HAL
open science

Bottleneck Channels Algorithm for Satellite Data Dimension Reduction: A Case Study for IASI

Victor Pellet, Filipe Aires

To cite this version:

Victor Pellet, Filipe Aires. Bottleneck Channels Algorithm for Satellite Data Dimension Reduction: A Case Study for IASI. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56 (10), pp.6069-6081. 10.1109/TGRS.2018.2830123 . hal-02376956

HAL Id: hal-02376956

https://hal.science/hal-02376956v1

Submitted on 22 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

"Bottheneck Channels" algorithm for satellite data dimension reduction: A case study for IASI

Victor Pellet[1] and Filipe Aires[1,2,3]

Abstract. The Infrared Atmospheric Sounder Interferometer (IASI) onboard the Metop-A and -B satellites is an essential instrument for numerical weather prediction centres. IASI observes surface and atmospheric geophysical variables twice a day around the world with its 8461 channels. This data volume will even increase in the future, stressing the need for practical and efficient dimension reduction techniques for storage, transmission, but also for inversion or assimilation. In a previous paper [*Aires et al.*, 2015a], a new dimension reduction method, the so-called Bottleneck-Channel (BC), has been introduced. BCs are a good compromise between compression and selection techniques. In this paper, we introduce several new technical configurations of the BC approach, comparing their advantages/drawbacks. Comparisons are also made with two classical methodologies widely used in the satellite community: Principal Component Analysis (PCA) compression and Entropy Reduction (ER) channel selection. The methodologies are tested on the temperature, water vapor and ozone retrieval application using IASI measurements. It is shown that the BCs present compression, de-noising and information content results analogue to the PCA, while preserving the physical meaning (i.e., actual channel TBs) of a channel selection approach. This has numerous advantages for inversion and assimilation: the mixing problem, contamination effects, minor constituants representation, or radiative transfer speed will be discussed.

1. Introduction

In order to improve weather forecasts, remote sensing technology has no ceased to evolve and improve. In particular, the number of satellite observations is constantly increasing because more satellite missions are launched, and because instruments have an increased time-, space- and/or spectral-resolution. For instance, infrared instruments are now hyper-spectral, while previous generation operated with up to 15 detectors, the Infrared Atmospheric Sounding Interferometer (IASI) launched in 2006 carries 8461 channels [*Hilton et al.*, 2012]. Microwave missions are also considered by space agencies [*Mahfouf et al.*, 2014; *Aires et al.*, 2015b; *Lipton*, 2003; *Pellet et al.*, 2015]. This increasing number of observations have improved weather forecast [*Collard et al.*, 2010] but this comes of course with practical difficulties. The data storage and transmission can become difficult for the current instruments: IASI or Atmospheric InfraRed Sounder (AIRS); and for the next generation: InfraRed Sounder (IRS) onboard Meteosat Third Generation (MTG) or IASI-Next Generation onboard MetOp-SG. Most importantly, retrieval and assimilation algorithms can hardly handle such amount of information.

The principle of dimension reduction of the satellite observations is to deliver to the inversion scheme (physical or statistical retrieval) the most relevant but compact information from the original observations. One way to reduce dimension is by compression by using the entire observations and extract its more pertinent characteristics. The classical Principal Component Analysis (PCA) and its derivatives is still the most widely used dimension reduction technique thanks to its speed, simplicity and performance [*Aires et al.*, 2002a; *Huang and Antonelli*, 2001; *Atkinson et al.*, 2010]. Despite their lack in physical meaning, PCA components are now used in assimilation scheme [*Collard and McNally*, 2009; *Matricardi and McNally*, 2014] and dedicated radiative transfert codes have been developed in the component space [*Matricardi*, 2010]). Another way to reduce the dimension is by channel selection, selecting only a part of the observation as inputs for retrieval algorithm. Channel selection is very useful when focusing on trace gases [*George et al.*, 2009; *Crevoisier et al.*, 2003; *Pommier et al.*, 2010] dealing with true channel (easily sampled) but induces information and redundancy loss. [*Collard*, 2007] shows that, typically, about 300 channels are needed to get a satisfactory information content from the observations.

A new alternative to channel selection or compression has been introduced: the so-called Bottleneck Channel (BC) [*Aires et al.*, 2015a; *Pellet et al.*, 2015]. BC is a good compromise between the two classical approaches: as for PCA, the whole observation is used to exploit the redundancy among the channels, but BCs are also pseudo-channel that keep a physical meaning. In this paper, some BC-based algorithm configurations are presented and tested on infrared high-spectral resolution observations from IASI for the retrieval of temperature, humidity and ozone atmospheric profiles.

The experiment tools are presented in section 2. Classic methods of dimension reduction then BC approach are presented in section 3. Compression and channels selection results are analyzed in section 4. Section 5 is a discussion on the practical advantages of BCs. Finally, section 6 provides conclusion and some perspectives.

2. IASI instrument and necessary tools

2.1. The IASI instrument

The IASI sounder is an infrared Fourier transform spectrometer based on a Michelson interferometer. It is part of the Metop-A (launched in 2006) and Metop-B (2012) missions [*Cayla et al.*, 1995]. The main IASI purpose is to support the Numerical Weather Prediction (NWP) centers [**?**].

[1]LERMA, Observatoire de Paris, France
[2]Estellus, Paris, France
[3]Columbia University, USA

Its spectral range goes from 645 to 2760 cm^{-1} with a bandwidth of 0.25 cm^{-1}. The IASI spectrum is divided into three main bands: B1 from 645 to 1210 cm^{-1}, B2 from 1210 to 2000 cm^{-1}, and B3 from 2000 to 2760 cm^{-1}. Tab. 1 indicates the geophysical variables measured by IASI. IASI is able to retrieve temperature and water vapor profiles in the troposphere and the lower stratosphere [*Huang and Antonelli*, 2001; *Aires et al.*, 2002b], as well as measure concentrations of ozone, carbon monoxide, methane and other compounds [*Clerbaux et al.*, 2009; *Turquety et al.*, 2004; *Scannell et al.*, 2012].

Table 1. Overview of the geophysical variables measured by IASI

Wavenumber (cm^{-1})	Geophysical variable
650-770	temperature (T)
770-980	cloud and surface prop.
1000-1070	ozone (O^3)
1080-1150	cloud and surface prop.
1210-1650	q, T, N_2O, CH_4, SO_2
2100-2150	total CO
2150-2250	T and total N_2O
2350-2420	T
2420-2700	cloud and surface prop.
2700-2760	total CH_4

Fig. 1(Upper) shows, for a given atmospheric situation (TIGR number 500), the Brightness Temperatures (TB) that should be measured by the IASI instrument. At a given wavenumber, the TB results from the emission and absorption of the different atmospheric components along the atmosphere at this wave number. The hotter the TB is, the less the signal has been absorbed through its path in the atmosphere (no absorption band of atmospheric components at this wave number) and the deeper in the atmosphere the IASI channel probes. For instance, the hottest TBs (280 K between 770-980, 1080-1150 and 2420-2700 cm^{-1}) correspond to the window channels, characterizing the surface properties. On the other hand, the spectral sampling of an absorption band (CO_2 or H_2O for instance, see Tab. 1) provides information on these gaz concentrations.

Figure 1. Simulated IASI TB-spectrum and corresponding instrumental noise over bands B1, B2 and B3

The IASI instrument noise has been modeled by CNES using a centered Gaussian noise distribution, characterized by its standard deviation $std(\nu, TB)$ with respect to ν, the wavenumber and the measured brightness temperature TB:

$$std(\nu, TB) = \frac{\frac{\partial B(TB=280,\nu)}{\partial TB}}{\frac{\partial B(TB=T,\nu)}{\partial TB}} * std_{280}, \quad (1)$$

where $B(TB, \nu)$ is the Planck function, and the std_{280} is the Noise Equivalent Temperature Difference (NEDT) at 280 K, which characterizes the IASI instrument. Fig. 1(Bottom) represents the standard deviation of the instrumental noise for a given atmospheric situation. Note the high level of noise in band B3 and, as a consequence, this band is generally not used for retrieval. We will not use B3 in this study either.

2.2. The TIGR database of atmospheric profile

The Thermodynamic Initial Guess Retrieval (TIGR) data set is a climatological library from the "Laboratoire de Météorologie Dynamique" designed to perform retrieval developments under clear-sky [*Chevallier et al.*, 2006]. It includes more than 2300 radiosonde profiles from all over the world. Fig. 2 shows TIGR spatial distribution over the globe. The profiles are selected to represent the maximum variability of atmospheric temperature and humidity conditions under clear-sky conditions. This means that, besides standard atmospheric profiles, TIGR focusses on extreme situations are also included to cover a wide range of atmospheric conditions, from the tropics to the polar regions. Each situation is characterized from the surface to the top of the atmosphere by its temperature (T), water vapor (q) and ozone (O_3) profiles, in a vertical grid of 101 pressure levels.

Figure 2. Spatial distribution of TIGR situations, the pixel color represents the Skin Temperature (T_s) in K

2.3. The Optimal Spectral Sampling (OSS) radiative transfer code

The OSS Radiative Transfer (RT) code is developed at Atmospheric and Environment Research (AER). In the infrared, OSS has been compared and validated with other RT models [*Moncet et al.*, 2010; *Lipton et al.*, 2009]. OSS is used here to calculate the IASI radiances for each atmospheric situation, from 645 cm^{-1} to 2760 cm^{-1} and with a spectral resolution of 0.25 cm^{-1}, under clear-sky conditions.

The OSS code has also been used here to estimate the RT Jacobians. The Jacobians provide, for a given situation, the channels sensitivity to a particular geophysical variable along the atmosphere. T Jacobians are expressed in K/K, and q and O_3 Jacobians in K/Kg/Kg. Fig. 3 represents the IASI Jacobians for T, q and O_3 for TIGR situation 500 (same as in Fig. 1). All IASI channels appear to be sensitive to T at various levels of the atmosphere: around the surface for the channels from 770 to 980 cm^{-1}, in the middle atmosphere from 1210 to 1650 cm^{-1}, and the top of the atmosphere from 650 to 770 cm^{-1} and from 2350 to 2420 cm^{-1}. The most sensitive channels to T are also the most sensitive ones to q. This is due to the dependency of the water vapor emission to temperature at different vertical layers, thus producing a high correlation between temperature and water vapor. Finally, fewer than four hundred channels (from 1000 to 1100 cm^{-1}) are sensitive to the O_3 concentration in the high atmosphere where O_3 is present (see Tab. 1).

Figure 3. From top to bottom: IASI Jacobians in temperature (T), humidity (q), and ozone (O_3), for TIGR situation 500

2.4. *A priori* error covariance matrix B

Most inversion methods use an *a priori* information on the state of the atmosphere in addition to the satellite observations to perform the retrieval [*Rodgers*, 1976, 1990]. It is therefore necessary to consider this type of information in the channel selection technique as it can have a strong impact on the most important information to extract in the observed spectra.

The quality of the *a priori* information is often characterized by the matrix **B** representing the covariance matrix of the *a priori* errors. This matrix **B** varies in space and time

in order to account for its complex state-dependency. In the following experiments, we have derived the Gaussian *a priori* error covariance matrix in a simple manner to facilitate the interpretation of the results. Covariance matrices $\mathbf{B}$ is defined as: $Cov(x, y) = Cor(x, y) \cdot \sigma(x) \cdot \sigma(y)$ where x and y represent two geophysical variables to be retrieved, σ the standard deviation. It is therefore necessary to decide on a correlation matrix and standard deviation for T, q and O_3 variables:

- First, it is often assumed that humidity (respectively ozone) error correlation (between the vertical layers) is close to the actual associated humidity (resp. ozone) correlation. It was chosen here to have the covariance matrices $\mathbf{B}$ for q and O_3 based on: (1) the "raw" humidity and ozone correlation matrices, and (2) an error variance of 20% of q (resp. 30% of O_3) on each layer.
- Temperature is highly correlated in the bottom of the atmosphere so using the T correlation matrix as the *a priori* error correlation matrix would mean that only one piece of information would be necessary to retrieve the lower part of the T profile. Obtaining a channel selection in these conditions would mean that no profile inversion could be measured in the lower part of the atmosphere only, and a limited number of channels would be selected in the tropopause. In order to avoid this difficulty, we decided to use an independent *a priori* errors along the vertical layers. In this way, the channel selection will try to obtain channels that are more uniformly distributed along the vertical. An error variance of 1 K is generally admitted and is chosen here.

3. Dimension reduction methods

In this section, two classical methodologies of dimension reduction are first presented, the Noise Adjusted Principal Component Analysis (NAPCA) compression and the ER channel selection. Both are widely used among the satellite remote sensing community and their principle(s) and advantage(s) are briefly highlighted here. Furthermore, a more recently developed methodology is also introduced-the Bottleneck Channels (BCs) [*Aires et al.*, 2015a; *Pellet et al.*, 2015].

3.1. Noise-Adjusted Principal Component Analysis (NAPAC)

Compression techniques are used to represent observation vector $\mathbf{TB}$ (i.e. IASI spectrum) of dimension m with a limited number of components r. The objective is to compress $\mathbf{TB}$, regardless of the potential use of $\mathbf{TB}$ for the retrieval of some geophysical variables. The observation vector is corrupted by an observation error linked to the instrumental noise: $\widetilde{\mathbf{TB}} = \mathbf{TB} + \epsilon$. The compression is based on an algebraic transformation of the original data, in which all the channels are mixed. In general, these transformations are linear, so the extracted components are a weighted sum of the original channels. It is therefore possible to reconstruct the full observation vector from these components.

The Noise-Adjusted PCA is an extension of the Principal Component analysis (PCA) approach [*Blackwell*, 2005; *Fan et al.*, 2014]. It is a technique widely used for compression/de-noising data (i.e., feature extraction). The NAPCA transformation is a linear, orthogonal operator $\mathbf{G_r}$ that projects a noisy m-dimensional observation vector $\widetilde{\mathbf{TB}} = \mathbf{TB} + \epsilon$ into an r-dimension subspace (where $r < m$), i.e. the Principal Components (PCs) space (a new set of variables). Assuming that ϵ is uncorrelated to $\mathbf{TB}$ and characterized by the noise covariance matrix $\mathbf{C}_{\epsilon\epsilon}$, the PCs of $\widetilde{\mathbf{TB}}$ from the NAPCA are defined as $\widetilde{\mathbf{P}} = \mathbf{G_r}^t \cdot \widetilde{\mathbf{TB}}$. The NAPCA transformation allows performing the noise-free reconstruction characterized by the cost function:

$$\mathbf{c}(\cdot) = E[(\widehat{\mathbf{TB}} - \mathbf{TB})^t(\widehat{\mathbf{TB}} - \mathbf{TB})], \quad (2)$$

where an estimation of $\mathbf{TB}$, $\widehat{\mathbf{TB}} = \mathbf{G_r} \cdot \mathbf{G_r}^t \cdot \widetilde{\mathbf{TB}}$ is the compression-decompression based on the r first PCs. $\mathbf{G_r}$ is a linear operator: $\mathbf{G_r} = \mathbf{C}_{\epsilon\epsilon}{}^{1/2} \cdot \mathbf{W_r} \cdot \mathbf{W_r}^t \cdot \mathbf{C}_{\epsilon\epsilon}{}^{-1/2}$. $\mathbf{W_r}$ contains the r first eigenvectors of the whitened noisy covariance matrix $\mathbf{C}_{WW} = \mathbf{C}_{\epsilon\epsilon}{}^{1/2} \cdot \mathbf{C}_{\widetilde{TB}\widetilde{TB}} \cdot \mathbf{C}_{\epsilon\epsilon}^{-1/2}$. Note that using NAPCA requires the *a priori* noise statistics. For detail, see [*Blackwell*, 2005; *Aires et al.*, 2015a].

3.2. Entropy Reduction (ER) channel selection

Channel selection is the process of selecting a subset of relevant channels optimized for a particular retrieval. The main assumption when using a feature selection technique is that the data contains many redundant (providing no more additional information after some channels are already selected) or irrelevant (no useful information for the considered retrieval) features. There are many channel selection techniques based on different algorithms, different criteria and different assumptions. Here we chose the iterative method [*Rodgers*, 1996; *Rabier et al.*, 2002] in which the channels are selected based on their capability to reduce the uncertainty of retrieving a geophysical variable. The ER uses the probabilities of the possible states and reaches a maximum when all the states have an equal probability:

$$\mathbf{ER} = \frac{1}{2} log_2 \left(\frac{\mathbf{B}}{\mathbf{A}}\right), \quad (3)$$

where $\mathbf{B}$ is the *a priori* covariance matrix for the geophysical state, and $\mathbf{A}$ is the theoretical retrieval covariance matrix. Eq. (3) is a classical retrieval accuracy metrics.

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^t \cdot \mathbf{R}^{-1} \cdot \mathbf{H}, \quad (4)$$

where $\mathbf{H}$ is the RT Jacobians, and $\mathbf{R}$ is the instrumental error covariance plus the contamination effect covariance [*Aires et al.*, 2015a]. This quantity is an estimation of the retrieval errors.

Figure 4. From top to bottom: wave number of the selected channels (in red) for the ER when selecting 20 channels for the Skin Temperature (T_s) retrieval and 100 channels for T retrieval, ER-constrained using 20 channels for T_s retrieval and 100 channels for T retrieval, ER-constrained 100 channels for q and ER-constrained 100 channels for O_3.

ER-constrained algorithm - This section highlights some physical issues to take into account. The ER-selected channel frequencies are represented for each variable to retrieve in Fig. 4: from top to bottom for ER-T, ER-q, ER-O_3. It can been seen that for an un-constrained selection (i.e. in which any channel of the band B1 and B2 can be selected), ER-T takes most of the channel in band B2 where channels are sensitive to q. This behavior results from the fact that the measure of the H_2O absorption is dependent on the temperature. Therefore, temperature information is also present in the q channels (in particular in the tropopause where humidity concentration is high). This explains why ER-T takes only a few channels sensitive only to T in the lower atmosphere. This has been observed since channel selection techniques have been used. In order to get a "pure" temperature selection, it is possible to constrain the ER selection into B1 only for T (i.e. ER-T) and only in B2 for

q (ER-q) to avoid this "crossing" problem. This is common practice in Numerical Weather Prediction centres.

Another issue is the impact of Skin Temperature (T_s) in the ER-T results: (1) the high correlation between T_s and the T profile in the tropopause (even if the *a priori* error covariance has been set to be uncorrelated, the Jacobians fully represent this correlation), and (2) the low instrument noise on the window channel. Therefore, the channel selection for T when including T_s will bias the selection towards the window channels. In order to avoid this difficulty, two independent selections of respectively 20 channels for T_s and 100 remaining channels for T profile are requested. The constrained and unconstrained selections are represented in Fig. 4. ER-q results are close to ER-q constrained (same for O_3) so they are not represented in this figure. It can be noted that the selection constrain is able to solve the two crossing difficulties (T with q and T with T_s).

3.3. Bottleneck Channels

This new approach has been developed in [*Aires et al.*, 2015a] and tested in the microwave domain in [*Pellet et al.*, 2015]. BC technique is a compromise between the classical compression and channel selection methods. Some experiments are presented in this section to justify the technical choices made for the results of section 4.

General principle - The goal is to find a limited number of "pseudo-channels" $\mathbf{s}$ representing, as well as possible, the ensemble of observations in the observed spectrum $\mathbf{TB}$. A linear projection $\mathbf{P}$ is used to link the noisy observations $\widetilde{\mathbf{TB}}$ to noise-free channels $\mathbf{TB}$ ($\tilde{\eta}_s = \mathbf{P} \cdot \widetilde{\mathbf{TB}}$), from which it is possible to reconstruct the whole spectrum. The selection of pseudo-channels $\mathbf{s}$ is chosen so that the reconstruction of the full spectrum $\mathbf{TB}$ from $\tilde{\eta}_s$ through a linear regression $\mathbf{TB} = \mathbf{Q} \cdot \tilde{\boldsymbol{\eta}}_s$ is as good as possible. For that purpose, the selected channels in $\boldsymbol{s}$ are iteratively chosen to minimise the least-square reconstruction errors over the whole dataset $\mathbf{D}$:

$$\sum\nolimits_{e \in \mathbf{D}} ||\mathbf{TB}^e - \mathbf{Q} \cdot \mathbf{P} \cdot \widetilde{\mathbf{TB}}^e|| \tag{5}$$

The BCs $\tilde{\boldsymbol{\eta}}_s$ are almost a channel selection, they can be interpreted as noise-free channels to some extent, but they result from a linear combination of the noisy observation vector $\widetilde{\mathbf{TB}}$ (as the PCA components). This allows the BCs to better exploit the redundancy in $\widetilde{\mathbf{TB}}$, and then to better reduce noise in observations than a strict channel selection. The BC method has some advantages. In the following, some BC-based algorithmic choices are presented.

Localized versus global BCs - It was mentioned previously that the BCs are a linear projection of the full observed spectrum: $\tilde{\boldsymbol{\eta}}_s = \mathbf{P} \cdot \widetilde{\mathbf{TB}}$. This means that a selected frequency s_i can result from frequency channels quite far in the IASI spectrum. For instance, a selected channel on T in the 660 cm^{-1} can be estimated from measurements in this band, but also from channels in the 2300 cm^{-1} region. This can be physically dangerous but it is possible to constrain the BCs s_i to only result from frequencies included in their surroundings. This version of BCs is referred here to the localized BCs. In addition to their additional frequency constraints, localized BCs are faster to estimate because the number of model inputs is limited in regression $\mathbf{P}$.

Nonlinear BCs - It was seen earlier that the BCs are estimated using a linear projection $\mathbf{P}$, and the full spectrum is then reconstructed using another linear projection $\mathbf{Q}$. It is possible to use a nonlinear regression instead of them, for instance by using a neural network model [*Aires et al.*, 2015a]. Indeed, some subtle features of the spectrum could be better reproduced by using a nonlinear model. In general, when a limited number of BCs is extracted, using a nonlinear model is beneficial, but when many BCs are extracted, the multiplicity of information overcomes the non-linear edge that disappears [*Aires et al.*, 2015a]. Since nonlinear models are computationally more expensive to obtain, they are limited to the localized BC version that limits the number of model inputs.

Constrained BCs - When performing a channel selection to retrieve multiple geophysical variables, some difficulties can appear. For instance, channels can be sensitive to multiple variables (e.g. T and q in B2) and these interactions are difficult to take into account [*Aires*, 2011]. In order to avoid such difficulties and also to speed up the channel selection process, the satellite remote sensing community generally selects channels for each geophysical variable one at a time. For example, T channels are selected only in the 645-800 cm^{-1} band. Then, q channels are selected in the 1200-1600 cm^{-1} region, and ozone in the 1000-1100 cm^{-1} band. This strategy can also benefit the BCs. Constrained-BCs use the same TB-related criteria (minimizing the mean squared error of reconstruction of the whole spectrum in Eq. (5)) than generic BCs, but the space of candidates to select $\boldsymbol{\eta}$ is reduced and focused around specific absorption bands for the various atmospheric variables of interest.

Fig. 5 represents the BC reconstruction errors (noise-free TB minus reconstructed spectrum from noisy observations after compression) for various configurations: the Localised-Linear, Localised NonLinear, and the global BC approaches. The localized linear BCs have several numbers of surrounding channels, it is compared to the linear and nonlinear global BCs. It can been seen that the Local Linear BC method presents a higher reconstruction error than the BC-global, even when using 1000 neighbor channels.

In this study, the more subtle compound we are focused on is the ozone. The ozone presents an absorption band around 1000-1150 cm^{-1}, meaning that IASI use about 750 channel to sample the ozone absorption band (with its accuracy of 0.25 cm^{-1}). Here, we chose to consider 1000 channels the limit of neighborhoods because for over 1000 inputs, a BC related to the ozone band would be constructed, by the local linear approach, from unrelated-ozone channels and then the meaning of local consideration would be lost.

We concluded that the linear locality, even if it is more physical in terms of channel, introduces a significant additional error that impacts the compression capacity of the BC method. The Local Non-Linear BC method presents a lower reconstruction error than the classic BC-global method and should be able to get a better compression rate whilst still being more realistic in terms of channel. Therefore, this method is very cost computing and has not been tested.

comp-premiervolet-bc-eps-converted-to.p

Figure 5. Evolution of the projection errors, $E[(\boldsymbol{\eta_s} - \mathbf{TB})^2]$, over a sample of 30 representative channels, with an increasing number of neighbor channels for the Localized Linear and Local Non-Linear BC approaches, compared to the projection error of the standard Global BCs.

4. Compression and retrieval results

4.1. Compression results

In this section BC refers to the standard global BC configuration.

In Fig. 6, the de-noising errors (i.e. noise-free spectrum $\mathbf{TB}$ minus the compressed-reconstructed noisy observed spectrum $\widehat{\mathbf{TB}}$) are represented for an increasing number of PCs or BCs, for the NAPCA and BC methodologies.

This de-noising diagnostic measures the ability of the data-reduction methodology to both compress the observations, and at the same time, to suppress part of the instrument noise. For NAPCA, the de-noising error decreases with increasing components (PCs) until an optimum is reached (about 47) and the error starts to increase. The decrease is due to a better compression when adding components, and the following increase results from the use of the higher PCs to represent the instrumental noise. Asymptotically, the de-noising error converges to about 0.45 K, the averaged instrument noise (i.e. perfect reconstruction of the noisy spectra). The optimum with 47 components represents the best compromise between the reduction of the compression error, requiring a large number of components, and a limitation of the number of components to avoid representing the instrument noise.

Figure 6. De-noising error (i.e. noise-free observation **TB** minus the reconstructed $\widehat{\mathbf{TB}}$ from noisy observation) with an increasing number of NAPCA components or BC channels. Instrumental noise averaged over each channel is represented in dashed black line for comparison purpose.

The BC approach is slightly less performant than NAPCA when a limited number of BC is used. This is to be expected: NAPCA is a very performant compression tool. Like NAPCA, BCs are a linear transformation of the observations aimed at compressing the signal, but the PCs have no constraint (except to represent as much signal as possible in each component), where the BCs are constrained to represent physical channels. This additional constrain slightly reduces their compression ability. However, the BC approach has a constantly decreasing de-noising error which converges towards the NAPCA optimum. With about 100 BCs, the de-noising error is equivalent to the optimal use of 47 PCs throughout the spectrum. Difference between spectral de-noising errors with NAPCA (47 components) and BC (100 BCs) is lower than $5{\cdot}10^{-3}$ along the spectrum.

For both methods and for each individual channel, it is possible to select a number of components/BCs for which the de-noising error is lower than the corresponding instrument noise. Fig. 7 represents the de-noising errors for the band B1, for three different numbers of PCs and BCs. For NAPCA, after the optimum number of components is reached, the de-noising error tends to increase, as commented earlier. For BC, as expected, the de-noising error decreases monotically with the number of channels. The same conclusions stand for the band B2. Note, however, that the optimal number of PCs/BCs over the whole band is not necessarily the optimum solution for each individual channel. Some particular spectral region of interest may be neglected in a statistical point of view compared to other predominant regions. Therefore even if 47 components seem to be the right compromise for compression/de-noising of the whole spectrum, it might be useful to use a lower or higher number of components for particular channels. This is especially true if the user is interested in localized particular channels. This could be the case, for instance, for trace gas retrievals.

Figure 7. Spectral de-noising error (i.e. noise-free observation **TB** minus the reconstructed $\widehat{TB}$) for the band B1 for NAPCA (top) and BC (bottom), for three different numbers of PCs or BCs: 20, 47 and 100.

4.2. Channel selection results

In this section, ER method refers to the ER-compositing approach where an ER-selection was made independently for temperature, humidity and ozone, and then gathered into a unique channel selection. BC is for the global linear solution.

4.2.1. Comparison of the channel selections

The ER objective is to select channels optimizing the retrieval of a given variable, whereas the BC method (as the NAPCA) intends to select channels from which it is possible to reproduce the full spectrum (i.e. a compression purpose). These different strategies impact the selection results.

The selected frequencies are presented in Fig. 8 for ER-compositing, BC and BC-constrained. For the ER-compositing method, by definition, the selection is located around the absorption bands of the variables to retrieve: CO_2 band for temperature, H_20 band for humidity, and O_3 band for ozone. For T, many channels are chosen in the 670-710 cm^{-1} region sensitive to the upper troposphere and lower stratosphere. This is partly a consequence of the difficulty to retrieve temperature in these layers (see following section 4.3) and the relatively higher instrumental noise levels for these channels compared to lower-level ones (see Fig. 1). The humidity selection is located in the H_2O absorption band centered at about 1600 cm^{-1} but channels are more gathered in the 1200-1600 cm^{-1} region that has a lower instrument noise level. The humidity information content is richer vertically than the ozone one and requires a large distribution along the absorption band to sample from the bottom to the top of the atmosphere. The ozone selection is well located in the O_3 absorption band. Since there is mostly an integrated information content in the ozone profile, there is no need to have a wide-spread sample of channels.

Like NAPCA, unconstrained BC intends to represent as well as possible the whole set of channels in the spectrum. It selects channels in all bands, regardless of the variable to retrieve (Fig. 8 middle). The selected BC channels are the chosen to be most useful to reconstruct the whole spectrum.

By definition, the BC-constrained solution (Fig. 8 bottom) is more localized than the generic BC and therefore closer to the ER selection. As in the standard BC, the constrained-BC is still focus on the window channels because of the lower instrument noise on these channels and the high correlation between the temperature in the tropospheric layers.

Figure 8. From top to bottom: ER-compositing, BC, and BC-constrained channels (in red).

4.2.2. Jacobians comparison

In this section and the following ones, BC will refer to BC-constrained method as it is more convenient for the retrievals than the standard BC version. Temperature (in K/K), humidity and ozone (in $K/Kg/Kg$) Jacobians are represented in Fig. 9 for the resulting BC and ER channels. Since NAPCA Jacobians have no real physical meaning, they are not represented here.

Pure channel selection method ER obtains typical channel Jacobians: their spread along the atmospheric vertical layers indicates how the T, q and O_3 sampling is performed along the vertical. BC Jacobians have a general behavior very close to the pure channel from selection method, this is one of its main advantages. The ER and BC selections

sample T along the atmosphere. For the T selection, the CO_2 and H_2O channels have different sensitivities. H_2O channels are sensitive to temperature up to 100 hPa with a higher sensitivity than CO_2 channels for the same pressure level. It can be noted that BC methods select more H_2O than CO_2 channels. The shape of the humidity (resp. ozone) Jacobians is highly dependent on the variables unit. Indeed, $\frac{K}{Kg/Kg}$ will artificially amplify the sensitivity where the concentration of H_2O (resp. O_3) is low. Since the ozone information content is very localized around 10 hPa, ozone Jacobians are represented from 100 to 0.1 hPa only. The sampling of humidity and ozone is less distributed along the vertical because of the lower profile information content in the observations for these variables.

Figure 9. Temperature Jacobian (K/K) (top row) humidity Jacobian ($K/Kg/Kg$)(middle row) and ozone ($K/Kg/Kg$) (bottom row) Jacobian, for the first 50 ER (left) and BC (right) channels.

4.3. Information content

For a further evaluation of the dimension reduction methodologies, an information content analysis is performed for the NAPCA compression and the ER and BC channel selection methods. The information content is measured using the metric presented by [*Rodgers*, 1996]: the covariance matrix of the retrieval errors is estimated using Eq. (4). It provides a way to estimate the quality of the various dimension reduction techniques.

Fig. 9 represents improvements for the three methodologies (i.e. NAPCA, ER-compositing, and BC-constrained) when 50 and 300 components/channels are selected, for the retrieval of T, q and O_3. The improvement is defined as:

$$Imp = 1 - \sqrt{\frac{diag(\mathbf{A})}{diag(\mathbf{B})}}, \tag{6}$$

where $\mathbf{B}$ is the *a priori* matrix and $\mathbf{A}$ the covariance of the retrieval error presented in Eq. (4). When only 50 pieces of information (components or channels) are used, NAPCA compression clearly outperforms the ER channel selection. There is barely an improvement between 50 and 300 components, which confirms that this number of components is close to the optimum. For the ER-compositing, the 50 first selected channels gather the 17 first channels for T, 17 first channels for q, and 16 first channels for O_3. BC results are almost as good as NAPCA. For both methods, the redundancy of the full set of channels in the spectrum is exploited to reduce the impact of instrumental noise, where ER approach needs more channels to reduce the instrumental noise. With only 50 channels ER-selection gets almost the maximum information in humidity. This can be explained by the fact that some CO_2 channels between 700 and 800 cm^{-1} are also sensitive to humidity. The same occurs for temperature: more than the 17 CO_2 channels are used since the q channels are also sensitive to T, but this is not enough because the T *a priori* error is quite low (1 K) and much information is needed to improve it. Note that for 300 channels, the ER method gets a good information content of the main observation (as expected based on [*Collard*, 2007]).

Since the humidity-sensitive channels for the higher atmosphere are also the noisiest ones, the improvement rates are lower than in the lower levels. NAPCA with 50 pieces of information does not get most of the information for these levels because 50 PCs is a compromised optimum over the full spectrum. This means that the TBs sensitive to humidity in the high atmosphere are not well represented by the first 50 PCs (neither by the first 50 BCs, for similar reason).

Figure 10. *A priori* error improvement rate (in %) for the retrieval of T (top), q (middle) and O_3 (bottom) profiles, for the NAPCA, ER-compositing, and BC-constrained methods, with 50 and 300 components/channels.

5. Discussion on the BC approach

As mentioned before, the compression and information content from the BC and NAPCA dimension reduction techniques are relatively similar. However, the BC representation intend to represent pseudo-channels with a reduced measurement noise, this brings a physical meaning that can facilitate their use in inversion or assimilation mode.

The mixing problem - PCA components mix various information into their components [*Aires et al.*, 2002c]. In Fig. 11, the correlation between the first 50 components and the temperature T, humidity q, and ozone O_3 profiles is represented. It can be seen that all the components have an information on T, q and O_3 constituants. As a consequence, it is not possible to use components only for one constituant, which can be useful for instance in hierarchical retrieval schemes where temperature is first retrieve, then humidity and ozone.

Figure 11. Correlation between the first 50 NAPCA PCs and the temperature (top), humidity (middle), and ozone (bottom) profiles.

On the contrary, it is easier with the BCs to select channels that are related to a single atmospheric constituant. For instance, the BCs can be focused on the ozone retrieval. As mentioned earlier, the BCs can be constrained to sample channels only in the 1000-1100 cm^{-1} O_3 absorption band. Furthermore, the de-noising statistics (i.e., noise-free spectrum $\mathbf{TB}$ minus the compressed-reconstructed observed spectrum $\widetilde{\mathbf{TB}}$) of the BCs can be as good as those from the NAPCA. Fig. 12 compares the NAPCA and BC-constrained de-noising errors for the 1000-1100 cm^{-1} O_3 absorption band. Since this subset of BC channels is focussing on the ozone band, its statistics are better than the general PCA components but this figure shows that it is possible to isolate in the BC channels an information that is specific to one constituant, with a competitive compression rate. This would be much more difficult to attain with a traditional channel selection technique.

Figure 12. Evolution of the de-noising error over the ozone absorption band with an increasing number of NAPCA components or BC channels in the ozone band. Constant mean instrumental noise is represented in dash for comparison purpose.

Contamination effects - We just saw that NAPCA mixes multiple constituents into a single component. When using these components in a retrieval, not considering all these

constituents is dangerous. Individual and independent retrievals should be avoided (see the un-mixing synergy concept described in [*Aires*, 2011; *Aires et al.*, 2011]). We define the "contamination effects" as the difficulties arising when the variations of other not considered/retrieved variables could have strong impact on the retrieval. Compared to channel selections that can focus on specific spectral domains, this general difficulty is emphasized when using NAPCA components for the retrieval. For instance, if clouds presence acts in a spectral region, they will impact most of the NAPCA components, except if these components have been computed in cloud-free spectral regions. BCs can propose an easier practical solution. They are more localized spectrally than NAPCA because they try to reproduce physical channels. As a consequence, some of the BCs will be contaminated by the presence of clouds but other will not. By comparing the TBs simulated using a radiative transfer model and the TBs from the observations, it is possible to detect the channels that will be cloud-contaminated (this is an usual procedure to detect clouds in operational centers) and then suppress them in the inversion or the assimilation scheme. On the contrary, with NAPCA, if one component is highly contaminated, it is difficult to suppress it without suppressing all of its lower-order components: each component is computed on the anomalies left by the first components, therefore each component is highly dependent on the higher-order components.

Minor constituents - Minor absorption bands for trace gazes can be negligible in terms of their statistical variability compared to major constituents such as CO_2 for temperature or H_2O absorption bands. Therefore, the information on trace gazes can be lost or mixed and disseminated in higher-order components of the NAPCA. With the BC strategy, even isolated channels can be chosen by the selection process (in particular if the BC quality criterion is to reduce uniformly the errors in all the channels of the spectrum). If this is not sufficient, it is easy to chose *a priori* some particular channels of particular interest during the channel selection stage of the BC algorithm. This flexibility of the BC allows to introduce automatically information to retrieve major constituents of the atmosphere, and manually channels dedicated to trace gases.

Calibration - When using satellite observations in assimilation mode or in a retrieval algorithm, the necessary radiative transfert simulations can have systematic biases with the real observations. A calibration scheme needs therefore to be put in place in order to correct this issue. When using NAPCA components, the calibration needs to involve all the TBs that are used to compute the components. Differences of simulated and observed TBs are more difficult to interpret. With BCs, the calibration can be done only on a small number of channels. This calibration compares the simulated TBs and the noise-reduced pseudo-channels of the BCs. Only a limited number of information is required and the monitoring of the dynamical calibration is easier to do.

Speeding up inversion or assimilation - As mentioned in the previous calibration issue, radiative transfer simulations need to be performed for most inversion or assimilation techniques. With NAPCA components, these simulations need to be done on the full spectral domain used to compute the components. A dedicated radiative transfert working in the component space can be developed *Matricardi* [2010]. However, this constrain is not necessary when using BCs since the radiative transfert simulations can be done only in the limited number of selected channels. This obviously speeds up the computation and facilitate the inversion or assimilation scheme.

6. Conclusion and perspectives

In this paper, we introduced an innovative dimension reduction technique, the Bottleneck channels (BCs), and compared it to two other classical approaches, the compression by NAPCA and channel selection by ER. Tests have been performed in the context of satellite remote sensing using the IASI instrument for the retrieval of temperature, humidity and ozone atmospheric profiles. Various experimental configurations have been investigated for the BC approach.

When using a limited number of pieces of information, compression outperforms channel selection approach since it better exploits the redundancy among the channel spectrum. However, channel selection techniques preserves a physical meaning of the compact representation which can be an advantage for their forthcoming use in inversion or assimilation. In this context, BCs appears to be a good compromise between compression and channel selection: It is shown that its compression statistics are similar to the PCA, but the compact representation gathers pseudo-channels that can be directly used in any retrieval or assimilation scheme.

Practical advantages of the BC approach are a better handling of the contamination effects by limitation of the mixing problem, an easier representation of the information for trace gazes, an easier calibration of observations, and a faster inversion or assimilation with simpler radiative transfert simulations.

Perspectives are numerous. (1) The BCs should be tested in an atmospheric dataset including cloudy situations to examine how the cloud contamination can be handled. Technical details need to be investigate to optimize the BCs in this context but the spectrally localized BC version should help isolate channels not cloud-contaminated. (2) The introduction of trace gases variability in the atmospheric dataset should allow testing the introduction of *a priori* chosen channels into the channel selection part of the BC algorithm. (3) Finally, retrieval and assimilation tests could be made to validate the BC approach on real-world conditions.

Acknowledgments. The authors grateful to CNES for supporting the study. Thank to Jean-Luc Moncet and Gennady Uymin (AER) for providing IASI simulation, thank to LMD for providing TIGR to AER simulation.Thank to Alannah Hope Crockatt for helping write this article.

References

Aires, F. (2011), Measure and exploitation of multisensor and multiwavelength synergy for remote sensing: 1. Theoretical considerations, *Journal of Geophysical Research: Atmospheres*, *116*(D), 2301, doi:10.1029/2010JD014701.

Aires, F., W. B. Rossow, N. A. Scott, and A. Chédin (2002a), Remote sensing from the infrared atmospheric sounding interferometer instrument 1. Compression, denoising, and first-guess retrieval algorithms, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *107*(D22), ACH 6–1–ACH 6–15, doi:10.1029/2001JD000955.

Aires, F., W. B. Rossow, N. A. Scott, and A. Chédin (2002b), Remote sensing from the infrared atmospheric sounding interferometer instrument 2. Simultaneous retrieval of temperature, water vapor, and ozone atmospheric profiles, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *107*(D22), ACH 7–1–ACH 7–12, doi:10.1029/2001JD001591.

Aires, F., W. B. Rossow, and A. ChéDin (2002c), Rotation of EOFs by the Independent Component Analysis: Toward a Solution of the Mixing Problem in the Decomposition of Geophysical Time Series., *Journal of Atmospheric Sciences*, *59*(1), 111–123, doi:10.1175/1520-0469(2002)059⟨0111:ROEBTI⟩2.0.CO;2.

Aires, F., M. Paul, C. Prigent, B. Rommen, and M. Bouvet (2011), Measure and exploitation of multisensor and multiwavelength synergy for remote sensing: 2. Application to the retrieval of atmospheric temperature and water vapor from MetOp, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *116*(D2), D02,302, doi:10.1029/2010JD014702.

Aires, F., V. Pellet, C. Prigent, and J.-L. Moncet (2015a), Dimension reduction of satellite observations forremote sensing, Part I: A comparison ofcompression, channel selection, and bottleneck channel approaches, *IEEE Trans. Geosc. Remote Sens.*, pp. 1–14.

Aires, F., C. Prigent, E. Orlandi, M. Milz, P. Eriksson, S. Crewell, C.-C. Lin, and V. Kangas (2015b), Microwave hyper-spectral measurements for temperature and humidity atmospheric profiling from satellite: the clear-sky case, *J Geophys Res-Atmos*, (in review), 1–45.

Atkinson, N. C., F. I. Hilton, S. M. Illingworth, J. R. Eyre, and T. Hultberg (2010), Potential for the use of reconstructed IASI radiances in the detection of atmospheric trace gases, *Atmospheric Measurement Techniques*, *3*(4), 991–1003, doi: 10.5194/amt-3-991-2010.

Blackwell, W. J. (2005), A neural-network technique for the retrieval of atmospheric temperature and moisture profiles from high spectral resolution sounding data, *IEEE Transactions on Geoscience and Remote Sensing*, *43*(11), 2535–2546, doi: 10.1109/TGRS.2005.855071.

Cayla, F., B. Tournier, and P. Hebert (1995), Performance budgets of IASI options.

Chevallier, F., S. Di Michele, and A. P. McNally (2006), Diverse profile datasets from the ECMWF 91-level short-range forecasts.

Clerbaux, C., P. F. Coheur, D. Hurtmans, and S. Turquety (2009), Measurements of total and tropospheric ozone from IASI: comparison with correlative satellite, ground-based and ozonesonde observations, *Atmos Chem* . . . .

Collard, A., B. Candy, and F. Hilton (2010), The importance of sounding instruments, *Weather*, *65*(8), 227–227, doi:10.1002/wea.598.

Collard, A. D. (2007), Selection of IASI channels for use in numerical weather prediction, *Quarterly Journal of the Royal Meteorological Society*, *133*(629), 1977–1991, doi:10.1002/qj.178.

Collard, A. D., and A. P. McNally (2009), Sounding interferometer radiances at ECMWF, *qjrms*, *135*, 1044–1058.

Crevoisier, C., A. Chédin, and N. A. Scott (2003), AIRS channel selection for CO2 and other trace-gas retrievals - Crevoisier - 2006 - Quarterly Journal of the Royal Meteorological Society - Wiley Online Library, *Quarterly Journal of the* . . . , doi:10.1256/qj.02.180/pdf.

Fan, J., Y. Liao, and W. Wang (2014), Projected Principal Component Analysis in Factor Models, *arXiv*.

George, M., C. Clerbaux, D. Hurtmans, S. Turquety, P. F. Coheur, M. Pommier, J. Hadji-Lazaro, D. P. Edwards, H. Worden, M. Luo, C. Rinsland, and W. McMillan (2009), Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors, *Atmos. Chem. Phys. Discuss.*, *9*, 9793–9822, doi:10.5194/acpd-9-9793-2009.

Hilton, F., R. Armante, T. August, C. Barnet, A. Bouchard, C. Camy-Peyret, V. Capelle, L. Clarisse, C. Clerbaux, P.-F. Coheur, A. Collard, C. Crevoisier, G. Dufour, D. Edwards, F. Faijan, N. Fourrié, A. Gambacorta, M. Goldberg, V. Guidard, D. Hurtmans, S. Illingworth, N. Jacquinet-Husson, T. Kerzenmacher, D. Klaes, L. Lavanant, G. Masiello, M. Matricardi, A. McNally, S. Newman, E. Pavelin, S. Payan, E. Péquignot, S. Peyridieu, T. Phulpin, J. Remedios, P. Schlüssel, C. Serio, L. Strow, C. Stubenrauch, J. Taylor, D. Tobin, W. Wolf, and D. Zhou (2012), Hyperspectral Earth Observation from IASI: Five Years of Accomplishments, *http://dx.doi.org/10.1175/BAMS-D-11-00027.1*, doi: 10.1175/BAMS-D-11-00027.1.

Huang, H. L., and P. Antonelli (2001), Application of principal component analysis to high-resolution infrared measurement compression and retrieval, *J. Appl. Meteorol.*, *40*, 365–388.

Lipton, A. E. (2003), Satellite sounding channel optimization in the microwave spectrum, *IEEE Transactions on Geoscience and Remote Sensing*, *41*(4), 761–781, doi:10.1109/TGRS.2003.810926.

Lipton, A. E., J. Moncet, S. A. Boukabara, G. Uymin, and K. J. Quinn (2009), Fast and Accurate Radiative Transfer in the Microwave With Optimum Spectral Sampling, *IEEE Transactions on Geoscience and Remote Sensing*, *47*(7), 1909–1917, doi:10.1109/TGRS.2008.2010933.

Mahfouf, J.-F., C. Birman, F. Aires, C. Prigent, E. Orlandi, and M. Milz (2014), Information content on temperature and water vapour from an hyper-spectral microwave sensor, *Q. J. Roy. Meteor. Soc.*, pp. 1–33.

Matricardi, M. (2010), A principal component based version of the RTTOV fast radiative transfer model, *Quarterly Journal of the Royal Meteorological Society*, *136*(652), 1823–1835, doi: 10.1002/qj.680.

Matricardi, M., and A. P. McNally (2014), The direct assimilation of principal components of IASI spectra in the ECMWF 4D-Var, *Quarterly Journal of the Royal Meteorological Society*, *140*(679), 573–582, doi:10.1002/qj.2156.

Moncet, J.-L., G. Uymin, A. E. Lipton, and H. E. Snell (2010), Infrared Radiance Modeling by Optimal Spectral Sampling, *http://dx.doi.org/10.1175/2008JAS2711.1*, doi:10.1175/2008JAS2711.1.

Pellet, V., F. Aires, and C. Prigent (2015), Dimension reduction of satellite observations for remote sensing, Part II: A comparison of compression, channel selection, and bottleneck channel approaches, *IEEE T. Geosci. Remote Sensing*, pp. 1–11.

Pommier, M., K. S. Law, C. Clerbaux, S. Turquety, D. Hurtmans, J. Hadji-Lazaro, P. F. Coheur, H. Schlager, G. Ancellet, J. D. Paris, P. Nédélec, G. S. Diskin, J. R. Podolske, J. S. Holloway, and P. Bernath (2010), IASI carbon monoxide validation over the Arctic during POLARCAT spring and summer campaigns, *Atmos. Chem. Phys. Discuss.*, *10*(10), 14,445–14,494, doi:10.5194/acpd-10-14445-2010.

Rabier, F., N. Fourrié, D. Chafäi, and P. Prunet (2002), Channel selection methods for Infrared Atmospheric Sounding Interferometer radiances, *Quarterly Journal of the Royal Meteorological Society*, *128*(581), 1011–1027, doi:10.1256/0035900021643638.

Rodgers, C. D. (1976), Retrieval of Atmospheric Temperature and Composition From Remote Measurements of Thermal Radiation, *Reviews of Geophysics and Space Physics*, *14*(4), 609, doi:10.1029/RG014i004p00609.

Rodgers, C. D. (1990), Characterization and error analysis of profiles retrieved from remote sounding measurements, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *95*(D5), 5587–5595, doi:10.1029/JD095iD05p05587.

Rodgers, C. D. (1996), Information content and optimisation of high spectral resolution measurements, *Proc SPIE*.

Scannell, C., D. Hurtmans, and A. Boynard (2012), Antarctic ozone hole as observed by IASI/MetOp for 2008–2010, *Atmospheric* . . . .

Turquety, S., J. H. Lazaro, and C. Clerbaux (2004), Operational trace gas retrieval algorithm for the Infrared Atmospheric Sounding Interferometer, *Journal of* . . . .

Corresponding author: Victor Pellet, LERMA observatoire de Paris, FRANCE . (victor.pellet@obspm.fr)