Analysis of MTG-IRS observations and general channel selection for Numerical Weather Prediction models

Olivier Coopmann, Nadia Fourrié, V. Guidard

Analysis of MTG-IRS observations and general channel selection for Numerical Weather Prediction models

O. Coopmann*, N. Fourrié and V. Guidard

CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

*Correspondence to: O. Coopmann, CNRM/GMAP, 42 Avenue Gaspard Coriolis, 31057 Toulouse cedex, France.
E-mail: olivier.coopmann@umr-cnrm.fr

IRS (infra-red Sounder) is an infra-red Fourier transform spectrometer that will be on board the Meteosat Third Generation series of the future EUMETSAT geostationary satellites and will have a unique 4D look at the atmosphere. After its planned launch in 2024, it will be able to measure the radiance emitted by the Earth at the top of the atmosphere using 1960 channels in two spectral bands between 680 - 1210 cm^{-1} (long-wave infra-red) and 1600 - 2250 cm^{-1} (mid-wave infra-red) at a spectral sampling of 0.6 cm^{-1}. It will perform measurements over the full Earth disk with high spatial and temporal resolution of 4 km at nadir and 30 minutes over Europe. Thus, the huge amount of data from IRS will present challenges, particularly in data transmission, data storage and assimilation into NWP models.
To reduce the volume of data, various methods are available including spatial sampling, principal component analysis and channel selection. The latter technique will be discussed in this paper by proposing general channel selection to provide Numerical Weather Prediction models. The objective of this selection is to improve essential variables for NWP such as temperature, humidity, skin temperature and ozone. This work has required the development of a large observation database and takes into account the main developments in assimilation techniques, including the use of full observation-error covariance matrices or the assimilation of ozone in global models for example. This study performs a specific analysis of the sensitivity of IRS observations and proposes a first general selection of 300 channels for NWP models. This selection allows us to reduce the analysis error in the troposphere by 48 % in temperature, 65 % in humidity and 17 % in ozone.

Key Words: MTG-IRS, Channel Selection, Numerical Weather Prediction, Radiative Transfer Model, Data Assimilation

Received . . .

1. Introduction

The benefits of infra-red hyperspectral sounders in improving Numerical Weather Prediction and monitoring of atmospheric composition are well established. Nowadays, the IASI instrument, on board the Metop-B and C satellites, is the most performing instrument. They allow a large coverage of the globe, but they can only observe the same place twice a day due to the nature of the polar satellites on which they are flown. There are currently two hyperspectral sounders in geostationary orbit: the Geostationary Interferometric Infrared Sounder (GIIRS-1 and 2), already aboard the FY-4A geostationary satellite over East Asia (Yin et al. 2020).

The rapid evolution of instrumental and spectral performance has enabled EUMETSAT (European Organization for the Exploitation of Meteorological Satellites) to propose its future geostationary satellite program, the Meteosat Third Generation (MTG) series. These platforms, whose first launch is planned for 2022, will host on board two types of instruments, an imaging platform (MTG-I) and a sounding platform (MTG-S). The latter, will have on board the Infra-Red Sounder (IRS) a imaging Fourier Transform Spectrometer with a spectral sampling of about 0.6 cm^{-1} in two spectral bands named Long-Wave Infra-red (LWIR or band 1) extending from 680 - 1210 cm^{-1} and Mid-Wave Infra-red (MWIR or band 2) extending from 1600 - 2250 cm^{-1}. After its launch, it will perform measurements over the full Earth disk with particular focus on Europe (revisited every 30 min), with a spatial resolution of 4 km at nadir.

Thus, IRS will have higher spatial and temporal sampling than its precursor GIIRS. In addition, it should also be the first to satisfy operational requirements, GIIRS-1 and 2 being experimental. IRS is expected to have a significant impact on NWP and nowcasting. In addition to the future Copernicus Sentinel-4 instrument which will be on the same platform, the duo will provide a set of

useful observations for atmospheric composition and air quality monitoring applications. The main challenges are:

- The rapidity of observations, where requirements on level 1 and 2 products are to be respectively made available within 15 and 30 minutes from sensing for any given dwell.
- The spatial and temporal resolution of the observations will be very important and will induce huge amount of data, which will require information compression. The spectra will be distributed as principal components using a so-called hybrid method. This method will make it possible to conserve the essential part of the atmospheric variability.
- The information content of IRS should be assessed and a selection of information can be made to reduce the amount and redundancy of information. The spectral sampling has been optimised from 0.625 to 0.6 cm^{-1}. An assessment of the cloud fraction and scene heterogeneity within each pixel will be available.
- Preparing users for the coming of IRS is essential in order to be ready when the first data is being acquired. Test data are provided by EUMETSAT to familiarise them with these future observation format.

The main objectives of this study are the analysis of the different sensitivities of IRS observations and a selection of information to provide a first list of IRS channels able to achieve the expected improvements for NWP. To carry out this work, we have developed a large database of synthetic IRS observations, background and "realistic" atmospheric profiles. In order to create this database, we place this work within a one-dimensional framework with synthetic observations. Finally, we have selected IRS channels, retaining the most informative ones using a method derived from Rodgers' information theory (Rodgers 1996) and used for the selection of AIRS (Fourrié and Thépaut 2003) or IASI (Collard 2007; Coopmann et al. 2020b) channels for example.

Section 2 describes in more details the instrumental and spectral characteristics of IRS and Section 3 presents a sensitivity analysis of IRS channels to thermodynamics and atmospheric chemistry, as well as their weighting functions and Jacobians. Then, Section 4 presents the different steps that led to the setting up of IRS observations and a synthetic observation database that can be used for the channel selection carried out in Section 5. Finally, Section 6 will deal with the conclusions and propose perspectives for future studies.

2. IRS caracteristics

The IRS instrument will simultaneously acquire a large number of spectral soundings using a bi-directional detector array in two spectral bands (Table 1).

	Band 1	Band 2
Number of channels	881	1079
Wave number [cm^{-1}]	679.7034 - 1210.4391	1599.7688 - 2250.5427
Wave length [μm]	8.3 - 14.7	4.4 - 6.3
Max Optical Path Difference [cm]	0.8290	0.8282
Unapodized sampling rate [cm^{-1}]	0.6031	0.6037

Table 1. IRS spectral specifications. Credit: EUMETSAT

Thus the Earth's disc will be measured by a sequence of square sub-images called Dwell. The coverage of the Dwells will be stepped in an east/west direction to form a sounding line, before moving northward to form a new one. The complete disc formed by these Dwells will be separated into 4 Local Area Coverage zones (LAC) as shown in Figure 1. These 4 LAC will be defined and scanned sequentially. Each Dwell will be measured in 10 s with a coverage of 640 x 640 km^2 at nadir. Each Dwell will contain 160 x 160 pixels with a spatial sampling distance of 4 x 4 km at nadir and approximately 4 x 7 km over Europe. Europe

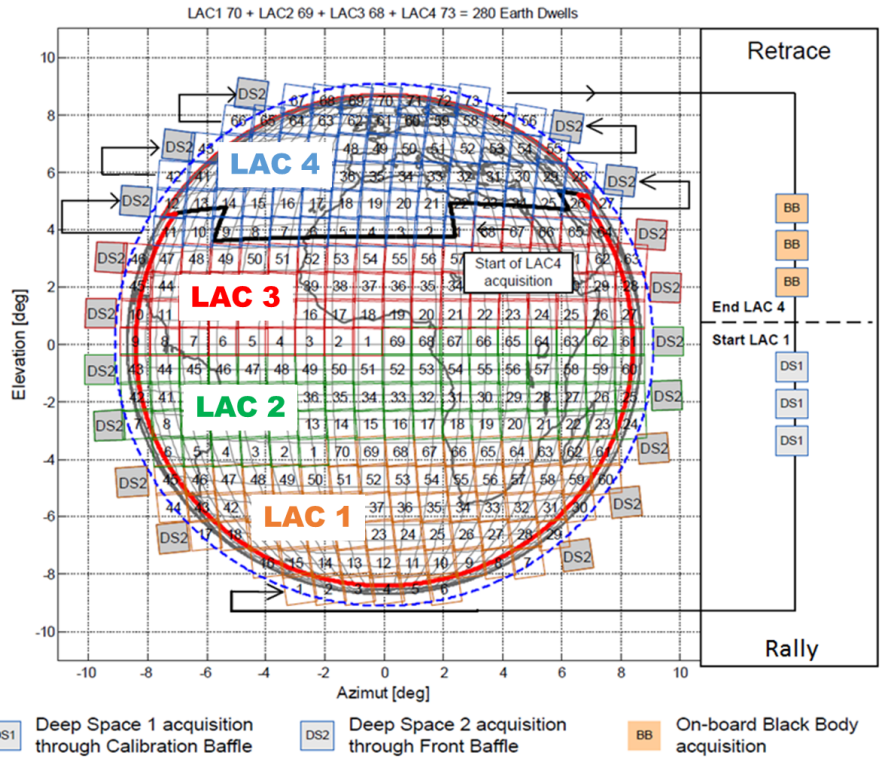

Figure 1. MTG Infra-Red Sounder dwell coverage. Credit: EUMETSAT

located in LAC 4 will be observed every 30 min. Below are the technical specifications of the instrument (EUMETSAT):

- Volume: 1.4 x 1.6 x 2.2 m^3
- Mass: 400 kg
- Power: 750 W
- Radiometric measurement ranges between 180 K and 313 K (equivalent black-body temperature)
- Spectral radiometric noise (excluding spectral calibration) at 280 K black-body: between 170 mK and 900 mK depending on the considered wave number inside the band of interest

3. Preliminary study

3.1. *Radiative Transfer Model*

In order to evaluate the IRS sensitivity, calculate the Jacobians and simulate IRS radiances, we used the Radiative Transfer Model (RTM) RTTOV which is developed and maintained by the Satellite Application Facility (SAF) of EUMETSAT for NWP. In the RTTOV algorithm, the input atmospheric profiles (temperature and humidity) are often variable and provided by the users, the other constituents such as O_3, CO_2, CH_4, CO, N_2O, etc. can also be provided or are assumed to be constant profiles in time and space (depending on the version of the coefficients). For our sensitivity study, all thermodynamic and chemical profiles will be variable.

3.2. *Sensitivity Analysis*

A spectral sensitivity study on the bands of the IRS spectrum was conducted using the RTTOV version 12 (Saunders et al. 2018).

We used the method developed by (Gambacorta and Barnet 2013) which consists in evaluating the brightness temperature response to a perturbation for each atmospheric constituent separately: temperature (T), skin temperature (T_{skin}), water vapour (q), ozone (O_3), carbon dioxide (CO_2), methane (CH_4), carbon monoxide (CO), nitrous oxide (N_2O) and sulfur dioxide (SO_2). The brightness temperature response (ΔBT) is calculated by the difference between simulations with perturbed profiles and simulations with unperturbed profiles. We used 6 atmospheric profiles spread over the globe from the (Matricardi 2008) database which we collocated with the space-time coordinates and instrumental characteristics of IRS. The vertical profiles of each variable have been modified by a constant perturbation typical of atmospheric variability whose values are:

- T: 1 K
- T_{skin}: 1 K
- q: 10 %
- O_3: 10 %
- CO_2: 1 %
- CH_4: 2 %
- CO: 1 %
- N_2O: 1 %
- SO_2: 1 %

Figure 2 shows the sensitivity analysis of the IRS channels (1960) by averaging the ΔBT over the 6 atmospheric profiles. It can be seen that the IRS instrument is potentially sensitive to all the variables evaluated here. However, the use of IRS to observe methane seems to be complex since the spectrum of band 1 stops at the beginning of the CH_4 absorption band. We considered that channels with a ΔBT greater than 0.01 K are sensitive to the studied species. Thus, IRS is a suitable candidate for the observation of atmospheric temperature and water vapour, surface properties and also for the observation and monitoring of the atmospheric composition of species such as O_3, CO_2, CO, N_2O and SO_2.

3.3. *Weighting function and Jacobians*

In order to evaluate the sensitivity of these variables according to the vertical in the atmosphere, we calculated the Weighting Functions and Jacobians using the RTTOV RTM with the 6 previous profiles. Figure 3 shows the average for these 6 profiles of the normalized Weighting Functions, the normalized Jacobians of water vapour, ozone, carbon dioxide, methane, carbon monoxide, nitrogen oxide, sulphur dioxide and the Jacobians of temperature as a function of pressure and skin temperature for the 1960 IRS channels. The Jacobian shows to which levels in the atmosphere the brightness temperature at given wavenumber is sensitive, with respect to temperature, humidity or concentrations of the different gases present in our case. The Jacobian matrix is denoted **H** in the following.

It can be seen in Figure 3, that IRS channels are mainly sensitive in the troposphere for all atmospheric variables except for temperature and carbon dioxide which has early band 1 channels sensitive in the stratosphere and UTLS (Upper Troposphere Lower Stratosphere). Ozone band has also sensitive channels in these same parts of the atmosphere. It can be noted that Jacobians of temperature and water vapour are mainly sensitive in the lower troposphere for channels located in the first atmospheric window between 770 and 970 cm^{-1}.

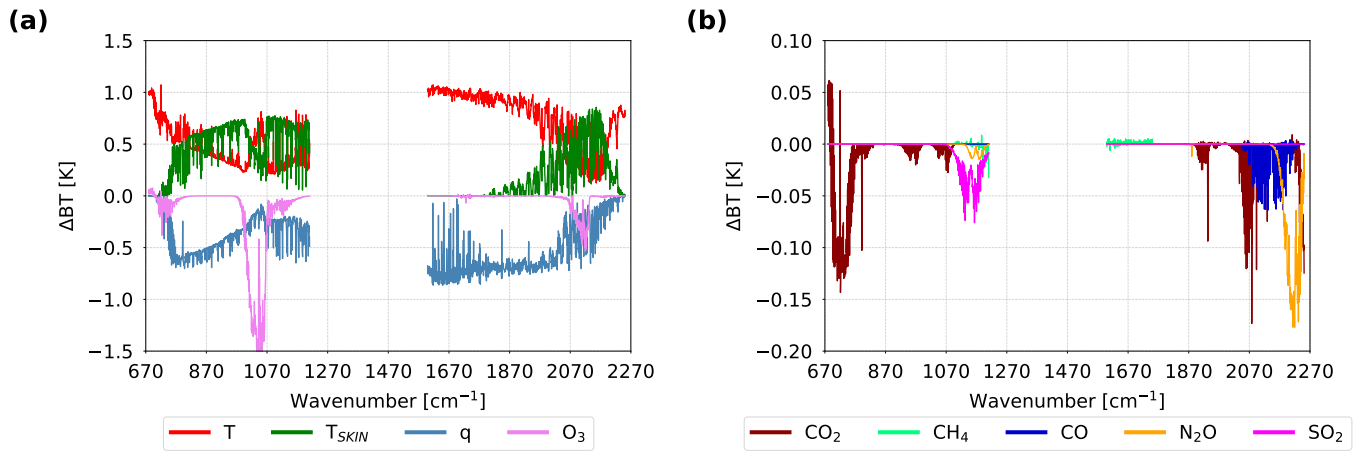

Figure 2. Sensitivity analysis of IRS brightness temperature to temperature, water vapour, ozone and skin temperature (a) and to carbon dioxide, methane, carbon monoxide, nitrous oxide and sulfur dioxide (b).

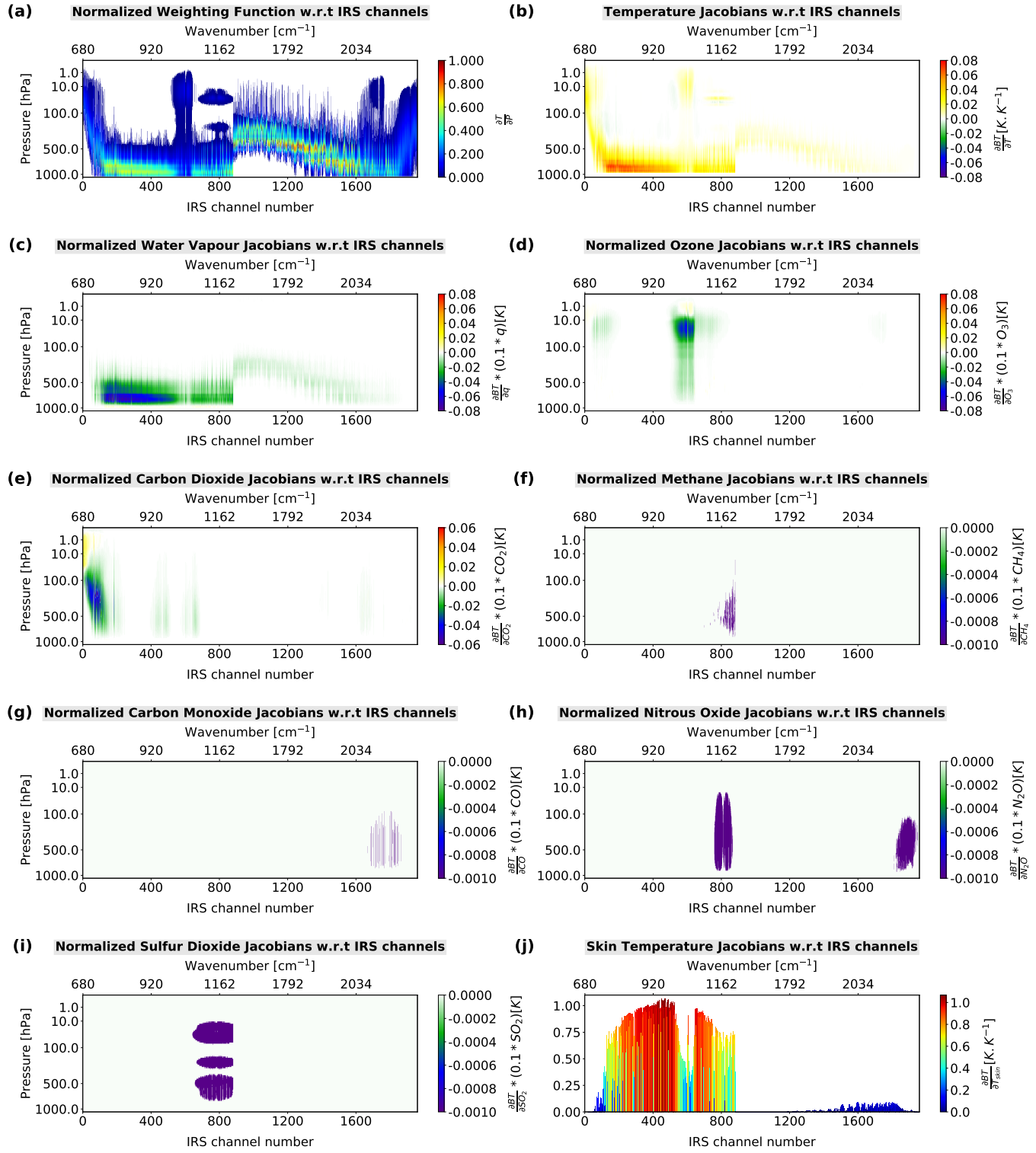

Figure 3. Mean of Normalized Weighting Function (a), Temperature Jacobians (b), Normalized Water Vapour (c), Ozone (d), Carbon Dioxide (e), Methane (f), Carbon Monoxide (g), Nitrous Oxide (h), Sulfur Dioxide (i) Jacobians with respect to pressure and Skin Temperature Jacobians (j) for IRS channels.

4. Database

4.1. Realistic atmospheric data

In order to carry out this study, we have set up a large database of atmospheric profiles spread over the IRS observation disc (max zenithal angle of 65 °). To do this, we randomly selected coordinates in the 4 LACs for 7458 cases for the dates of 15 July and November 2016, day and night, on land and sea. Longitudes, latitudes, zenith and azimuth angles for the IRS instrument were provided by EUMETSAT and the solar zenith and azimuth angles have been calculated for these study cases. Finally, we extracted atmospheric profiles (temperature, humidity, ozone, cloud fraction, cloud liquid water, cloud solid water) and surface parameters (surface pressure, skin temperature, temperature at 2 metres, humidity at 2 metres, zonal and meridional wind at 10 metres) for the 7458 case studies. The profiles are extracted from a research configuration of the ARPEGE NWP global model at Météo-France (Coopmann et al. 2020a). This database will be used for the construction of the background profiles used as a priori and the synthetic observations. It will be considered as the reality ($\mathbf{x}^t$) in the rest of the study. Since the ARPEGE profiles are on 105 levels, we have interpolated onto 101 levels in order to have the same pressure levels as the IRS coefficients provided for the RTTOV RTM.

4.2. *Background atmospheric data*

In order to produce the background profiles and IRS observations, we have chosen the method used in an OSSE (Observing System Simulation Experiment) context. In a first step, the background profiles were generated from the database of "true" profiles perturbed with the addition of background-errors such as:

$$\mathbf{x}^{\mathrm{b}} = \mathbf{x}^{\mathrm{t}} + \mathbf{B}^{1/2}.\eta^{\mathrm{b}} \quad (1)$$

where $\mathbf{x}^{\mathrm{b}}$ is the perturbed background profile, $\mathbf{x}^{\mathrm{t}}$ is considered as the "true" profile, η^{b} is a random vector drawn from a Gaussian distribution with zero mean and unit standard deviation and $\mathbf{B}^{1/2}$ is the square root of the background-error standard deviations determined from a multi-variate background-error covariance matrix **B**. The latter is derived from the Coopmann et al. (2020b) study and calculated using the NMC (The National Meteorological Center) method (Parrish and Derber 1992) over one year of data for temperature, humidity, ozone and skin temperature.

4.3. *Observation data*

A major difficulty of this study is that no IRS observations are already available. To compensate for this lack of data, a framework is proposed for generating synthetic observations that are close to reality. As for background profiles, IRS observations are calculated from perturbed simulations with the addition of observation-errors such as:

$$\mathbf{y} = \mathcal{H}(\mathbf{x}^{\mathrm{t}}) + \mathbf{R}^{1/2}_{\mathrm{NE\Delta T}}.\eta^{\mathrm{o}} \quad (2)$$

where $\mathbf{y}$ is the perturbed IRS observations, $\mathcal{H}(\mathbf{x}^{\mathrm{t}})$ is the simulated observations from "true" profiles, η^{o} is a random vector drawn from a Gaussian distribution with zero mean and unit standard deviation and $\mathbf{R}^{1/2}_{\mathrm{NE\Delta T}}$ is the square root of the observation-error covariance matrix (**R**).

As a first approximation, the diagonal R matrix was constructed using the instrumental noise values provided by EUMETSAT. The IRS noise is given in terms of the equivalent noise of temperature difference at 280 K $\mathrm{NE\Delta T(T_{ref})}$. For the 7458 case studies, **R** matrices were constructed independently from the converted $\mathrm{NE\Delta T(T)}$ for the corresponding scene temperature T with the formula:

$$\mathrm{NE\Delta T(T)} = \mathrm{NE\Delta T(T_{ref})}.\frac{\partial \mathrm{B}/\partial \mathrm{T(T_{ref})}}{\partial \mathrm{B}/\partial \mathrm{T(T)}} \quad (3)$$

where $(\mathrm{T_{ref}})$ denotes values at 280 K, (T) denotes values at the scene temperature and (B) is the Planck function. This conversion was applied for each channel and each study case.

To simulate IRS observations, we chose to use the RTTOV RTM version 12. Atmospheric profiles and surface parameters from $\mathbf{x}^t$ were used as input to the RTM. Moreover, having cases on land and sea and in order to increase the realism of the simulations, we used the land emissivity values from the Combined ASTER MODIS Emissivity over Land (CAMEL) (Borbas et al. 2018) and the emissivity values over sea from a surface emissivity model (IREMIS) (Saunders et al. 2017) available in RTTOV. Finally the simulations were performed in clear and cloudy skies. To take clouds into account, we used the microphysical properties assembled in the Optical Properties of Aerosols and Clouds (OPAC) software package (Hess et al. 1998) for liquid water considering maritime and continental stratus and the optical properties of ice clouds provided by Baran 2018 for solid water (Vidot and Brunel 2018). Here, the simulations are performed without apodisation.

Thus, we simulated IRS observations for the 7458 case studies. Figure 4 shows the mean (a) and standard deviation (b) of the simulated spectra for the 1960 channels in clear (blue) and cloudy (grey) sky. As expected, it can be seen that the means are lower in cloudy sky for the atmospheric window channels and near-surface channels. Inversely, these same channels have higher standard deviations than in clear sky indicating a much greater variability induced by the presence of clouds.

4.4. *Cloud detection*

One of the main complexities in the assimilation of infra-red observations is the consideration of clouds. Indeed, a large part of the infra-red spectrum is sensitive to clouds. This can be problematic because, as (Fourrié and Rabier 2004; McNally 2002) show, 90 % of the pixels of the IASI instrument are totally or partially contaminated by clouds. In order to avoid this under-use of observation, the major NWP centres have developed methods to assimilate clear channels above clouds. Among them, the (McNally and Watts 2003) Cloud Detection Scheme allows the detection of channel affected by the presence of a cloud.

Within the framework of this study, it is important to assess the sensitivity of IRS observations to clouds as well as the capacity of the scheme to detect them. To achieve this, we used the ECMWF Aerosol and Cloud Detection Software which uses the method given by (McNally and Watts 2003) available on the NWPSAF website. To reveal the radiative effect of the cloud and to separate

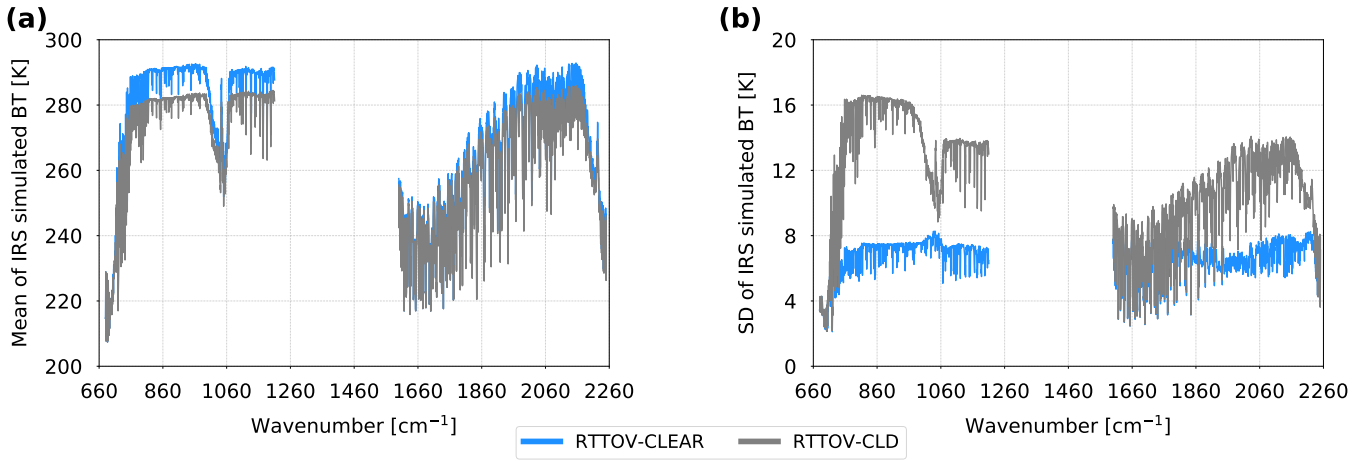

Figure 4. Means (a) and standard deviations (b) of IRS spectra simulated in brightness temperature from RTTOV in clear (blue) and cloudy (grey) sky for the 1960 channels.

it from other contributions, the cloud detection algorithm works by taking the background departures (i.e., the difference between the observed and simulated clear brightness temperature (BT)) and looking for the signature of opacity that is not included in the clear-sky calculation (i.e. cloud). To do this, the channels are first ranked in the vertical, according to their height assignments (Eresmaa 2020).

In this section, we have two main objectives;

- Step 1: to evaluate the ability of the algorithm to correctly classify the clear/cloudy channels based only on the ideal IRS simulations
- Step 2: to estimate the impact of the instrumental noise on the cloud detection when creating the synthetic IRS observation

For this first step we provide the cloud detection software with the ideal IRS simulations in clear ($\mathcal{H}_{\text{clear}}(\mathbf{x}^{\text{t}})$), cloudy ($\mathcal{H}_{\text{cloud}}(\mathbf{x}^{\text{t}})$) and the associated height assignments calculated in the previous section for the 1960 channels and 7458 pixels. Thus, the algorithm provides a clear/cloudy identification for all channels and the full data set based on the difference:

$$\delta_{\text{step1}} = \mathcal{H}_{\text{cloud}}(\mathbf{x}^{\text{t}}) - \mathcal{H}_{\text{clear}}(\mathbf{x}^{\text{t}}) \tag{4}$$

The figure 5 shows the frequency of IRS channels flagged clear for the 1960 channels and 7458 pixels. The behaviour is consistent with the results obtained by (McNally and Watts 2003) which show that channels located in the strongly absorbing parts of the spectrum (CO_2 absorption band from 680 to 740) are generally clear and are correctly evaluated. Also that channels in the atmospheric window are rarely clear and larger increases in absorption due to water vapour result in a marked increase in the occurrence and detection of clear channels. Nevertheless, we notice that our simulated observations appear clearer overall. This can be explained by the ideal framework used in this study, as explained by (Masutani et al. 2010), by a degree of cloud contamination of the measurements that may be underestimated.

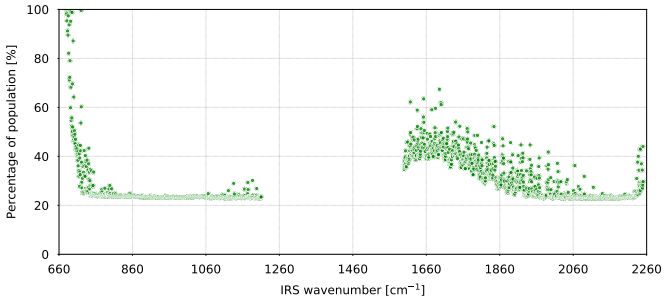

Figure 5. Frequency of IRS channels flagged clear.

In the same step, we can evaluate the clear/cloudy flagged channels as a function of the geophysical parameters used to simulate the all-sky observations in order to identify a consistency between the sensitivity level of the flagged channel and the cloud height. Thus, Figure 6 provides an example of the cloud fraction for three different cases; high (a), medium (b) and low (c) clouds associated with the normalized weighting functions of the identified clear (green) and cloudy (grey) channels. It can be seen that the cloud detection algorithm seems to work well by identifying as clear only those channels with a sensitivity peak above the cloud. In case (a) 42 channels are identified clear, 206 in case (b) and 588 in (c). Despite an overestimation of clearly identified channels, the algorithm is able to consistently classify IRS channels.

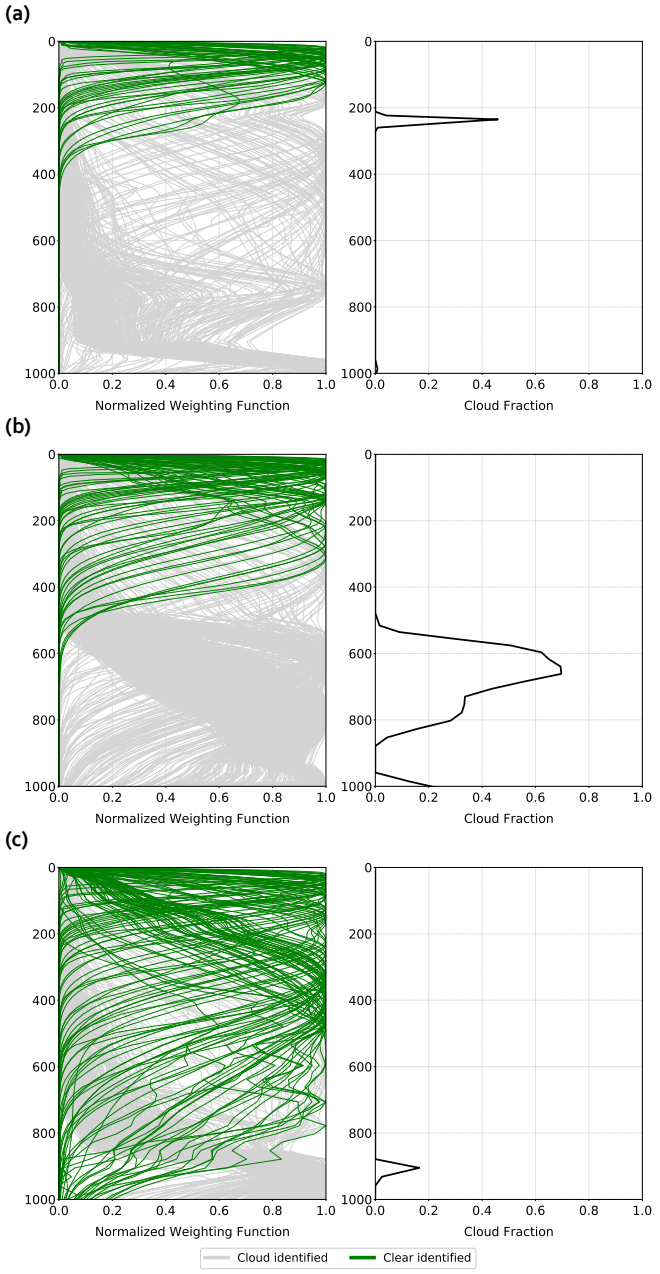

Figure 6. Normalized Weighing Functions of the IRS channels (identified clear in green and cloudy in grey) of band 1 (left) and cloud fraction (right) as a function of pressure for cases of high (a), medium (b) and low (c) clouds.

For the second step, we proceed in the same way with the difference that the cloud simulation is perturbed by the IRS instrumental noise (Equation 2). This allows us to evaluate the effect of the noise on the ability of the software to correctly classify the clear/cloudy channels:

$$\delta_{\text{step2}} = \text{y}_{\text{cloud}} - \mathcal{H}_{\text{clear}}(\mathbf{x}^{\text{t}}) \tag{5}$$

Figure 7 shows the agreement rate between step 1 and step 2 according to the 1960 channels for the 7458 pixels considered. It can be seen that the calculation of the synthetic observations used in step 2, slightly introduce cloudy signal in the channels by 5 % over the whole spectrum (in yellow) and clears (in orange) the channels between 5 and 10 % in the CO_2 and water vapour absorption bands. Thus, the calculation of synthetic observations changes very slightly the nature of the observation in terms of cloud detection. These results will not have any impact in our study since we have chosen to work only in clear sky conditions. However, these results show that in a more operational assimilation framework, a misclassification of clear1-cloudy2

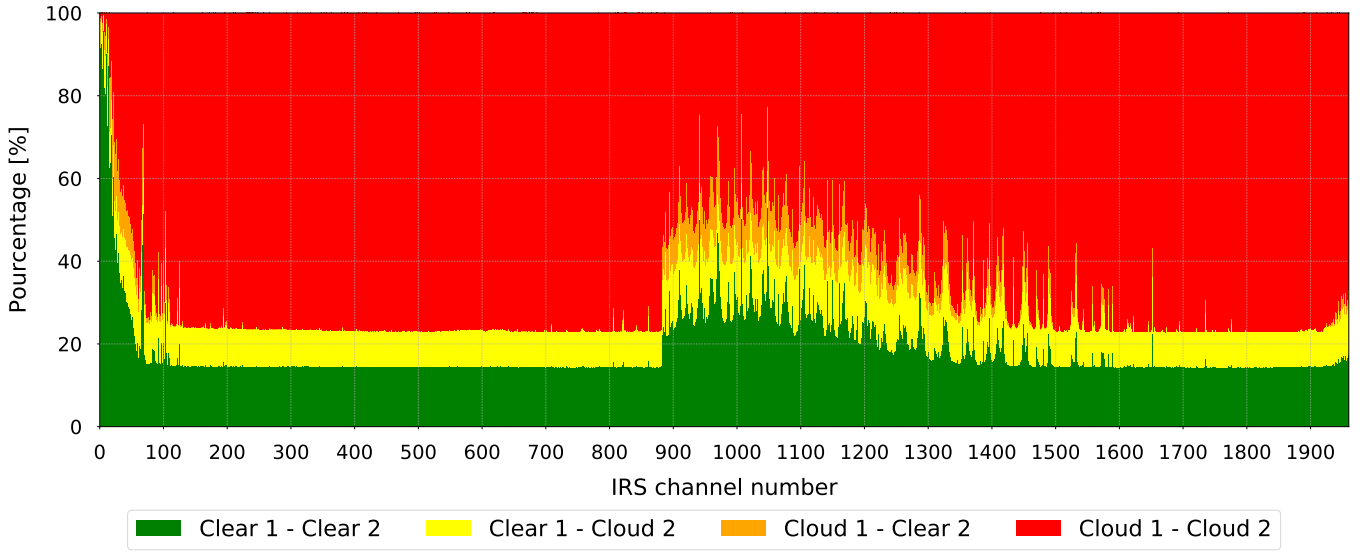

Figure 7. Agreement rate between step 1 and step 2 as a function of the IRS channel numbers for the 7458 study cases.

(in yellow) would prevent the assimilation of some valuable information, while cloudy1-cloudy2 (in orange) would introduce wrong information in the system. The cloud detection algorithm allows to highlight 1700 pixels of which all channels are identified as clear.

4.5. Observation-errors

The impact of the observation-errors in the data assimilation process is very important. This is why it is essential to estimate accurately the observation-error covariances **R** in this context which uses synthetic observations (Vittorioso et al. 2021). For several years, innovative techniques have been used to achieve this by deriving estimates of the real observation-error from the departure statistics from assimilation systems (e.g. Hollingsworth and Lönnberg 1986; Desroziers et al. 2005). These methods make it possible to diagnose both the variances and covariances of observation-error, including cross-channel error correlations allowing an accurate estimate of the total errors that characterise the observation (instrumental noise, error of spatial representativeness, error in the calculation of radiative transfer, etc.).

This work is exploratory since we try to estimate real observation-errors for synthetic observations. This is a difficult task but achievable within a precise synthetic observation framework. Indeed, (Privé et al. 2013) have shown that an accurate representation of observation-errors is also important in a synthetic observation context. Nevertheless, the estimated error in this case will be smaller than the error of the real observations.

As a first step, we estimated the potential bias of these synthetic observations. We calculated the innovations of each IRS channel for the 1700 profiles by computing the difference between the synthetic observations ($\mathbf{y}$) and the simulations of observations using the background profiles within the RTTOV. The average of the first-guess departures was used to make a bias correction applied to the observations ($\mathbf{y}$). This bias correction is nevertheless slight since the mean of ($\mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{b}})$) values are between 0.2 and -0.2 K.

In order to accurately estimate the variances and covariances of IRS observation-error, as well as inter-channel correlations, we have chosen to use the method developed in (Coopmann et al. 2020b). It consists in using the values of the standard deviations of the first-guess departures previously calculated as a first estimate for standard deviations of observation-error. This diagonal R matrix determined is the starting point of the computation. Then, we have used the NWP SAF One Dimensional Variational (1D-Var) data assimilation algorithm (Havemann and Smith 2020) to apply the Desroziers method to our database and thus to precisely estimate the IRS observation error-covariance matrix **R**.

Thus the 1D-Var experiment was carried out on the 1700 cases previously considered as clear (see above). We used the background profiles ($\mathbf{x}^{\mathrm{b}}$ on 101 levels) calculated from Equation 1, associated with the multi-variate **B** matrix mentioned previously. The synthetic IRS observations ($\mathbf{y}$ on 1960 channels) are calculated from Equation 2, associated with the diagonal **R** matrix estimated by the standard deviations of the First-Guess departure. 1D-Var data assimilation experiment minimized on temperature between 1000 and 0.1 hPa, humidity between 1000 and 100 hPa, ozone between 1000 and 0.1 hPa and skin temperature.

To diagnose the observation error covariance matrices R, the Desroziers method is commonly used in NWP centres (Weston et al. 2014; Stewart et al. 2014; Bormann et al. 2016; Coopmann et al. 2020b;). Since the real state of the atmosphere is not known, the method shows that it is possible to estimate in the space of the observations, the **R** matrix with the deviations of the observations from the background and the analysis. Given several assumptions, the previous 1700 analyses allow to statistically diagnose the observation-error covariance matrix such as:

$$\mathbf{R} = \mathrm{E}[\mathbf{d}_{\mathrm{a}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] \tag{6}$$

where $\mathrm{E}[.]$ is the expectation operator, $\mathbf{d}_{\mathrm{a}}^{\mathrm{o}} = \mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{a}})$ is the Analysis departure and $\mathbf{d}_{\mathrm{b}}^{\mathrm{o}} = \mathbf{y} - \mathcal{H}(\mathbf{x}^{\mathrm{b}})$ is the First-Guess departure. The diagnosed **R** matrix contains the variances and covariances of IRS observation-error taking into account the cross-channel errors for the 1960 channels considered.

The instrumental noise values of IRS at 280 K are shown in grey, the values of First-Guess departures standard deviation (used as a first approximation of observation-errors) in red and the values of diagnosed observation-errors in blue in Figure 8 (a). Then Figure 8 (b) shows the diagnosed observation-error correlation matrix. The First-Guess departures standard deviation (in red) have values close to instrumental noise (0.3 to 1.0 K) up

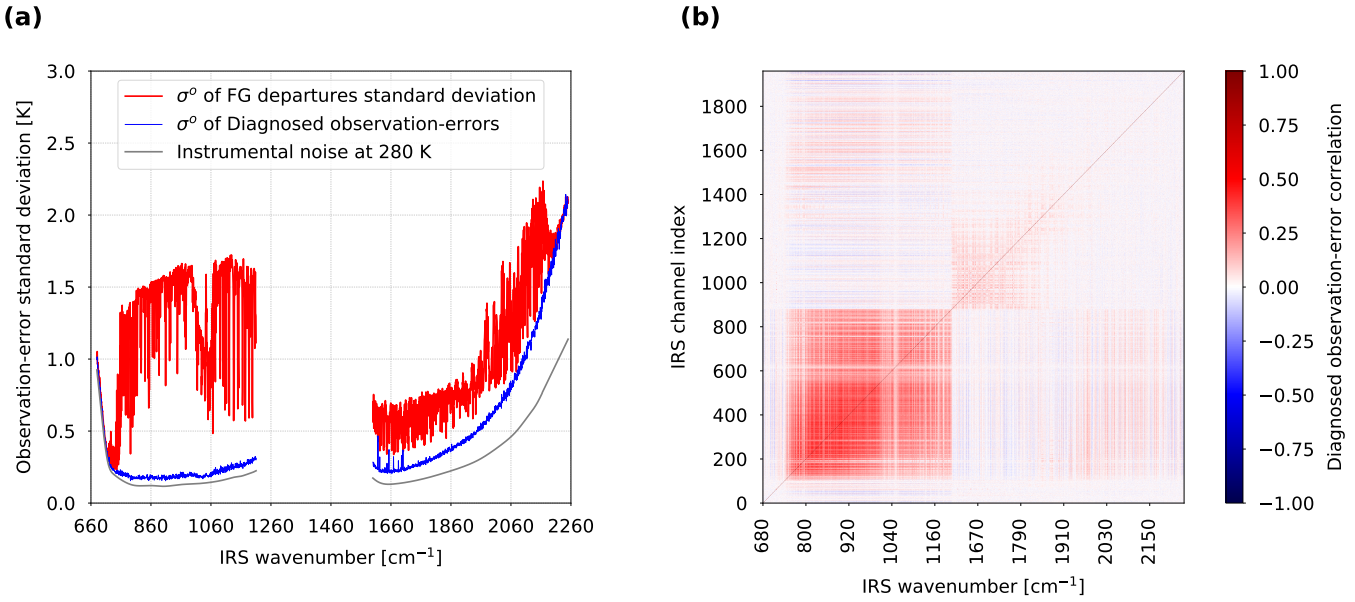

Figure 8. Observation-error standard deviation from FG departures standard deviation in red line, diagnosed observation-errors from Desroziers method using 1D-Var data assimilation system in blue line and instrumental noise at 280 K in grey with respect to 1960 IRS channel number and wave-number [cm^{-1}] (bands 1 & 2) over the set of 1700 clear atmospheric profiles (a). Diagnostic of IRS observation-error correlation with respect to the same channels as before (b).

to 700 cm^{-1} in the CO_2 absorption band. They are then always higher than instrumental noise and are between 0.5 and 1.7 K in atmospheric windows (700 - 1000 cm^{-1} and 1060 - 1250 cm^{-1}), between 0.5 and 1.4 K in the ozone absorption band (1000 - 1060 cm^{-1}) and between 0.5 and 2.0 K in the water vapour absorption band (1600 - 2250 cm^{-1}).

It is noticeable that the First-Guess departures standard deviation for atmospheric window channels are high values. This indicates a presence of a incorrect signal in some pixels of the 1700 clear cases considered and is probably due to the perturbation of skin temperatures used to obtain the background value as indicated in the cloud detection study previously performed. However, the diagnostic of observation-errors allows this effect to be smoothed out and consistent values between instrumental noise and observation-errors of FG departures standard deviation can be found. These diagnosed observation-errors are close to the errors generally obtained for infra-red sounders. Nevertheless, these errors are probably lower than those we would have with real observations. Moreover, the estimation of these errors takes into account this instrumental noise and radiative transfer errors but does not take into account errors related to spatial representativeness for example. In this context, the diagnostic of the **R** matrix allows a first estimation of the observation-errors for the IRS channels. The observation errors diagnosed for the channels of the ozone absorption band are lower than those of IASI (between 0.5 - 1.5 K) for example. This is probably due to the limitation of the ozone variability within the IRS disk which does not provide measures over the poles.

The Figure 9 (b) shows higher error correlation values (0.2 - 0.4) for the channels and inter-channels used for temperature sounding in the lower atmosphere between 700 and 750 cm^{-1}, as well as for the first atmospheric window between 750 and 950 cm^{-1}. The error correlation values for the channels in the ozone absorption band are about 0.2 and then 0.25 for the second atmospheric window between 1060 et 1200 cm^{-1}. Finally, the error correlation values are low for the channels of the water vapour band. These error correlation values are lower than expected in the same way that the errors are generally lower for these synthetic observations compared to real observations (Privé et al. 2013). This diagnosed **R** matrix will be used in the following to perform IRS channel selection.

5. Channel selection

5.1. *Methodology*

The observations measured by the IRS instrument will represent an impressive amount of data. This high volume of information constitutes a challenge for data storage and transmission. The computational cost of assimilating all IRS channels would be too high and inefficient due to redundant information. Typically, EUMETSAT disseminate a smaller quantity of channels especially selected to improve the NWP models of weather centres. The aim is to make a precise selection of all IRS channels to provide the optimal information for the NWP models. Currently, the channels that can provide information on atmospheric temperature, water vapour and surface temperature represent the most useful observations for improving weather forecasts. In addition, several NWP centres have included ozone as a variable in the models because it improves the assimilation of specific channels in the infra-red, which are sensitive to ozone, temperature and water vapour, and it allows better accounting of radiative feedback, which improves weather forecasts (Derber and Wu 1998; Lahoz et al. 2007; John and Buehler 2004; Ivanova et al. 2017; Coopmann et al. 2018; Dragani et al. 2018; Coopmann et al. 2020a). Thus, we have selected the most informative channels in temperature, water vapour, skin temperature and ozone.

To carry out this selection, we applied a method that is widely used in the scientific community. This is the Degrees of Freedom for Signal (DFS), which allows to select a set of optimal channels having the largest information content for each atmospheric profile as described by (Rodgers 1996, 2000). The DFS is based on the information theory and provides a measure of the gain in information gathered by the observations according to the formula:

$$\mathrm{DFS} = \mathrm{Tr}(\mathbf{I} - \mathbf{A}\mathbf{B}^{-1}) \tag{7}$$

where Tr denotes the trace, $\mathbf{I}$ the identity matrix, $\mathbf{B} \in \mathbb{R}^{n \times n}$ (n parameters to be retrieved) is the background-error covariance matrix and $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the analysis-error covariance matrix

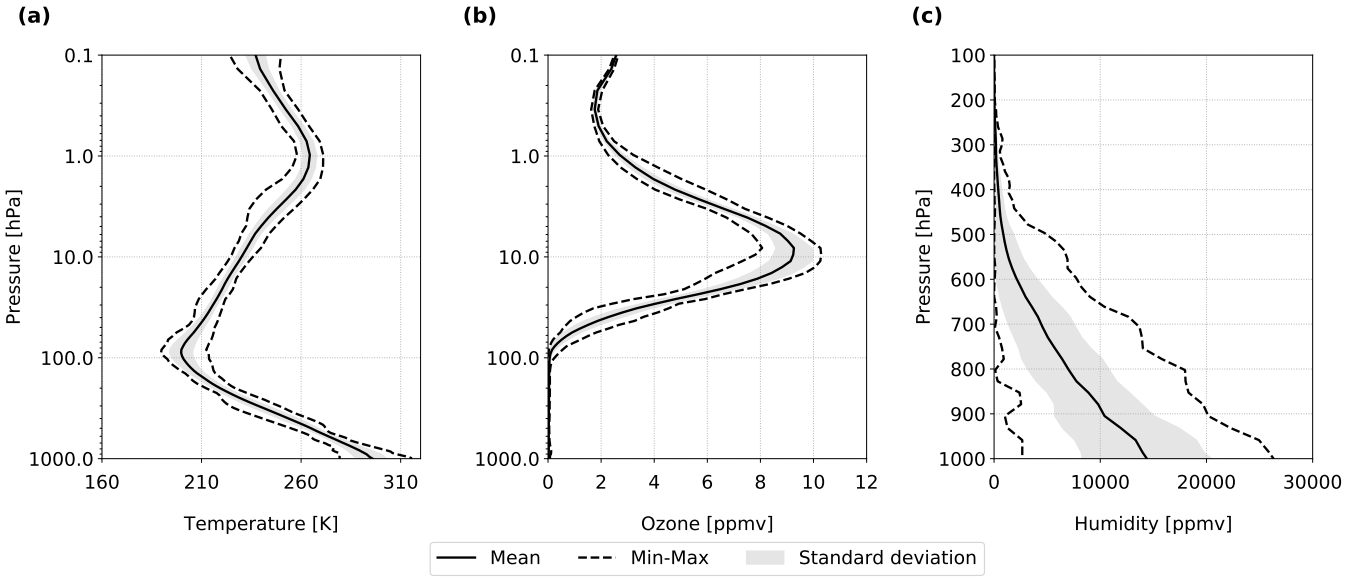

Figure 9. Mean ± standard deviation and min-max values of temperature (a), ozone (b) and humidity (c) profile with respect to pressure over the subset of 60 atmospheric database. Note that humidity statistics are shown between 1000 and 100 hPa

which is calculated as follow:

$$\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1} = (\mathbf{I} - \mathbf{B}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{B} \quad (8)$$

where $\mathbf{R} \in \mathbb{R}^{\mathrm{mxm}}$ (m channels considered) is the observation-error covariance matrix and $\mathbf{H} \in \mathbb{R}^{\mathrm{mxn}}$ (the derivatives of each channel with respect to each parameter) represent Jacobians matrix for all IRS channels.

As we have seen previously, some IRS channels, like others in the infra-red, are sensitive to several variables (multi-sensitive channels). As we have a diagnosed observation-error covariance matrix including inter-channel errors, we have carried out this selection by using as a figure of merit the total DFS taking into account the DFS of temperature, water vapour, skin temperature and ozone, in order to select the most sensitive or poly-sensitive channels that provide the maximum amount of information, such as:

$$\mathrm{DFS_{TOT}} = \mathrm{DFS_T} + \mathrm{DFS_q} + \mathrm{DFS_{O_3}} + \mathrm{DFS_{T_{skin}}} \quad (9)$$

Finally, we chose 60 profiles among all the 1700 clear cases considered and distributed over the full IRS disk (15 profiles per LAC) in order to carry out the channel selection and to take into account different scenarios of atmospheric variability (see Figure 9). The use of cross-channel error correlations allows us to consider all channels during selection as well as adjacent channels. We used the diagnosed observation-error covariance matrix **R**, the multi-variate background-error covariance matrix **B** and the Jacobian matrix **H** containing for each profile the Jacobians of the 1960 IRS channels. For channel selection, the DFS calculation considers for each profile, 93 levels for temperature and ozone from 1000 to 0.1 hPa and 53 levels for humidity from 1000 to 100 hPa.

The first selection step consists in selecting for each of the 60 atmospheric profiles the most informative channel with the highest total DFS among all IRS channels using a **R** matrix $\mathbb{R}^{1\mathrm{x}1}$. For each profile, the first selected channel is fixed and a new combination of this with a second channel from the (1959 - 1) channels is searched using a **R** matrix $\mathbb{R}^{2\mathrm{x}2}$. The combination giving the highest total DFS allows to select and fix the second channel. This operation is repeated as many times as necessary. We have chosen to stop this process after the selection of 500 IRS channels for each of the 60 profiles since EUMETSAT is limited in the dissemination of data by the Global Telecommunications System (GTS).

5.2. *First step of selection*

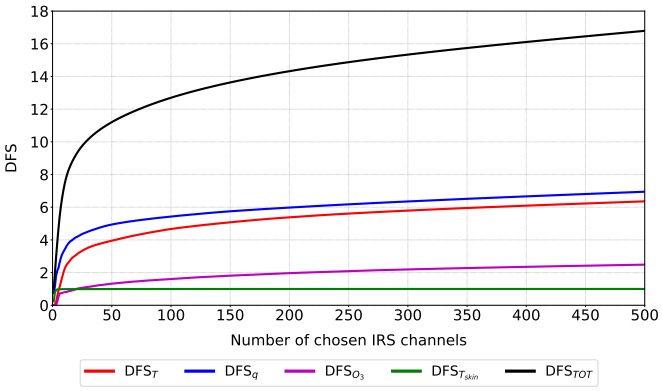

Figure 10. Evolution of mean DFS for temperature (red), humidity (blue), ozone (violet), skin temperature (green) and total (black) during the channel selection over the subset of 60 atmospheric profiles.

Figure 10 shows that the information content increases very quickly with only 50 channels which represents a mean total DFS of 11.20 or 66.7 % of the mean total DFS available with 500 channels (total DFS of 16.79). It is noticeable that the mean DFS of humidity is slightly higher than the mean DFS of temperature. This is consistent with the previous sensitivity study which showed that a very large number of IRS channels are sensitive to water vapour even those located in the atmospheric window bands. Two other hypotheses can explain this behaviour. Contrary to other hyperspectral sounders, IRS observations start further out in the spectrum at around 680 cm^{-1} unlike IASI and CrIS which start at 645 and 650 cm^{-1} respectively. This deprives IRS of many strongly absorbing and temperature sensitive channels. In addition, IASI and CrIS observe in other parts of the spectrum that are highly sensitive to temperature between 1200 -1600 cm^{-1} and 2250 - 2400 cm^{-1} which IRS does not possess. Finally, these results can also be explained by the background errors in humidity. Indeed, the models are more wrong in humidity than in temperature, so the humidity errors are proportionally stronger

than the temperature errors. Finally, we observe that the average ozone DFS increases very rapidly to reach a plateau, while the mean skin temperature DFS reaches with only 3 channels more than 90 % of the mean skin temperature DFS available with 500 channels.

In order to analyse more specifically the selection process, we have plotted in Figure 11 a histogram of the percentage of the number of channels selected on the 60 considered atmospheric profiles. This means that if a channel is selected for all profiles, it achieves 100 % selection. Conversely, a channel never selected among 60 profiles reaches 0 % selection. These percentages are characterised by spectral band and main applications for the NWP (temperature sounding in the CO_2 absorption band, humidity sounding in the water vapour absorption band, ozone absorption band and atmospheric windows) and for all channels. This histogram shows that approximately 434 channels are always selected between 50 and 100 % of the time (for 30 to 60 profiles) and that 1526 channels are only selected between 0 and 50 % of the time (for 0 to 30 profiles). This shows an information redundancy of a large number of channels. Thus, a selection from the 434 channels should be sufficient to extract the maximum amount of useful information for numerical weather prediction.

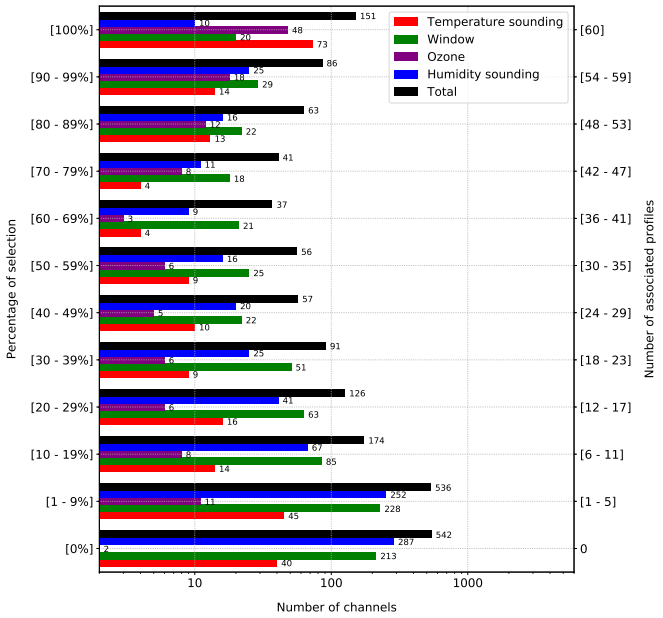

Figure 11. Percentage of the number of channels selected (up to 500 channels) on the subsets of 60 atmospheric profiles separeted by spectral group and main application to temperature sounding (in red), humidity sounding (in blue), ozone (in violet), atmospheric window (in green) and total (in black) with respect to number of related profiles.

Among the 1960 considered IRS channels, 151 channels are always selected from the 60 profiles (100 %), among them, 73 for temperature sounding, 48 in the ozone absorption band, 20 in atmospheric windows and 10 for humidity sounding (in green in Figure 12). Overall, there is a decrease in the number of channels selected up to 60 to 69% selection percentage and then an increase up to 0 % selection percentage. The number of channels always selected in the ozone absorption band and in the atmospheric windows is significant and shows the multi-sensitive properties of these channels, which can provide a large amount of information. Finally, 542 channels are never selected, mainly in the water vapour absorption band and in the atmospheric windows (in red in Figure 12).

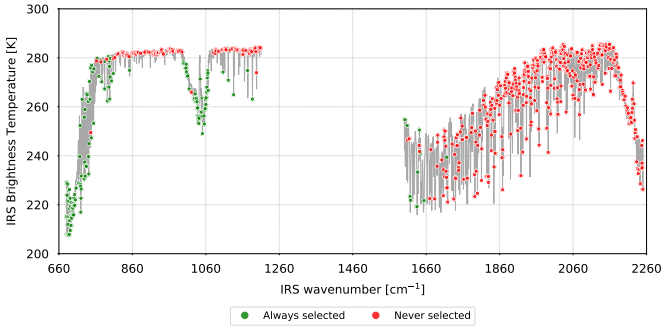

Figure 12. Location on a typical IRS spectrum in brightness temperature of the always (green) and never (red) selected channels.

5.3. *Results*

Once this first step was completed, we sorted these 500 pre-selected channels using two different methods to complete the final channel selection. The first method consists of selecting increasingly the most often selected channels while the second method selects the first rank selected channels. For both methods, we have separated the selections by packages of 75, 150, 300 and 500 channels. For each of these selections we have calculated the mean total DFS over 60 profiles. We notice that the first 75 channels selected have a higher mean total DFS than the 75 most selected. Then the selections with 150 channels have an equivalent mean total DFS. Finally, the selections with 300 and 500 most selected channels have higher mean total DFS than the first selected channels as shown in Figure 13. To ensure the robustness of these selections, the latter were used to calculate the mean total DFS for the 1700 clear profiles. The results are similar to the values shown in Figure 13. We have therefore chosen to evaluate the most often selected channel selections in the following.

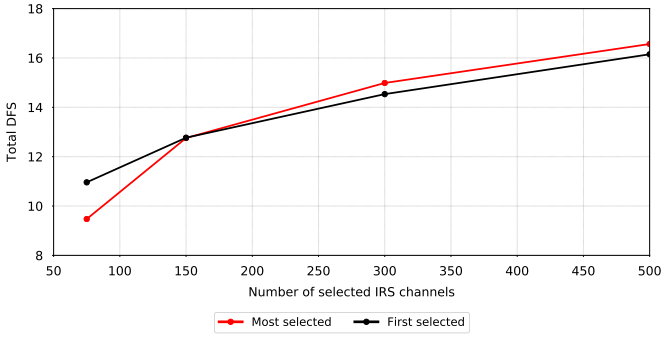

Figure 13. Evolution of the mean total DFS over 60 profiles as a function of the number of selected channels for the 75, 150, 300 and 500 first selected channels (in black) and the most selected channels (in red).

Table 2 summarizes the values of mean total DFS for temperature, humidity, ozone and skin temperature over 60 profiles for the 75, 150, 300, 500 most selected channels and all channels. It can be seen that the percentage of mean total DFS available increases rapidly between 75 and 150 channels (9.2 %), then 6.2 % between 150 and 300 channels and 4.4 % between 300 and 500 channels. A large part of the mean total DFS value is provided by the mean DFS of temperature and humidity, followed by the ozone and finally skin temperature components.

To assess the impact of these channel selections on weather analyses, we performed 1D-Var data assimilation experiments on the 1700 profiles considered clear. We used the diagnosed **R** matrix and the multi-variate **B** matrix. The data assimilation process allows to minimize on temperature, humidity, ozone and skin temperature. These experiments allowed us to calculate

	Temperature DFS	Humidity DFS	Ozone DFS	Tskin DFS	Total DFS
75 most selected channels	3.44 (27.7 %)	3.50 (25.0 %)	1.55 (22.2 %)	0.98 (98 %)	9.47 (26.4 %)
150 most selected channels	4.98 (35.8 %)	5.01 (35.8 %)	1.78 (25.5 %)	0.99 (99 %)	12.76 (35.6 %)
300 most selected channels	5.71 (41.0 %)	6.12 (43.7 %)	2.16 (30.9 %)	1.00 (100 %)	14.99 (41.8 %)
500 most selected channels	6.29 (45.2 %)	6.79 (48.5 %)	2.48 (35.5 %)	1.00 (100 %)	16.56 (46.2 %)
All channels	13.91	13.99	6.98	1.00	35.88

Table 2. Mean of Total, temperature, humidity, ozone and skin temperature DFS over 60 profiles and percentage of available mean DFS for the 75, 150, 300, 500 most selected channels and all channels.

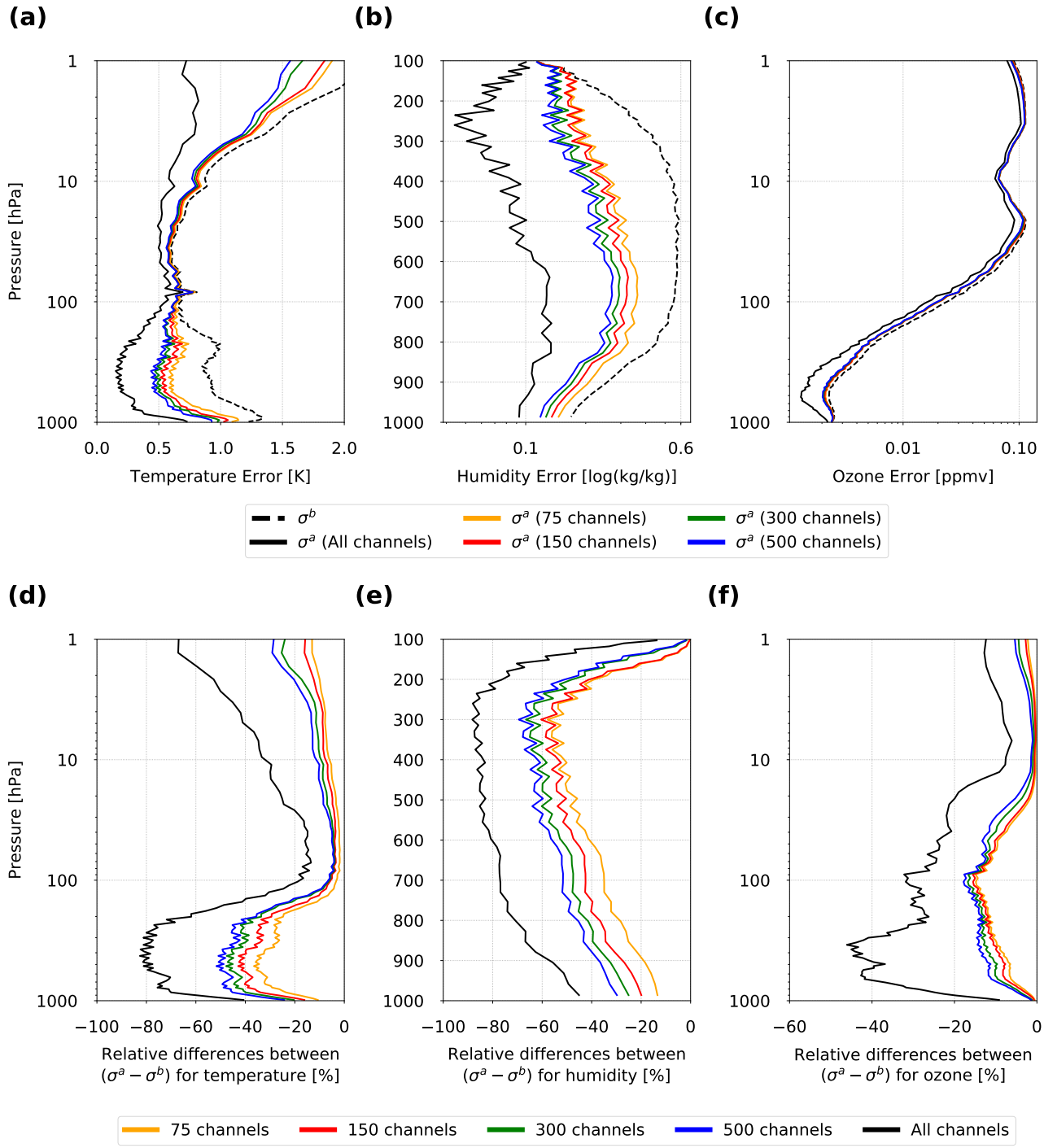

Figure 14. Vertical profiles of background-error standard deviation (σ^b) and mean analysis-error standard deviation (σ^a) for temperature (a), humidity (b) and ozone (c) and vertical profiles of relative différence between σ^a and σ^b for temperature (d), humidity (e) and ozone (f) with respect to pressure. This results are derived from 1D-Var data assimilation experiments over a set of 1700 atmospheric profiles with different channel selections (75, 150, 300, 500, all channels).

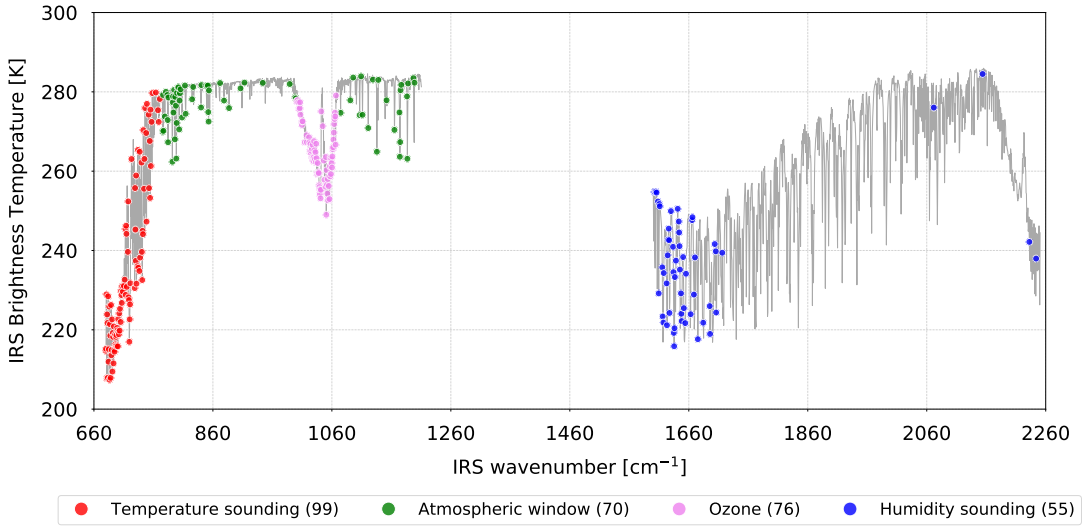

Figure 15. Location of the 300 selected IRS channels on a typical spectrum by main sensitivity.

analysis-error covariance matrices averaged over the 1700 study cases. We expect these selections minimize the analysis-error mainly in temperature and humidity for optimal use in numerical weather prediction.

We have represented on Figure 14 the vertical profiles of background-error standard deviation (σ^b) and mean analysis-error standard deviation (σ^a) for temperature (a), humidity (b) and ozone (c) and the vertical profiles of relative difference between σ^a and σ^b for temperature (d), humidity (e) and ozone (f) with respect to pressure. The mean of the analysis-error standard deviations is calculated from the analysis-error covariance matrices **A** determined by the 1D-Var experiments over the 1700 cases considered. Data assimilation experiments were performed for selections of 75, 150, 300, 500 and all channels.

In Figure 14 (a, b, c), the reduction in analysis error is the greatest in the troposphere and upper stratosphere for temperature and over the entire troposphere for humidity. The vertical profiles of analysis error reduction are close between selections of 75 - 150 channels and 300 - 500 channels. Thus, a good compromise between the quality of the analysis error and the quantity of used channels would be to select the 300 IRS channel selection. In Figure 14 (d, e, f) are represented the relative analysis error reductions compared to the background errors. By selecting 300 channels, the analysis error compared to the background error can be reduced by up to 48 % for temperature in the troposphere, up to 65 % for humidity in the Mid-troposphere and up to 17 % for ozone in the UTLS. These results show the potential for improvement in weather forecasts by selecting 300 IRS channels for the NWP.

This selection of 300 channels is shown in Figure 15 on a typical IRS spectrum per main application for NWP. There are 99 channels mainly selected in the CO_2 absorption band and used to provide information on atmospheric temperature, 55 in the water vapour absorption band for information on atmospheric humidity, 70 in the atmospheric windows to provide information on surfaces but also on temperature and humidity in the lower troposphere and 76 channels in the ozone absorption band to provide information on ozone but also on temperature in the troposphere and stratosphere and humidity in the troposphere. As seen previously, the ozone and surface channels have multiple sensitivities making them usable for several applications. This use is conditioned by the use of an observation error covariance matrix taking into account cross-channel error correlations. Thus, in order to extract the maximum amount of information from the selected channels, a precise estimation of the **R** matrix is necessary, which we have tried to do in this study.

6. Conclusions and perspectives

In order to make a first general selection of IRS channels, we were inspired by the synthetic observation method to build a database for this work. Our realistic atmospheric state is provided by the global model ARPEGE (research experiment). A database of more than 7000 atmospheric profiles was extracted from this model in order to simulate the satellite observations of IRS using the RTTOV (version 12) radiative transfer model. The simulations were carried out under clear and cloudy skies, allowing us to evaluate the cloudy detection capacity of the IRS channels. A database of more than 7000 synthetic IRS observations has been built up and collocated with atmospheric profiles. The latter were perturbed to obtain background profiles for temperature, humidity, ozone and skin temperature. This method makes it possible to obtain an initial database (observations and background) which can then be used for various studies such as the estimation of observation errors or the selection of the information content of IRS.

In this study we used a multi-variate background error covariance matrix in temperature, humidity, ozone and skin temperature, calculated in a previous study using the NMC method over one year of data. In order to extract the maximum of information from the IRS channels, it is essential to take into account the cross-channel error correlation. Thus, we calculated the full observation error covariance matrix taking into account the inter-channel errors using the Desroziers diagnostic. The estimation of this R-matrix was performed in a 1D-var data assimilation experiment using the clear-sky cases. The estimation of the observation errors shows results consistent with what is expected for infra-red instruments. The diagnosed observation error value is between the instrumental noise and the standard deviations of the first-guess departures (synthetic observations - simulated observations).

The results obtained for the observation error correlation matrix show similar behaviour to the other hyperspectral sounders, but with lower values compared to the correlations obtained for IASI or CrIS. The largest observations error correlations (0.25 to 0.75 value) were found for CO_2 channels sensitive in the lower troposphere (720-740 cm $^{-}1$), the sensitive channels in the

atmospheric windows (742-1010 cm^{-}1 and 1070-1210 cm^{-}1), in the ozone absorption band (1010-1060 cm $^{-}$1). Another group of channels with large observation correlation errors but in a lesser extent is located in the beginning of the water vapour sensitive band (1600-1900 cm $^{-}$1). This estimation of **R** provides interesting results but it is important to take into account the underestimation of errors induced by the use of synthetic observations.

From this database, we have made a general selection of IRS channels that can be used for NWP models. The objective is to provide a first list of the most informative channels before the real arrival of IRS observations. This selection will also be useful in the OSSE framework for the regional AROME model, giving us a channel selection to be simulated and assimilated in the 3D-Var data assimilation system. The selection of channels was carried out over the full IRS spectrum using 60 profiles representative of the atmospheric variability in the IRS disk measurement area. A selection algorithm was implemented to select the most informative channels from the DFS (Degree of Freedom for Signal) method. This selection is carried out iteratively by selecting the most informative channel in temperature, humidity, ozone and skin temperature, then the combination of the first with the second most informative, and so on. We stopped the selection at 500 channels, which is generally the broadcast limit on the GTS at EUMETSAT.

The mean evolution of the DFS shows that the IRS channels are slightly more informative in humidity than in temperature. This is due to a broad band of sensitivity to water vapour in the IRS spectrum and also to a high sensitivity of some channels to water vapour in the atmospheric windows. Two other hypotheses may explain this behaviour; the non-observation of parts of the infrared spectrum that are highly sensitive to temperature compared to other hyperspectral sounders and the higher humidity background errors than temperature. The information content reaches more than 66 % of the mean total DFS available over 500 channels (T, q, O_3 and T_{skin}) with only 50 channels. Over the 60 profiles used for selection, 151 channels are always selected and 542 are never selected. The always selected channels consist of 73 channels that are mainly sensitive to temperature, 48 to ozone, 20 in atmospheric windows and 10 to humidity. The sensitivity analysis carried out at the beginning of the study shows that the channels in the atmospheric windows are also sensitive to temperature and humidity in the troposphere and some channels in the ozone absorption band are also sensitive to temperature and humidity in the troposphere and to temperature in the stratosphere. The use of an **R**-matrix taking into account inter-channel error correlations allows the extraction of all the information that these channels can provide.

We evaluated the information-providing capacity of several channel selections (75, 150, 300, 500, all channels) from 1D-Var data assimilation experiments on 1700 clear cases. The channel selections were evaluated by their ability to reduce the analysis error relative to the background error. The results show that channel selections mainly reduce the analysis error in the troposphere in temperature and humidity and more slightly in the upper stratosphere. It can be seen that the (75 and 150) and (300 and 500) channel selections provide close results in analysis error reduction to each other.

Our objective is to make the best compromise between quantity and quality of information. We have thus chosen the selection of 300 IRS channels which allows to reduce on mean the analysis error in temperature in the troposphere by 48 % compared to the background error, by 65 % for humidity in the mid-troposphere and by 17 % for ozone at the UTLS. This selection consists of 99 channels located in the CO_2 absorption band primarily used for atmospheric temperature sounding, 55 in the vapour absorption band primarily for atmospheric humidity sounding, 76 in the ozone absorption band that can provide information for temperature and humidity and 70 in the atmospheric windows for skin temperature, surfaces, temperature and humidity in the lower troposphere.

This first selection of channels has been made within a general framework for global NWP. We have taken into account general atmospheric information over the full IRS disk at surface up to 0.1 hPa. In order to be able to use this selection of 300 IRS channels for the AROME model, it is important to identify the sensitive channels in the upper stratosphere since these will not be usable in AROME having a model top at 10 hPa. An additional channel sensitivity study from RTTOV shows that about 280 IRS channels are sensitive above 10 hPa and 58 are present in our channel selection (red circles in Figure 16). These are channels located in the CO_2 and ozone absorption band as shown in Figure 26. Thus, 242 channels of our selection will be potentially usable in our 3D-Var AROME data assimilation system (black circles in Figure 16).

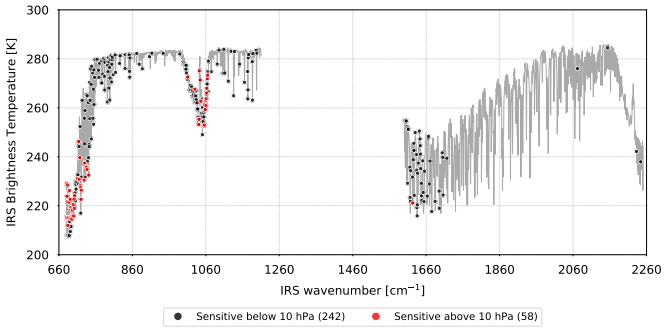

Figure 16. Representation of the sensitive IRS channels above 10 hPa (in red) among the 300 channels selected on a typical spectrum.

The IRS observations are expected to be distributed in the format of principal components (PCs). As we do not yet have the necessary information, we have carried out this study based on brightness temperatures (raw radiances). Data compression from PCA-based approaches has the advantage of drastically reducing data volume and reduces noise. However, the variations of atmosphere signals are more correlated across hyperspectral channels. This implies higher observation error correlations for reconstructed radiances. Hence the need for a specific study of the observation error diagnostics from these reconstructed radiances to take this behaviour into account. A study will be conducted on this subject.

A. List of the selection of the 300 new IRS channels

Channel number	Wave-number	Main sensitivity
0001	679.70	Temperature
0002	680.31	Temperature
0003	680.91	Temperature
0004	681.51	Temperature
0005	682.12	Temperature
0006	682.72	Temperature
0007	683.32	Temperature
0008	683.92	Temperature
0009	684.53	Temperature
0010	685.13	Temperature
0011	685.74	Temperature
0012	686.34	Temperature
0013	686.94	Temperature
0014	687.54	Temperature
0015	688.15	Temperature
0016	688.75	Temperature
0017	689.35	Temperature
0018	689.96	Temperature
0019	690.56	Temperature
0020	691.16	Temperature
0021	691.77	Temperature
0022	692.37	Temperature
0023	692.97	Temperature
0024	693.58	Temperature
0025	694.18	Temperature
0026	694.78	Temperature
0027	695.38	Temperature
0028	695.99	Temperature
0029	696.59	Temperature
0030	697.19	Temperature
0031	697.80	Temperature
0032	698.40	Temperature
0033	699.00	Temperature
0034	699.61	Temperature
0035	700.21	Temperature
0036	700.81	Temperature
0037	701.41	Temperature
0038	702.02	Temperature
0039	702.62	Temperature
0040	703.23	Temperature
0041	703.83	Temperature
0043	705.03	Temperature
0044	705.64	Temperature

Channel number	Wave-number	Main sensitivity
0045	706.24	Temperature
0046	706.84	Temperature
0047	707.45	Temperature
0049	708.65	Temperature
0052	710.46	Temperature
0054	711.67	Temperature
0056	712.87	Temperature
0057	713.48	Temperature
0058	714.08	Temperature
0059	714.68	Temperature
0060	715.29	Temperature
0063	717.10	Temperature
0064	717.70	Temperature
0065	718.30	Temperature
0066	718.91	Temperature
0067	719.51	Temperature
0068	720.11	Temperature
0069	720.72	Temperature
0070	721.32	Temperature
0073	723.13	Temperature
0082	728.55	Temperature
0083	729.16	Temperature
0084	729.76	Temperature
0085	730.37	Temperature
0086	730.97	Temperature
0087	731.57	Temperature
0090	733.38	Temperature
0091	733.98	Temperature
0092	734.59	Temperature
0095	736.40	Temperature
0096	737.00	Temperature
0097	737.60	Temperature
0101	740.01	Temperature
0102	740.62	Temperature
0103	741.22	Temperature
0104	741.82	Temperature
0105	742.43	Temperature
0106	743.03	Temperature
0108	744.24	Temperature
0109	744.84	Temperature
0111	746.04	Temperature
0114	747.86	Temperature
0115	748.46	Temperature

Channel number	Wave-number	Main sensitivity
0116	749.06	Temperature
0121	752.08	Temperature
0122	752.68	Temperature
0124	753.89	Temperature
0125	754.49	Temperature
0126	755.09	Temperature
0127	755.70	Temperature
0129	756.90	Temperature
0131	758.11	Temperature
0141	764.14	Temperature
0147	767.76	Temperature
0149	768.96	Temperature
0152	770.77	Temperature
0160	775.60	Window
0161	776.20	Window
0162	776.80	Window
0166	779.22	Window
0169	781.03	Window
0173	783.44	Window
0174	784.04	Window
0175	784.64	Window
0176	785.25	Window
0187	791.88	Window
0188	792.49	Window
0189	793.09	Window
0190	793.69	Window
0192	794.90	Window
0194	796.10	Window
0195	796.71	Window
0197	797.91	Window
0198	798.52	Window
0199	799.12	Window
0200	799.72	Window
0202	800.93	Window
0204	802.13	Window
0206	803.34	Window
0207	803.94	Window
0208	804.55	Window
0210	805.75	Window
0214	808.17	Window
0222	812.99	Window
0224	814.20	Window
0242	825.05	Window

Channel number	Wave-number	Main sensitivity
0245	826.86	Window
0267	840.13	Window
0268	840.73	Window
0281	848.57	Window
0284	850.38	Window
0286	851.59	Window
0287	852.19	Window
0288	852.80	Window
0289	853.40	Window
0320	872.10	Window
0331	878.73	Window
0345	887.17	Window
0377	906.47	Window
0388	913.11	Window
0439	943.87	Window
0514	989.10	Window
0529	998.14	Window
0531	999.35	Window
0533	1000.56	Ozone
0536	1002.37	Ozone
0538	1003.57	Ozone
0540	1004.78	Ozone
0542	1005.99	Ozone
0543	1006.59	Ozone
0544	1007.19	Ozone
0548	1009.60	Ozone
0549	1010.21	Ozone
0550	1010.81	Ozone
0556	1014.43	Ozone
0561	1017.44	Ozone
0562	1018.05	Ozone
0568	1021.67	Ozone
0569	1022.27	Ozone
0570	1022.87	Ozone
0572	1024.08	Ozone
0573	1024.68	Ozone
0574	1025.28	Ozone
0576	1026.49	Ozone
0577	1027.09	Ozone
0578	1027.70	Ozone
0579	1028.30	Ozone
0580	1028.90	Ozone
0581	1029.51	Ozone

Channel number	Wave-number	Main sensitivity
0582	1030.11	Ozone
0583	1030.71	Ozone
0584	1031.32	Ozone
0586	1032.52	Ozone
0587	1033.13	Ozone
0588	1033.73	Ozone
0589	1034.33	Ozone
0590	1034.93	Ozone
0592	1036.14	Ozone
0593	1036.74	Ozone
0594	1037.35	Ozone
0595	1037.95	Ozone
0596	1038.55	Ozone
0597	1039.16	Ozone
0598	1039.76	Ozone
0599	1040.36	Ozone
0600	1040.97	Ozone
0601	1041.57	Ozone
0602	1042.17	Ozone
0606	1044.58	Ozone
0609	1046.39	Ozone
0611	1047.60	Ozone
0614	1049.41	Ozone
0615	1050.01	Ozone
0616	1050.62	Ozone
0617	1051.22	Ozone
0618	1051.82	Ozone
0619	1052.42	Ozone
0620	1053.03	Ozone
0621	1053.63	Ozone
0623	1054.84	Ozone
0624	1055.44	Ozone
0625	1056.04	Ozone
0626	1056.65	Ozone
0627	1057.25	Ozone
0628	1057.85	Ozone
0629	1058.46	Ozone
0630	1059.06	Ozone
0631	1059.66	Ozone
0632	1060.26	Ozone
0633	1060.87	Ozone
0634	1061.47	Ozone
0635	1062.07	Ozone
0636	1062.68	Ozone
0637	1063.28	Ozone
0638	1063.88	Ozone
0639	1064.49	Ozone
0640	1065.09	Ozone
0641	1065.69	Ozone
0642	1066.30	Ozone
0643	1066.90	Ozone
0656	1074.74	Window
0683	1091.02	Window
0691	1095.85	Window
0709	1106.70	Window
0713	1109.12	Window
0717	1111.53	Window
0733	1121.18	Window
0746	1129.02	Window
0757	1135.65	Window
0761	1138.07	Window
0784	1151.94	Window
0806	1165.21	Window
0820	1173.65	Window
0821	1174.25	Window
0822	1174.86	Window
0823	1175.46	Window
0825	1176.66	Window
0841	1186.31	Window
0842	1186.92	Window
0844	1188.12	Window
0859	1197.17	Window
0862	1198.98	Window
0886	1602.18	Water-vapour
0890	1604.60	Water-vapour
0892	1605.81	Water-vapour
0896	1608.22	Water-vapour
0898	1609.43	Water-vapour
0900	1610.64	Water-vapour
0901	1611.24	Water-vapour
0908	1615.46	Water-vapour
0909	1616.07	Water-vapour
0911	1617.28	Water-vapour
0912	1617.88	Water-vapour
0920	1622.71	Water-vapour
0921	1623.31	Water-vapour

Channel number	Wave-number	Main sensitivity
0923	1624.52	Water-vapour
0924	1625.12	Water-vapour
0926	1626.33	Water-vapour
0927	1626.93	Water-vapour
0928	1627.54	Water-vapour
0932	1629.95	Water-vapour
0938	1633.58	Water-vapour
0939	1634.18	Water-vapour
0940	1634.78	Water-vapour
0941	1635.39	Water-vapour
0942	1635.99	Water-vapour
0943	1636.59	Water-vapour
0946	1638.40	Water-vapour
0951	1641.42	Water-vapour
0954	1643.23	Water-vapour
0955	1643.84	Water-vapour
0956	1644.44	Water-vapour
0957	1645.05	Water-vapour
0960	1646.86	Water-vapour
0961	1647.46	Water-vapour
0962	1648.06	Water-vapour
0966	1650.48	Water-vapour
0968	1651.69	Water-vapour
0972	1654.10	Water-vapour
0974	1655.31	Water-vapour
0987	1663.16	Water-vapour
0991	1665.57	Water-vapour
0992	1666.17	Water-vapour
0996	1668.59	Water-vapour
0999	1670.40	Water-vapour
1007	1675.23	Water-vapour
1022	1684.28	Water-vapour
1040	1695.15	Water-vapour
1041	1695.75	Water-vapour
1054	1703.60	Water-vapour
1057	1705.41	Water-vapour
1058	1706.02	Water-vapour
1075	1716.28	Water-vapour
1664	2071.85	Window
1800	2153.95	Window
1930	2232.43	Water-vapour
1949	2243.90	Water-vapour

Acknowledgements

Olivier Coopmann is funded by the EUMETSAT Research Fellowship programme. The authors thank Bertrand Théodore of EUMETSAT for providing me with the geometric specification of the IRS observations and Jérôme Vidot of CEMS (Centre d'Études en Météorologie Satellitaire) for the rapid availability of a new set of IRS coefficients on the new spectral resolution.

References

Borbas, E. E., Hulley, G., Feltz, M., Knuteson, R., and Hook, S. (2018). The combined ASTER MODIS emissivity over land (CAMEL) part 1: Methodology and high spectral resolution application. *Remote Sensing*, 10(4):643.

Bormann, N., Bonavita, M., Dragani, R., Eresmaa, R., Matricardi, M., and McNally, A. (2016). Enhancing the impact of iasi observations through an updated observation-error covariance matrix. *Quarterly Journal of the Royal Meteorological Society*, 142(697):1767–1780.

Collard, A. (2007). Selection of IASI channels for use in numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 133(629):1977–1991.

Coopmann, O., Guidard, V., Fourrié, N., and Josse, B. (2020a). Use of variable ozone in a radiative transfer model for the global météo-france 4d-var system. *Quarterly Journal of the Royal Meteorological Society*, 146(733):3729–3746.

Coopmann, O., Guidard, V., Fourrié, N., Josse, B., and Marécal, V. (2020b). Update of infrared atmospheric sounding interferometer (IASI) channel selection with correlated observation errors for numerical weather prediction (NWP). *Atmospheric Measurement Techniques*, 13(5):2659–2680.

Coopmann, O., Guidard, V., Fourrié, N., and Plu, M. (2018). Assimilation of iasi ozone-sensitive channels in preparation for an enhanced coupling between numerical weather prediction and chemistry transport models. *Journal of Geophysical Research: Atmospheres*, 123(21):12–452.

Derber, J. C. and Wu, W.-S. (1998). The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Monthly Weather Review*, 126(8):2287–2299.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P. (2005). Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3385–3396.

Dragani, R., Benedetti, A., Flemming, J., Balsamo, G., Diamantakis, M., Geer, A. J., Hogan, R., Stockdale, T., Ades, M., Agusti-Panareda, A., Barré, J., Bechtold, P., Bozzo, A., Hersbach, H., Hólm, E., Kipling, Z., Inness, A., Letertre-Danczak, J., Massart, S., Matricardi, M., McNally, T., Parrington, M., Sandu, I., Soci, C., and Vitart, F. (2018). Atmospheric composition priority developments for numerical weather prediction. (833).

Eresmaa, R. (2020). Ecmwf aerosol and cloud detection software user guide. volume NWPSAF-EC-UD-015. Numerical Weather Prediction Satellite Application Facilities.

Fourrié, N. and Rabier, F. (2004). Cloud characteristics and channel selection for IASI radiances in meteorologically sensitive areas. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 130(600):1839–1856.

Fourrié, N. and Thépaut, J.-n. (2003). Evaluation of the AIRS near-real-time channel selection for application to numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 129(592):2425–2439.

Gambacorta, A. and Barnet, C. D. (2013). Methodology and information content of the NOAA NESDIS operational channel selection for the cross-track infrared sounder (CrIS). *IEEE Transactions on Geoscience and Remote Sensing*, 51(6):3207–3216.

Havemann, S. and Smith, F. (2020). *NWPSAF 1D-Var User Manual, Software Version 1.2*. Met Office, Exeter, UK.

Hess, M., Koepke, P., and Schult, I. (1998). Optical properties of aerosols and clouds: The software package OPAC. *Bulletin of the American meteorological society*, 79(5):831–844.

Hollingsworth, A. and Lönnberg, P. (1986). The statistical structure of short-range forecast errors as determined from radiosonde data. part i: The wind field. *Tellus A*, 38(2):111–136.

Ivanova, I., de Grandpré, J., Rochon, Y. J., and Sitwell, M. (2017). Impact of ozone radiative feedbacks on global weather forecasting. In *AGU Fall Meeting Abstracts*, volume 2017, pages A33A–2332.

John, V. and Buehler, S. (2004). The impact of ozone lines on AMSU-B radiances. *Geophysical research letters*, 31(21).

Lahoz, W. A., Errera, Q., Swinbank, R., and Fonteyn, D. (2007). Data assimilation of stratospheric constituents: a review. *Atmospheric Chemistry and Physics*, 7(22):5745–5773.

Masutani, M., Schlatter, T., Errico, R., Stoffelen, A., Andersson, E., Lahoz, W., Woollen, J., Emmitt, G., Riishojgaard, L. P., and Lord, S. (2010). Observing system simulation experiments. *Data Assimilation: Making Sense of Observations*, pages 647–679.

Matricardi, M. (2008). *The generation of RTTOV regression coefficients for IASI and AIRS using a new profile training set and a new line-by-line database*. European Centre for Medium-Range Weather Forecasts.

McNally, A. (2002). A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 128(585):2551–2556.

McNally, A. and Watts, P. (2003). A cloud detection algorithm for high-spectral-resolution infrared sounders. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 129(595):3411–3423.

Parrish, D. F. and Derber, J. C. (1992). The national meteorological center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8):1747–1763.

Privé, N., Errico, R., and Tai, K.-S. (2013). The influence of observation errors on analysis error and forecast skill investigated with an observing system simulation experiment. *Journal of Geophysical Research: Atmospheres*, 118(11):5332–5346.

Rodgers, C. D. (1996). Information content and optimization of high-spectral-resolution measurements. In *Optical spectroscopic techniques and instrumentation for atmospheric and space research II*, volume 2830, pages 136–147. International Society for Optics and Photonics.

Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: theory and practice*, volume 2. World scientific.

Saunders, R., Hocking, J., Rundle, D., Rayer, P., Havemann, S., Matricardi, M., Geer, A., Lupu, C., Brunel, P., and Vidot, J. (2017). RTTOV v12 science and validation report, 78 pp. Numerical Weather Prediction Satellite Application Facilities.

Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., Vidot, J., Roquet, P., Matricardi, M., Geer, A., et al. (2018). An update on the RTTOV fast radiative transfer model (currently at version 12). *Geoscientific Model Development*, 11(7).

Stewart, L., Dance, S. L., Nichols, N. K., Eyre, J., and Cameron, J. (2014). Estimating interchannel observation-error correlations for iasi radiance data in the met office system. *Quarterly Journal of the Royal Meteorological Society*, 140(681):1236–1244.

Vidot, J. and Brunel, P. (2018). Comparison of the RTTOV-12 ice cloud models for hyperspectral IR instruments using the A-Train. In *Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques and Applications VII*, volume 10780, page 1078004. International Society for Optics and Photonics.

Vittorioso, F., Guidard, V., and Fourrié, N. (2021). An infrared atmospheric sounding interferometer–new generation (iasi-ng) channel selection for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 147(739):3297–3317.

Weston, P., Bell, W., and Eyre, J. (2014). Accounting for correlated error in the assimilation of high-resolution sounder data. *Quarterly Journal of the Royal Meteorological Society*, 140(685):2420–2429.

Yin, R., Han, W., Gao, Z., and Di, D. (2020). The evaluation of fy4a's geostationary interferometric infrared sounder (giirs) long-wave temperature sounding channels using the grapes global 4d-var. *Quarterly Journal of the Royal Meteorological Society*, 146(728):1459–1476.