Super-Channel Selection for IASI Retrievals

Peter Schlüssel

EUMETSAT, Am Kavalleriesand 31, 64295 Darmstadt, Germany

Abstract

The Infrared Atmospheric Sounding Interferometer (IASI), to be flown on Metop as part of the EUMETSAT Polar System (EPS), will be used to derive a number of atmospheric parameters. A challenging task will be to make proper use of the 8461 spectral radiance samples provided by IASI. Despite the development of fast radiative transfer models it will not be possible to make direct use of all samples in a variational retrieval scheme or to assimilate them all in numerical weather forecasts because of the huge amount of data. As many of the spectral radiances are well correlated with each other it seems straightforward to combine highly correlated ones to so-called super channel clusters. The advantages are reduced noise of the super channels, when compared to that of measured single spectral samples, and the possibility to chose only one of the samples to represent each cluster in radiative transfer calculations. The composition of the super channels and their usefulness for the retrieval of temperature and water vapour profiles in diverse atmospheric situations is studied by means of RTIASI-5 simulations for globally distributed sets of atmospheric and surface situations. A variational retrieval that makes use of super channels has been implemented in the core ground segment of the EPS for the generation of level 2 products.

Introduction

The Infrared Atmospheric Sounding Interferometer (IASI), that has been built under EUMETSAT- CNES co-operation, will be flown on the Metop satellites as part of the EUMETSAT Polar System (EPS) from 2006 onwards. Details of the IASI instrument are described by Cayla (1993). IASI will deliver high-resolution radiance spectra that allow to retrieve atmospheric temperature and humidity profiles for numerical weather prediction and climate research at accuracies of 1 K and 5%, respectively, at high vertical resolution. Trace gases to be derived from IASI include ozone columnar amounts in deep layers, and columnar amounts of carbon monoxide, nitrous oxide, methane, and carbon dioxide. Cloud parameters measured from IASI are cloud fraction, cloud top temperature, cloud height, and cloud phase.

The IASI Level 2 processor (Schlüssel et al., 2005) makes use of different retrieval techniques. These include statistically based methods like regression on principal component scores (Schlüssel and Goldberg, 2002; Goldberg et al., 2003) and artificial neural networks (Turquety et al., 2004). Alternatively, a variational retrieval (e.g. Rodgers, 2000) can be included either as stand-alone retrieval or be used in combination with a statistical retrieval, where the latter provides the first guess for the former one. The variational retrieval implemented is a simultaneous iterative retrieval seeking the maximum posterior probability solution for the minimisation of a cost function

$$J = (\mathbf{y}(\mathbf{x}) - \mathbf{y}^{\mathbf{m}}) \cdot \mathbf{E}^{-1} \cdot (\mathbf{y}(\mathbf{x}) - \mathbf{y}^{\mathbf{m}})^{\mathbf{T}} + (\mathbf{x} - \mathbf{x}^{\mathbf{b}}) \cdot \mathbf{C}^{-1} \cdot (\mathbf{x} - \mathbf{x}^{\mathbf{b}})^{\mathbf{T}} \quad , \qquad (1)$$

where $\mathbf{x}$ is the atmospheric state vector as calculated iteratively, $\mathbf{x}^{\mathbf{b}}$ is the background atmospheric state, $\mathbf{C}$ is the covariance matrix associated with the background, $\mathbf{y}^{\mathbf{m}}$ is the measurement vector, $\mathbf{y}(\mathbf{x})$ is the forward model operator at a given state $\mathbf{x}$, and $\mathbf{E}$ is the combined measurement and forward model error covariance. The minimisation is achieved by the Marquardt-Levenberg method. However, despite the availability of fast radiative transfer models, it is not possible to perform a variational retrieval that includes separately all 8461 spectral samples in $\mathbf{y}$ and $\mathbf{y}^{\mathbf{m}}$. Constraints of the near-real time processing with today's computer capabilities require that the number of radiances included in the iterative retrieval is restricted to about 500. Although this is a strong limitation, the fact that the IASI spectrum contains much redundant information and many of the spectral samples are highly correlated with each other, enables the inclusion of the entire IASI spectrum in the variational retrieval by the use of super-channel clusters.

Super-Channel Composition

Redundancy in the IASI spectrum allows the composition of super-channel clusters, in which highly correlated spectral samples are combined. Single clusters can be populated by one to many (often hundreds) single samples, of which one in each cluster, so-called lead sample, is chosen to represent the cluster in the calculation of the forward model operators $y_L(\mathbf{x})$ and the related Jacobians. The measured radiance of the clusters are given by a weighted average of the contributing samples, where the average has to represent the radiance of the lead channel. The more samples contribute to a cluster the lower the effective measurement noise will be in the averaged radiance of the cluster. The weighted average is best achieved by linear regression of the lead channel radiance against all radiances in the same cluster. The weights consist of regression coefficients, taking into account the correlation and the noise of the respective samples. The "measured" radiance of a lead channel is calculated as

$$y_L^m = \frac{1}{N} \sum_{i=1}^{N} a_i + b_i y_i^m \ , \qquad (2)$$

where N is the number of spectral samples in a cluster, y_i^m is the measured radiance of spectral sample i, and a_i, b_i are regression coefficients (for the lead channel we have a_1=0 and b_1=1). The errors in the measured radiance y_L^m of the lead channel will include measurement and regression errors.

Data Set and radiative Transfer Calculations

For the generation and analysis of super-channel clusters IASI spectra have been simulated for a globally representative set of 53980 atmospheric and surface situations. The basis for this set is the sub-sampled

ERA-40 (40 years ECMWF Re-Analysis) as described by Chevallier (2001). The rather smooth temperature profiles have been randomly perturbed with 1K/1km double-Dirac dipoles in order to represent fine structure in the profiles. Land surface types have been randomly composed of up to three types (out of 14 possible ones) and corresponding random fractions. According to the surface types emissivity spectra have been composed according to Snyder et al. (1998), where the mean emissivity spectra have been randomly varied between the possible limits of each type. Trace gas profiles have been taken from climatology, they have been randomly varied between 50 and 150% of the respective trace gas amounts. Aerosol types have been randomly selected, and their climatological profiles have been randomly varied between 0 and 200%. About half of the data set has been set up to simulate oceanic conditions with sea surface conditions where the surface emissivity is parameterised according to the wind-roughened sea surface.

For the entire set of 53980 cloud-free situations simulations have been carried out with the fast radiative transfer model RTIASI-5 (Matricardi, 2004a, 2004b). The instrument viewing angles for each simulation was chosen randomly among the possible values. The resulting Gaussian apodised radiance spectra are normalised with the standard deviation of the respective noise spectra before entering the correlation analysis.

Correlation and Regression Analysis

A correlation analysis has been done between spectral samples. Starting at the low wavenumber boundary of 645 cm^{-1} the spectrum has been scanned for samples with a high correlation. All pairs with a correlation higher than a threshold are retained in the respective cluster, irrespective of the spectral distance between the samples, and each sample must be a member of exactly one cluster. The assumed thresholds for the minimum correlation vary between 0.95 and 0.999. The number of super-channel clusters obtained varies according to the threshold and generally decreases with rising threshold as seen in table 1. At a low correlation of 0.95 the number of super-channel cluster is as low as 47 and it is clear that the spectral information which was originally in the IASI spectrum is lost in a sense comparable to a broadband filter radiometer. The number of samples within single clusters can reach several hundreds, but can also be as low as one. At higher correlation thresholds the clusters are more sparsely populated and often contain only single spectral samples with pronounced, individual spectral signatures (Figure 1).

Table 1: Number of super-channel clusters for given correlation thresholds.

Correlation	Number of Clusters
0.95	47
0.98	113
0.99	222
0.995	417
0.999	1633

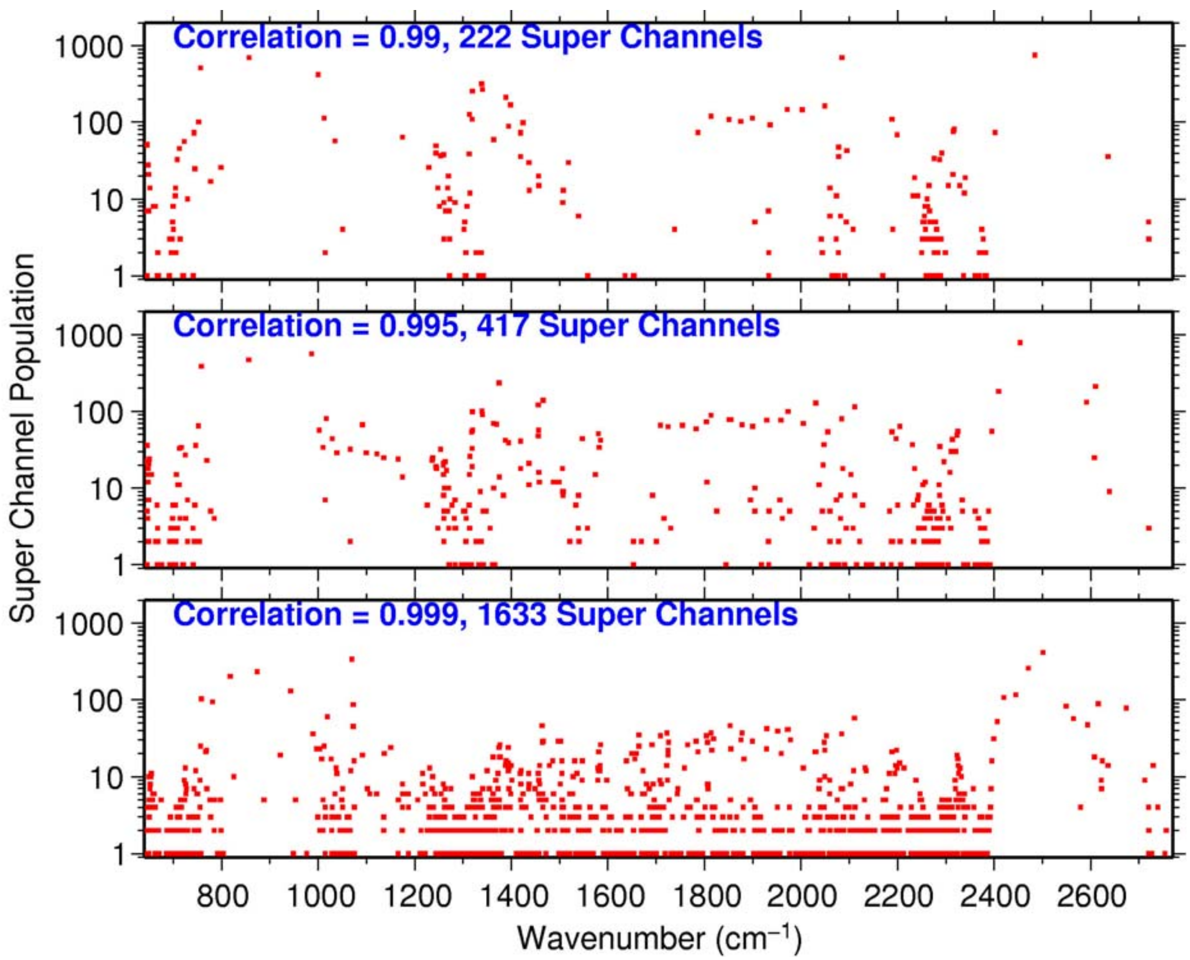

Fig. 1: Population of super-channel clusters at correlation thresholds of 0.99 (top), 0.995 (middle), and 0.999 (bottom).

The spectral range covered by single clusters is shown in figure 2. It is seen that at the lower correlation a spectral range of almost 1000 cm^{-1} can be covered by single clusters. At increasing correlation the spectral range in a single cluster is limited to about 300 cm^{-1}. As mentioned above, the population of the super-channel clusters also reduces the measurement noise entering the variational retrieval. The noise figures to be considered for the lead channels consist of measurement noise and regression errors. At lower correlation the latter prevail, while the measurement noise becomes the major part at high correlation implying a low regression error. Figure 3 illustrates for a correlation of 0.99 (222 super-channel clusters) the effective noise on the lead channel together with the population of the respective super channel. The noise (shown for a noise-normalised radiance) is reduced for many clusters by a factor greater than 5. For clusters populated with single samples only the normalised noise stays at 1, corresponding to the plain measurement noise.

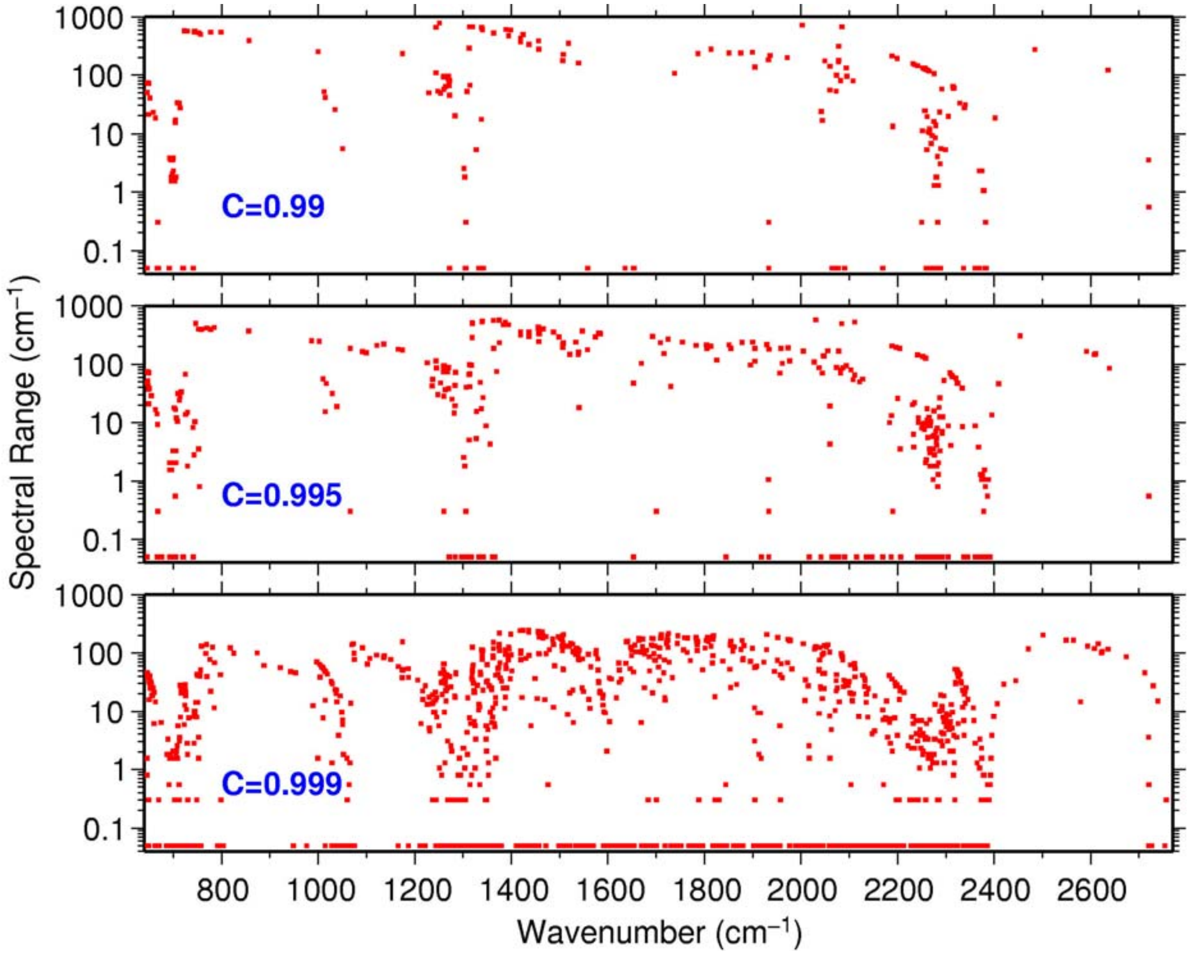

Fig. 2: Spectral range covered by single super-channel clusters at correlation thresholds of 0.99 (top), 0.995 (middle), and 0.999 (bottom).

Two examples for super-channel populations at a correlation threshold 0.995 are shown in figure 4 together with the portion of a IASI radiance spectrum. They demonstrate the big differences that are possible between super channels and the variety of spectral spacing between different samples within single clusters.

Conclusion and Discussion

IASI spectral samples can be efficiently combined to super-channel clusters in order to reduce the number of measurements entering a variational retrieval or an assimilation of IASI spectra into numerical weather forecast models, whilst retaining redundant information to reduce the measurement noise of the radiance samples entering the retrieval or assimilation scheme.

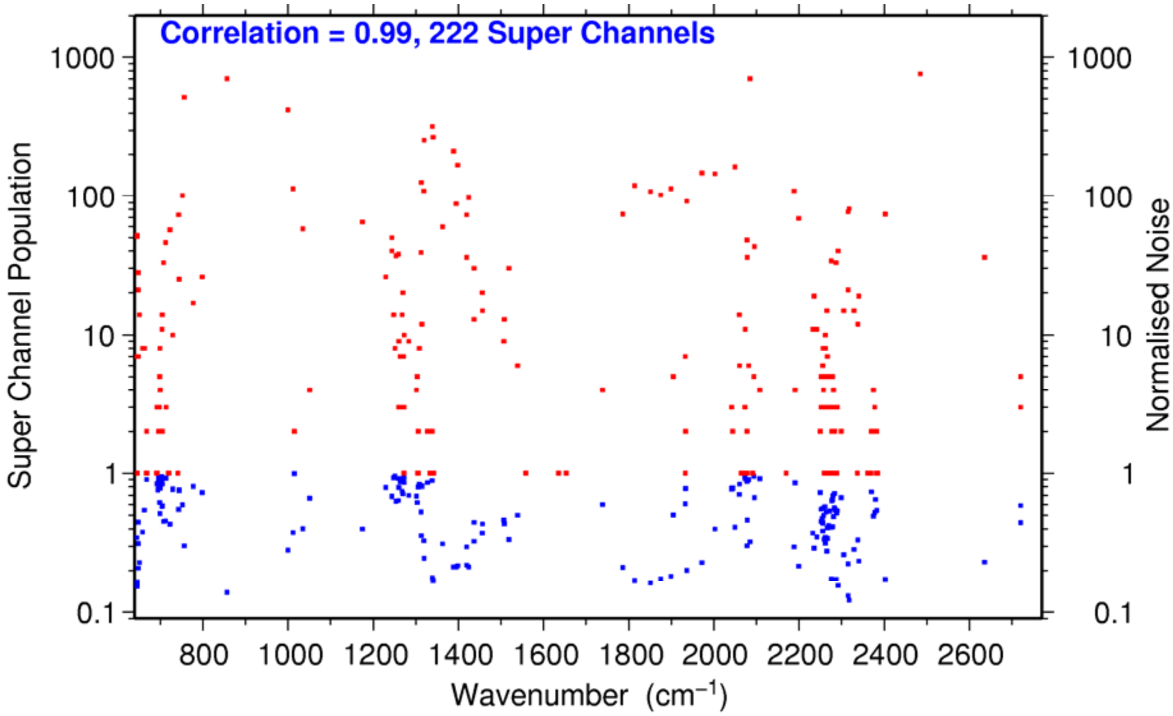

Fig. 3: Normalised noise (blue) of the lead channel radiance after weighted averaging and the corresponding population (red) of super channels for a correlation of 0.99.

Lead channels are selected to represent the clusters in the radiative transfer simulations used as forward model operators. The choice of lead samples made in this study was based on a plain search along the wavenumber scale for the next sample that was not already member of a cluster. Likewise, the lead sample selection could be optimised with respect to forward model errors, selecting among those samples of a cluster the one as lead sample which has the lowest forward model error.

The number of clusters can range from as low as 47 to more than 1633, assuming sample correlations ranging from 0.95 to higher than 0.999. The right choice for the number of clusters is to be optimised, trading off the noise reduction against spectral information. Operational retrieval or assimilation systems clearly require an upper limit of clusters that is suitable for near real-time processing. A number of 222 to 417 clusters, corresponding to correlations of 0.99 and 0.995, respectively seems to be realistic for such systems at present computing capabilities.

The super-channel clusters derived for high correlation thresholds contain many clusters which are populated with single samples only. This is desirable when spectral information is needed that is unique in single samples and which is not to be smeared by the creation of a highly-populated super-channel cluster. Likewise, it is always possible to remove single samples from a cluster and include it as single sample in

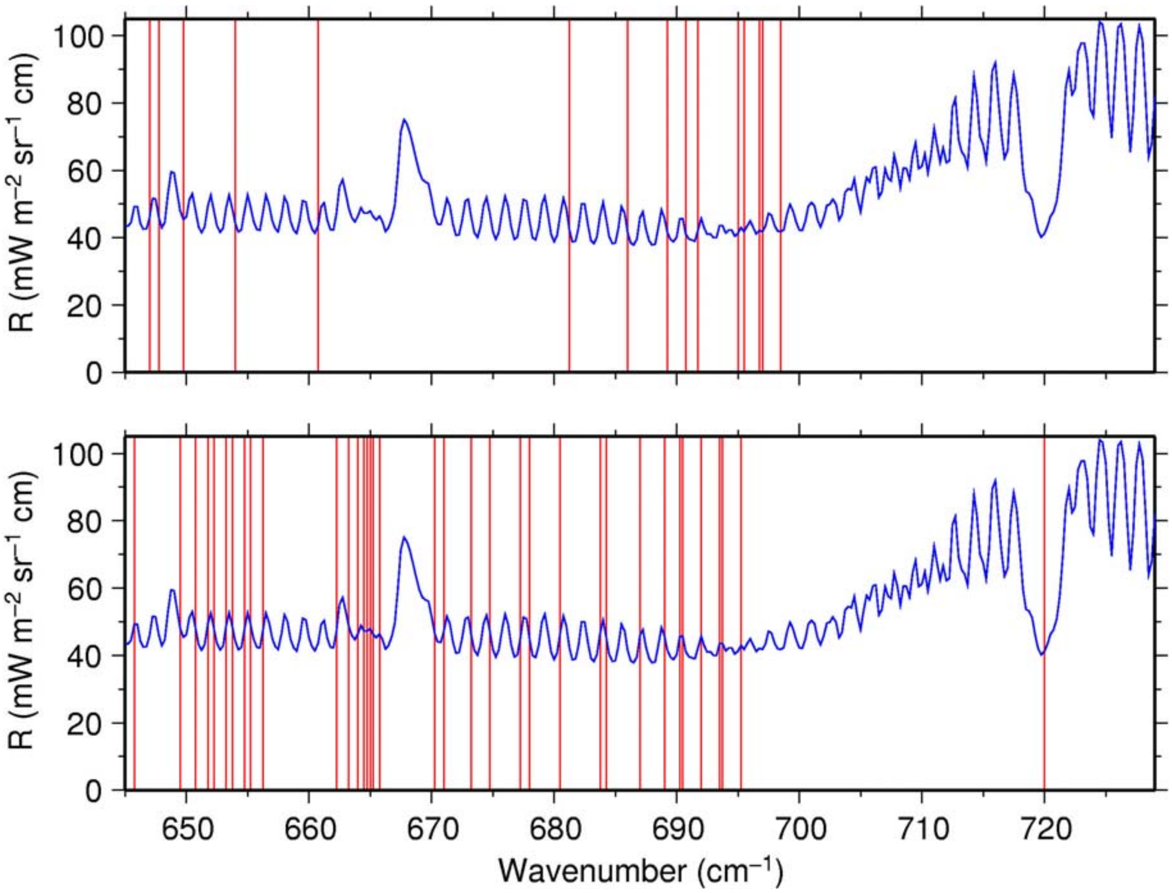

Fig. 4: Examples of super-channel clusters at a correlation threshold of 0.995. Top: cluster 6 with lead sample at 647 cm^{-1} including 15 samples; bottom: cluster 3 with lead sample at 645.75 cm^{-1} including 36 samples. The blue curve is an atmospheric reference spectrum, the red lines indicate the positions of the spectral samples.

the retrieval or assimilation, if the particular spectral information of this sample is needed (e.g. retrieval of trace gases).

The spectral range covered by single clusters is not limited to narrow spectral regions, but can cover a major part of the IASI spectrum. The drastic change in order of magnitude of the radiances across the spectrum is handled efficiently by using noise-normalised radiances.

For a modest number of 222 super channel clusters the noise is reduced substantially, a factor of 5 is achieved for many clusters. A variational retrieval that makes use of this set of super channel clusters has been implemented and tested in the EPS core ground segment at EUMETSAT, clearly demonstrating the usefulness of this approach.

References

Cayla, F.-R., 1993: IASI infrared interferometer for operations and research, in: Chedin, A., Chahine, M.T., Scott, N.A. (Eds.), High Spectral Resolution Infrared Remote Sensing for Earth's Weather and Climate Studies. *NATO ASI Series*, **I 9**, Springer Verlag, Berlin-Heidelberg, 9-19.

Chevallier, F., 2001: Sampled Databases of 60-level Atmospheric Profiles from the ECMWF Analyses. Research Report **4**, EUMETSAT/ECMWF.

Goldberg, M.D., Qu, Y., McMillin, L.M., Wolf, W., Zhou, L., Divakarla, M., 2003: AIRS Near-Real-Time Products and Algorithms in Support of Numerical Weather Prediction. *IEEE Trans. Geosci. Remote Sens.*, **41**, 379-389.

Matricardi, M., 2004a: The inclusion of Aerosols and Clouds in RTIASI, Report EUMETSAT Contract EUM/CO/02/989/PS, ECMWF, Reading, 57 pp.

Matricardi, M., 2004b: RTIASI-5 User's Guide, Report EUMETSAT Conract EUM/CO/02/989/PS, ECMWF, Reading, 47 pp.

Rodgers, C.D., 2000: Inverse Methods for Atmospheric Sounding - Theory and Practice. *World Scientific*, Singapore, pp 238.

Schlüssel, P. and Goldberg, M. 2002: Retrieval of atmospheric temperature and water vapour from IASI measurements in partly cloudy situations, *Adv. Space. Res.*, **29(11)**, 1703-1706.

Schlüssel, P., Hultberg, T.H., Phillips, P.L., August, T., Calbet, C., 2005: The operational IASI level 2 processor. *Adv. Space Res.*, **in press**.

Snyder, W.C., Wan, Z., Zhang, Y., and Feng, Y-Z., 1998: Classification-based emissivity for land surface temperature measurement from space, *Int. J. Remote Sensing*, **19**, 2753-2774.

Turquety, S., Hadji-Lazaro, J., Clerbaux, C. Hauglustaine, D.A., Clough, S.A., Cassé, V., Schlüssel, P., Mégie, G., 2004: Operational trace gas retrieval algorithm for the Infrared Atmospheric Sounding Interferometer, *J. Geophys. Res.*, **109**, D21301.