transform spectrometer covering the range 645–2760 $cm^{-1}$ (15.5–3.6 $\mu m$) sampled at 0.25 $cm^{-1}$, with 0.5 $cm^{-1}$ resolution after apodisation. There is also an associated imaging system primarily for cloud detection.

MetOp A was launched on 19 October 2006, MetOp B on 17 September 2012 and the third is due for launch in 2016. Thus IASI instruments are expected to provide measurements over at least a decade.

In many respects, IASI is the successor to the High-resolution InfraRed Sounder (HIRS) instrument flown on the NOAA satellites but, while HIRS had just 20 broadband channels, each IASI spectrum contains a total of 8461 spectral points, conventionally also designated 'channels'. Using the full IASI spectrum is not practical for operational retrieval and data assimilation schemes, which typically only have to deal with a few tens of measurements per location from other satellite instruments. Choosing an optimal subset of the IASI channels is an established method of data reduction, the aim being to identify the smallest number of channels which convey the essential information on the target atmospheric and surface parameters.

Current channel selection schemes tend to be based on a correct treatment of random errors but ignore, or represent only crudely, spectrally correlated errors which often limit the accuracy of the radiative transfer modelling.

Taking a standard algorithm used to select IASI channels for numerical weather prediction (NWP) by Collard (2007), this article describes a modification which allows the influence of spectrally correlated errors to be modelled properly without significant computational overheads. The relative performance of the two algorithms is then compared.

The proposed scheme is a simplification of a solution already developed by Dudhia *et al.* (2002) for the related, but more

complex, problem of microwindow selection for a limb-viewing interferometer.

Although the channel selection here is expressed as an explicit one-dimensional (1D) retrieval problem, it is also assumed to be applicable to channels selected for direct radiance assimilation within a three- or four-dimensional variational (3D- or 4D-Var) system.

The theoretical basis of the selection algorithms and underlying optimal estimation framework are explained in section 2 with their application to a selection of channels for the IASI instrument described in section 3. Analysis of temperature, water vapour and ozone error profiles are examined in section 4 using both the original and modified algorithms, and the maximum information available from an IASI spectrum is assessed in section 5. The results are summarised in section 6.

## 2. Channel selection algorithms

Much of the theoretical background of the following is explained in more detail in Rodgers (2000) but, in keeping with many articles for this particular application, NWP symbol conventions are used.

### 2.1. Optimal estimation framework

The optimal estimation approach defines a notional state vector $\mathbf{x}$, containing $n$ parameters to be retrieved (e.g. an atmospheric temperature profile) and an associated *a priori* covariance matrix $\mathbf{B}$ (dimension $n \times n$) representing the knowledge of the state before any satellite measurements are included (e.g. model prediction error or climatological variability).

This is combined with the information from a set of $m$ measurements with error covariance $\mathbf{R}$ ($m \times m$) and, assuming no intercorrelations between the $\mathbf{R}$ and $\mathbf{B}$ matrices, the resulting *a posteriori* covariance is given by

$$\mathbf{A} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1}\right)^{-1}, \tag{1}$$

where $\mathbf{H}$ is the Jacobian matrix ($m \times n$) containing the derivatives of each measurement with respect to each element of the state vector. The matrix $\mathbf{R}$ includes errors inherent in the measurements themselves (noise, calibration errors) as well as those associated with the radiative transfer model used to predict the measurement from a given state and representativeness errors.

The aim of the channel selection is to find either the 'best' subset of $m$ channels from the measurement domain or, alternatively, to find the minimum number of channels required to meet a target retrieval accuracy.

In either case, a 'figure of merit' has to be specified: a scalar quantity which allows different $\mathbf{A}$ matrices to be compared. The 'degrees of freedom for signal' (DFS) is one such measure, defined as

$$d_{\mathrm{s}} = \mathrm{Tr}\left(\mathbf{I} - \mathbf{A}\mathbf{B}^{-1}\right), \tag{2}$$

where $\mathbf{I}$ is the identity matrix (order $n$). The DFS is more loosely defined as the number of independent parameters that can be retrieved. A value $d_{\mathrm{s}} = 0$ corresponds to $\mathbf{A} = \mathbf{B}$, i.e. the case before any measurements have been used or when the measurements have provided no information on the state vector. In the limit of a 'perfect' retrieval $\mathbf{A} \to 0$, hence $d_{\mathrm{s}} \to \mathrm{Tr}(\mathbf{I}) = n$, the number of elements in the state vector.

### 2.2. Conventional algorithm

With over 8000 IASI channels to choose from, it might be feasible to work through every possible two- or even three-channel combination and evaluate the DFS for each, but for selecting a 100 channels, or even just 10, another approach is required. Although ideally we would like to solve the problem of which is the best set of $m$ channels to use, in practice it is necessary to use a sequential selection which solves a slightly different, but more tractable problem: given $(m-1)$ channels already selected, which of the remaining channels is the best one to add?

Rabier *et al.* (2002) discuss various alternative schemes but prefer that originally proposed by Rodgers (1996). This exploits a particular property of optimal estimation: provided that the matrix $\mathbf{R}$ is diagonal (i.e. measurement errors are uncorrelated), Eq. (1) can also be evaluated sequentially:

$$\mathbf{A}_m = \left\{\mathbf{h}^{\mathrm{T}}(\sigma^2)^{-1}\mathbf{h} + \mathbf{A}_{m-1}^{-1}\right\}^{-1}, \tag{3}$$

where $\mathbf{A}_m$ is the solution for Eq. (1) after $m$ measurements have been used, $\mathbf{h}$ is the row of $\mathbf{H}$ corresponding to the $m$th channel and $\sigma^2$ (often just taken as the square of the instrument noise) is the equivalent diagonal element of $\mathbf{R}$. This has the advantage that the matrix inversion in Eq. (3) is of order $n$, whereas in Eq. (1) there is an additional, and expanding, matrix inversion of order $m$ ($\mathbf{R}^{-1}$). Typically $n$ is a few tens while $m$ can exceed 100.

Before beginning the selection, a radiative transfer model is used to compute the $\mathbf{H}$ matrix for all channels in the measurement domain, assuming a typical atmospheric profile (or set of profiles).

The selection begins ($m = 1$) by setting $\mathbf{A}_0 = \mathbf{B}$, solving Eq. (3) for each channel individually and evaluating the corresponding DFS. The channel with the largest DFS is selected as the first channel.

Then ($m = 2$) replacing $\mathbf{B}$ with the solution covariance $\mathbf{A}_1$ from the first selected channel, Eq. (3) and the corresponding DFS (still using the original $\mathbf{B}$) are re-evaluated for all remaining channels to determine the second channel. This is repeated until the required number of channels, or target DFS value, has been met.

### 2.3. Application to IASI

In developing the above algorithm into a practical scheme for IASI channel selection, Collard (2007) notes that it contains no mechanism for handling spectrally correlated errors. Such errors are commonly associated with the radiative transfer model. For example, these can arise from assuming an incorrect concentration of an interfering (i.e. a non-retrieved) molecule; an underestimate of the concentration generally leads to an overestimate of the radiance at all spectral points overlapping the molecule's absorption bands.

The adopted strategy was to estimate the magnitude of such error contributions at each spectral point. If larger than some threshold ($\sim$1 K in brightness temperature) the channel was 'blacklisted', i.e. excluded from selection. If smaller, the contribution was added to the noise and treated as a random error.

This *ad hoc* approach is unsatisfactory in that it does not properly allow for the cumulative effect of correlated errors in the retrieval accuracy. (While the random noise impact reduces as the square root of the number of measurements, the correlated error remains the same so may become dominant when $\sim$100 channels are used.) On the other hand, setting the blacklist threshold even lower may exclude channels whose benefit in signal/noise outweighs their correlated error contribution.

### 2.4. Spectrally correlated errors

Any correlated error within a set of $m$ spectral measurements $\mathbf{y}$ can be represented as a combination of one or more statistically independent error vectors $\delta\mathbf{y}^j$, each contributing an additive term to the overall error covariance matrix:

$$\mathbf{R}^j = (\delta\mathbf{y}^j)(\delta\mathbf{y}^j)^{\mathrm{T}}. \tag{4}$$

For example, a 0.1 K radiometric uncertainty in the onboard black-body calibration target would be a single vector containing $m$ values of 0.1 K. More complex cases such as a ($k \times k$) covariance matrix representing the uncertainty in the assumed profile of

 

a contaminating species can be decomposed into $k$ empirical orthogonal functions (EOFs) in profile space and, using an appropriate Jacobian matrix, mapped into $k$ independent error vectors in spectral space.

A total measurement error covariance matrix can then be represented as a sum of random (spectrally uncorrelated) and spectrally correlated components

$$\mathbf{R}^{\mathrm{tot}} = \mathbf{R}^{\mathrm{rnd}} + \sum_j (\delta\mathbf{y}^j)(\delta\mathbf{y}^j)^{\mathrm{T}}. \tag{5}$$

The distinction between random and correlated components is convenient rather than fundamental: a diagonal $\mathbf{R}^{\mathrm{rnd}}$ matrix could also be represented as a set of $m$ error vectors each containing a single non-zero element.

Walker *et al.* (2011) use this approach to construct a total error covariance as part of a fast linear retrieval scheme for IASI. Their proposed solution for a sequential channel selection is simply to compute the figure of merit for each potential new channel using the full inversion of Eq. (1), with $\mathbf{R}^{\mathrm{tot}}$ for $\mathbf{R}$ (and in this case without $\mathbf{B}$). This is feasible since they only deal with small matrices: just two elements in the state vector ($n = 2$) and a limited spectral domain for measurements ($m \ll 100$).

In the conventional selection method, repeated inversion of the full $\mathbf{R}$ matrix was avoided by assuming that it was diagonal, so, at first sight, any generalisation to allow for a correlated covariance $\mathbf{R}^{\mathrm{tot}}$ would seem to require a significant increase in computation time. A solution lies in the distinction between how the channels are *selected*, and how they will be *used* in the retrieval or data assimilation. Although recent work has been carried out to include more realistic $\mathbf{R}$ matrices, including correlated errors, within 4D-Var assimilation systems (Bormann *et al.*, 2010; Weston, 2011), here we discuss only the use of a diagonal $\mathbf{R}$ matrix.

### 2.5. Linear error propagation

In linearised form the optimal estimation solution is a combination of measurement $\mathbf{y}$ and *a priori* information $\mathbf{x}_{\mathrm{b}}$ represented by

$$\mathbf{x} = \mathbf{K}\mathbf{y} + (\mathbf{I} - \mathbf{K}\mathbf{H})\,\mathbf{x}_{\mathrm{b}}\,, \tag{6}$$

where $\mathbf{K}$ (the 'gain matrix') is given by

$$\mathbf{K} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}. \tag{7}$$

The measurement covariance $\mathbf{R}$ which appears in the above expression for $\mathbf{K}$ is that used by the retrieval, which is not necessarily (and not usually) the total error covariance. It will be assumed that the retrieval just uses the random measurement error covariance $\mathbf{R}^{\mathrm{rnd}}$, hence uses Eq. (1):

$$\mathbf{K} = \mathbf{A}^{\mathrm{rnd}}\mathbf{H}^{\mathrm{T}}\left(\mathbf{R}^{\mathrm{rnd}}\right)^{-1}, \tag{8}$$

where $\mathbf{A}^{\mathrm{rnd}}$ is the random error covariance of the solution $\mathbf{x}$. From this it follows that any measurement error vector $\delta\mathbf{y}^j$ is mapped into a corresponding retrieval error vector $\delta\mathbf{x}^j$ by

$$\delta\mathbf{x}^j = \mathbf{K}\delta\mathbf{y}^j + (\mathbf{I} - \mathbf{K}\mathbf{H})\delta\mathbf{x}^j_{\mathrm{b}}\,, \tag{9}$$

where $\delta\mathbf{x}^j_{\mathrm{b}}$ represent components of the same error $j$ which contribute to the *a priori* estimate (i.e. via a previously incorporated measurement). Since these errors are independent, they can then be combined in retrieval space, by analogy with Eq. (5), to give a total retrieval error

$$\mathbf{A}^{\mathrm{tot}} = \mathbf{A}^{\mathrm{rnd}} + \sum_j (\delta\mathbf{x}^j)(\delta\mathbf{x}^j)^{\mathrm{T}}. \tag{10}$$

Thus, if the retrieval process can be expressed as a sequential application of Eq. (6) –which basically means that the retrieval regards the measurement errors as uncorrelated –then spectrally correlated errors can be mapped into the solution as vectors without having to invert, or even construct, the full measurement covariance matrix $\mathbf{R}^{\mathrm{tot}}$.

### 2.6. Modified selection algorithm

The conventional algorithm can now be modified by

(a) correctly modelling the measurement error correlations during the selection process itself, while
(b) allowing for the selected channels being subsequently used in a retrieval as if uncorrelated

It is assumed that spectrally correlated errors have been modelled as a set of independent error spectra $\delta\mathbf{y}^j$ and that the initial *a priori* covariance $\mathbf{B}$ contains no contributions from these errors.

Starting with $\mathbf{A}^{\mathrm{rnd}}_0 = \mathbf{B}$ and $\delta\mathbf{x}^j_0 = 0$, the modified channel selection then proceeds from $m = 1$ as follows:

$$\mathbf{A}^{\mathrm{rnd}}_m = \left\{\mathbf{h}^{\mathrm{T}}(\sigma^2)^{-1}\mathbf{h} + (\mathbf{A}^{\mathrm{rnd}}_{m-1})^{-1}\right\}^{-1}, \tag{11}$$

$$\mathbf{k} = \mathbf{A}^{\mathrm{rnd}}_m\mathbf{h}^{\mathrm{T}}(\sigma^2)^{-1}, \tag{12}$$

$$\delta\mathbf{x}^j_m = \mathbf{k}\delta y^j + (\mathbf{I} - \mathbf{k}\mathbf{h})\,\delta\mathbf{x}^j_{m-1}, \tag{13}$$

$$\mathbf{A}^{\mathrm{tot}}_m = \mathbf{A}^{\mathrm{rnd}}_m + \sum_j (\delta\mathbf{x}^j_m)(\delta\mathbf{x}^j_m)^{\mathrm{T}}, \tag{14}$$

$$d^{\mathrm{tot}}_{\mathrm{s}}(m) = \mathrm{Tr}\left(\mathbf{I} - \mathbf{A}^{\mathrm{tot}}_{\mathrm{m}}\mathbf{B}^{-1}\right), \tag{15}$$

where $\mathbf{k}$ is the gain matrix $\mathbf{K}$ from Eq. (7) reduced to a $(n \times 1)$ vector for a single measurement. Eqs (12)–(14) are new, but computationally fast compared to the inversion in Eq. (11) (which is the same as in the original scheme). As well as the random error covariance $\mathbf{A}^{\mathrm{rnd}}_m$, it is now also necessary to save the set of vectors $\delta\mathbf{x}^j_m$, representing the various systematic error contributions to the next *a priori* estimate.

At this point it is noted that apodisation introduces noise correlations between IASI channels and therefore the selection described in section 3 does not allow for adjacent channels to be chosen. However, if selection of adjacent channels *is* to be allowed, the sequential procedure explained above must be modified.

If the retrieval still assumes an uncorrelated $\mathbf{R}$, then the 'true' $\mathbf{A}^{\mathrm{rnd}}_m$ in Eq. (14) (as opposed to the uncorrelated version assumed by the retrieval in Eqs (11) and (12)) would have to be modified to

$$\mathbf{A}^{\mathrm{rnd}}_m = \left\{\mathbf{H}^{\mathrm{T}}_m(\mathbf{R}^{\mathrm{rnd}}_m)^{-1}\mathbf{H}_m + \mathbf{B}^{-1}\right\}^{-1}, \tag{16}$$

where $\mathbf{R}^{\mathrm{rnd}}_m$ is the correlated noise covariance matrix for the $m$ measurements selected so far. Given $\mathbf{R}^{\mathrm{rnd}}_m$ is now a sparse tridiagonal matrix (of growing dimension $m \times m$), this could still be practical.

On the other hand, if the retrieval also models correlated errors, it may be adequate to apply the above to select each channel by replacing Eq. (11) with Eq. (16) and subsequently updating Eqs (12) and (13) with the equivalent for the full measurement vector so far:

$$\mathbf{K} = \mathbf{A}^{\mathrm{rnd}}_m\mathbf{H}^{\mathrm{T}}\left(\mathbf{R}^{\mathrm{rnd}}_m\right)^{-1}, \tag{17}$$

$$\delta\mathbf{x}^j_m = \mathbf{K}\delta\mathbf{y}^j_m\,, \tag{18}$$

where $\delta\mathbf{y}^j_m$ is the error vector for the $m$ selected measurements.

## 3. Selection of IASI channels

In order to compare the original and modified algorithms, we first attempt to reproduce the selection of 300 channels carried out by Collard and then repeat the selection using the modified algorithm. This ensures that differences are due to the algorithms themselves rather than details of the calculations. For example, in this study the radiative transfer calculations are performed using the Reference Forward Model, RFM (AOPP, 2002) while Collard used RTIASI-4 (Matricardi and Saunders, 1999; Matricardi,

 

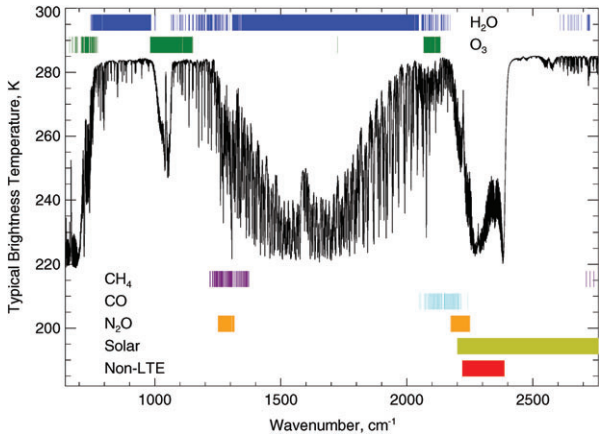


**Figure 1.** A typical IASI spectrum indicating the channels blacklisted due to the atmospheric variability of $H_2O$, $O_3$, $CH_4$, CO and $N_2O$ along with those affected by non-LTE effects and solar irradiance. Figure reproduces that by Collard (2007).

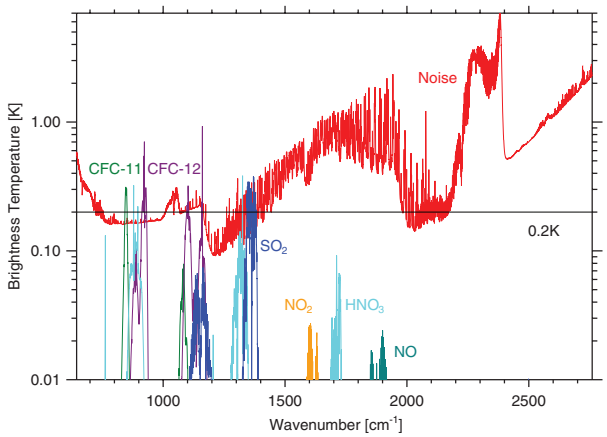


**Figure 2.** The error in IASI brightness temperature due to the climatological variability, in a mid latitude daytime atmosphere, of each minor species included in **R** in the original selection and included as an independent error source in the modified selection. The instrument noise and 0.2 K assumed forward model uncertainty are also plotted.

2003); the background-error covariance matrices are based upon different operational 1 D-Var matrices, and different assumptions are made on the variability of minor species. However, the original Collard channels are also included in the comparisons to give some indication of the robustness of the original selection to supposedly minor differences in implementation.

Given a set of channels selected by either method, the formalism of the modified algorithm can then be applied to evaluate both DFS for total error, $d_s^{\mathrm{tot}}$, (i.e. the quantity that the modified algorithm aims to maximise) or the DFS for random error, $d_s^{\mathrm{rnd}}$ (maximised by the original algorithm).

### 3.1. *Reproducing original selection*

Collard selected 300 channels for distribution to NWP centres in near-real time. This means that the only species of interest are water vapour, ozone, and temperature (from the $CO_2$ bands).

The same flags used by Collard (personal communication) are used to blacklist channels where the climatological variability of $CH_4$, CO and $N_2O$ cause a brightness temperature uncertainty of $> 1$ K, as well as channels influenced by non-LTE effects (i.e. where the usual assumption of local thermodynamic equilibrium (LTE) breaks down, chiefly at shorter wavelengths and high altitudes), solar irradiance and, in certain selection runs, water vapour and ozone. The locations of affected channels are shown in Figure 1.

The random error on each measurement is taken as a combination of the following terms (in quadrature):

1. Instrument noise from the Centre National d'Etudes Spatiales (CNES) 2008 post-launch estimate (F. Smith, personal communication);
2. The climatological variability (using data from the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) IG2 by Remedios *et al.*, 2007) for each minor species ($CCl_4$, CFC-11, CFC-12, CFC-14, $HNO_3$, $NO_2$, OCS, NO and $SO_2$);
3. An additional forward modelling uncertainty, assumed to be a constant 0.2 K.

These terms are plotted in Figure 2.

In order to create a 'global selection' rather than be specific to one particular location, Jacobian spectra are precomputed for two different viewing angles (nadir and 40°) and for six different FASCODE atmospheres (US standard atmosphere, Tropics, midlatitude summer, midlatitude winter, sub-arctic summer, sub-arctic winter). The DFS is computed for each and the figure of merit $d_s^{\mathrm{rnd}}$ is taken as the average of the twelve DFS values.

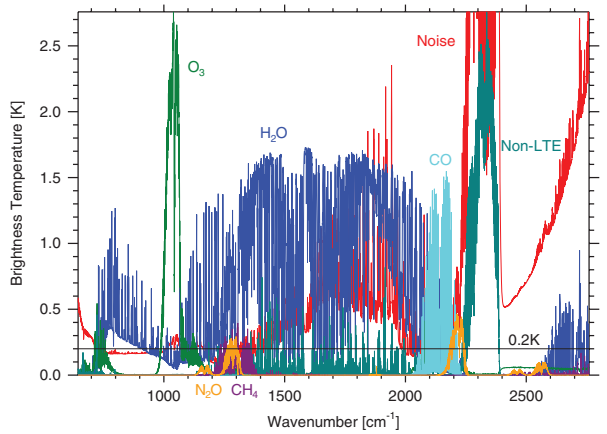


**Figure 3.** The additional error spectra (to those in Figure 2) included in the modified channel selection algorithm. The largest water vapour error vector is plotted as an example. The instrument noise and 0.2 K assumed forward model uncertainty are also plotted.

Finally, to avoid noise correlations introduced by the apodisation, once a channel is selected, the two adjacent channels are blacklisted.

### 3.2. *Modified channel selection*

A set of error vectors was created representing the uncertainty in concentration for each molecular absorber – the species represented by blacklists in the original scheme as well as each of the minor species. In most cases these were defined by a single error spectrum representing the climatological variability, which can be seen in Figures 2 and 3.

However, for water vapour, the relevant components of the Met Office 1D-Var background covariance were decomposed into 27 independent error vectors. Non-LTE was also modelled as an error spectrum created by running the RFM with and without non-LTE modelling switched on. However, to avoid solar effects, only channels up to 2200 $\mathrm{cm}^{-1}$ were considered, as with the original scheme.

The random component of measurement error was represented by the instrument noise (with the additional 0.2 K for consistency).

The DFS, this time based on the total error, $d_s^{\mathrm{tot}}$ was again taken as the average over the same six atmospheres and two viewing angles.

As with the original algorithm, once a channel was selected, the adjacent channels were blacklisted.

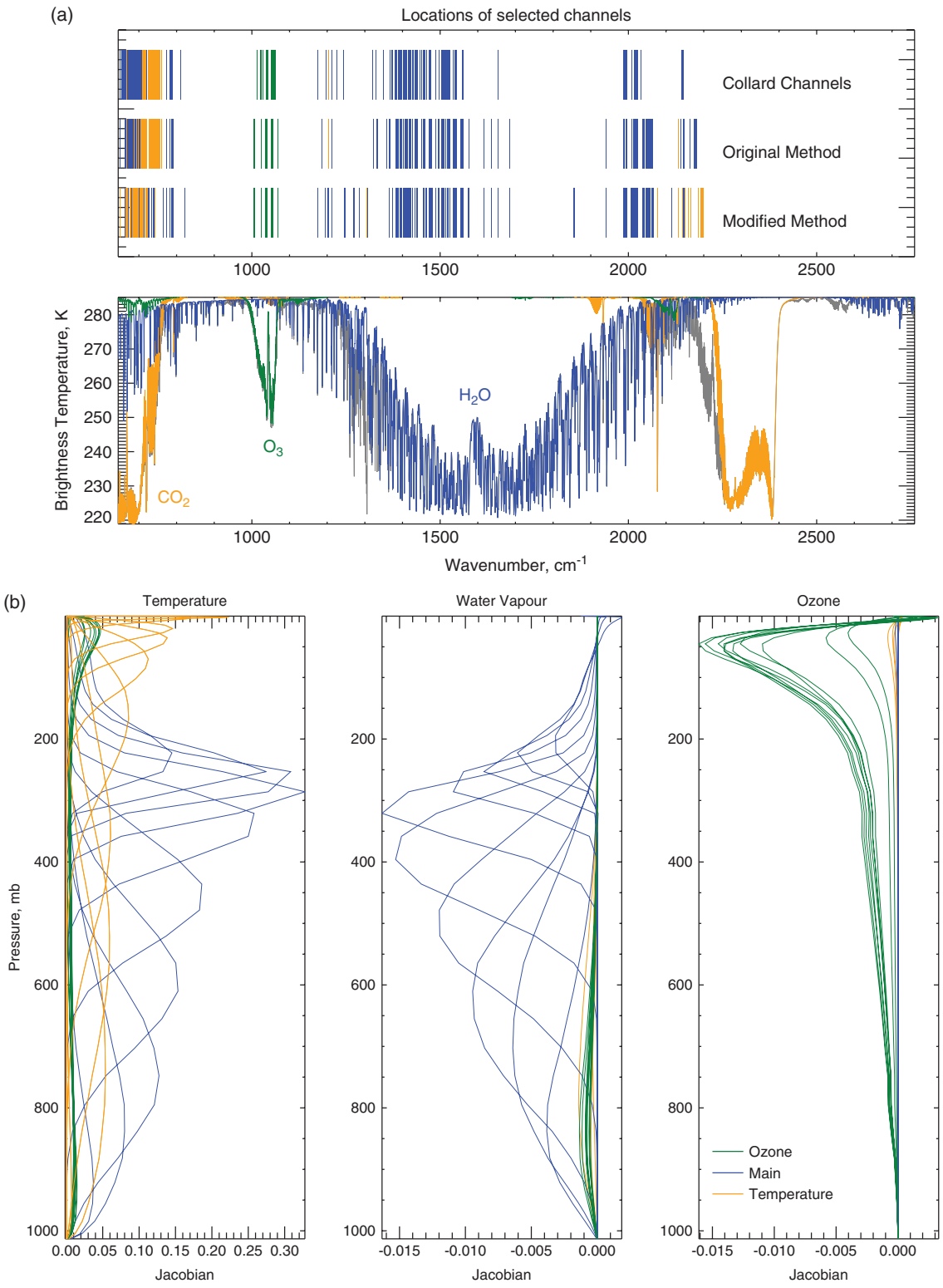


**Figure 4.** a) The locations of the 267 channels selected by the original and modified algorithms; the 66 temperature channels, 186 temperature and water vapour channels and 15 ozone channels. The lower plot shows a typical IASI spectrum, with the spectral features for $CO_2$, $H_2O$ and $O_3$ over plotted. (b) The temperature, water vapour and ozone Jacobians (given by the change in brightness temperature due to a perturbation of 1 K, 1%VMR and 1%VMR respectively) associated with the first ten channels selected in each of the temperature, main and ozone selection runs using the modified algorithm. The Jacobians shown are for the mid-latitude daytime atmosphere.

## 4. Results

### 4.1. *Temperature channels*

Following Collard, the selection starts with channels for the temperature retrieval. For this a state vector is defined containing 44 variables: temperature at 43 atmospheric levels plus the surface temperature. (For the calculation of Jacobians, it is assumed that the surface is at the same temperature as the lowest level of each of the six modelled atmospheres.)

In the original algorithm, the $H_2O$ and $O_3$ blacklist masks are also applied to ensure that $CO_2$ features are targeted. The first 30 channels are selected, followed by another 36 limited to the spectral range 707.25–759.75 cm$^{-1}$, intended to improve the information in the troposphere.

For the modified algorithm, the same state vector was defined and the first 66 channels were selected without any restriction. A selection limited to the 707.25–759.75 cm$^{-1}$ region after the first 30 channels was investigated, also weighting the figure of merit to favour information in the troposphere, but negligible

 

Table 1. Degrees of freedom for signal per profile of 43 levels plus surface term from 66 temperature channels selected by Collard, and by the original and modified algorithms.

| | | Collard | Original | Modified |
|---|---|---|---|---|
| $T_{\{43\}}$ | $d_s^{rnd}$ | 6.22 | 6.43 | 6.35 |
| | $d_s^{tot}$ | 3.86 | 3.71 | 5.98 |
| $T_s$ | $d_s^{rnd}$ | 0.97 | 0.99 | 0.99 |
| | $d_s^{tot}$ | 0.93 | 0.97 | 0.97 |

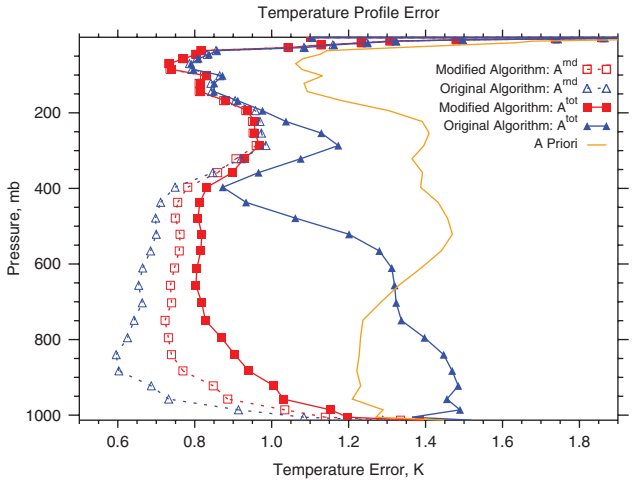


**Figure 5.** Temperature error profiles for the US standard atmosphere viewed at nadir from 66 channels for the original and modified algorithms. The error due to both the random and total errors are shown.

improvement was found in the profile error for this region, suggesting that most of the available tropospheric information had already been extracted within the first 30 channels of the modified algorithm.

The locations of the selected channels are shown in Figure 4 along with the Jacobians associated with the first ten selected channels. The largest difference between the temperature selections is that the channels from the original algorithm are almost entirely between 645 and 800 $cm^{-1}$, whereas the modified algorithm chooses a large number in the short-wave end of the spectrum. This is also the case with the first channel; for the original algorithm this is 1204 $cm^{-1}$, whereas for the modified version the first channel is at 2150 $cm^{-1}$. In both cases they are obtaining surface information and in fact obtain almost all the surface information in just this one channel. Favouring the shorter wavelength is thought to be due to the effect of the water vapour errors.

The resulting DFS values for the original and modified algorithms are listed in Table 1 alongside the values for the channels selected by Collard for comparison. The Collard values show that the reproduced set of channels manages to provide similar information and select similar channels, as can be seen in Figure 4. The small discrepancies arise from the unavoidable differences in the reproduction mentioned earlier.

As expected, the original algorithm has a better DFS for random error (since this is the figure which is optimised) but not by much. However the big difference is in the DFS evaluated for total error. For the modified algorithm it is only slightly worse than for the random error, suggesting that the spectrally correlated errors have little net contribution, but for the original algorithm the figure is almost halved. It turns out that most of this is due to the water vapour-induced errors so, in principle, recoverable if water vapour is jointly retrieved, as discussed in the next section.

The vertical profiles of random and total error for the US standard atmosphere, viewed at nadir, are shown in Figure 5. Here it can be seen that the main differences are at altitudes below 200 hPa: the original algorithm has a smaller random error but a larger total error, again consistent with using channels sensitive to uncertainties in the water vapour concentration and, between 200 and 400 hPa, the ozone uncertainty. It can also be seen that total error for the original selection algorithm is larger than the *a priori* uncertainty in the lower troposphere: here the measurements are contributing *negative* DFS.

Table 2. Degrees of freedom for signal per profile of 43 levels from 252 temperature and water vapour channels selected by the original, reproduced, and modified algorithms.

| | | Collard | Original | Modified |
|---|---|---|---|---|
| $H_2O_{\{43\}}$ | $d_s^{rnd}$ | 5.86 | 6.06 | 6.04 |
| | $d_s^{tot}$ | 5.68 | 5.67 | 5.93 |
| $T_{\{43\}}$ | $d_s^{rnd}$ | 7.53 | 7.62 | 7.62 |
| | $d_s^{tot}$ | 7.32 | 7.17 | 7.44 |
| $T_s$ | $d_s^{rnd}$ | 0.99 | 1.00 | 1.00 |
| | $d_s^{tot}$ | 0.99 | 0.99 | 1.00 |

### 4.2. *Main channels*

Once the temperature channels have been selected, the 'main run' is carried out. The aim is to find channels mainly sensitive to water vapour but which also provide additional temperature information.

A water vapour volume mixing ratio profile is added to the state vector (an extra 43 elements) and the *a priori* covariance is taken from the Met Office 1D-Var scheme. For the original algorithm, the channels initially excluded by the water vapour 'blacklist' are reinstated, and for the modified algorithm the error vectors representing water vapour uncertainty are removed.

In line with the Collard example, a further 186 channels are selected in this manner, giving a total of 252 channels. The locations of these channels are also plotted in Figure 4 along with the Jacobians associated with the first ten channels selected. The locations of the channels appear very similar; despite this, ~ 40% of the 252 are in differing locations. This in part depends upon the channels previously chosen and information already gained. Table 2 shows the DFS values for the various retrieval components. The addition of the new channels increases the temperature DFS by over 1 but, more noticeably, the random and total error values have also converged, as expected if the primary source of correlated error – the water vapour profile – is now jointly retrieved. This can be seen visually in Figure 6, which shows the vertical temperature error profile after the 'main run'. The difference between the random and total error components is much smaller than previously but, nevertheless, the modified algorithm retains a clear advantage in total error over the original algorithm, while the original algorithm has only a small advantage in random error.

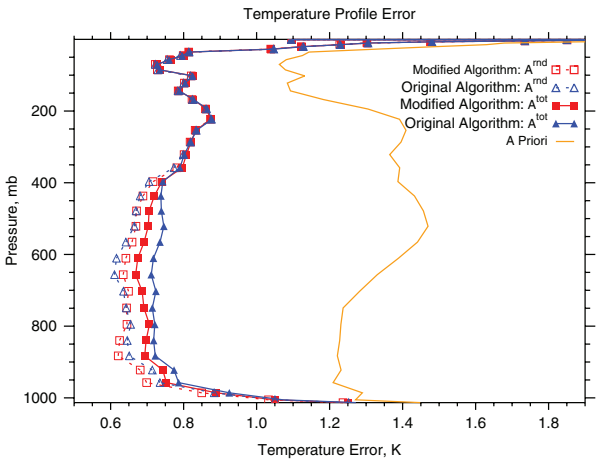


**Figure 6.** As in Figure 5 but showing the temperature error profiles after a joint retrieval with water vapour using 252 channels.

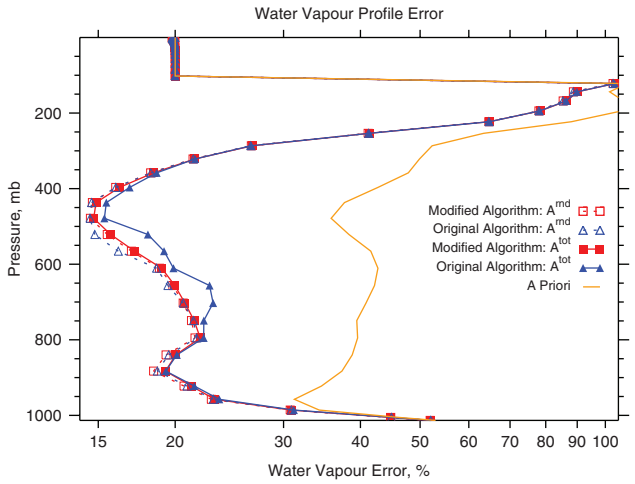


**Figure 7.** Water vapour error profiles from 252 channels for the original and modified algorithms. The error due to both the random and total errors are shown.

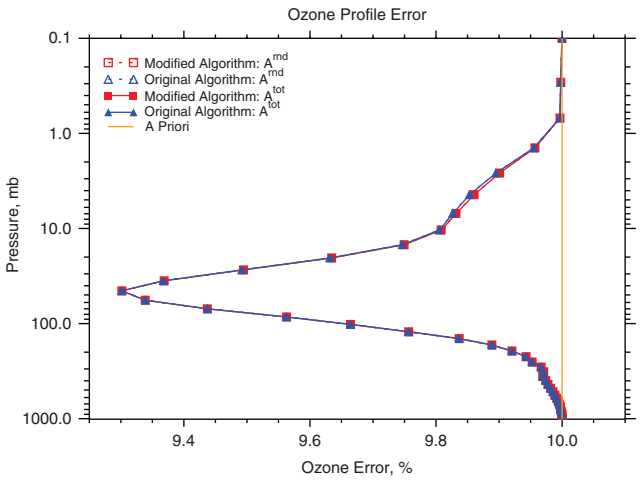


**Figure 8.** Ozone error profiles from 267 channels for the original and modified algorithms. The error due to both the random and total errors are shown, but almost completely coincide.

Table 3. Degrees of freedom for signal per profile of 43 levels from the 267 channels selected by the original, reproduced, and modified algorithms.

| | | Collard | Original | Modified |
|---|---|---|---|---|
| $O_{3\{43\}}$ | $d_s^{rnd}$ | 1.05 | 1.16 | 1.15 |
| | $d_s^{tot}$ | 1.05 | 1.16 | 1.15 |

Noticeably, the temperature Jacobians for the first ten channels of the 'main run' are larger than those from the channels chosen in the temperature selection. This is due to the removal of the error caused by water vapour, allowing the channels to be selected.

For the water vapour component, the conclusions are similar: the original algorithm has a small advantage in random error but the modified algorithm has a large advantage in total error. Again the total error reduction is small. However this is most likely because water vapour has the largest signal in this spectral region, so it will not be as adversely affected by the application of the other errors. The vertical water vapour error profile can be seen in Figure 7, which shows that the profile for the modified case is barely changed whereas the original algorithm is worse between 400 and 800 hPa. This is caused by the second largest contributing error –ozone. However, note that there is no improvement at altitudes above the 100 mb surface or in the very lowest levels, where IASI has limited sounding capability (at least for cloud-free atmospheres with no surface temperature contrast).

### 4.3. *Ozone channels*

Following the selection of temperature and water vapour channels, a further 15 channels are chosen to maximise the ozone information. In this instance, the channel selection is for ozone only and starts from the 252 channels already selected. A correlated background-error covariance matrix is assumed with a diagonal of error $(10\%)^2$. Figure 8 shows the expected error profiles for ozone having selected all 267 channels. The two algorithms result in the same 15 channels (although not selected in the same order), hence the effect of the total error is the same and the profile errors are almost identical. This is due to there being no large correlated errors (water vapour having been excluded) in the ozone band between 1000 and 1100 $cm^{-1}$. The slightly lower errors seen in the original algorithm between 10 and 1 mb, and similarly the larger degrees of freedom shown in Table 3, are due to the previously selected channels; the modified algorithm reduces the use of channels sensitive to ozone and hence there is no information already gained on ozone in the selection of the 252 channels.

Table 4. Maximum degrees of freedom for signal available from the spectral region 645–2200 $cm^{-1}$ per profile of 43 levels plus surface term considering an $\mathbf{R}^{tot}$ with and without the correlated error components. Separated into the three retrieval types: temperature only, temperature and water vapour, and ozone only.

| | $\mathrm{Max}(d_s^{rnd})$ | $\mathrm{Max}(d_s^{tot})$ |
|---|---|---|
| $T_{\{43\}}$ | 12.63 | 12.37 |
| $T_s$ | 1.00 | 1.00 |
| $H_2O_{\{43\}}$ | 7.52 | 7.50 |
| $T_{\{43\}}$ | 10.01 | 9.74 |
| $T_s$ | 1.00 | 1.00 |
| $O_{3\{43\}}$ | 2.25 | 2.24 |

## 5. Maximum DFS

So far the modified algorithm has been applied on the assumption that the measurements are treated by the retrieval as random, which implies a limit will be reached after which any further channels will start to reduce the DFS for total error. However the question arises: what value DFS could be reached if all the available measurements were combined optimally using the total measurement covariance? This requires solving

$$\mathbf{A}^{tot} = \left\{\mathbf{H}^{T}(\mathbf{R}^{tot})^{-1}\mathbf{H} + \mathbf{B}^{-1}\right\}^{-1}, \qquad (19)$$

where $\mathbf{R}^{tot}$ (Eq. (5)) is now the total measurement error covariance, including the correlated noise between adjacent channels, for the entire set of IASI measurements (or at least all those below 2200 $cm^{-1}$). This is a large matrix ($m = 6220$) but, for this calculation, it only has to be inverted once.

The resulting maximum DFS available are shown in Table 4. Both cases achieve similar DFS although $\max(d_s^{tot})$, which includes the effect of the correlated error components, is always slightly reduced.

Figure 9 shows the evolution of the DFS as the channel selection progresses. In this case the 'maximum available DFS' is $\max(d_s^{tot})$; it is separately calculated for the three runs, accounting for temperature only, the main run of temperature and water vapour, and ozone only. The steep initial curve in all cases shows that most of the information is obtained within the first few channels selected and that, towards the end of the selection, the channels contain a high fraction of the total information available, with only small improvements being made with the addition of each channel. Similar behaviour is observed for the original algorithm. Initially the algorithms agree very well, but as the selection progresses the original algorithm is adversely affected by its sensitivity to the included error sources and less information

 

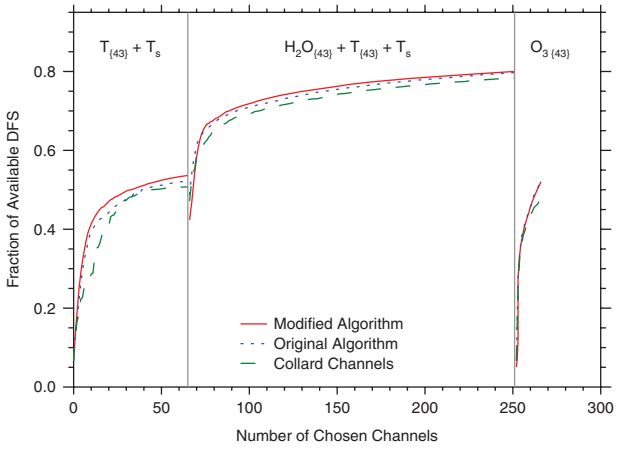


**Figure 9.** The evolution of the DFS during the channel selection, applying the systematic error component to both channel sets. The 'maximum available DFS' is calculated for; purely a temperature analysis, a temperature and water vapour analysis and a purely ozone analysis. The channels selected using the modified algorithm are shown in red, using the original algorithm are shown in blue and the original Collard channels are shown in green.

is gained (compared to the modified algorithm) on the addition of new channels. This is more pronounced when looking at the original Collard channels, initially, although some of these discrepancies arise from the differences in the reproduction of the original algorithm, such as the forward model and B matrices used.

Switching to the 'main run' for temperature and water vapour, both the original algorithm and Collard's channels begin with a larger fractional DFS than the modified version due to no longer considering the errors in water vapour on the previously selected channels. Hence, initially the original algorithm continues to have the greater value but, as selection continues, the modified algorithm gains more DFS in the first channels and soon has the larger fractional value. Again the Collard channels have a reduced improvement to the DFS, compared to the reproduced channel set, by the selection of certain channels, again highlighting the difference in reproducing the method. However, it can be seen that, towards the end of the selection, the improvement to the DFS by all methods on the inclusion of each new channel is approximately the same.

Similar behaviour can be seen in the ozone selection, again the original algorithm begins at a larger value than the modified version due to previously considering ozone as a correlated error and therefore there is minimal ozone information already available in the modified case, whereas the channels from the original method have already gained some information from the 252 channels selected. However, after the channel selection, the two sets contain almost identical amounts of ozone information.

## 6. Summary

A new method for the selection of IASI channels has been presented. Through comparison with the algorithm used to select the currently operational NWP channels, it has been shown that the new algorithm maintains the same number of DFS whilst reducing the sensitivity of the channel set to unknown (or unspecified) spectral correlations. The methods are intrinsically the same in that the channels are selected based upon their contribution to the DFS; however, where the previous method only considers the random component of the error and fully excludes certain channels from being chosen, the new method incorporates, for example, the effects of species' atmospheric variability within the selection process as a correlated error. Although carried out within an NWP context to allow comparison, the new selection method can be generalised to the selection of channels for the retrieval of any species.

## Acknowledgements

The authors would like to thank Andrew Collard for providing his original channel masks and John Eyre, Fiona Smith (née Hilton) and Joanne Walker for their useful comments on this article.

## References

AOPP. 2002. 'RFM, Reference forward model'. Atmospheric, Oceanic and Planetary Physics, Oxford University: UK. http://www.atm.ox.ac.uk/RFM, accessed 7 November 2013.

Bormann N, Collard AD, Bauer P. 2010. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II: Application to AIRS and IASA data. *Q. J. R. Meteorol. Soc.* **136**: 1051–1063, doi: 10.1002/qj.615.

Collard AD. 2007. Selection of IASI channels for use in numerical weather prediction. *Q. J. R. Meteorol. Soc.* **133**: 1977–1991, doi: 10.1002/qj.178.

Dudhia A, Jay VL, Rodgers CD. 2002. Microwindow selection for high-spectral-resolution sounders. *Appl. Opt.* **41**: 3665–3673, doi: 10.1364/AO.41.003665.

Matricardi M. 2003. 'RTIASI-4: An improved version of the ECMWF fast radiative transfer model for the Infrared Atmospheric Sounding Interferometer'. In *Proceedings of International TOVS Study Conference XIII*, Sainte Adèle, Canada.

Matricardi M, Saunders R. 1999. Fast radiative transfer model for simulation of infrared atmospheric sounding interferometer radiances. *Appl. Opt.* **38**: 5679–5691.

Rabier F, Fourrié N, Chafäi D, Prunet P. 2002. Channel selection methods for Infrared atmospheric sounding interferometer radiances. *Q. J. R. Meteorol. Soc.* **128**: 1011–1027.

Remedios JJ, Leigh RJ, Waterfall AM, Moore DP, Sembhi H, Parkes I, Greenhough J, Chipperfield M, Hauglustaine D. 2007. MIPAS reference atmospheres and comparisons to V4.61/V4.62 MIPAS level 2 geophysical data sets. *Atmos. Chem. Phys. Discuss.* **7**: 9973–10017, doi: 10.5194/acpd-7-9973-2007.

Rodgers CD. 1996. Information content and optimization of high-spectral-resolution measurements. *Proc. SPIE* (Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II) **2830**: 136–147, doi: 10.1117/12.256110.

Rodgers CD. 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific Publishing: London.

Walker JC, Dudhia A, Carboni E. 2011. An effective method for the detection of trace species demonstrated using the MetOp Infrared Atmospheric Sounding Interferometer. *Atmos. Meas. Tech.* **4**: 1567–1580, doi: 10.5194/amt-4-1567-2011.

Weston P. 2011. 'Progress towards the implementation of correlated observation errors in 4D-Var'. Forecasting Research Technical Report 560. Met Office: Exeter, UK.

 