

# IASI Channel Selection in Cloudy Conditions

P.Martinet<sup>1</sup>, L.Lavanant<sup>2</sup>, N.Fourrié<sup>1</sup>, F.Rabier<sup>1</sup>, F.Faijan<sup>2</sup>

<sup>1</sup>*Météo France & CNRS/CNRM-GAME*

<sup>2</sup>*Météo France, Centre de Météorologie Spatiale*

## Introduction

The Infrared Atmospheric Sounding Interferometer (IASI) on-board the MetOp satellite belongs to a new generation of advanced infrared sounders and follows the launch of the Atmospheric InfraRed Sounder (AIRS) on-board the Aqua satellite in 2002. AIRS with 2378 channels and IASI with 8461 channels provide information about atmospheric temperature and humidity with a far better spectral resolution compared to previous instruments such as the High Resolution InfraRed Sounder (HIRS).

To reduce the significant computational costs, among the 8461 IASI channels, about 300 channels are monitored by most of the numerical weather prediction (NWP) centres.

The channel reduction was processed by a channel selection based on clear profiles of temperature, humidity, ozone, carbon dioxide and surface temperature (Collard 2007 for IASI and Susskind et al 2003 for AIRS).

These selections were satisfactory for the assimilation of clear observations which represent most of the assimilated radiances. However, the high correlation between cloud cover and meteorologically sensitive areas underlines the need to use infrared observations in presence of clouds (McNally 2002, Fourrié and Rabier 2004). In order to progress in the use of cloud-affected radiances and with the view to add the cloud variables (liquid water content, ice water content) in the state vector of the assimilation system, we investigated the potential benefit of a new channel selection in cloudy conditions. This work is in line with the HyMeX campaign (Hydrological cycle in the Mediterranean EXperiment) aiming at a better understanding, quantification and modelling of the hydrological cycle over the Mediterranean. The main source of information over sea is satellite data and about 80% of the radiances are cloudy. It is important to constrain the analysis in cloudy areas to improve forecast of heavy precipitation events. In this framework, the French convective scale model AROME (Application of Research to Operations at MESoscale) is used to provide cloud profiles over the Mediterranean.

Several information content studies have been evaluated to select the most useful channels in order to minimize the total loss of information for each state variable. Different techniques have been compared by Rabier et al (2002), who conclude that the channel selection method of Rodgers (1996,2000) is the most nearly optimal. This method was then applied to the selection of AIRS channels by Fourrié and Thépaut (2003) and IASI channels by Collard (2007).

This study attempts to assess the relevance of the selection of the 366 IASI channels used operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF) in the context of cloud-affected profiles. This selection (Collard and McNally 2009) called CM2009 hereafter is evaluated in terms of Degrees of Freedom for Signal (DFS) to evaluate its capability to retrieve information about microphysical variables. Then, a subset of new channels selected on cloud profiles is added to the CM2009 selection. For this new selection, two methods considering different state vectors are evaluated by the DFS of each atmospheric component. After the choice of the best channel selection, its robustness and sensitivity to different parameters are quantified.

## Experimental framework

The general framework of this channel selection is the linear optimal estimation in the context of NWP. In the following section, we summarize the main elements presented at length by Rabier et al (2002). The IASI measurements are represented by the vector  $\mathbf{y}$  and the atmospheric profiles of temperature, humidity, liquid water content and ice water content by

the vector  $\mathbf{x}$ . The observations are linked to the atmospheric state by an observation operator including vertical and horizontal interpolations in addition to a radiative transfer model:

$$\mathbf{y} = \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}_O + \boldsymbol{\varepsilon}_F$$

where the measurement and the forward model errors  $\boldsymbol{\varepsilon}_O$  and  $\boldsymbol{\varepsilon}_F$  are assumed to be gaussian noises with error covariance matrices  $\mathbf{O}$  and  $\mathbf{F}$ . We will denote  $\mathbf{R} = \mathbf{O} + \mathbf{F}$  the resulting observation error covariance matrix. The error covariance matrix of the background state  $\mathbf{x}_b$  is denoted  $\mathbf{B}$ . The radiative transfer equation is assumed to be weakly nonlinear in the vicinity of the background state, making the tangent linear assumption valid:

$$\mathcal{H}(\mathbf{x}) = \mathcal{H}(\mathbf{x}_b) + \mathbf{H}(\mathbf{x} - \mathbf{x}_b)$$

where  $\mathbf{H}$  is the tangent linear model of the radiative transfer model  $\mathcal{H}$ .

The best linear approximation of the true atmospheric state is given by:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{y}_b)$$

with  $\mathbf{K} = \mathbf{A}\mathbf{H}^T\mathbf{R}^{-1}$  the Kalman gain matrix and  $\mathbf{A} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$  the analysis-error covariance matrix.

### Channel selection methodology

The channel selection that was chosen for this study is based on the methodology described by Rodgers (2000). The method consists in performing successive analyses considering only one channel at a time. The impact of the addition of single channels is evaluated by the DFS which is used as the figure of merit of the channel selection. The background-error covariance matrix is updated at the next step by the analysis-error covariance matrix previously calculated. In order to take into account the gain brought by the previously selected channels, this update of the  $\mathbf{B}$  matrix is important.

The starting point of the study is the CM2009 selection composed of 366 IASI channels. Firstly, the analysis-error covariance matrix  $\mathbf{A}$  considering the 366 IASI channels of the CM2009 selection is evaluated. Then, new channels are selected updating the  $\mathbf{B}$  matrix by the previously calculated  $\mathbf{A}$  matrix. The atmospheric components for which information is expected are temperature, humidity, liquid water content (ql) and ice water content (qi). Two methods of selection were compared: using the four atmospheric variables in the state vector of the selection (SELC4) or focusing on cloud information considering only ql and qi (SELC2) in the state vector.

### Models and Data

Profiles of temperature, humidity, liquid water content and ice water content were extracted from short-range forecasts (3-hour) of the convective scale model AROME WMED. AROME WMED is a research version of the operational model AROME dedicated to the HyMeX campaign. AROME is a limited area model with a 2.5 km grid based on non-hydrostatic equations (Seity et al 2011). Our study is performed over a period of 30 days (7 October 2010 to 7 November 2010) on a domain centred on the Mediterranean Sea (-0.5°W to 17°E, 36°N to 44.5°N).

In order to reduce the number of AROME profiles, a preselection was achieved considering each interpolated profile within a IASI field of view (FOV). The FOVs were chosen to focus on homogeneously covered scenes in both the observation and the forecast. A radiance analysis (Cayla 2001) of co-located AVHRR (Advanced Very High Resolution Radiometer) pixels inside each IASI FOV was used for this screening stage as described in Martinet et al 2012. In order to cover most of the cloud variability and to limit the computation cost, 20 profiles were chosen: 5 of ice opaque clouds, 5 of semi-transparent ice clouds, 5 of low liquid clouds and 5 of mixed phase clouds.

The version 10 of the fast radiative transfer model RTTOV (Hocking et al, 2010) has been used for this study to compute the jacobian matrix. Profiles of liquid water content, ice water content and cloud fraction are provided to the advanced interface RTTOVCLD for a better modelling of clouds with possibility of multi-layer clouds and two cloud types per layer.

The measurement-error covariance matrix  $\mathbf{O}$  was constructed with the values of the instrumental noise provided by CNES. These values are valid at a temperature of 280 K and are converted at the scene temperature for each wave number and each profile. A constant error is added to take into account the radiative transfer model error. We chose a value of 0.2 K for liquid cloud profiles which are assumed to be better simulated and 0.5 K for ice clouds and multiphase clouds.

As we are interested in a channel selection in cloudy conditions, a specific  $\mathbf{B}$  matrix has been computed. The background-error statistics were derived from an AROME ensemble assimilation, that considers explicit observation perturbations and implicit background perturbations through the cycling, coupled with the operational ensemble assimilation at global scale (Desroziers et al 2008). They were calculated from a set of 18 convective cases observed during the months of July, August and September 2009. Geographical cloud masks (Montmerle et Berre 2010), based on values of simulated vertically integrated cloud contents, were then applied to differences of background perturbations in order to gather data from which the statistics are performed. As in Michel et al (2011), the multivariate formalism proposed by Berre (2000) has been extended to  $q_l$  and  $q_i$ , allowing couplings with forecast errors of temperature and specific humidity. The standard deviations for this matrix are shown in figure 1.

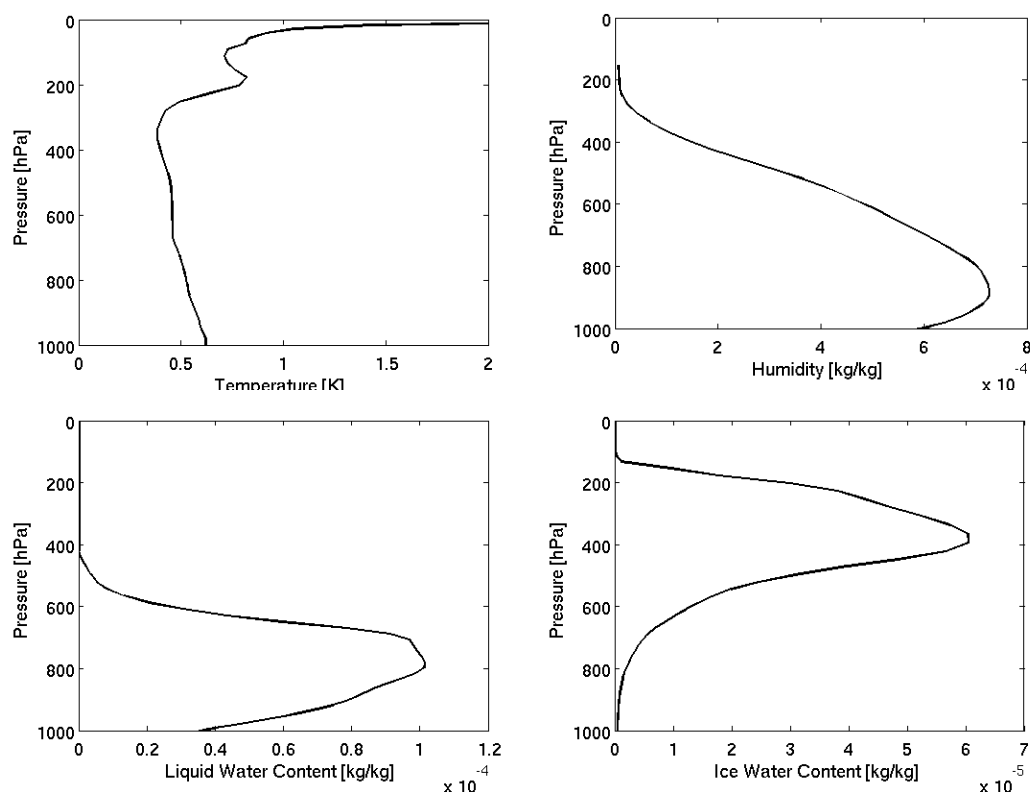


Fig 1: Standard deviations of temperature, humidity, liquid water content, ice water content taken from the AROME  $\mathbf{B}$  matrix computed in cloudy conditions.

### Channel Selection Results

For each of the 20 profiles, 200 channels were added considering different variables in the state vector. The two channel selections will be evaluated comparing a ‘constant’ selection instead of an optimal channel selection for each profile. The ‘constant’ selection is an average of 20 optimal selections. At each channel picked by the channel selection, a rank

from 1 to 200 is given corresponding to its rank in the iterative selection process. Finally, a ‘global’ set of 200 channels is chosen considering the channels with the lowest average rank. Another method would be to choose the 200 channels which are picked the most often. The selection based on the average rank was preferred because it spreads the channels all over the spectrum whereas the second option favours some specific spectral regions. Figure 2 shows the two global selections considering the four atmospheric variables in the state vector (SELC4) or the cloud variables only (SELC2).

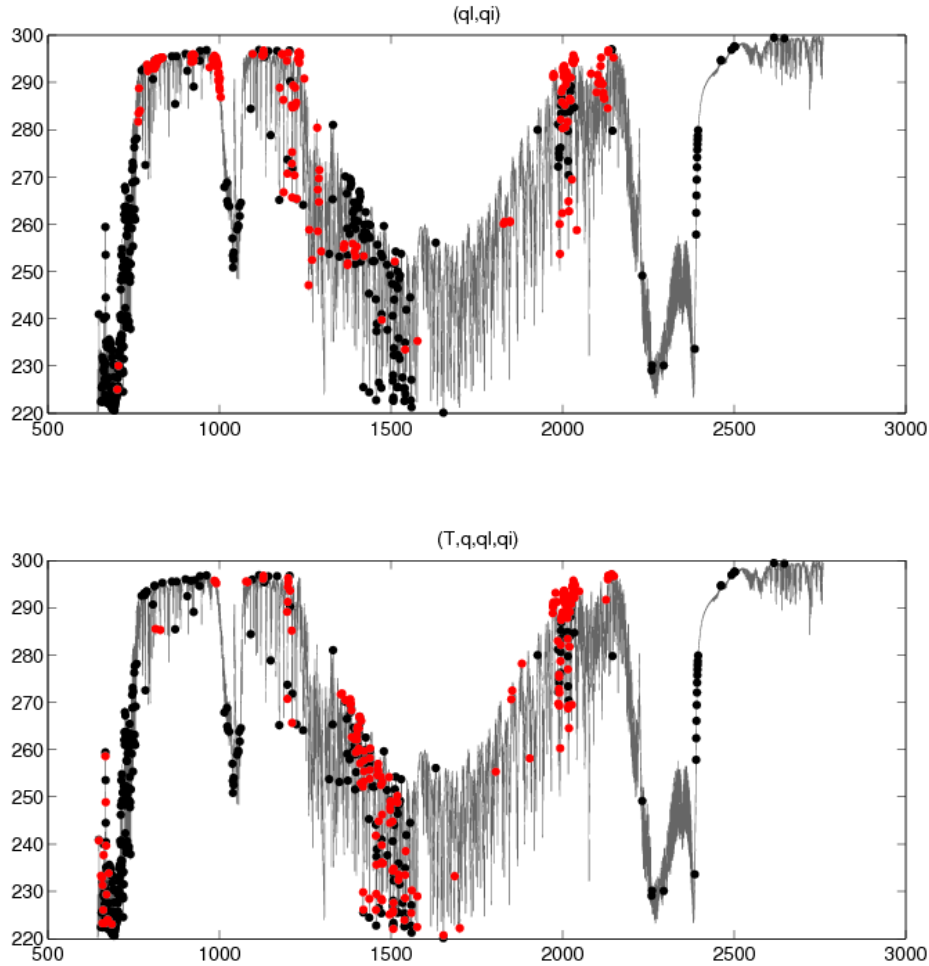


Fig 2: Location of the selected channels averaged over all the profiles. The CM2009 selection of ECMWF is represented with black points and the new channels in red points. The selection considering  $q_l$  and  $q_i$  in the state vector (top) is compared to the selection considering the four variables in the state vector (bottom).

The number of channels selected in the long-wave part of the water vapour band is higher for the SELC4 selection. When focusing on  $q_l$  and  $q_i$ , the surface window channels sensitive to the cloud top are preferred. This means that even if the four variables are introduced in the state vector of the selection, the DFS value is mainly dominated by temperature and humidity. However, the channels selected for temperature and humidity may bring information about the microphysical variables.

Table 1 compares the values of DFS for each variable and each channel selection. Each selection method is evaluated by the percentage of DFS explained by the subset of channels compared to the maximal values of DFS considering all the 8461 channels. Firstly, it is worth

noting that even with 8461 channels, the extraction of information is dominated by temperature and humidity. The DFS for temperature represents 50% of the total DFS, humidity 33%, ice water content 15% and liquid water content only 2%. Despite the predominance of temperature and humidity, the results show a potential gain in information about ice water content. Secondly, the experiment shows the robustness of the channel selection algorithm employed by Collard (2007) to provide the main part of the CM2009 selection. Despite the fact that this channel selection was developed on clear profiles, 70% of the total DFS calculated with all the channels is reached. The DFS is particularly high for temperature, humidity and ice water content (70%) and reasonable for liquid water content (60%). Finally, we can note that the addition of 200 channels improves the results but the impact on the total DFS is better with the SELC4 approach (gain of 5% instead of 3%). These improvements show that temperature and water vapour channels are informative not only for temperature and humidity but also for liquid and ice water contents. This is certainly due to nonlinearity in the radiative-transfer coupling temperature and humidity jacobians with  $q_l$  and  $q_i$  variables but also to the coupling between humidity and microphysical variables in the **B** matrix. Both approaches were compared with the standard deviations of analysis-errors (diagonal elements of the analysis-error covariance matrix **A**). Almost no difference was observed to favour one method over the other.

Table1: Values of Degrees of Freedom for Signal (DFS) for both selection methods and for the CM2009 selection. The percentage of available DFS is indicated.

Number of Channels	Experiment	Temperature	Humidity	Liquid Water Content	Ice Water Content	Total
366	CM2009	4.11 (68%)	2.79(69%)	0.17(59%)	1.45(76%)	8.52(70%)
566	SELC2 (q <sub>l</sub> ,q <sub>i</sub> )	4.24(71%)	2.93(73%)	0.20(66%)	1.59(84%)	8.95(73%)
566	SELC4 (T,q <sub>l</sub> ,q <sub>i</sub> )	4.41(73%)	3.08(76%)	0.19(63%)	1.54(81%)	9.22(75%)
8461	All channels	6.00	4.04	0.30	1.90	12.24

### Dependence of the selection to the cloud phase

The selections previously described were obtained averaging the selections over all the profiles. A good channel selection must be usable for different cloud types in an operational context. The different channel selections after an averaging over one cloud type are shown in figure 3 considering the SELC4 (left panel) and SELC2 (right panel) selections. The selections adapted to each cloud type are relatively close to the one averaged over all the different cloud scenes except for liquid clouds in the context of the SELC2 method. In fact, when temperature and humidity are not included in the state vector of the selection, water vapour channels are not chosen any more when concentrating on liquid clouds because of jacobians peaking in the mid troposphere. The ‘global’ selection derived from the SELC4 method shares about 66% of its channels with each cloud type dedicated channel selection whereas the SELC2 method shares 53% of its channels.

This study has proved the SELC4 method to be more robust considering different cloud phases. In the view of operational assimilation where both clear scenes and cloudy scenes are mixed and considering the computation cost of the two methods, SELC4 seems to be a good compromise.

Another question was investigated: What happens to the SELC4 selection if it is based on a cloud type and applied to a different one ? A cloud type dedicated channel selection provides better DFS for each air mass than the global selection averaged over all the cloud types but the gain is weak (around 1 or 2%). The DFS obtained when applying a cloud type dedicated channel selection to a different one is similar to the one obtained with the ‘global’ selection. Altogether, the SELC4 selection seems to be insensitive to the cloud-type category.

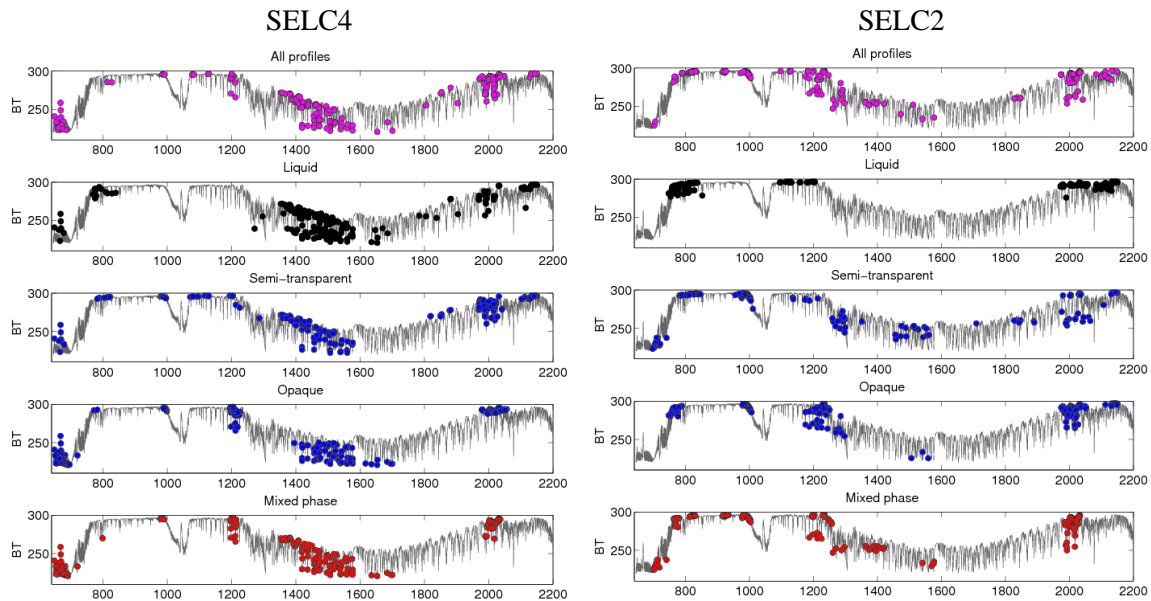


Fig 3: Location of the selected channels depending on the cloud phase. From top to bottom: averaged over all the profiles, only liquid cloud profiles, semi-transparent clouds, opaque clouds and mixed phase clouds.

Figure 4 shows the evolution of the DFS values with respect to the number of selected channels for each variable. For temperature, humidity (left panel) and liquid water content (right panel), the selection of 566 channels seems sufficient to reach the DFS saturation. For ice water content, the asymptotic value of the DFS would be reached from using at least up to 2000 channels but the selection of 566 channels explains about 80% of the maximum value of the ice water content DFS.

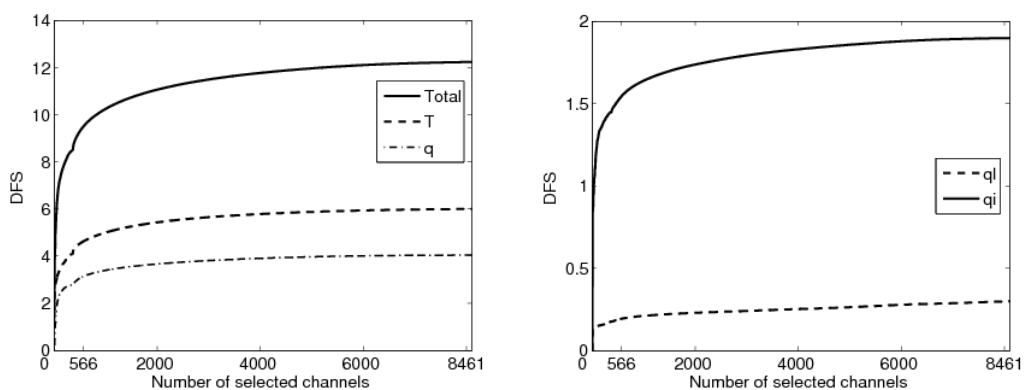


Fig 4: Evolution of the degrees of freedom for signal (DFS) for the global selection SELC4 with respect to the number of selected channels, averaged over the 20 profiles.

It is also interesting to illustrate the variation of the reduction analysis when selecting a different number of channels. For that purpose, we examine the standard deviations of

background errors and analysis errors for temperature, humidity, liquid water content and ice water content averaged over each cloud type (figure 5).

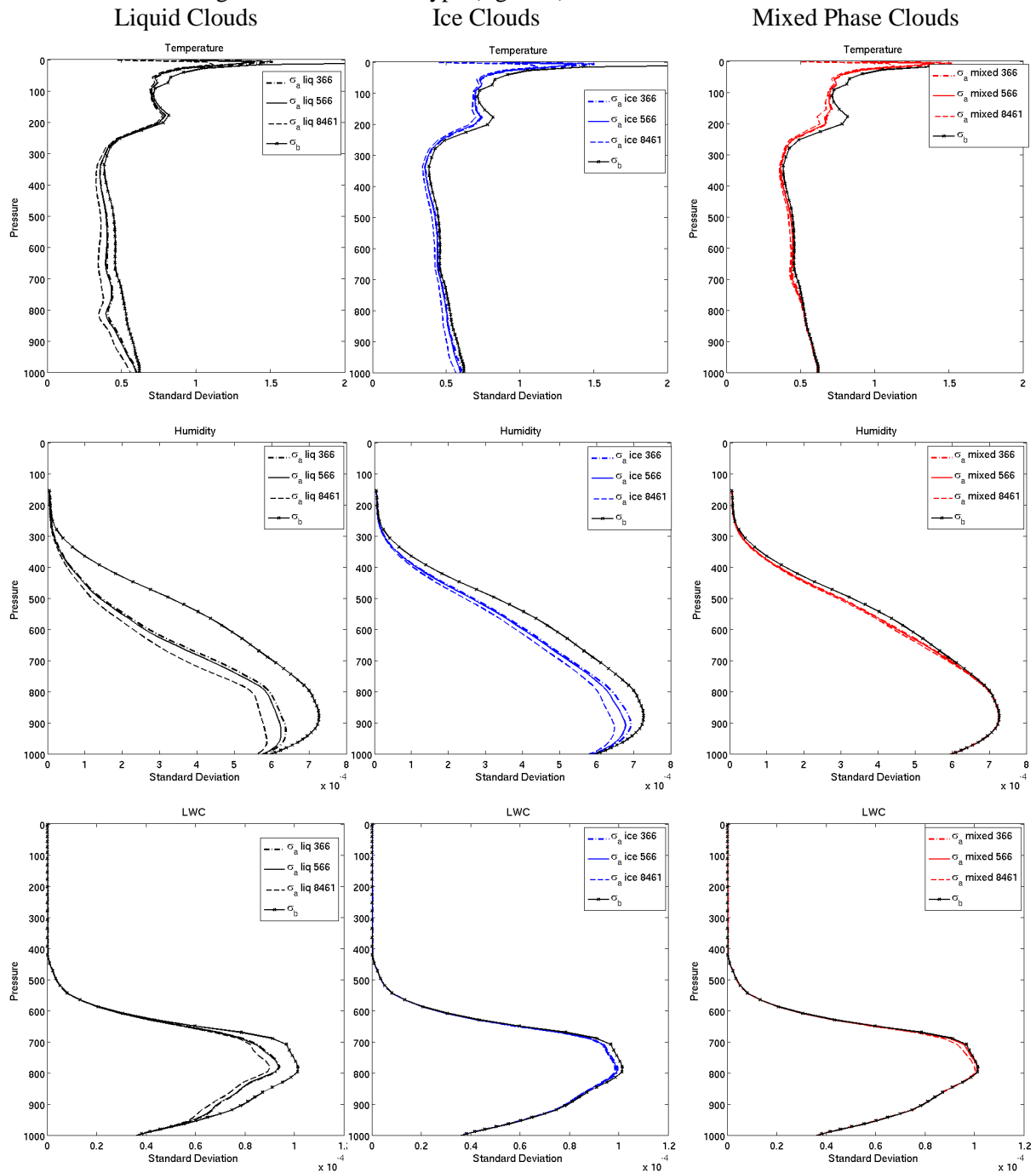


Fig 5: Analysis-Error Standard Deviations for each cloud type: liquid clouds in black lines, ice clouds in blue lines, mixed phase clouds in red lines. For each cloud type, the analysis-error standard deviations obtained with the new selection of 566 channels (solid line) is compared to the ones obtained with the CM2009 (dash dot line) selection of 366 channels and the best performance using the 8461 channels (dotted line).

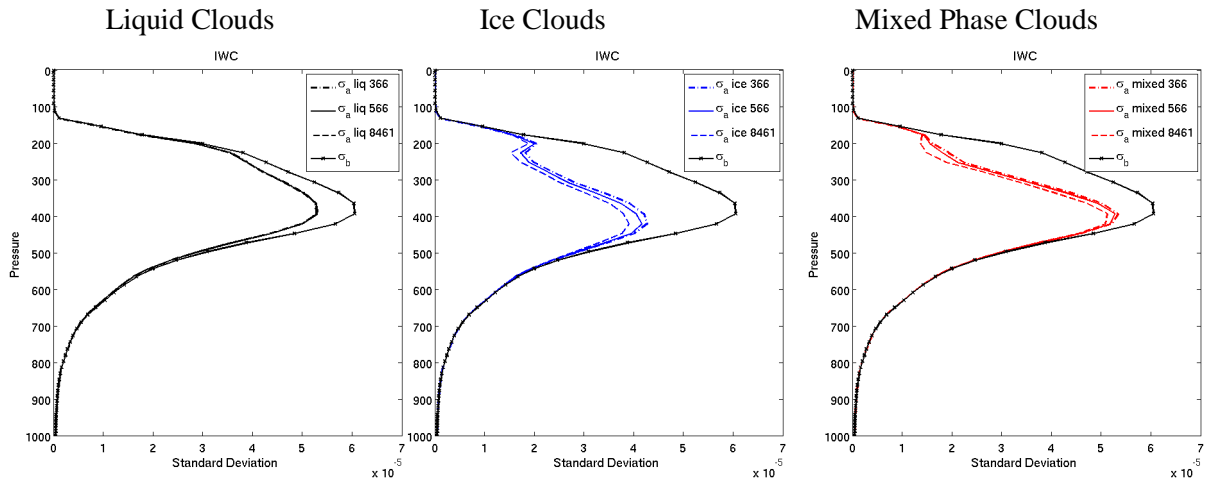


Fig 5 (continued): results for ice water content.

For temperature and humidity, the performances of the CM2009 and the SELC4 selections are equivalent. For ice water content and humidity, the new selection of 566 channels outperforms the CM2009 selection of 366 channels in terms of analysis-reduction error. It is interesting to note that the more transparent the cloud is, the highest the improvement of the analysis over the background is. As mixed phase clouds are mainly represented by opaque clouds with an ice cloud layer over a liquid cloud layer, they bring little information about temperature and humidity. The improvement of the analysis over the background is better with ice clouds than mixed-phase clouds because the ice cloud set was divided between opaque and semi-transparent clouds.

Liquid clouds which are located in the low troposphere make possible the improvement of the background errors through the entire atmospheric column. The improvement of the temperature background error is decreased by almost 0.2 K around 800 hPa where liquid clouds are mainly located. In the same way, the reduction of ice water content analysis error is performed best around 400 hPa where most of the selected ice clouds peak.

The use of all the 8461 channels reduces the temperature analysis error of about 0.1 K in the stratosphere with a maximum of 0.43 K for the first two levels easily explained by the large values of background errors in the stratosphere. The reduction is less significant for the troposphere (about 0.02 K). An improvement in the other analysed variables is also observed but the loss of information content seems reasonable. The results of the new selection are very comparable with the all channel analysis in terms of linear 1D-Var retrieval error.

We conclude that despite the loss of information content expected when using a sample of the total number of IASI channels, the new channel selection provides reasonable performance.

## Robustness of the selection to the optical parameter parameterizations in RTTOV

In the radiative transfer model RTTOV, the ice cloud optical parameters are parameterized as a function of the effective diameter of the size distribution. Consequently, for ice clouds the user must choose what assumption to use to parameterize the effective diameter and must also specify which shape is to be used for the ice crystals since optical parameters are available for hexagonal ice crystals and ice aggregates (Faijan et al 2012). Each parameterization was tested for the SELC4 channel selection. Figure 6 shows the selection with each parameterization (left panel) assuming hexagonal ice crystals. The right panel shows the same comparison but assuming ice aggregates instead of ice crystals. Almost 82 % of the selected



channels are common to the different parameterizations. However, this number is turned down to 63 % when comparing a given parameterization applied to the two crystal shapes (hexagonal or aggregates). In figure 6, we can notice that much more window channels are selected with the aggregate shape. The variation of the channel selection is thus more sensitive to the ice crystal shape than the ice parameterization. This difference with the two ice shapes is attenuated considering the SELC4 method because the selection is dominated by temperature and humidity. When considering only liquid water content and ice water content in the state vector, the difference between the channel selections with different ice shapes is emphasized.

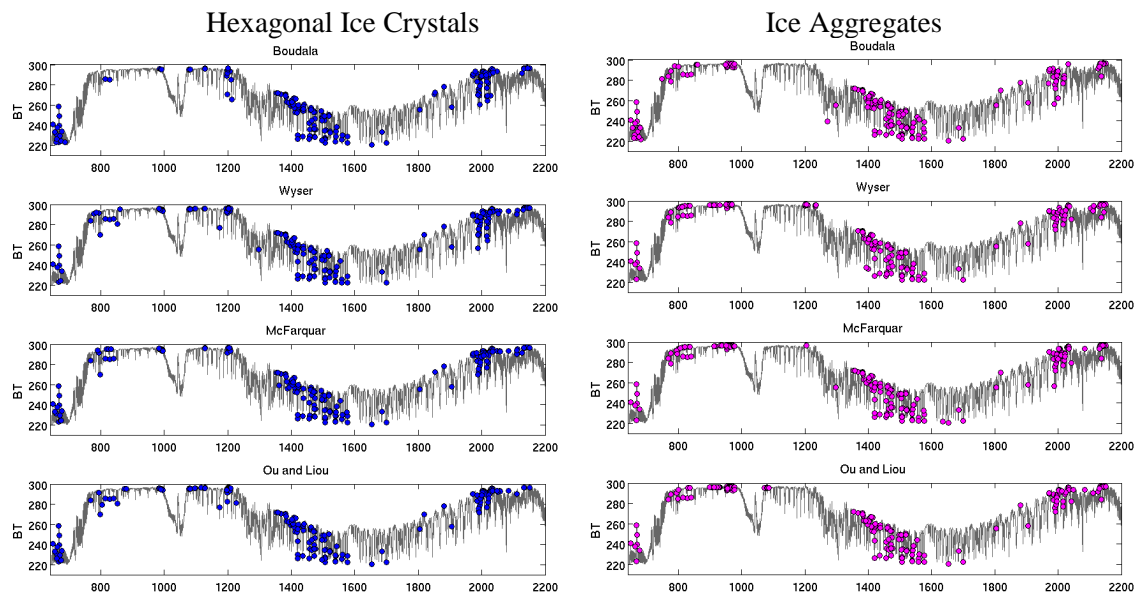


Fig 6: Location of the selected channels depending on the RTTOV parameterization for hexagonal ice crystals (left panel) and aggregate ice crystals (right panel).

## Conclusions

In this study, we evaluated the optimality of the IASI selection (CM2009) used operationally at ECMWF in the context of cloud-affected profiles. This selection has been performed by Collard 2007 on clear atmospheric profiles representative of standard atmospheres. The ECMWF selection has been evaluated on cloud profiles of temperature, humidity, liquid water content and ice water content extracted from short-range forecasts of the operational convective scale model AROME. It was first shown that the CM2009 selection of 366 channels is able to reach about 70% of the total information that would be gained using all of the 8461 channels. It was then decided to add about 200 channels to the CM2009 selection but chosen specifically on cloudy conditions. An iterative method following Rodgers (2000) was used to evaluate two selection methods considering different state vector. The first method SELC4 considers four variables which are temperature, humidity, liquid water content and ice water content in the state vector of the selection whereas SELC2 considers only liquid water content and ice water content. Even if the two ‘global’ selections share only 60 channels of the 200 selected channels, they lead to similar results in terms of DFS and reduction of analysis-errors. However, the SELC4 selection is slightly better considering the total DFS of the four atmospheric variables and it is more robust with respect to the cloud

type. The channel selections dedicated to each cloud type are quite similar with the ‘global’ SELC4 selection and the impact of applying a cloud type dedicated channel selection to a different one is negligible.

The impact of the ice parameterization in the radiative transfer model was also studied. It was shown that the channel selection is more sensitive to the ice crystal shape (hexagonal or aggregates) than the parameterization scheme itself.

The loss of information considering only 566 channels was quantified comparing the analysis-error reduction with an ‘all channel’ retrieval. Even if an improvement of the analysis over the background is observed with the 8461 channels the new selection performs reasonably well.

This study is limited by the optimal linear estimation theory. In the future, the limitations of the study because of the non-linearity of the observation operator will be investigated.

The robustness of the selection with different air-mass types will also be studied. Cloud profiles representative of tropical, polar or mid-latitude air-mass types from a global NWP model will be used to perform the same channel selection and evaluate its consistency with our selection on Mediterranean profiles.

## References

- Berre L, 2000. Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Monthly Weather Review* , **128**(3), 644-667.
- Cayla F-R, 2001. AVHRR Radiance Analysis inside IASI FOV's. IA-TN-0000-2092-CNE, *CNES internal report*.
- Collard A, 2007. Selection of IASI channels for use in numerical weather prediction. *Quart.J.Roy.Meteor.Soc.*, **133**, 1977-1991.
- Collard A, McNally A.P, 2009. The assimilation of Infrared Atmospheric Sounding Interferometer radiances at ECMWF. *Quart.J.Roy.Meteor.Soc.*, **135**, 1044-1058.
- Desroziers G, Berre L, Pannekoucke O, Stefanescu S-E, Brousseau P, Auger L, Chapnik B, Raynaud L, 2008. Flow-dependent error covariances from variational assimilation ensembles on global and regional domains. *HIRLAM Technical Report No. 68*.
- Fajjan F, Lavanant L, Rabier F, 2012. Towards the use of microphysical properties to simulate IASI spectra in an operational context. *JGR*, submitted.
- Fourrié N, Thépaut J.N, 2003. Evaluation of the AIRS near-real-time channel selection for application to numerical weather prediction. *Quart.J.Roy.Meteor.Soc* , **129**, 2425-2439.
- Fourrié N, Rabier F, 2004. Cloud characteristics and channel selection for IASI radiances in meteorologically sensitive areas. *Quart.J.Roy.Meteor.Soc*, **128**, 2551-2556.
- Hocking J, Rayer P, Saunders R, Matricardi M, Geer A, Brunel P, 2010. RTTOV v10 Users Guide. NWPSAF-MO-UD-023, EUMETSAT, Darmstadt, Germany.
- Martinet P, Fourrié N, Guidard V, Rabier F, Montmerle T and Brunel P, 2012. Towards the use of microphysical variables for the assimilation of cloud-affected infrared radiances. *Quart.J.Roy.Meteor.Soc*, submitted.
- McNally A.P, 2002. A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Quart.J.Roy.Meteor.Soc* , **128**, 2551-2556.
- Michel Y, Auligné T, Montmerle T, 2011. Heterogeneous Convective-Scale Background Error Covariances with the Inclusion of Hydrometeor Variables. *Monthly Weather Review*, **139**, 2994-3015.
- Montmerle T, Berre L, 2010. Diagnosis and formulation of heterogeneous background-error covariances at the mesoscale. *Quart.J.Roy.Meteor.Soc* , **136**, 1408-1420.

- Rabier, F., Fourrié, N., Chafaï D. and P. Prunet, 2002: Channel selection methods for infrared atmospheric sounding interferometer radiance., *Quart.J.Roy.Meteor.Soc* , **128**, 1011-1027 .
- 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: Theories and practice. *World Scientific Publ., Singapore*, 238pp.
- Seity Y, Brousseau P, Malardel S, Hello G, Benard P, Bouttier F, Lac C, Masson V, 2011. The AROME-France convective-scale operational model. *Monthly Weather Review*, **139**, 976-991.
- Susskind J, Barnet C.D, Blaisdell J.M, 2003. Retrieval of Atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds. *IEEE Remote Sensing*, **136**, 1064-1074.