

CHANNEL SELECTION METHODS FOR INFRARED ATMOSPHERIC SOUNDING INTERFEROMETER RADIANCES

By FLORENCE RABIER *, NADIA FOURRIÉ, DJALIL CHAFAÏ, PASCAL PRUNET

Météo-France, France

September 18, 2001

SUMMARY

Advanced infrared sounders will provide thousands of radiance data at every observation location. The number of individual pieces of information is not usable in an operational NWP context, and we have investigated the possibilities of choosing an "optimal" subset of data. These issues have been addressed in the context of optimal linear estimation theory, using simulated IASI data. Several methods have been tried to select a set of the most useful channels for each individual atmospheric profile. These are two methods based on the Data Resolution Matrix (**DRM**), one method based on the Jacobian matrix, and one iterative method selecting sequentially the channels with largest information content. The Jacobian method and the iterative method were found to be the most suitable for the problem. The iterative method was demonstrated to always produce the best results, but at a larger cost than the Jacobian method. To test the robustness of the iterative method, a variant has been tried. It consists in building a mean channel selection aimed at optimizing the results over the whole database, and then applying to each profile this "constant" selection. Results show that this "constant" iterative method is very promising, with results of intermediate quality between the ones obtained for the optimal iterative method and the Jacobian method. The practical advantage of this method for operational purposes is that the same set of channels can be used for various atmospheric profiles.

KEYWORDS: Remote-sensing Channel selection IASI high spectral measurements

1. INTRODUCTION

Numerical Weather Prediction (NWP) systems have been using satellite data in the assimilation for some time, with positive impact demonstrated on their forecast performance (eg Bouttier and Kelly, 2001). This type of data is particularly beneficial in data-void areas such as the oceans or the southern hemisphere. With new advanced technologies, the information content of the data will be enhanced and one can expect further improvements in the analysis and forecast of the atmosphere. Noticeable improvements have already taken place after the launch of the new AMSU (Advanced Microwave Sounding Unit) data on board the NOAA polar-orbiting satellites (eg English et al, 2000). In the infrared part of the spectrum, new technologies will provide thousands of channels in each sounding. The first instrument with kilo-channel data will be AIRS (Advanced Infrared Sounder) on the Aqua satellite to be launched by NASA in 2002. On the European side, the French space agency CNES and EUMETSAT are developing the IASI instrument to be on board the first operational European polar-orbiting satellites METOP. IASI is an Infrared Atmospheric Sounding Interferometer covering the spectral range 645cm^{-1} to 2760cm^{-1} with a spectral sampling of 0.25cm^{-1} , which leads to 8461 channels in each IASI spectrum.

IASI data contain information on surface temperature, temperature humidity and ozone profiles, as well as CO_2 , N_2O , CO and CH_4 total column amounts (Camy-Peyret and Eyre, 1998). For operational NWP, the main focus will be on the assimilation of temperature, humidity, and to a lesser extent ozone information. To insert this information in the assimilation, one can either use retrievals (temperature and humidity profiles reconstructed from the radiances and prior information) or assimilate the radiances directly. Information content studies have been performed (Prunet et al., 1998, Amato et al., 1997, Collard, 1998) showing the potential benefit of IASI data. However, the large

* Corresponding author: CNRM/GMAP, Météo-France, 42 av. G. Coriolis, 31057 TOULOUSE Cedex, France. e-mail: florence.rabier@meteo.fr

volume of data and amount of computations required to treat all IASI channels is prohibitive in an operational context. The problem arising is then the choice of an "optimal" subset of data to be inserted in the assimilation. Solutions proposed to solve this problem include the selection of relevant limited spectral bands (Aires et al., 2001), the grouping of highly correlated channels in the same spectral area into super-channels (McMillin and Goldberg, 1997), the use of a partial eigen-decomposition of the radiance data (Joiner and da Silva, 1998), and the selection of individual channels based on objective criteria (Rodgers, 1996, Lerner et al., 2001, Aires et al., 2001, AIRS ATBD, 2000, and Goldberg, pers. com.). Fast radiative transfer models being mainly geared to the computation of radiances from individual channels, it is simpler to investigate the selection of individual channels for the time being. This is the approach we have pursued, and the questions we want to address are dealing with finding the best way to select a subset of IASI channels for our NWP application.

The first question is: among the various ways of selecting a limited number of channels for each individual atmospheric profile, which one leads to the best analysis accuracy? The second question is: once we have selected an "optimal" channel selection for each individual atmospheric profile, how can this method be applied robustly in practice to all atmospheric conditions? We started addressing these issues in the context of linear optimal estimation theory, using simulated IASI data computed from a set of about 500 representative atmospheric profiles for clear sky and night-time conditions. The experimental context is described in section 2, and the comparison of different channel selection methods is presented in section 3. Then, results of the use of a degraded "non-optimal" set of channels for all profiles are shown in section 4 and the influence of correlated errors and of the number of channels selected is studied in section 5. Finally, section 6 concludes this preliminary study and discusses some perspectives.

2. EXPERIMENTAL FRAMEWORK

The general framework of this study is linear optimal estimation theory in the context of Numerical Weather Prediction. The notations, definitions and data used are briefly explained in the following. The atmospheric profile in temperature and humidity at a given location is represented by a vector x and the satellite radiances by a vector y . The observations are linked to the atmospheric state by the radiative transfer equation

$$y = \mathcal{H}(x) + \varepsilon_{\mathbf{O}} + \varepsilon_{\mathbf{F}} \quad (1)$$

where the measurement and forward model errors $\varepsilon_{\mathbf{O}}$ and $\varepsilon_{\mathbf{F}}$ are assumed to be Gaussian noises with covariance matrices \mathbf{O} and \mathbf{F} . We will denote $\mathbf{R} = \mathbf{O} + \mathbf{F}$ the resulting observation error covariance matrix. Our previous knowledge of the atmospheric state is represented by the background state vector x_b . Its error covariance matrix is denoted \mathbf{B} . All errors are assumed to be unbiased (zero mean). The radiative transfer equation is assumed to be weakly non-linear, making the tangent linear assumption valid

$$\mathcal{H}(x) = \mathcal{H}(x_b) + \mathbf{H}(x - x_b) \quad (2)$$

The optimal analysed state is given by $x_a = x_b + \mathbf{K}(y - y_b)$, with $\mathbf{K} = \mathbf{A}\mathbf{H}^{\top}\mathbf{R}^{-1}$ and $\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^{\top}\mathbf{R}^{-1}\mathbf{H})^{-1}$. \mathbf{K} is the Kalman gain matrix, and \mathbf{A} is the analysis error covariance matrix. \mathbf{K} can be interpreted as the generalized inverse of \mathbf{H} , allowing one to reconstruct the atmospheric profile from the data. One can introduce the so-called Data Resolution Matrix (**DRM**) and Model Resolution Matrix (**MRM**). If one takes the ensemble average over "observation" noise ($\varepsilon_{\mathbf{O}} + \varepsilon_{\mathbf{F}}$) of the analysis equation, one

gets $x'_a = x_b + \mathbf{K}(y_t - y_b)$ where $y_t = \mathcal{H}(x_t)$ represents the true vector of observations without noise (the mean noise is zero) and x'_a is a "noise-free" analysis without any contribution from the observation error. This can be re-written as $x'_a - x_b = \mathbf{KH}(x_t - x_b) = \mathbf{MRM}(x_t - x_b)$. In observation space, the equivalent is $y'_a - y_b = \mathbf{H}(x'_a - x_b) = \mathbf{HK}(y_t - y_b) = \mathbf{DRM}(y_t - y_b)$. These two matrices (**DRM** and **MRM**) link the "noise-free" analysis to the true values, either in data or parameter space.

Some useful information theory concepts can be used to quantify the gain in information brought by the data. These are in particular the Degrees of Freedom for Signal (DFS)

$$\text{DFS} := \text{Tr}(\mathbf{I} - \mathbf{AB}^{-1}) \quad (3)$$

or the Entropy Reduction (ER)

$$\text{ER} := -\frac{1}{2} \log_2 \det(\mathbf{AB}^{-1}) \quad (4)$$

Both quantities give a global measure of the reduction of uncertainty brought by the analysis, and are well explained in Rodgers (2000). For more detailed diagnostics of the quality of the analysis, one can use the standard-deviations of analysis errors (from the diagonal components of \mathbf{A}) and an estimate of vertical resolution associated with the resulting profiles. The vertical resolution at each level in the vertical, following Purser and Huang (1993), is given by the ratio of the layer thickness with the corresponding diagonal element of the **MRM**. In an ideal case for which the analysis is equal to the true atmospheric state, **MRM** is equal to the identity matrix and the vertical resolution of the analysis is given by the resolution of our layering.

For this study, we have used a set of representative atmospheric situations. These are part of the SODA database (provided by F. Karcher, Météo-France) and form a set of 492 profiles of temperature and humidity. These data are from a wide range of meteorological conditions from various sites and at dates spanning the whole year. These are real observed atmospheric temperature profiles for all stations, and real humidity profiles for a few stations. For most stations, the humidity profile is artificially built alternating between dry and wet profiles below 600 hPa, (respectively 50% and 95% relative humidity) and fixing a 75% humidity value above 600 hPa. The data have been extended up to 80 km from the results of a photochemical model. IASI data have been simulated from each profile, using the fast radiative transfer model RTIASI (Matricardi and Saunders, 1999) over sea for nadir views. The scene is assumed to be cloud-free. Noise has been added to the simulated data using the covariance matrix \mathbf{O} provided by CNES in January 2000. Values of noise range from 0.1K-0.22K at wavenumbers 645 to 2250 cm^{-1} and increase up to 1.9K at 2760 cm^{-1} . These values are valid at a temperature of 280K, and are converted at the appropriate scene temperature for each wavenumber and each profile. An example of the observation error spectrum is given in Fig. 1 for a Biscarosse (SW France) profile.

A constant error of 0.2K coming from the radiative transfer model noise (matrix \mathbf{F}) is added to the measurement error (matrix \mathbf{O}). Radiances are assumed to be apodised, which introduces some correlations between adjacent channels. However these correlations are ignored in the observation error covariance matrix.

The parameter space is the temperature and humidity profiles over the 43 RTIASI levels. For the time being the surface temperature has not been included in the control variable. The background error covariance matrix is typical of a state-of-the-art NWP model. It is built from the 60-level ECMWF covariance matrix representing short-range forecast errors projected on our 43 levels and is broadly similar to the one used for instance in Prunet *et al.* (1998).

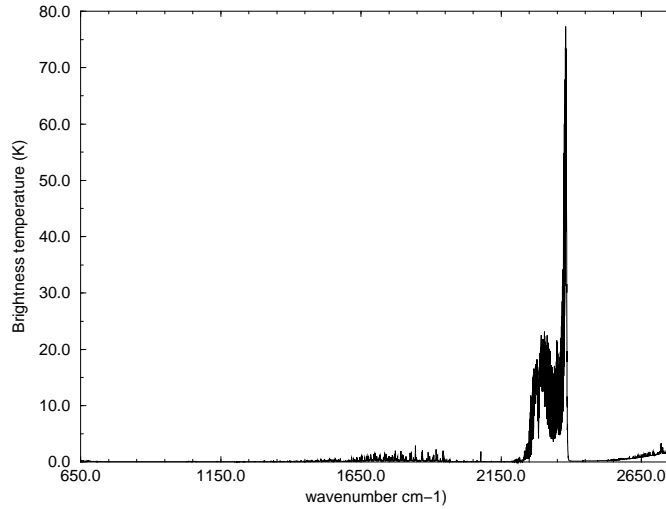


Figure 1. Typical spectrum of observation error values

3. CHOICE OF A CHANNEL SELECTION METHOD

In the framework of linear estimation theory, a few methods have been proposed to compute the information brought by each observation in the analysis. In particular, it might help us choose the channels which are the most relevant to perform a retrieval. A few methods are described below. These will then be compared to select the most efficient in terms of numbers of channels per amount of information content.

(a) Method description

i *Method based on the Data Resolution Matrix (DRM)*. Following Menke (1984), one uses the relationship $y'_a - y_b = \mathbf{DRM}(y_t - y_b)$ to interpret the **DRM** matrix. The resolution matrix **DRM** shows how much weight each observation has in influencing the analysis, and because the diagonal elements of this matrix indicate how much weight a datum has in its own analysis, these diagonal elements can be singled out and called the "importance" of the data. This interpretation suggests a simple way of selecting the channels most useful in a retrieval. It consists in picking up the channels with the largest "importance".

ii *Method based on Singular Value Decomposition (SVD) of the DRM*. A variant of the method described above can be proposed, based on the work of Wunsch (1978) and Prunet et al. (1996). A further refinement to the standard **DRM** method is that a singular value decomposition (Menke, 1984) of the normalised Jacobian matrix

$$\mathbf{G} = \mathbf{R}^{-\frac{1}{2}} \mathbf{H} \mathbf{B}^{\frac{1}{2}} \quad (5)$$

can be performed

$$\mathbf{G} = \mathbf{U} \mathbf{A} \mathbf{V}^{\top} \quad (6)$$

where the matrices \mathbf{U} and \mathbf{V} are orthogonal with respective dimensions $m \times m$ and $n \times n$. n is the dimension of the atmospheric vector and m the dimension of the observation vector. $\mathbf{\Lambda}$ is a matrix of dimension $n \times m$, containing the singular values of \mathbf{G} on the main diagonal (same indices for rows and columns). These are ordered by decreasing values. They are also equal to the square-root of the eigenvalues of $\mathbf{G}^T \mathbf{G}$. Following Prunet et al (1998), the eigenvalues of the matrix

$$\mathbf{G}^T \mathbf{G} = \mathbf{B}^{\frac{1}{2}} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{B}^{\frac{1}{2}} \quad (7)$$

can be interpreted as $\frac{\sigma_a^2}{\sigma_o^2}$. Furthermore, $\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$ can also approximately be interpreted as

$$\frac{1}{\sigma_a^2} \approx \frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}, \quad (8)$$

where σ_a , σ_b and σ_o are respectively the analysis, background and observation error standard-deviations. If one wants to use only the observations contributing by more than an arbitrary threshold of 10% to the analysis

$$\frac{\sigma_a^2}{\sigma_o^2} \geq \frac{10}{100} \quad (9)$$

this can be translated to a threshold

$$\frac{\sigma_b^2}{\sigma_o^2} \approx \frac{1}{\frac{\sigma_a^2}{\sigma_o^2} - 1} \geq \frac{1}{\frac{100}{10} - 1} = \frac{1}{9} \quad (10)$$

on the eigenvalues of the $\mathbf{G}^T \mathbf{G} = \mathbf{B}^{\frac{1}{2}} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{B}^{\frac{1}{2}}$ matrix. This leads to a truncation in $\mathbf{\Lambda}$

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (11)$$

One can write the generalized inverse \mathbf{G}^{-g} of \mathbf{G}

$$\mathbf{G}^{-g} = \mathbf{V}_p \mathbf{\Lambda}_p^{-1} \mathbf{U}_p^T \quad (12)$$

where \mathbf{V}_p and \mathbf{U}_p represent the truncated vectors corresponding to \mathbf{V} and \mathbf{U} . The matrix

$$\mathbf{G} \mathbf{G}^{-g} = \mathbf{U}_p \mathbf{U}_p^T \quad (13)$$

is the Data Resolution Matrix, for this truncated problem retaining only the significative information from the observations. As in the previous method, the diagonal elements of this matrix can be interpreted as the degree to which the corresponding channel is independently contributing useful information to the system.

iii *Iterative Method.* This method follows Rodgers (1996). Let us first perform a normalisation of the Jacobian matrix by $\mathbf{R}^{-\frac{1}{2}}$, $\mathbf{H}' = \mathbf{R}^{-\frac{1}{2}} \mathbf{H}$.

The use of a single channel at a time changes the error covariance matrix by

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{h} \mathbf{h}^T \quad (14)$$

where \mathbf{h} is the vector corresponding to a line in \mathbf{H}' linked to this channel. Following the classical relationships

$$\mathbf{A} = \mathbf{B} - \mathbf{KHB} = \mathbf{B} - \mathbf{BH}^\top (\mathbf{R} + \mathbf{HBH}^\top)^{-1} \mathbf{HB} \quad (15)$$

one can also write

$$\mathbf{A} = \mathbf{B} - \frac{1}{1 + (\mathbf{Bh})^\top \mathbf{h}} \mathbf{Bh}(\mathbf{Bh})^\top = \mathbf{B} \left(\mathbf{I} - \frac{1}{1 + (\mathbf{Bh})^\top \mathbf{h}} \mathbf{h}(\mathbf{Bh})^\top \right) \quad (16)$$

Using the following relationships between vectors \mathbf{a} and \mathbf{b}

$\text{Tr}(\mathbf{ab}^\top) = \mathbf{b}^\top \mathbf{a} = \mathbf{a}^\top \mathbf{b}$ and $\det(\mathbf{I} + \mathbf{ab}^\top) = 1 + \mathbf{b}^\top \mathbf{a} = 1 + \mathbf{a}^\top \mathbf{b}$, one gets the following expressions for the DFS and the entropy reduction ER

$$\text{DFS} := \text{Tr}(\mathbf{I} - \mathbf{AB}^{-1}) = \frac{\mathbf{h}^\top \mathbf{Bh}}{1 + \mathbf{h}^\top \mathbf{Bh}} \quad (17)$$

and

$$\text{ER} := -\frac{1}{2} \log_2(\det(\mathbf{AB}^{-1})) = \frac{1}{2} \log_2(1 + \mathbf{h}^\top \mathbf{Bh}). \quad (18)$$

The method consists in performing a series of successive analyses, each one using only one channel at a time. The analysis error covariance matrix is updated accordingly, and it is used at the next step. It thus takes into account the information brought by all previous channels before selecting the next one. The channel selection at each step is based on optimizing the DFS or the ER. In this study, the ER has been chosen as figure of merit, but both criteria give similar results. A variant of an application of this method can be found in Lerner et al (2001) where it is applied separately for each retrieval level. This method is quite expensive in terms of CPU time because one has to scan all remaining channels before performing the selection of a single channel. In the evaluation of the methods, an interesting variation of the method is also tried. It is the iterative method for which the error covariance matrix is not updated at each intermediate analysis step. The background error covariance matrix \mathbf{B} is kept throughout the iterative analysis. It corresponds to selecting the channels based on information content figures of merit, but not taking into account the gain in information brought by previously selected channels.

iv *Jacobian Method.* Another way of selecting channels which are relevant for the retrieval problem is to perform a selection based on the characteristics of their weighting functions (Aires et al, 2001, Goldberg, pers. com.). Here, the Jacobian matrix has first been normalised by the square-roots of \mathbf{R} and \mathbf{B} . Then, for each retrieved parameter, at each level in the vertical, one selects channels among those peaking next to the level with the largest ratio: amplitude of the peak / "width" of the weighting function. The "width" is actually computed in our application as the square-root of the sum of the absolute values of the components of the weighting function relative to the parameter being treated (temperature or humidity) over the 43 levels. A few variants have been tried before adopting this particular criterion. However, this method is rather "ad-hoc" and results might benefit from further tuning of the function to be optimised. The cost of this method is quite modest. The jacobian method is about 3 times to 20 times faster than the iterative method. This cost ratio increases with an increasing number of channels to be selected. A typical example is a factor of 3.5 for the selection of 300 channels.

(b) *Results for various methods*

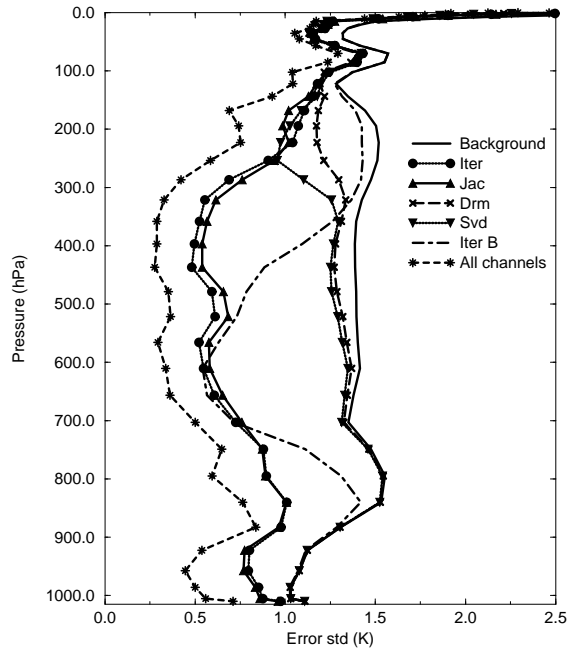


Figure 2. Analysis error standard-deviations based on various methods used to select 43 channels for the retrieval of temperature. "Background" corresponds to the standard deviation of background errors. "Iter", "Jac", "svd" and "drm" are related to the standard deviation of analysis errors obtained with the iterative, jacobian, Singular Value Decomposition and Data Resolution Matrix methods. "Iter B" corresponds to the standard deviation obtained with the iterative method for which no update of the error covariance matrix **B** is made. "All channels" corresponds to an experience where all IASI channels are used in the analysis.

Before selecting some useful channels among all the spectrum, it seemed relevant to eliminate beforehand those which might not be adequately simulated by the radiative transfer model. Similar to Lerner *et al.* (2001), spectral bands sensitive to foreign gases are eliminated. It was not found necessary to exclude the main area of sensitivity to O_3 ($1000\text{-}1080\text{ cm}^{-1}$) as channels in this waveband were never picked by the selections. The excluded spectral bands are crudely defined as $1267\text{-}1312\text{ cm}^{-1}$ for N_2O and CH_4 and $2092\text{-}2355\text{ cm}^{-1}$ for N_2O and CO . One can note that some information on temperature might be lost in the relatively broad band $2092\text{-}2355\text{ cm}^{-1}$, whereas some other aspects have not been taken into account (CFCs, channels that are sensitive to effects which are not in Local Thermodynamic Equilibrium,...). The excluded bands should be updated more precisely in the future. This pre-selection reduces the total number of channels from 8461 to 7227 for a given spectrum. It is used before each selection. In a first step, the selection methods are compared in a configuration allowing us to emphasize the differences. The framework is the retrieval of temperature values for a given mid-latitude profile by using only 43 channels. In this special case, the control variable is limited to atmospheric temperature on the 43 pressure levels. The profile was chosen among those available for the Biscarosse station (SW France). Results on other profiles have also been obtained but they are not reproduced here, as they did not bring any message different

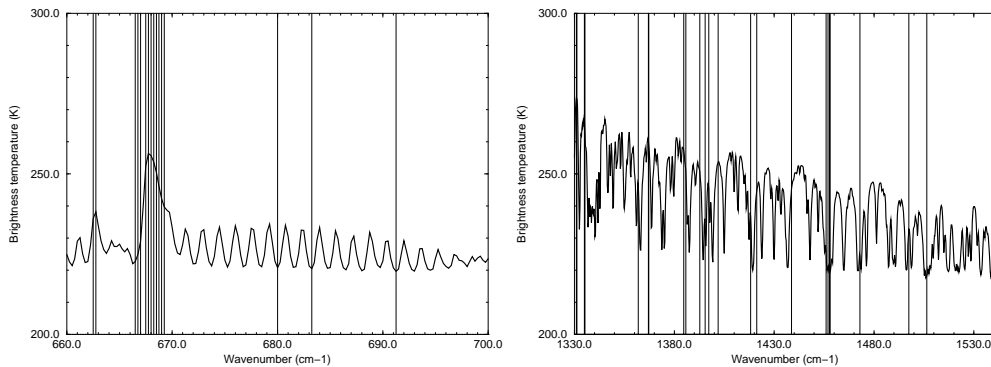


Figure 3. Location of the 43 channels selected for the retrieval of temperature by the iterative method. The left-hand panel is a zoom in the 15 micron band. The right-hand panel is a zoom in the water vapour band.

from those obtained for this particular profile. Figure 2 shows the standard-deviation of analysis error, as estimated by the linear estimation analysis, for the Biscarosse profile and for the various methods. The background errors are between 1 K and 1.5K for most of the pressure levels, and reach larger values for the top of the atmosphere. The analysis errors when all channels are used are shown as stars: they are well below 1K, and even below 0.5K for parts of the profile in the upper troposphere. The iterative and Jacobian methods using 43 channels lead to very reasonable results below 1K in the troposphere. Both methods are "forced" to improve the whole atmospheric profile, by conception: the iterative method updates its error covariance indicating where the intermediate analysis has already been improved, and the Jacobian method works in a sequential manner, level by level. On the other hand, the iterative method with fixed error covariance matrix \mathbf{B} leads to an analysis of good quality only in the mid-troposphere, around 600hPa. This indicates that the channels with largest information content are in the mid-troposphere, which is not surprising as the weighting functions have a large amplitude in this area. One clearly sees the degradation brought by this version of the iterative method, which illustrates the benefit gained by updating the analysis error covariance matrix by taking into account the information coming from the channels already assimilated. As for the DRM and SVD methods, their channel selections lead to a good analysis accuracy in the stratosphere, but not below. The fact that these methods, similarly to the degraded iterative method, are not able to scan all levels in the vertical is not surprising. Both types of methods do not take into account the redundancy between channels. It might be more surprising to note that the favoured area is not similar between the DRM and SVD methods on the one hand, and the degraded iterative on the other hand. The ingredients included in the computation of the information criteria are the weighting functions, the observation error matrix \mathbf{R} in both cases. The difference is that the DRM and SVD methods use \mathbf{A} in

$$\mathbf{DRM} = \mathbf{H}\mathbf{K} = \mathbf{H}\mathbf{A}\mathbf{H}^T\mathbf{R}^{-1} \quad (19)$$

whereas the degraded iterative method uses \mathbf{B} in

$$\text{ER} := -\frac{1}{2} \log_2(\det(\mathbf{A}\mathbf{B}^{-1})) = \frac{1}{2} \log_2(1 + \mathbf{h}^T\mathbf{B}\mathbf{h}). \quad (20)$$

For the degraded iterative method using \mathbf{B} , 600hPa is an area where the weighting functions have a large amplitude and \mathbf{B} has a reasonably large amplitude. Whereas for

the SVD or DRM methods, the analysis errors being quite large above 300hPa, it shifts the emphasis to the stratosphere. These different results arise because of slightly different philosophies in the channel selection. The question being asked by the degraded iterative method is: where would a channel bring the most information knowing where our errors in the background are? In a similar way, the question being asked by the DRM and SVD methods is: which is the most useful channel in the analysis knowing that all channels have been inserted in the analysis? For a posteriori diagnostics the DRM and SVD methods are very relevant, but it seems that for channel selection purposes the other approach based on \mathbf{B} is more appropriate. However, none of these methods is comparable to the iterative or Jacobian methods in terms of analysis accuracy. It is interesting to look at where the channels are picked in the spectrum. The iterative method giving the best results, the channel selection deduced from this method has been illustrated for various bands in Fig. 3. Apart from three channels being picked at 2021.5, 2022 and 2022.25 cm^{-1} , these two panels show all the selected channels. As already shown by Prunet *et al.* (1998) and Amato *et al.* (1997), most information on temperature comes from the 15micron band (a zoom is available in the left-hand panel), the noise being too large in the 4.3micron band for IASI to be the area preferred for the selection of the most useful channels. One can see in Fig. 3 that adjacent channels are often picked by the method (as clearly shown in the left-hand panel for the CO_2 ν_2 Q-branch), and that channels tend to be picked in between lines (as seen in the right-hand panel). This location optimizes the sharpness of the weighting functions (Kaplan *et al.*, 1977). One can note that water vapour channels are preferably selected on the long wavelength side of the 6.7 micron band. This certainly has to do with increasing noise on the short wavelength side of this band (see Figure 1).

Results are generalised to the selection of 300 channels for the 24 profiles of the Biscarosse station. The control variable now consists of the temperature and humidity profiles on the 43 pressure levels. Different channel selections are performed separately on each of the 24 profiles: there are as many channel selections as profiles and they are adapted to each individual profile. Average results for the various methods are shown in Fig. 4. These results broadly confirm those previously obtained for 43 channels. However, the characteristics of the methods tend to be less striking as they use more channels. The two best methods are still the iterative and the Jacobian methods, with the iterative being now significantly better. This might be explained by the fact that the Jacobian method is not as optimal, being forced by construction to optimize sequentially all 43 levels for temperature (T) and humidity (Q). It selects sequentially the channel most suited for the retrieval of temperature on the first pressure level, then on the second one and so on, and proceeds similarly with humidity. In particular, a number of channels are devoted to the retrieval of Q in the upper levels of the atmosphere in exactly the same way as for the lower levels for temperature for instance: there is no hierarchy between the various components of the parameter space.

After this preliminary study, the intermediate conclusion might be that two methods can be considered as suitable methods to be used for channel selection. These are the iterative method, which systematically gives the best results and the Jacobian method which is slightly worse but more economical.

4. USE OF A NON-OPTIMAL SET OF CHANNELS

Considering that the iterative selection is too expensive to be performed for each atmospheric profile, one might envisage two options for operational purposes: either a Jacobian selection adapted to each profile, or a "constant" selection based on off-line

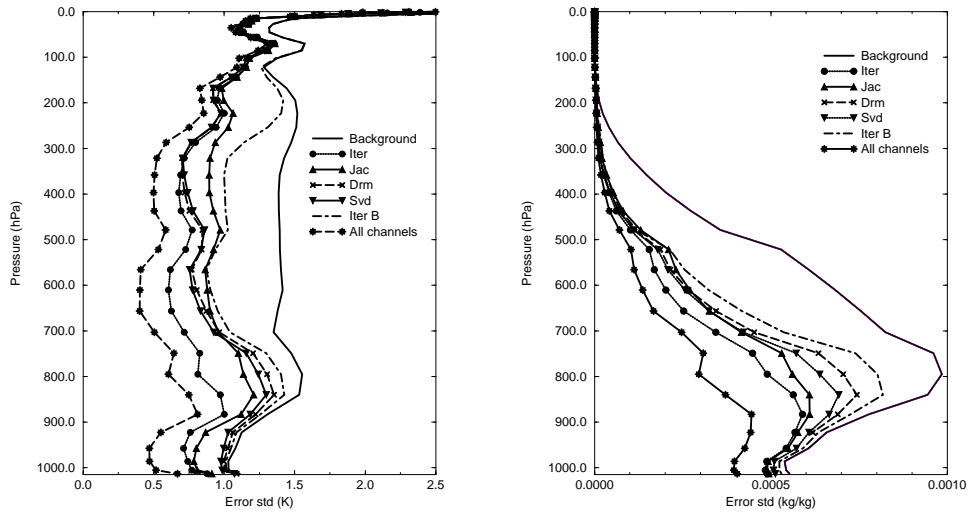


Figure 4. Standard-deviation of errors averaged over 24 profiles, for various channel selections using 300 channels to retrieve T (left-hand panel) and Q (right-hand panel). "Background" corresponds to the standard deviation of background errors. "Iter", "Jac", "svd" and "drm" are related to the standard deviation of analysis errors obtained with the iterative, jacobian, Singular Value Decomposition and Data resolution matrix methods. "Iter B" corresponds to the standard deviation obtained with the iterative method for which no update of the error covariance matrix \mathbf{B} is made. "All channels" corresponds to an experience where all IASI channels are used in the analysis.

studies with the iterative method. Both approaches are interesting a priori, but one needs to know by how much the information content is reduced with the "constant" iterative selection method before judging the relative quality of the two methods. Our goal with a "constant" selection is to find one channel selection to be applied to all the profiles to obtain the best average performance. This is investigated in the following. Firstly, a "constant" channel selection is computed as an average selection based on the set of 492 different channel selections (one per profile) provided by the iterative approach. For each of the 492 profiles, each channel picked by the iterative method is given a rank from 1 to 300 which corresponds to its rank in the iterative selection process. All the others channels which are not selected are given the largest possible rank corresponding to the number of channels on which the selection is performed (7227). Two options were then tried for the average selection over all profiles: taking the 300 channels with lower average rank or those which are picked the most often, whatever their rank between 1 and 300. Both options were found to give very similar results, with a slight advantage to the method based on averaged rank, which is then chosen for "constant" selection. This constant selection is presented in Fig. 5. The beginning of the spectrum (15 micron band) is very well sampled by the selection, compared to the other major CO_2 band (4.3 micron) at the end of the spectrum for which the IASI noise is large. The H_2O band is also an area adequately sampled: 172 selected channels are regularly picked up between 1310 cm^{-1} - 1610 cm^{-1} . When applying this constant selection to the whole set of profiles, one loses around 0.7 units in terms of DFS, compared to the optimal selection varying with each profile. This is about a third of the loss in terms of DFS (2.2) when going from the iterative to the Jacobian method for which the values are respectively

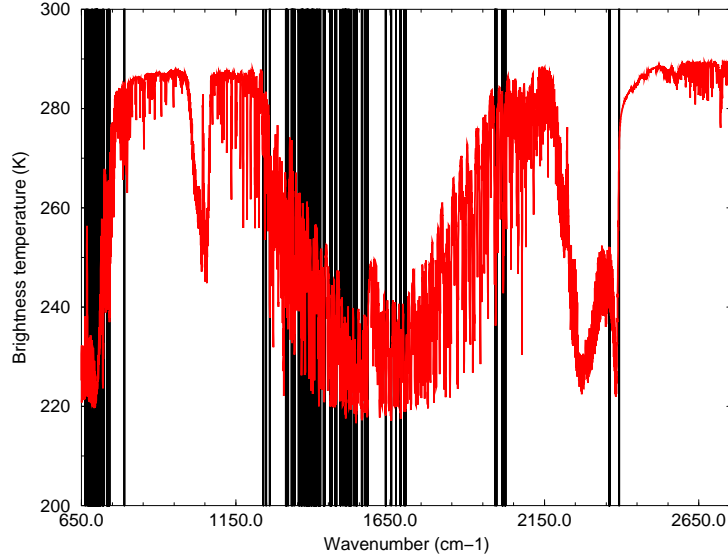


Figure 5. Location of the 300 channels selected for the retrieval of temperature and humidity by the "constant" iterative method.

19.09 and 16.87. Results in terms of analysis errors and vertical resolution are shown in Fig. 6 and 7. Comparing the iterative method with the "constant" iterative method, one can see a degradation, but the "constant" selection is still better than the Jacobian method. The degradation brought by the "constant" iterative method is more marked for the humidity profiles, because of more natural variability in this variable. In any case, the conclusion is that the "constant" iterative selection approach is a good candidate for operational purposes.

5. INFLUENCE OF CORRELATED ERRORS AND OF THE NUMBER OF CHANNELS SELECTED

(a) *Influence of correlated errors*

Some of the assumptions made in this study are known to be rather crude. For instance, the assumption of uncorrelated errors between channels is not realistic, in particular because of apodisation. We have tried to take the correlation between adjacent channels into account in an additional experiment (not shown). The selection algorithm was forced to discard any adjacent channels once one was picked, for the subsequent choices. This led to a solution where no adjacent channels were present in the selection, in order to minimize the risk of correlation between channels to be used in the retrieval/assimilation. The results in terms of analysis accuracy were almost identical to those obtained with the unconstrained iterative method, although the selections were

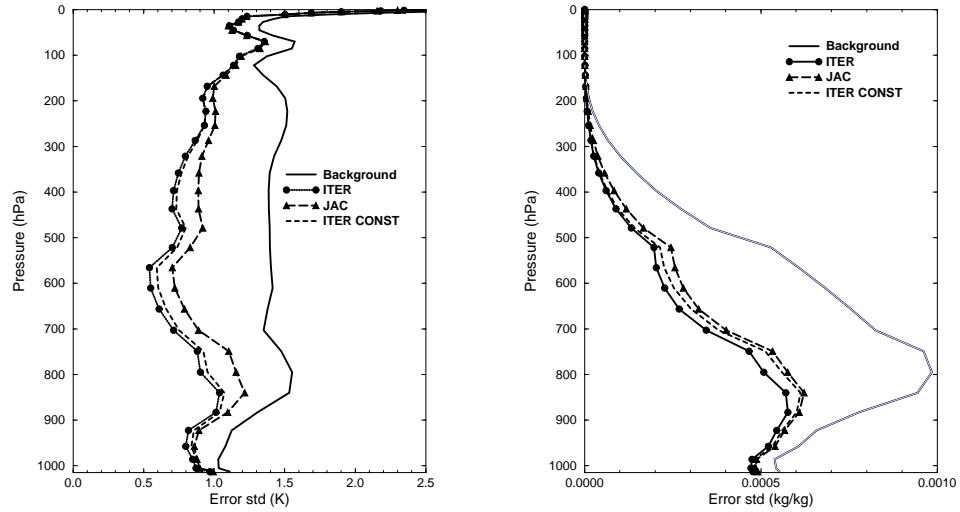


Figure 6. Standard-deviation of errors averaged over 492 profiles, for various channel selections using 300 channels to retrieve T (left-hand panel) and Q (right-hand panel). "Background" corresponds to the standard deviation of the background errors, "iter" and "jac" corresponds respectively to the one obtained with the iterative and the jacobian method and "iter const" to the one derived from a constant selection deduced from the average of the 492 individual selections.

different. This experiment seems to indicate that, as long as the inter-channel error correlation is quite localized, the iterative selection algorithm can be made to find a channel selection that picks up uncorrelated channels and leads to a satisfactory analysis accuracy. The problem should be investigated further with some more realistic error correlations which might be less localized in the spectrum due to radiative transfer model errors for instance (Sherlock, 2000).

(b) *Influence of the number of channels selected*

The various comparisons performed in this study have assumed that 300 channels would be a typical number of pieces of information one can introduce in the operational assimilation systems in the near-future. It is however interesting to illustrate the variation of analysis quality when this number is varied from less than 100 to 8461. Results are shown in Fig. 8 and 9 for respectively the analysis error standard-deviations and the analysis vertical resolution (averaged over the Biscarosse profiles). The gain brought by the first 86 channels is remarkable compared to the background. Increasing the number of channels gradually increases the information content and analysis quality. This improvement in analysis quality is more obvious in the tropospheric temperature than in the stratospheric temperature, and very significant in the lower humidity profile. Although 300 channels provide a considerable achievement in terms of analysis quality compared to what can be expected from other instruments, it is obvious that the information content has not yet reached an asymptotic value and one would benefit from using at least up to 2000 channels. For instance, going from 300 to 8461 channels brings a decrease in maximum error in the tropospheric temperature profile from 1.5K to 1K, and in the humidity profile from 0.6 g/kg to 0.4 g/kg. Similarly, the vertical resolution typical for temperature

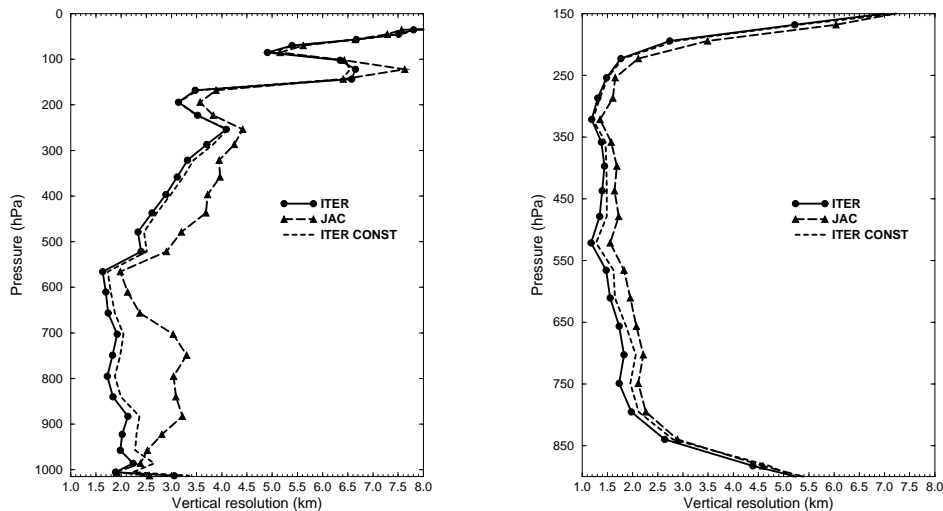


Figure 7. Vertical resolution of the analysis averaged over 492 profiles, for various channel selections using 300 channels to retrieve T (left-hand panel) and Q (right-hand panel). The abbreviations of the curves correspond to those of Fig 6.

and humidity is improved from around 2km to 1.5km. One should note that these results are very dependent on the background error covariance matrix \mathbf{B} used in this study, and cannot be compared easily to other results using very different background information.

6. SUMMARY AND PERSPECTIVES

In this paper, the issue of reducing the number of IASI channels to be used in data assimilation has been addressed in the context of optimal linear estimation theory, using simulated IASI data computed from a set of about 500 representative atmospheric profiles, with observation errors taken from current estimated values and background errors typical of a state-of-the-art NWP system. The selection of an "optimal" subset of data can be performed in various ways and different methods have been tried in this study. These are two methods based on the Data Resolution Matrix (**DRM**) for which one picks up channels corresponding to the largest diagonal elements of the **DRM**, one method based on the Jacobian matrix selecting the channels with most appropriate weighting functions for retrieving each parameter at each level, and one iterative method selecting sequentially the channels with largest information content and updating the error covariance matrix of the current analysis according to this new information. The methods based on the **DRM** were found to be more relevant for a posteriori diagnostics than for prior channel selection, whereas the Jacobian method and the iterative method were suitable for the problem. The iterative method was found to always produce the best results, but at a larger cost than the Jacobian method. The iterative method is 3.5 times (for 300 channels) to 18 times (for 7227 channels) slower than the jacobian method. Two variants of the iterative method have also been tried in an attempt to reduce its cost. One is an iterative method for which the update of the current analysis error estimate is not performed. The other one is an iterative method for which a "constant" channel

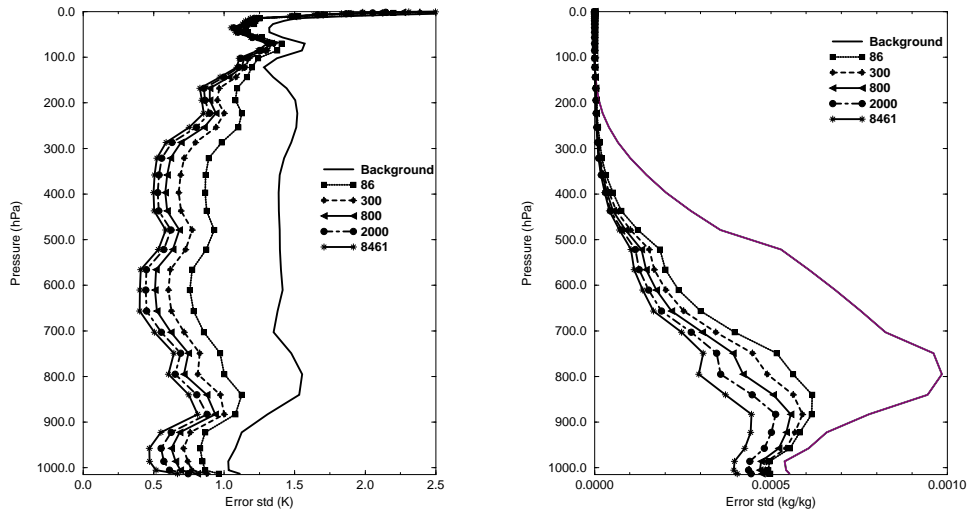


Figure 8. Standard-deviation of errors averaged over the Biscarosse profiles, for various number of channels selected with the iterative method to retrieve T (left-hand panel) and Q (right-hand panel).

selection is applied to each profile, this mean channel selection having been computed beforehand on the database. The iterative method with no update of the error covariance matrix leads to very degraded results, highlighting the need for taking into account the possible redundancy between channels. In contrast, results show that the "constant" iterative method is very promising, with results of intermediate quality between the ones obtained for the optimal iterative method and the Jacobian method. Furthermore, this method is thought to be quite robust: computing the "constant" channel selection on a subset of profiles (62, representing 13% of the 492 profiles of the whole database) does not drastically change the selection (with 84% of channels in common). The practical advantage of this method for operational purposes is that the same set can be used for various atmospheric profiles.

The channel selection problem has been tackled here in the context of a large set of representative profiles with different air-masses, but only for clear-sky, nadir views over sea. An obvious limitation is that the information content is known to depend on cloud amount (Eyre, 1990), on surface characteristics and scan angle. The state vector should be expanded to include other parameters such as surface temperature and emissivity to investigate further the channel selection issue over land. The channel selection for cloudy pixels should also be studied carefully. For low clouds, IASI may contain useful information down to cloud level. In these cases, it may be particularly important to include channels with sharp weighting functions, and this needs to be further studied. This is particularly important for forecast performance quality, as Fourrié (2000) and McNally (2001) have shown that sensitive areas are widely covered by clouds. These sensitive areas are parts of the globe where initial errors in the atmospheric structure might develop into significant errors (Rabier et al, 1996). If sounders such as IASI have to lead to a very significant impact on the forecast quality, one should address in more details the question of retaining the most information possible in these sensitive atmospheric

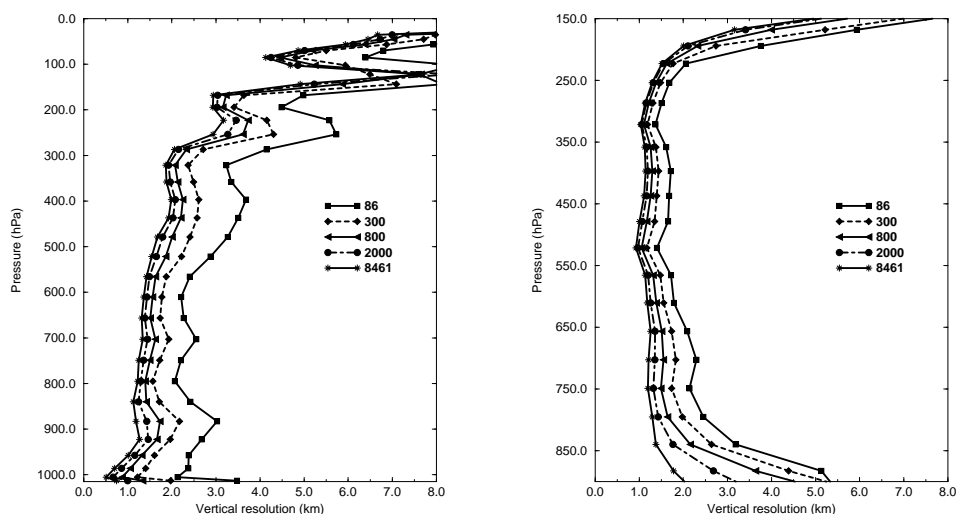


Figure 9. Vertical resolution of the analysis averaged over the Biscarosse profiles, for various number of channels selected with the iterative method to retrieve T (left-hand panel) and Q (right-hand panel).

structures leading to major forecast failures. Studies such as those of Prunet *et al* (1998) and Collard (1998) should now be re-addressed to see how a limited number of channels can adequately represent fine structures in the vertical in the cases most relevant for forecast quality, at least down to cloud level.

In practice, the channel selection could be performed in two steps. Firstly, one would extract a rough pre-selection based on monitoring statistics, keeping only the channels for which both the observation and the forward model errors are small. Secondly, one would compute several sets of channel selections for various configurations of scan angle, type of surface and cloud conditions. These sets of channel selections might have a core of common channels and additional ones tailored to the specific configuration. If the total number of channels likely to be used in any of the configurations is still significantly less than the total number of useful IASI channels, the selection process could also be used to reduce transmission costs and delays for real-time data transfer to NWP centres.

ACKNOWLEDGEMENTS

This study was partly supported by CNES through post-doctorate fellowships of Nadia Fourrié and Pascal Prunet. It also benefitted from the MASSIF contract between CNRM and EUMETSAT.

REFERENCES

- | | | |
|---|------|--|
| Amato, U., Cuomo, V., Rizzi, R. and Serio, C. | 1997 | Evaluating the effect of inter-relationships among the different spectral bands on IASI performance, <i>Q. J. R. Meteorol. Soc.</i> , 123 , 2231–2244 |
| Aires, F., Chédin, A. and Scott, N.A. | 2001 | A Regularized Neural Net approach for Retrieval of Atmospheric and Surface Temperatures with the IASI Instrument, <i>To appear in the Journal of Applied Meteorology</i> |

- AIRS ATBD 2000 Algorithm Theoretical Basis Documents, available at <http://eosps0.gsfc.nasa.gov/atbd/airstables.html>
- Bouttier, F. and Kelly, G. 2001 Observing system experiments in the ECMWF 4D-Var data assimilation system, *Q. J. R. Meteorol. Soc.*, **127**, 1469–1488
- Camy-Peyret, C. and Eyre, J.R. 1998 The IASI Science Plan, ISSWG report available from <http://www-projet.cst.cnes.fr:8060/IASI/index.html>
- Collard, A.D. 1998 Notes on IASI performance, *NWP Technical Report No. 253, Met Office, Bracknell*
- English, S.J., Renshaw, R.J., Dibben, P.C., Smith, A.J., Rayer, P.J., Poulsen, C., Saunders, F.W. and Eyre, J.R. 2000 A comparison of the impact of TOVS and ATOVS satellite sounding data on the accuracy of numerical weather forecasts, *Q. J. R. Meteorol. Soc.*, **126**, 2911–2931
- Eyre, J.R. 1990 The information content of data from satellite sounding systems: A simulation study, *Q. J. R. Meteorol. Soc.*, **116**, 401–434
- Fourrié, N. 2000 Intérêt des observations du sondeur satellitaire TOVS pour l'analyse et la prévisions des dépressions pendant FASTEX. *Phd dissertation, Université paris VI*
- Joiner, J. and da Silva, A.M. 1998 Efficient methods to assimilate remotely sensed data based on information content, *Q. J. R. Meteorol. Soc.*, **124**, 1669–1694
- Kaplan, L.D., Chahine, M.T., Susskind, J. and Searl, J.E. 1977 Spectral band passes for a high precision satellite sounder, *Applied Optics*, **Vol 16, No.2**, 322–325
- Lerner, J. A., Weisz, E. and Kirchengast, G. 2001 Temperature and humidity retrieval from simulated Infrared Atmospheric Sounding Interferometer (IASI) measurements, *Submitted to Journal of Geophysical Research*
- Matricardi, M. and Saunders, R. 1999 A fast radiative transfer model for Infrared Atmospheric Sounding Interferometer radiances, *J. Appl. Optics*, **38**, 5679–5691
- McMillin, L., M. and Goldberg, M., D. 1997 The use of super channels for high resolution soundings, *Proceedings of the Ninth International TOVS Study Conference, Igls, Austria, 20-26 February 1997* 327–330
- McNally, A., P. 2001 The occurrence of cloud in meteorologically sensitive areas and the implication for advanced infrared sounders, *Proceedings of the Eleventh International TOVS Study Conference, Budapest, Hungary, 20-26 September 2000* xxx–xxx
- Menke, W. 1984 *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press, New York, 289pp
- Prunet, P., Minster, J.-F., Ruiz-Pino, D. and Dadou, I. 1996 Assimilation of surface data in a one-dimensional physical-biogeochemical model of the surface ocean. 1. Method and preliminary results *Global Biogeochem. Cycles*, **10**, 111–138
- Prunet, P., Thépaut, J.-N. and Cassé, V. 1998 The information content of clear sky IASI radiances and their potential for numerical weather prediction, *Q. J. R. Meteorol. Soc.*, **124**, 211–241
- Purser, R.J. and Huang, H.-L. 1993 Estimating effective data density in a satellite retrieval or an objective analysis, *J. Appl. Meteor.*, 1092–1107
- Rabier, F., Klinker, E., Courtier, P. and Hollingsworth, A. 1996 Sensitivity of forecast errors to initial conditions, *Q. J. R. Meteorol. Soc.*, **122**, 121–150
- Rodgers, C. D. 1996 Information content and optimisation of high spectral resolution measurements, *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II, SPIE, Volume 2830*, 136–147
- Rodgers, C. D. 2000 Inverse methods for Atmospheres: Theory and Practice, *World Scientific Publ., Singapore*, 238pp
- Sherlock, V.J. 2000 Impact of RTIASI fact radiative transfer model error on IASI retrieval accuracy, *Forecasting Research Technical Report No. 319, Met Office, Bracknell*

- Wunsch, C. 1978 The general circulation of the north Atlantic west of 50W determined from inverse methods, *Rev. Geophys.*, **16**, 583-620