

Optimal flow-dependent selection of channels from advanced sounders in the presence of cloud

Stefano Migliorini

European Centre for Medium-Range
Weather Forecasts, Research Department
and
National Centre for Earth Observation,
University of Reading, UK

Submitted to Monthly Weather Review to be included in the
AMS Special Collection on the 6th WMO DA Symposium

July 2014

This paper has not been published and should be regarded as an Internal Report from ECMWF.
Permission to quote from it should be obtained from the ECMWF.



European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:

<http://www.ecmwf.int/publications/>

Contact: library@ecmwf.int

©Copyright 2014

European Centre for Medium-Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, England

Literary and scientific copyrights belong to ECMWF and are reserved in all countries. This publication is not to be reprinted or translated in whole or in part without the written permission of the Director-General. Appropriate non-commercial use will normally be granted under the condition that reference is made to ECMWF.

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error, omission and for loss or damage arising from its use.

Abstract

This study aims to illustrate a general procedure based on well-known information theory concepts to select the channels from advanced satellite sounders which it is most advantageous to assimilate in all-sky conditions – i.e., both in clear sky and in the presence of cloud – using a flow-dependent estimate of forecast uncertainty. To this end, the standard iterative channel selection method, which is used to select the most informative channels from advanced infrared sounders for operational assimilation, was revisited so as to allow its use with measurements that have correlated errors. The method is here applied to determine a small set (namely, 24) – relatively to a total of 8461 channels that are available on the Infrared Atmospheric Sounding Interferometer (IASI) on board the EUMETSAT Polar System Metop satellites – of humidity-sensitive channels, which can be used to perform all-sky data assimilation experiments, in addition to those currently used for operational data assimilation of IASI data at ECMWF. Care was taken to use in the channel selection procedure a realistic specification of forecast error uncertainty, which was determined from an ensemble of data assimilation (EDA) forecast fields for a case study in July 2012. Also, (cumulative) weighting functions that provide a vertically-resolved picture of the (total) number of degrees of freedom for signal expressed by a given set of measurements were introduced, which allow us to define a novel channel selection merit function that can be used to select measurements that are most sensitive to variations of a given parameter over a given atmospheric region (e.g., in the troposphere).

1 Introduction

Over the last decade or recent decades there has been a formidable increase in the amount of data that is being acquired by satellite sounding instruments and disseminated to operational meteorological centres for assimilation, particularly in the infrared spectral region. At ECMWF, the infrared sounding instruments that are currently monitored and assimilated are: the High Resolution Infrared Radiation Sounder (HIRS), on the EUMETSAT Polar System Metop-A polar orbiting satellite, with twenty channels; the Advanced InfraRed Sounder (AIRS) on board AQUA and measuring over 2378 channels; the Infrared Atmospheric Sounding Interferometer (IASI) also onboard on Metop with 8461 channels and the Cross-track Infrared Sounder (CrIS) on board the Suomi NPP satellite, with 1305 channels. In each case only a subset of channels are assimilated.

In order to be able to exploit such wealth of data, operational centres had to overcome numerous technological and scientific challenges, including making appropriate choices about which subset of channels from each instrument to consider for assimilation. The problem was put on firm theoretical grounds by Rodgers [23], who described an iterative method to determine an optimal set of channels by maximizing a figure of merit based on their Shannon information content. In particular, according to this method, a new channel is selected if it provides the largest information increment with respect to the information content already provided by all previously selected channels. A subsequent study [21] showed that the iterative method provided a more effective channel selection – i.e., a larger state estimate error reduction – than other existing methods based on a non-iterative approach (see, e.g., [20]). As recognized in [21], a likely reason for the shortcomings of the non-iterative methods is their difficulties in providing a selection of channels that are informative over atmospheric partial columns at different height ranges, i.e., that are representative of different spectral regions.

The iterative channel selection method – after appropriate pre-screening of channels with too large forward-model uncertainty or with characteristics that make them more sensitive to mis-specifications of forecast error uncertainty in observation space (e.g., with multiple gas sensitivities or with jacobians that have multiple peaks or long tails) – was used at ECMWF [5] to determine an optimal set of (currently 191) IASI channels sensitive to atmospheric temperature, water vapour, ozone and surface conditions for

operational assimilation. The impact on analyses and forecasts of the selected IASI channels with respect to the ECMWF operational system at the time of the experiments is discussed in [6]. The same channel selection methodology was also used to select the most informative channels from AIRS [8].

In order to reduce the amount of computational resources needed by the iterative method, the channels to be selected are usually assumed to have independent errors so that the retrieval error covariance can be calculated sequentially. This is, however, not often the case as radiance data calculated with fast forward models used for operational assimilation can have spectrally correlated errors (see [18], his section 6 and Figure 13) with length scales that can be significantly larger than those due to apodization, which only involves adjacent channels. Other relevant sources of inter-channel error correlation include: the variations of atmospheric species such as water vapour or ozone when selecting temperature sensitive channels; errors arising from shortcomings in accounting for cloud as well as surface emissivity errors; representativeness errors [4].

A recent channel selection study [25] investigated a way to account for observation error correlations arising from imperfect knowledge of the concentration of unconstrained (i.e., not retrieved) atmospheric constituents with absorption lines in the spectral regions that are sampled by the set of channels considered for selection. Observation errors are expressed in [25] as a combination of random and systematic components, with the random component being assumed as spectrally uncorrelated and as the only observation error component that is relevant to update the retrieval error covariance calculated using the previously selected channels. In this way it is still possible to make use of the sequential method to update the retrieval error covariance, with some computational savings. At the same time, both observation error components – the diagonal random error and the spectrally-correlated systematic error components – are considered to compute the information-content based figure of merit used for channel selection. Another recent study [16] investigated the use of the iterative channel selection technique to complement the IASI channels, which are already in use for operational assimilation of IASI data in clear sky conditions (see [6]), with additional channels that are most effective for the joint retrieval of ice and liquid water content using IASI data without solar contamination. The impact of the additional channels on water vapour estimates was also assessed. The authors found that the additional channels provided at best only marginal improvements with respect to the case when only the standard channels are used in the retrieval.

In this study, the iterative channel selection methodology was revisited and modified to be used in a consistent way with observations having correlated errors. This novel formulation of the iterative non-sequential selection method was then used to select the most effective IASI channels for the estimation of atmospheric water vapour profiles both in clear sky and overcast conditions. To this end, a flow-dependent estimate of forecast errors, derived from an ECMWF’s “ensemble of data assimilations” (EDA) run on a 91-level and 50-member configuration, was used for a case study during summer 2012. It is important to note that the main aim of this work is not to replace existing sets of IASI channels selected for assimilation in clear sky but rather to determine a relatively small number of additional channels that can provide the largest impact on meteorological analyses in all-sky conditions. This means that the IASI channels selected in this study are considered to be best suited to assimilate water-vapour sensitive observations of radiation emerging from either a clear-sky or a cloud-affected scene with a single observation operator that includes a parametrization of multiple scattering by clouds and no need for cloud detection (see, e.g., [2]).

The paper is structured as follows. Section 2 provides a detailed description of the channel selection methodology and a step-by-step algorithm. Also in this section, the standard information-based figure of merit used for selection is extended to allow a selection that is optimal for estimation over a subspace of the state space (e.g., over a given height range or a given parameter). In section 3 a description of the

case study, with particular attention to the quantification of the sensitivity of the radiation emerging from the atmosphere to the presence of cloud at a given location is provided. Section 4 discusses the effects of the chosen forecast and observation error specifications on the signal-to-noise characteristics of the satellite instrument, while section 5 provides details on the selection of optimal channels for atmospheric humidity estimation in all-sky conditions as resulting from the use of the selection method described in this paper, including a list of the selected channels. Finally, a summary of the work and its main conclusions are given section 6.

2 Iterative channel selection with correlated observation errors

The channel selection method as described in [23] is based on finding the channel that, at each iteration, provides the largest increment to the number of degrees of freedom for signal (DFS) already provided by the previously selected channels. This procedure is repeated until the required number of channels have been selected. To reduce the computational costs of the iterative selection process the original algorithm also assumes that the measurement error covariance for the considered channels is diagonal. In this way it is possible to calculate the maximum-a-posteriori retrieval error covariance found when making use of a set of k measurements (see, e.g., [24], his section 5.4) as an update of the retrieval error covariance valid for a set of $k-1$ measurements. This simplifying assumption is also used in [25], while observation error correlations are instead taken into account within the DFS-based figure of merit used for channel selection.

The iterative channel selection method used here avoids the use of the sequential retrieval error covariance update formula and can then consistently be used in the presence of correlated observation errors. The increased computational costs incurred when avoiding the sequential update proved to be affordable, also considering that the method is not supposed to be used for near-real-time applications. But before discussing the algorithm further let us define some relevant quantities.

In order to retrieve an estimate of $\mathbf{x}^t \in \mathbb{R}^n$, the true state of the system, we can make use of a set of measurements that are assumed to be components of the measurement vector $\mathbf{y}^o \in \mathbb{R}^m$. The relationship between the measurement vector and the state vector can be written as

$$\mathbf{y}^o = H(\mathbf{x}^t) + \boldsymbol{\varepsilon}^o \quad (1)$$

where $H(\mathbf{x}^t) \in \mathbb{R}^{m \times n}$ is the observation operator and $\boldsymbol{\varepsilon}^o \in \mathbb{R}^m$ is an additive observation error vector, assumed to be unbiased and Gaussian-distributed with nonsingular covariance \mathbf{R}_m . If the observation operator is approximately linear in a given region of the state space around \mathbf{x}_0 it is meaningful to approximate Eq. (1) as

$$\mathbf{y}^o \simeq H(\mathbf{x}_0) + \mathbf{H}(\mathbf{x}^t - \mathbf{x}_0) + \boldsymbol{\varepsilon}^o. \quad (2)$$

where $\mathbf{H} \equiv (\partial H / \partial \mathbf{x})$ calculated at \mathbf{x}_0 is the Jacobian matrix. In this case, we can define \mathbf{y} as (see [24], his section 8.3)

$$\mathbf{y} \equiv \mathbf{y}^o - H(\mathbf{x}_0) + \mathbf{H}\mathbf{x}_0 \simeq \mathbf{H}\mathbf{x}^t + \boldsymbol{\varepsilon}^o. \quad (3)$$

The algorithm used in this work is based on maximizing a figure of merit given by the number of DFS. The total number of DFS using a given set of measurements, d_s is given by (e.g., [23], his section 2.5)

$$d_s = \text{tr}(\mathbf{I}_n - \mathbf{P}^a \mathbf{B}^{-1}) \equiv \text{tr}(\mathbf{A}) \quad (4)$$

where $\mathbf{P}^a \in \mathbb{R}^{n \times n}$ is the retrieval error covariance matrix when a measurement vector with m component is used in the retrieval, $\mathbf{B} \in \mathbb{R}^{n \times n}$ is the forecast (also denoted as prior or background) error covariance

matrix, \mathbf{I}_n is the identity matrix of order n and where \mathbf{A} is the so-called averaging kernel matrix. Let us now introduce an extension of the DFS-based figure of merit that may be used to maximize the number of DFS over a given number of state vector components, that is, over a given range of model levels (e.g., in the troposphere) or a given state vector parameter (e.g., specific humidity or the profile of a given atmospheric chemical compound of interest).

2.1 DFS weighting functions and the effective DFS

If now \mathbf{B} is expressed as $\mathbf{B} = \mathbf{X}\mathbf{X}$, where $\mathbf{X} \in \mathbb{R}^{n \times n}$ is the symmetric square root of \mathbf{B} , we can define the signal-to-noise matrix \mathbf{S} of the considered measurements as $\mathbf{S} = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{X}$. The signal-to-noise matrix can be expressed in terms of its singular value decomposition as $\mathbf{S} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ are orthogonal matrices whose columns are the left and right singular vectors of \mathbf{S} , respectively, and where $\mathbf{\Lambda} \in \mathbb{R}^{m \times n}$ is the matrix whose nonzero elements have the same row and column indexes and are given by the singular values λ_j of \mathbf{S} , with $j = 1, \dots, r$ where r is the rank of \mathbf{S} , with $r \leq \min(m, n)$.

It is possible to show (see [19], his section 5) that the retrieval error covariance \mathbf{P}^a can be expressed in term of the singular values and the right singular vectors of \mathbf{S} as

$$\mathbf{P}^a = \mathbf{X}\mathbf{V} \begin{pmatrix} (\mathbf{\Lambda}_r^2 + \mathbf{I}_r)^{-1} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{I}_{n-r} \end{pmatrix} \mathbf{V}^T \mathbf{X} = \sum_{j=1}^r \mathbf{X} \frac{\mathbf{v}_j \mathbf{v}_j^T}{1 + \lambda_j^2} \mathbf{X} + \sum_{i=r+1}^n \mathbf{X}^f \mathbf{v}_i \mathbf{v}_i^T \mathbf{X} \quad (5)$$

Noting that $\sum_{j=1}^n \mathbf{v}_j \mathbf{v}_j^T = \mathbf{I}_n$, Eq. (5) can be written as

$$\mathbf{P}^a = \mathbf{B} - \sum_{j=1}^r \frac{\lambda_j^2}{1 + \lambda_j^2} \mathbf{X}^f \mathbf{v}_j \mathbf{v}_j^T \mathbf{X}^f. \quad (6)$$

From Eqs. (6) and (4) we can write

$$d_s = \text{tr} \left(\sum_{j=1}^r \frac{\lambda_j^2}{1 + \lambda_j^2} \mathbf{X}^f \mathbf{v}_j \mathbf{v}_j^T \mathbf{X}^f \mathbf{B}^{-1} \right) \equiv \text{tr} \left(\sum_{j=1}^r d_{s_j} \mathbf{X}^f \mathbf{v}_j \mathbf{v}_j^T \mathbf{X}^f \mathbf{B}^{-1} \right) = \sum_{j=1}^r d_{s_j} \text{tr}(\mathbf{X}^f \mathbf{v}_j \mathbf{v}_j^T \mathbf{X}^{f-1}), \quad (7)$$

as d_s can also be expressed as (see [19], his section 4) $d_s = \sum_{j=1}^r \frac{\lambda_j^2}{1 + \lambda_j^2} \equiv \sum_{j=1}^r d_{s_j}$. As the trace of a matrix is invariant under similarity transformations (e.g., [10], their section 7.1.1) Eq. (7) can be written as

$$d_s = \sum_{j=1}^r d_{s_j} \text{tr}(\mathbf{v}_j \mathbf{v}_j^T) = \sum_{j=1}^r d_{s_j} \mathbf{v}_j^T \mathbf{v}_j \quad (8)$$

Noting that $\mathbf{v}_j^T \mathbf{v}_j = 1$ as \mathbf{v}_j is an orthonormal basis vector, the expression in Eq. (7) linking the degrees of freedom for signal (DFS) in observation space with the DFS in state space represents a trivial identity that follows from the fact that the two quantities are equivalent (e.g., [24], his section 2.5). However, it is possible to make use of the diagonal elements of the matrix $d_{s_j} \mathbf{v}_j \mathbf{v}_j^T$ given by

$$\mathbf{s}_j \equiv \text{diag}(d_{s_j} \mathbf{v}_j \mathbf{v}_j^T) = d_{s_j} (v_{1j}^2, v_{2j}^2, \dots, v_{nj}^2) = d_{s_j} \mathbf{v}_j \circ \mathbf{v}_j \quad (9)$$

where \circ denotes the element-wise (or Schur) product, in order to determine how the number of DFS d_{s_j} are vertically distributed for each of the variables included in the state vector. Note that $\mathbf{s} = \sum_{j=1}^r \mathbf{s}_j =$

$\text{diag}(\mathbf{A})$. From Eq. (8) it also follows that the vertical distribution of the total number of DFS for a given model variable is given by

$$\mathbf{s} = \left(\sum_{j=1}^r d_{s_j} v_{1j}^2, \sum_{j=1}^r d_{s_j} v_{2j}^2, \dots, \sum_{j=1}^r d_{s_j} v_{nj}^2 \right) = \sum_{j=1}^r d_{s_j} \mathbf{v}_j \circ \mathbf{v}_j. \quad (10)$$

In other words, the quantity $d_{s_j} v_{lj}^2$ represents the fraction of d_{s_j} achieved on a given model variable (e.g. temperature) at model level l while $\sum_{j=1}^r d_{s_j} v_{lj}^2$ is the total number of DFS achieved on the same model variable at model level l . The components of \mathbf{s}_j define the j -th *DFS weighting function* for all relevant model variables, while the components of \mathbf{s} define the *cumulative DFS weighting function* for all relevant model variables. Note that the elements (always non-negative) of the cumulative DFS weighting function coincide with the diagonal elements of the averaging kernel matrix \mathbf{A} .

The sum d_s^{eff} of a subset of the components of \mathbf{s} defined over a given region of the state space defined as

$$d_s^{\text{eff}} \equiv \sum_{k=k_1}^{k_2} s_k = \sum_{j=1}^r d_{s_j} \sum_{k=k_1}^{k_2} v_{kj}^2 \quad (11)$$

represents the *effective degrees of freedom for signal* achieved by a set of measurements over that region of the state space, with k_1 and k_2 being the lower and upper model levels defining an atmospheric (partial) column of interest, for a given parameter of interest (e.g., atmospheric humidity). As discussed in the next section, the (effective) degrees of freedom for signal can be used, for example, as a figure of merit for channel selection in order to select the set of n_{sel} channels that provide most information about the whole (selected region of the) state space. Note that $d_s^{\text{eff}} \equiv d_s$ when $k_1 = 1$ and $k_2 = n$. It is also important to note that the iterative channel selection results obtained using the number of DFS as merit function do not change if another monotonically increasing merit function of a set of λ_i is used instead. In particular, the same set of channels are selected when using the Shannon information content I defined as (see, e.g., [24], his section 2.5.2) $I = 0.5 \sum_i \ln(1 + \lambda_i^2)$.

Finally, it is useful to compare the newly-introduced DFS weighting function with the familiar jacobian defined as a given row of \mathbf{H} in Eqs. 1 and 2. To determine the region of the state space a given instrument can sense it is necessary to explore the jacobians for all instrument channels of interest to check where they are significantly different from zero and possibly where they peak. This region of the state space, however, in general does not coincide with that on which a given sounder can provide most of its information, which is also a function of the observation and forecast error covariance matrices used for assimilation. Also, the jacobians can be negative and this may cause confusion when the jacobians are used as a measure of vertical resolution for a given variable. The cumulative DFS weighting function \mathbf{s} includes contributions from the jacobians of all considered channels and then it provides a concise depiction of the sensitivity of a given instrument on different vertical atmospheric layers for different model variables. Also, from the fact that the sum of the components of \mathbf{s} is equal to d_s it follows that the sum of the elements of \mathbf{s} over a given atmospheric partial column and model variable quantifies the number of DFS that the instrument can provide on the chosen model variable over that region. As explained above, however, the characteristics of the DFS weighting functions vary according to the forecast error uncertainty used for retrieval or assimilation so that they are not only linked to the instrument specifications.

2.2 Description of the selection algorithm

An iterative, non-sequential, channel selection algorithm can be devised as follows. At the first iteration step $l = 1$ the instrument channel i is considered and the signal-to-noise matrix $\mathbf{S}_{l=1,i} \in \mathbb{R}^{1 \times n}$ is calculated

as

$$\mathbf{S}_{l=1,i} = \frac{\mathbf{h}_i^T}{\sigma_i^o} \mathbf{X}, \quad (12)$$

where \mathbf{h}_i^T and σ_i^o are the row of \mathbf{H} and the observation error standard deviation corresponding to the measurement channel i , respectively. Then the Gramian $g_{l=1,i} \in \mathbb{R}$ of $\mathbf{S}_{l=1,i}^T$, given by

$$g_{l=1,i} \equiv \mathbf{S}_{l=1,i} \mathbf{S}_{l=1,i}^T \equiv \lambda_{l=1,i,j=1}^2, \quad (13)$$

is computed, where the index j orders the singular values of $\mathbf{S}_{l=1,i}$. Then the number $d_{s_{l=1,i}}$ of DFS for channel i is calculated as

$$d_{s_{l=1,i}} = \frac{\lambda_{l=1,i,j=1}^2}{1 + \lambda_{l=1,i,j=1}^2}. \quad (14)$$

Note that from Eqs. (12), (13) and (14) it follows that $d_{s_{l=1,i}}$ can also be expressed as

$$d_{s_{l=1,i}} = \frac{\mathbf{h}_i^T \mathbf{B} \mathbf{h}_i}{\sigma_i^{o^2} + \mathbf{h}_i^T \mathbf{B} \mathbf{h}_i} \quad (15)$$

and this shows the equivalence between Eq. (14) and Eq. (17) in [21].

At iteration step $l = 1$ the Gramian $g_{l=1,i}$ and the number $d_{s_{l=1,i}}$ of DFS are calculated m times, with $i = 1, \dots, m$. It is important to note that these computations can be performed in parallel. The channel selected at $l = 1$ is the channel i_1 with $d_{s_{l=1,i_1}} = \max(d_{s_{l=1,i}})$, with $i = 1, \dots, m$.

The iteration step $l = 2$ consists in calculating $\mathbf{S}_{l=2,i} \in \mathbb{R}^{2 \times n}$ as

$$\mathbf{S}_{l=2,i} = \mathbf{R}_{l=2,i}^{-1/2} \begin{pmatrix} \mathbf{h}_{i_1}^T \\ \mathbf{h}_i^T \end{pmatrix} \mathbf{X} \quad (16)$$

where $\mathbf{R}_{l=2,i} \in \mathbb{R}^{2 \times 2}$ is a submatrix of the observation error covariance \mathbf{R} defined as

$$\mathbf{R}_{l=2,i} = \begin{pmatrix} \sigma_{i_1}^{o^2} & r(i_1, i) \sigma_{i_1}^o \sigma_i^o \\ r(i_1, i) \sigma_{i_1}^o \sigma_i^o & \sigma_i^{o^2} \end{pmatrix} = \mathbf{L}_{l=2,i} \Gamma_{l=2,i}^2 \mathbf{L}_{l=2,i}^T \quad (17)$$

and $\mathbf{R}_{l=2,i}^{-1/2} = \mathbf{L}_{l=2,i} \Gamma_{l=2,i}^{-1} \mathbf{L}_{l=2,i}^T$. The Gramian matrix $\mathbf{G}_{l=2,i} \in \mathbb{R}^{2 \times 2}$, with $i \neq i_1$, is then given by

$$\mathbf{G}_{l=2,i} \equiv \mathbf{S}_{l=2,i} \mathbf{S}_{l=2,i}^T = \mathbf{U}_{l=2,i} \Lambda_{l=2,i}^2 \mathbf{U}_{l=2,i}^T \quad (18)$$

Then the number of DFS for channel $i \neq i_1$ is calculated as

$$d_{s_{l=2,i}} = \sum_{j=1}^r \frac{\lambda_{l=2,i,j}^2}{1 + \lambda_{l=2,i,j}^2} \quad (19)$$

with $r \leq \min(2, n)$. At iteration step $l = 2$ the Gramian matrix $\mathbf{G}_{l=2,i}$ and the number of DFS $d_{s_{l=2,i}}$ are calculated $m - 1$ times, with $i = 1, \dots, m \neq i_1$. The channel selected at $l = 2$ is the channel i_2 with $d_{s_{l=2,i_2}} = \max(d_{s_{l=2,i}})$, with $i = 1, \dots, m \neq i_1$.

The algorithm is iterated until $l = n_{\text{sel}}$ channels are selected. Note that if the state vector \mathbf{x} is composed of two state vectors \mathbf{x}_1 and \mathbf{x}_2 – say temperature and humidity – and the background error covariance matrix \mathbf{B} is block diagonal over the subspaces defined by \mathbf{x}_1 and \mathbf{x}_2 , then the signal-to-noise matrix \mathbf{S} can be calculated as $\mathbf{S}_{\mathbf{x}} = (\mathbf{S}_{\mathbf{x}_1}, \mathbf{S}_{\mathbf{x}_2})$.

Finally, note that when there is the need to select channels that are particularly suited to reduce retrieval errors over a given region of the state space, it is possible to replace $d_{s,i}$ with $d_{s,i}^{\text{eff}}$ as a channel selection figure of merit. This can be achieved by replacing $\mathbf{G}_{l,i} \in \mathbb{R}^{l \times l}$ in Eqs. 13 and 18 with $\mathbf{G}_{l,i}^T \in \mathbb{R}^{n \times n}$, the Gramian of $\mathbf{S}_{l,i}$, given by

$$\mathbf{G}_{l,i}^T \equiv \mathbf{S}_{l,i}^T \mathbf{S}_{l,i} = \mathbf{V}_{l,i} \Lambda_{l,i}^2 \mathbf{V}_{l,i}^T \quad (20)$$

and then by using the j -th column $\mathbf{v}_{l,i,j}$ of $\mathbf{V}_{l,i}$ and the j -th diagonal element $\lambda_{l,i,j}^2$, with $j = 1, \dots, r$ to calculate d_s^{eff} from Eq. 11 over a selected region of the state space. It is then possible to replace d_s with d_s^{eff} in Eqs. 14 and 19 and use it as a figure of merit for channel selection over a target subspace of the state space. Note that with this modifications the computational expense of the selection algorithm becomes significantly larger, in the typical case when $n > n_{\text{sel}} \geq l$, as at each iteration it is necessary to compute the eigenvector decomposition of a $n \times n$ rather than of a $l \times l$ matrix. This can still be computationally affordable, e.g., when only temperature is considered in the state vector and there is the need to find the best channels for temperature estimation, e.g., in the troposphere.

3 Description of the case study

In this work, the channel selection method discussed above is used to check whether the informative potential on key atmospheric variables of a set of instrument channels changes when cloud is present in the instrument field of view. The answer to this question may lead to the selection of a set of channels that provide significant information on temperature or water vapour in all-sky conditions.

To study channel selection strategies that are effective and robust both in clear sky and in the presence of cloud a case study was selected on 30th June 2012 at 2100 UTC. A 50-member ensemble of short-range forecasts – including also cloud liquid and ice water content and cloud fraction – generated from an ensemble of data assimilations (EDA) at ECMWF was used in this study to define a flow-dependent forecast error covariance over 91 model levels. Care was taken to inflate the variance of the forecast ensemble in order to be approximately equal to the EDA mean square error (routinely estimated from operational ECMWF analyses) so as to lead to more reliable forecasts [3]. For each ensemble member at each location it is possible to calculate the cumulative cloud cover N_{tot} , defined as

$$N_{tot} = 1 - (1 - N_1) \prod_{l=2}^n \frac{1 - \max(N_{l-1}, N_l)}{1 - N_{l-1}} \quad (21)$$

when all cloud layers between the top ($l = 1$) and the bottom layer ($l = n$) are considered. Here we have assumed cloud layers in the column to have maximum-random overlap [22] [17]. Figure 1 shows the ensemble-mean values of N_{tot} for the whole atmospheric depth as well as the 135 locations of atmospheric columns over ocean that are cloud-free ($N_{tot} = 0$, marked with a red cross) or the 169 overcast ($N_{tot} = 1$, marked with a blue cross). Given that cloud fraction is a non-negative quantity, the locations that are cloud-free or overcast in the forecast ensemble mean are also such for all ensemble members. We can then denote the cross-marked columns as either “almost surely” (a.s.) – i.e., with probability 1, as defined by the considered forecast ensemble – cloud free or a.s. overcast. Hereafter, when referring to clear-sky or overcast conditions, it is assumed to consider them in an almost-surely sense. Also, in Figure 2 is shown the vertical distribution of ensemble mean cloud quantities at the a.s. overcast locations.

In this study the radiation emerging from the atmosphere was simulated using version 11 of the RTTOV package (Radiative Transfer model for Television Infrared Observation Satellite Operational Vertical sounder) [12] in the “scattering parametrization” configuration [17] to account for cloud radiative effects. In order to investigate the sensitivity of IASI observations to temperature and water vapour at

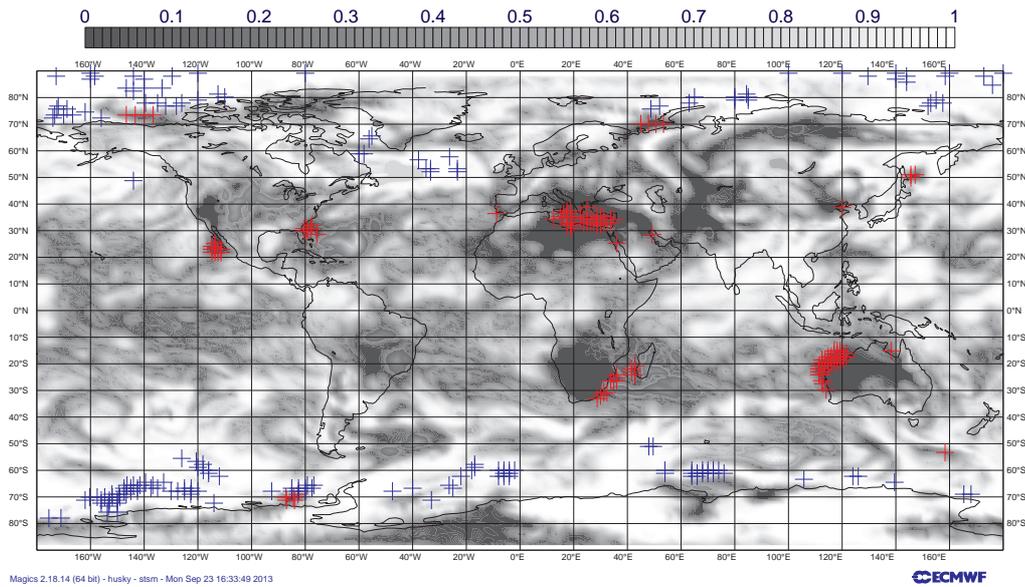


Figure 1: Ensemble mean cumulative cloud fraction N_{tot} over all 91 model levels. Cloud-free columns (with $N_{tot} = 0$) and overcast columns (with $N_{tot} = 1$) are marked with a red or a blue cross, respectively.

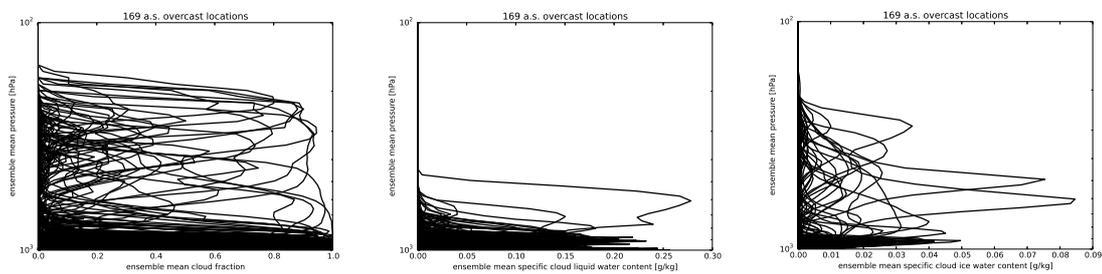


Figure 2: Ensemble mean cloud fraction (left panel), specific cloud liquid water content (mid panel) and specific cloud ice water content over all 91 model levels at the 169 a.s. overcast locations shown in Figure 1.

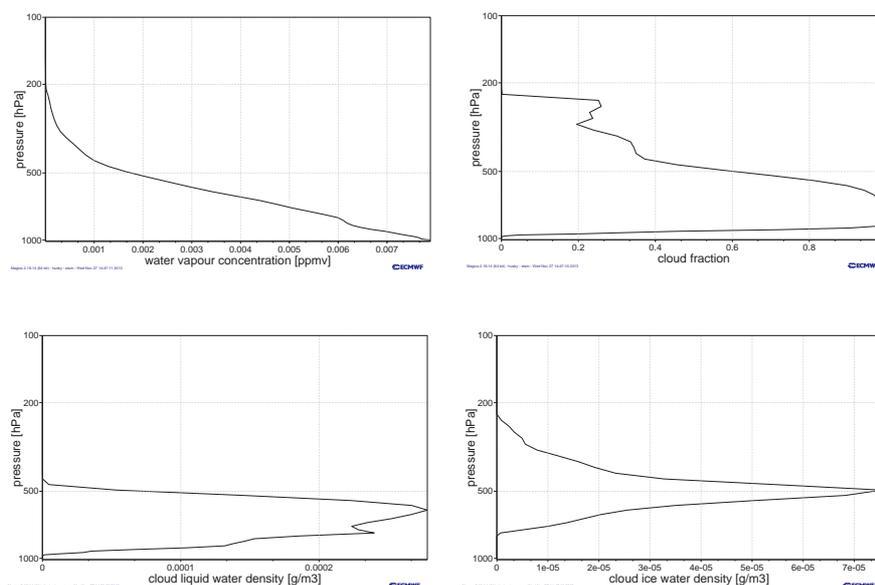


Figure 3: Water vapour and cloud ensemble mean profiles at a selected overcast location over the ocean.

a given channel depending on cloud conditions it is possible to calculate the Jacobian matrix, i.e. the linearized observation operator \mathbf{H} (see Eq. 2) about the ensemble mean forecast. Figure 3 shows water vapour and cloud ensemble-mean profiles at an overcast location over the Atlantic ocean. In Figure 4 are shown the elements corresponding to atmospheric temperature and humidity of a row of the Jacobian matrix linearized about the ensemble mean forecast corresponding to IASI channel 921 (centred at 875 cm^{-1} in the infrared atmospheric absorption window). In particular, it is interesting to check how the atmospheric temperature and humidity jacobians change when cloud is removed from the instrument field of view. Figure 4 shows that the height of the peak of the temperature jacobian in the presence of cloud is very close to that where cloud ice water density reaches its maximum value, while in clear sky the measurements of T_b in the considered atmospheric window channel are mainly sensitive to temperature variations in the lower troposphere, as expected. Also, the measurement's vertical resolution is considerably higher when cloud is present in the instrument field of view and the peak temperature sensitivity is more than 20 times larger than that experienced in clear sky. The changes in the water vapour jacobian due to the presence of cloud are largely similar, although the altitude of the peak of the jacobian in the presence of cloud is higher and its width is larger than in the case of the temperature jacobian when cloud is present.

4 Evaluation of flow-dependent signal-to-noise characteristics of IASI channels

As recognized by previous channel selection studies cited in this paper, from the discussion presented in section 2 it follows that information-content-based channel selection results depend critically on the signal-to-noise characteristics of a given instrument as expressed in a particular data assimilation or retrieval system. In particular, a meaningful expression for the number of DFS of a set of measurements requires full-rank expressions for the vertical forecast error covariance \mathbf{B} and observation error covari-

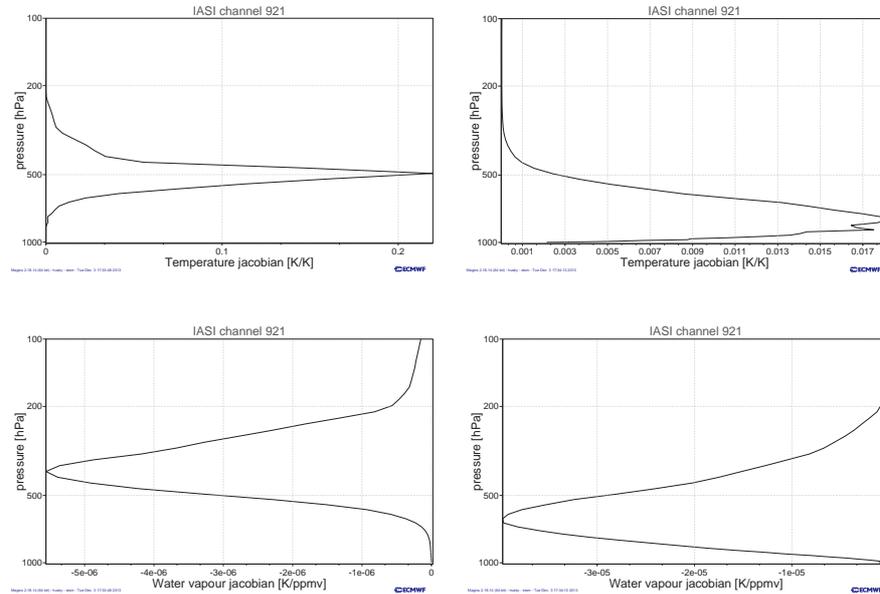


Figure 4: Temperature and humidity jacobians at a selected overcast location over the ocean (the same location chosen in Figure 3) in cloudy conditions and when cloud is removed from the field of view.

ance **R**.

A forecast ensemble of large size K (i.e., with $K > n$), for example, can be used to provide a full-rank approximation of **B**. The dimension n of the state space, however, is usually significantly larger than the ensemble size (even when the state describes the vertical profile of a single model variable!) so that the directions spanned by the forecast error vectors derived from the ensemble explore only a subset of the state space at observation location. This means that the vertical forecast error covariance estimated from the ensemble is rank deficient and it is well known that this may lead to spurious long-range correlations (e.g., [11]). To avoid this problem, two complementary strategies are here adopted: a) vertical localization of the state space to suppress correlations beyond a given threshold distance from each model level by multiplying element-by-element the forecast error covariance with a correlation matrix with compact support (e.g., [14] [15]); b) to combine the flow-dependent EDA forecast error variances with climatological (i.e., full rank) vertical forecast error correlation matrices derived from EDA forecasts over two seasons and geographically varying.

Figure 5a shows the vertical temperature forecast error correlation from EDA at the previously selected overcast location. The figure shows that in the boundary layer below about model level 78 (at about 870 hPa) the temperature error correlations at different model levels are relatively large. This is consistent with a well-mixed boundary layer that is decoupled from the above free troposphere. Also evident is the presence of spurious long range correlations. An eigenvector decomposition of the correlation matrix shows that the rank of the matrix is insufficient (i.e., less than $n = 91$) and equal to its theoretical maximum value ($= K - 1 = 49$). As anticipated, we try to address this shortcoming by localizing the raw correlation function with a correlation function (as described in [9]) that is set to zero beyond a given distance, here chosen to be reached when the ratio r of the distance between two different model levels and the atmospheric scale height is equal to 2.0. The localized correlation matrix is shown in Figure 5b. The localized correlation matrix looks reasonable as it retains the physically-consistent large correlations in the short-range while it suppresses the spurious ones at distance larger than two atmospheric scale

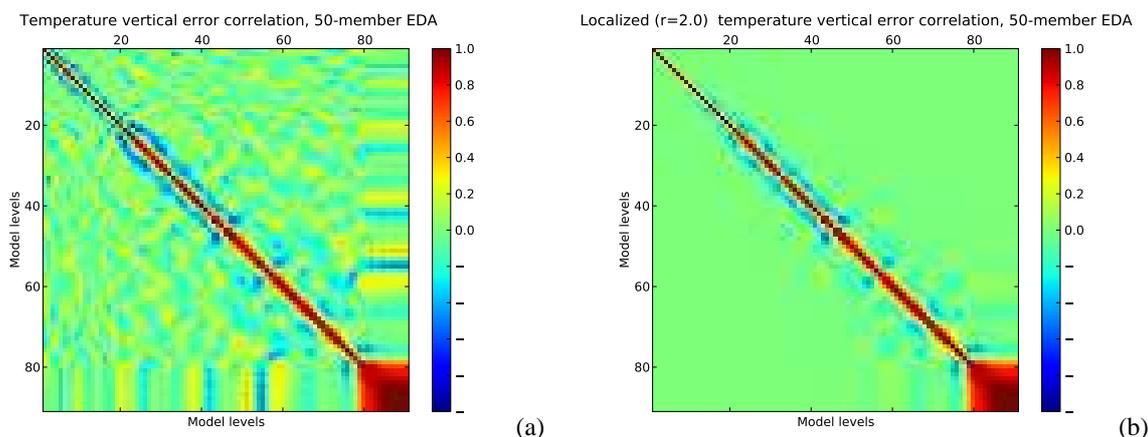


Figure 5: Temperature vertical forecast error correlation from a 50-member EDA. Panel (a): raw EDA correlations; panel (b): localized correlation matrix with $r = 2.0$ (see text).

heights. The rank of the localized correlation matrix is now full and equal to 91, the dimension of the state space at observation location when only the atmospheric temperature profile is included in the state vector.

As discussed above, it is also possible to consider a regional climatology of vertical forecast error correlations – in addition to the EDA-derived variances – to calculate the signal to noise matrix for a given linearized observation operator. Vertical correlations are available for temperature, humidity and ozone over regions of 625-km grid size and averaged over a month to a season [1]. Recent investigations [13] show that the seasonal dependence of the correlations is small with respect to their geographical variability. In Figure 6 (left panel) the climatological vertical temperature forecast error correlation over 91 model levels, interpolated at the selected location is shown. A comparison between Figures 5b and 6 (left panel) shows that the localization procedure applied to the raw EDA vertical error correlation matrix can make the correlation length scales of the raw matrix comparable to those characterizing the climatological covariance, with still some differences in the upper stratosphere above model level 20. An evaluation of the eigenvalues of the correlation matrices determined above (see Figure 6, right panel) also shows that the absolute differences between the 49 largest eigenvalues of the climatological and localized EDA correlation matrices are dramatically reduced with respect to the corresponding difference when the raw rather than the localized EDA correlation matrix is considered. The localized EDA correlation matrix, however, is less conditioned than the climatological one due to the lower magnitude of the eigenvalues corresponding to eigenvectors of the localized-EDA correlation matrix spanning the subspace of the state space that is not represented by the raw forecast ensemble.

Overall, the comparison of the characteristics of the localized version of the forecast error covariance based on EDA and of that from a regional climatology shows that it is reasonable to make use of a localized flow-dependent forecast error covariance for a set of model fields to provide an estimate of the information content of a number of measurements that is as consistent as possible with the actual information content provided by the same number of measurements when assimilated in an operational data assimilation system. In view of these results, the channel selection method applied to IASI data in this work always made use of a localized EDA-based forecast error covariance to determine an expression for \mathbf{B} to be used in the calculation of \mathbf{G} , as discussed in section 2. In particular, the state vector as defined in this study includes temperature, humidity and ozone components. The temperature components are defined over 91 atmospheric model levels and one surface level (surface skin temperature), while the

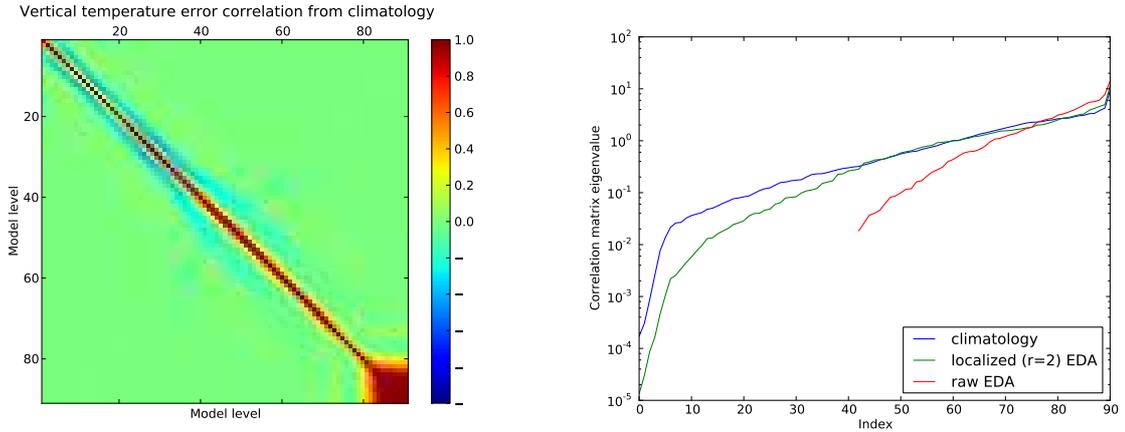


Figure 6: Temperature vertical forecast error correlation from a regional climatology, interpolated to the selected overcast location (left panel); eigenvalues of temperature vertical forecast error correlation from a regional climatology, localized-EDA and raw EDA at cross-hair location (right panel).

humidity and ozone components are defined over 91 atmospheric model levels. Note that errors on the temperature (including surface skin temperature), humidity and ozone components of the state are assumed to be mutually uncorrelated.

The IASI observation error covariance \mathbf{R} used in this study – in brightness temperature units – is that provided with the 1DVar scheme developed by the Met Office for the Satellite Application Facility for numerical weather prediction (NWP-SAF) and includes forward model error. The apodization process applied during the radiometric calibration of the observations from IASI introduces correlations between adjacent channels. As the channel selection method used in this work allows for the presence of correlations, there was no need to exclude adjacent channels once a given channel was selected, as done when sequential methods are used. More generally, the selection method presented here allows the inclusion of systematic error components that present relatively long-range spectral correlations, which may not be compatible with channel exclusion procedures. Of this kind are errors arising from imperfect knowledge of “contaminant” species affecting the estimate of a “target” component of the state vector, such as errors due to incorrect specification of water vapour or ozone concentrations within a temperature retrieval (see, e.g., [24], his section 4.1.2, [7] and [25], provided that the contaminant and the target species have independent forecast errors. In this case, the total observation error covariance \mathbf{R}^{tot} to be used in the place of \mathbf{R} to select channels that are best suited to estimate a target species using the method described in section 2 can be calculated as

$$\mathbf{R}^{\text{tot}} = \mathbf{R} + \sum_i \mathbf{H}_{c_i} \mathbf{B}_{c_i} \mathbf{H}_{c_i}^T, \quad (22)$$

where \mathbf{H}_{c_i} and \mathbf{B}_{c_i} are the jacobian and background error matrices, respectively, for the contaminant c_i . Note that the use of the exact expression for the “systematic” components $\mathbf{H}_{c_i} \mathbf{B}_{c_i} \mathbf{H}_{c_i}^T$ of the observation error covariance makes the procedure, used in [7] and [25] to represent the systematic error covariance by means of an ensemble of perturbed measurement vectors, here unnecessary. Figure 7 shows the spectral dependence of the standard deviation of the components of the IASI observation error covariance used in this work for channel selection purposes (left panel) and their spectral correlations for channels below 2200 cm^{-1} (right panel).

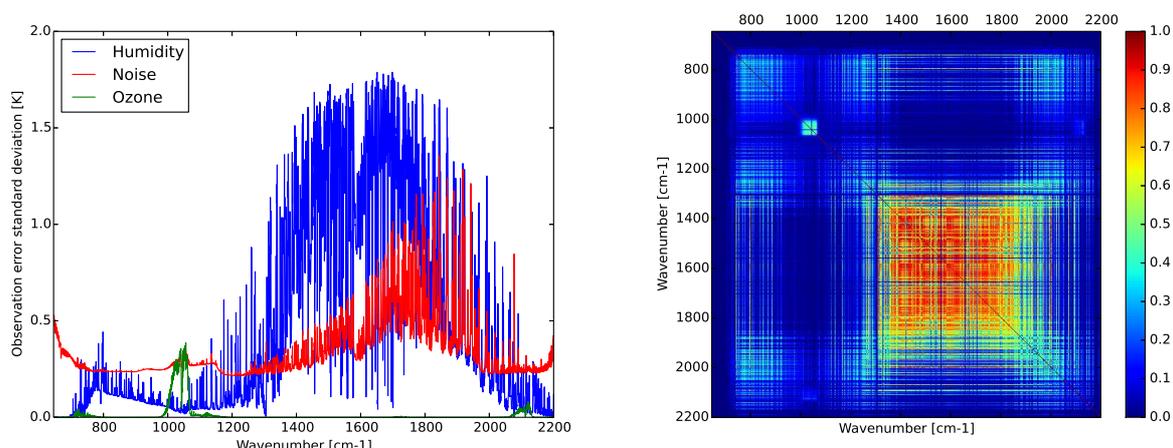


Figure 7: Standard deviation of IASI observation error components used in this work. Left panel: noise (including forward model error contributions) as provided with the 1DVar scheme developed by the Met Office for the NWP-SAF (red solid line), interference from humidity (blue solid line) and from ozone (green solid line). Right panel: IASI observation error correlation for channels below 2200 cm⁻¹, including contributions due to apodization as well as humidity and ozone (see Eq. 22).

5 Channel selection results

The channel selection method discussed in section 2 was applied to each of the 135 clear-sky and 169 overcast columns in our case study, with the aim of selecting a number of humidity-sensitive IASI channels to be used for all-sky data assimilation experiments in addition to the temperature- humidity- and ozone-sensitive IASI channels already assimilated operationally in clear sky. Note that the channel selection figure of merit used here is the number of DFS expressed by a set of measurement channels, but a figure of merit given by the number of effective DFS (see Eq. 11) could have been used instead if the aim was to select a number of humidity-sensitive IASI channels over a given atmospheric region.

Similarly to the previous studies cited in section 1, the first step was to select channels primarily sensitive to atmospheric temperature profile variations located in the 15 μm carbon dioxide band, in a way to minimize contaminations from atmospheric species such as water vapour, ozone and carbon monoxide that are radiatively active in the infrared, as well as to avoid non-local-thermodynamic-equilibrium (non-LTE) effects and solar contributions. An additional benefit of this channel pre-screening procedure is that it reduces the nonlinearity of the observation operator, which could make the temperature jacobians dependent on the state of the contaminant species and potentially lead the data assimilation analysis to be critically dependent on the minimization first guess: a temperature jacobian, for example, may result in having its peak at an incorrectly lower height in the troposphere when the short-range model's forecast underestimates the mixing ratio of the contaminant species (e.g., water vapour). To this end, 100 temperature-sensitive IASI channels were selected among those with wavenumber less than 900 cm⁻¹ (i.e., out of a total of 1020 IASI channels) at each considered location. For temperature channel selection, the total observation error covariance matrix included systematic contributions (see Eq. 22) to account for contaminations due to uncertainty on humidity and ozone while for humidity channel selection the only additional systematic contribution was that due to ozone uncertainty.

Once the 100-th channel was added to the list of those maximizing the number of DFS for temperature, at each location the temperature state vector was augmented with the 91 components of the specific

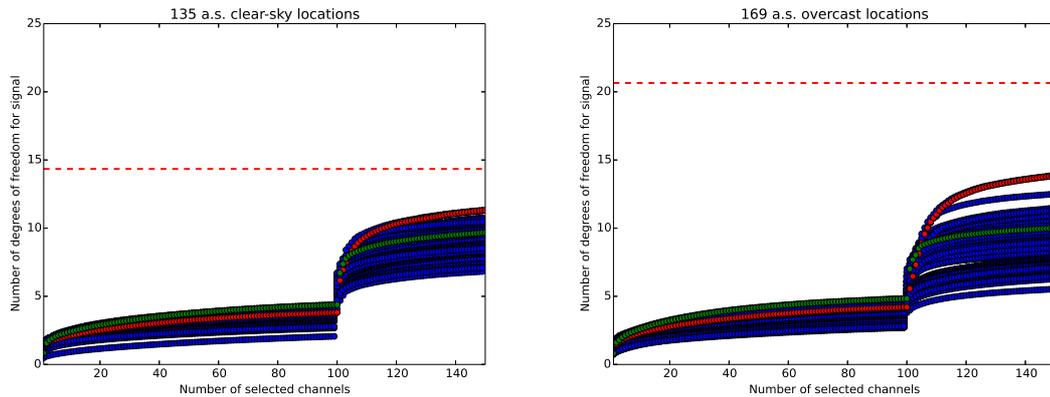


Figure 8: Number of degrees of freedom for signal (DFS) achieved in clear-sky (left panel) and in overcast (right panel) conditions when selecting first up to 100 temperature-sensitive IASI channels and then when selecting up to additional 50 channels that are sensitive to water vapour. The red dashed line represents the number of DFS obtained when all 8461 IASI channels are considered, at the clear-sky and overcast locations where the largest values of DFS are achieved for the maximum number of selected channels. The number of DFS for the channels selected at these locations are denoted by red dots, while the number of DFS for the channels selected at the locations where the largest values of DFS for temperature are achieved for the maximum number of selected temperature channels are denoted by green dots.

humidity vertical profile. A further set of 50 channels at each location were chosen this time among those with wavenumber between 1100 cm^{-1} and 2200 cm^{-1} (i.e., out of a total of 4399 IASI channels) to exclude the channels already selected for temperature as well as to avoid solar contamination and non-LTE effects. The number of DFS achieved by the selected channels are shown in Figure 8. Note that the two locations (one in clear-sky and one in overcast conditions) where the overall maximum number of DFS is achieved are different from the locations where the maximum number of DFS for temperature is captured. Note also that the maximum number of DFS with 150 selected channels in clear-sky (overcast) conditions is 74.93% (64.52%) of the 15.10 (21.48) DFS achieved when all 8461 IASI channels are considered, which is still only 8.2% (11.7%) of the value that would be necessary to achieve the ideal goal of a direct and error-less joint estimate of the whole 183-component state vector.

5.1 Channel selection dependence on the presence of cloud

It is now interesting to discuss the different channel selection results obtained in clear sky and in overcast conditions. Considering the cloud vertical distribution at overcast locations shown in Figure 2, which indicates that overcast conditions are reached below about 800 hPa, it is reasonable to expect that the most informative water vapour channels selected in overcast conditions have jacobians that are mainly different from zero above about 800 hPa. In clear sky, however, it is expected that the selected channel also provide information about humidity in the lower troposphere. In Figure 9 are shown the water vapour jacobians for the ten most informative humidity-sensitive channels at four selected clear sky and overcast locations. Figure 9 indeed confirms that in clear sky the selected channels can provide an estimate of water vapour mixing ratio over a wider vertical range, although the largest contributions to the total humidity DFS both in clear-sky and overcast conditions come from channels that are sensitive to water vapour in the middle and upper troposphere.

The IASI water vapour channel that, in combination with the previously selected 100 temperature channels, is mostly selected (over 28 out of 135 clear sky locations) to provide the largest number of DFS

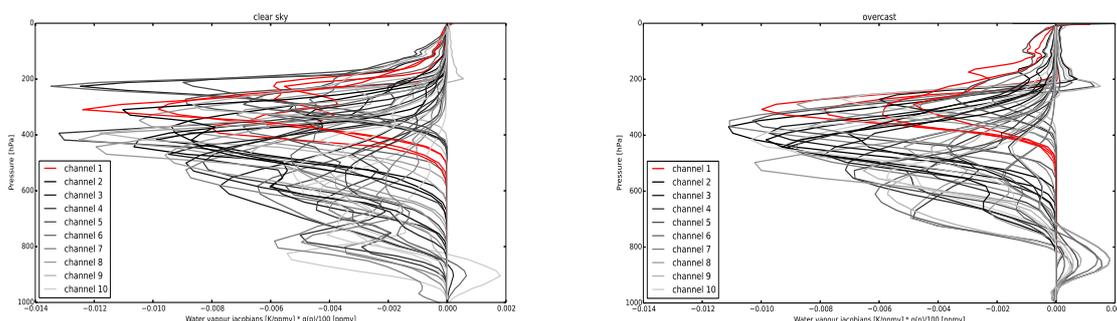


Figure 9: Water vapour jacobians – here defined as the vertical profiles of brightness temperature perturbations due to a 1% water vapour volume mixing ratio profile perturbation – for the ten most informative humidity-sensitive channels at four selected clear sky (left panel) and overcast (right panel) locations. In the clear sky case, the four locations are at $(73.46^{\circ}\text{N}, 146.67^{\circ}\text{W})$, $(33.08^{\circ}\text{N}, 25.0^{\circ}\text{E})$, $(16.26^{\circ}\text{S}, 120.38^{\circ}\text{E})$ and $(26.36^{\circ}\text{S}, 34.88^{\circ}\text{E})$, while in the overcast case at $(89.14^{\circ}\text{N}, 100.0^{\circ}\text{E})$, $(74.58^{\circ}\text{N}, 162.0^{\circ}\text{W})$, $(62.24^{\circ}\text{S}, 64.0^{\circ}\text{E})$ and $(68.97^{\circ}\text{S}, 149.33^{\circ}\text{W})$. Note that the first selected channel – whose jacobians at the four selected locations are denoted with a red solid line – is responsible for an average value over the four selected clear sky (overcast) locations of 39.1% (38.6%) of the total number of water vapour DFS when all the 50 selected channels are considered.

in clear sky is channel 3446 (centred at 1506.25 cm^{-1}), while in overcast conditions is channel 3244 (centred at 1455.75 cm^{-1}) to be mostly selected (over 25 out of 169 overcast locations). Note that IASI channel 3244 – whose water vapour jacobian when a 1% humidity mixing ratio perturbation is considered peaks at about 300 hPa – is also selected at 16 out of 135 clear sky locations as the most informative humidity-sensitive channel in clear sky and it is as well the most important water vapour channel selected during the “main run” in [5]. This confirms that IASI is most effective in estimating water vapour in the upper-middle troposphere even in clear sky conditions.

5.2 A strategy to select additional channels for all-sky data assimilation

From section 5.1 it follows that it is important to determine how many times the 50 channels selected at a given clear-sky (overcast) location – regardless their selection ranking – are also selected at the other 134 clear-sky (168 overcast) locations. Figure 10 shows the number of times (in percentage) that a given channel is selected at the considered clear-sky (overcast) locations relative to the total number of clear-sky (overcast) locations. The 24 humidity-sensitive channels that are selected over at least 40% of the clear-sky (overcast) locations and that are also selected over at least 40% of the overcast (clear-sky) locations are denoted in Figure 10 by red dots. These 24 channels – out of the 6750 (8450) non-unique channels that are selected at all the clear-sky (overcast) locations – accounting for about 3.5 DFS for humidity at a given clear sky location, populate the final channel selection shortlist. It is important to note that other criteria may be used to select the final shortlist of n'_{sel} humidity-sensitive channels. For example, it is possible to apply again the iterative channel selection procedure to the n_{comm} channels that are selected both at clear-sky and overcast locations (with $n_{\text{comm}} \geq n'_{\text{sel}}$) so as to pick the n'_{sel} common channels that provide the largest DFS increments. A difficulty with this strategy, however, is that the DFS calculation should be performed by considering the appropriate forecast error uncertainty at each relevant location and it may be difficult to account for the fact that not all common n_{comm} channels are selected at the same locations.

As discussed above, the importance of a selected channel depends on both its selection frequency f_i over the considered locations and the iteration step p_i (in %) in which the channel was selected – the earlier the

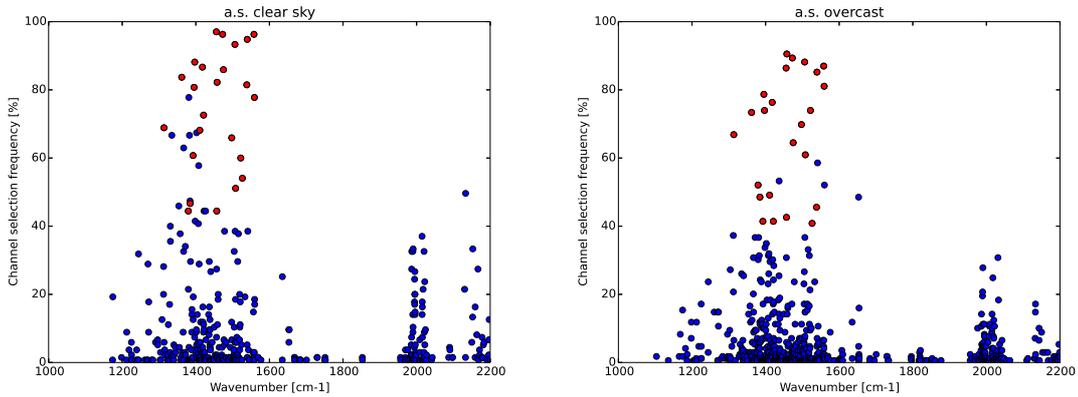


Figure 10: The left (right) panel shows the channel selection frequency at the 135 (169) clear-sky (overcast) locations. Selection frequency values are denoted with red dots for the 24 channels that are selected both in clear-sky and overcast conditions among the set of channels that are selected over at least 40% of all clear-sky or overcast locations.

step the more important the channel i , defined as $p_i = 100 \times (n_{\text{sel}} - l_i) / (n_{\text{sel}} - 1)$ where n_{sel} is the number of selected channels (in this case 50) and l_i is the iteration in which channel i was selected. For example, $l_i = 1$ means that channel i provides the largest increase in DFS on temperature and humidity given the previously selected i – in this case 100 – temperature sensitive channels, leading to a value of p_i given by $p_i = 100\%$. The ranking of the selected channel i is then calculated as the average between f_i and p_i . A final ranking q_i of the shortlisted channels (see Figure 11 and Table 1) can be obtained by calculating the weighted average between the ranking of each channel in clear sky and in overcast conditions, where the weights reflect the fact that the number of clear sky locations in general (as in this case) differs from the number of overcast locations in the considered case study. For sake of illustration, let us calculate step by step the final ranking for one of the shortlisted channels, i.e., channel 2675. This channel was selected at 68.89% (66.86%) of all clear-sky (overcast) locations and its iteration step percentage averaged over all considered clear-sky (overcast) locations was 67.46% (75.08%). This means that channel 2675 has a weighted average selection frequency given by $\bar{f}_i = (135 \times 68.89 + 169 \times 66.86) / (135 + 169)$ and a weighted average iteration step percentage given by $\bar{p}_i = (135 \times 67.46 + 169 \times 75.08) / (135 + 169)$ corresponding to a final ranking q_i for channel 2675 given by $q_i = (\bar{f}_i + \bar{p}_i) / 2 = 69.73\%$.

In Figure 12 are shown the humidity jacobians for the channels listed in Table 1 calculated using RTTOV v11 – with coefficients based on the LBLRTM line-by-line model over 101 vertical levels – at a non-isolated clear-sky location (i.e. surrounded by other clear-sky columns) over the Mediterranean sea at 36.45° N , 17.5° W . It is interesting to note that channel 3248, characterized by a humidity jacobian with the largest (in magnitude) peak at about 200 hPa, is the lowest ranking channel, presumably due to the signal-offsetting effects of its significant secondary peak with opposite sign in the stratosphere where the temperature gradient is positive. The relatively low importance of this channel may be used as an objective justification for excluding the channel from the shortlist, in addition to the practical consideration that the assimilation of channels with sensitivities over a wide range of heights may be more challenging due to imperfect knowledge of vertical forecast error correlations.

Finally, it is also interesting to calculate the DFS weighting functions for the selected 24 channels, which are shown in Figure 13 in the case when the atmospheric temperature and humidity standard deviations are and are not calibrated (see section 3). A comparison between Figures 12 and 13 shows that the width of the region where the jacobians of the selected channels are significantly different from zero (between

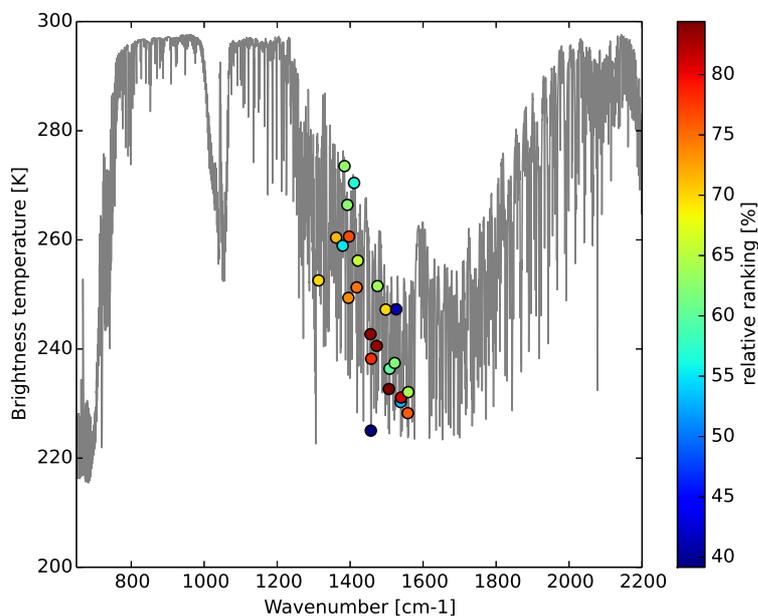


Figure 11: Spectral locations (indicated by coloured circles) of a set of 24 humidity-sensitive IASI channels as selected using the iterative, non-sequential channel selection procedure described in this work. The colour associated to the marker used to show the spectral location of channel i reflects the ranking value q_i for that channel (see text).

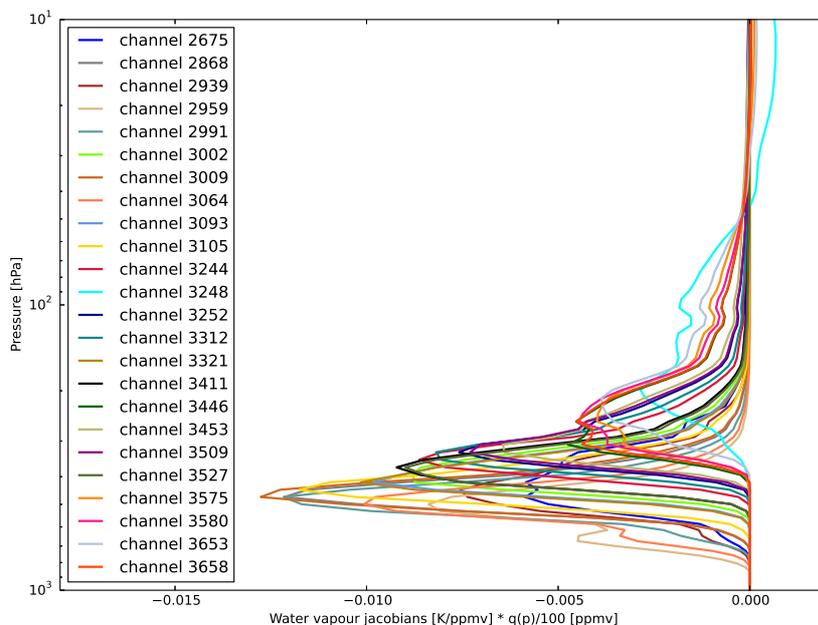


Figure 12: Humidity jacobians – here defined as the vertical profiles of brightness temperature perturbations due to a 1% water vapour volume mixing ratio profile perturbation – of the 24 selected channels listed in Table 1, at a clear sky location.

*Table 1: Humidity-sensitive IASI channels selected using the procedure described in the text. The IASI channel numbers in the left-most column shown in *italic (bold)* are currently operationally monitored (assimilated). See text for a definition of channel ranking.*

IASI channel number	wavenumber (cm-1)	channel ranking (%)	ranking order
2675	1313.5	69.73	12
2868	1361.75	71.50	10
2939	1379.5	55.15	21
2959	1359.5	62.90	16
2991	1392.5	62.55	17
3002	1395.25	73.75	9
3009	1397	76.63	6
3064	1410.75	56.39	20
3093	1418	74.70	8
3105	1421	65.72	13
3244	1455.75	83.36	2
3248	1456.75	39.16	24
3252	1457.75	78.31	5
3312	1472.75	83.25	3
3321	1475	63.87	15
3411	1497.5	70.04	11
3446	1506.25	84.39	1
3453	1508	59.73	19
3509	1522	61.98	18
3527	1526.5	41.16	23
3575	1538.5	53.11	22
3580	1539.75	80.71	4
3653	1558	75.76	7
3658	1559.25	64.21	14

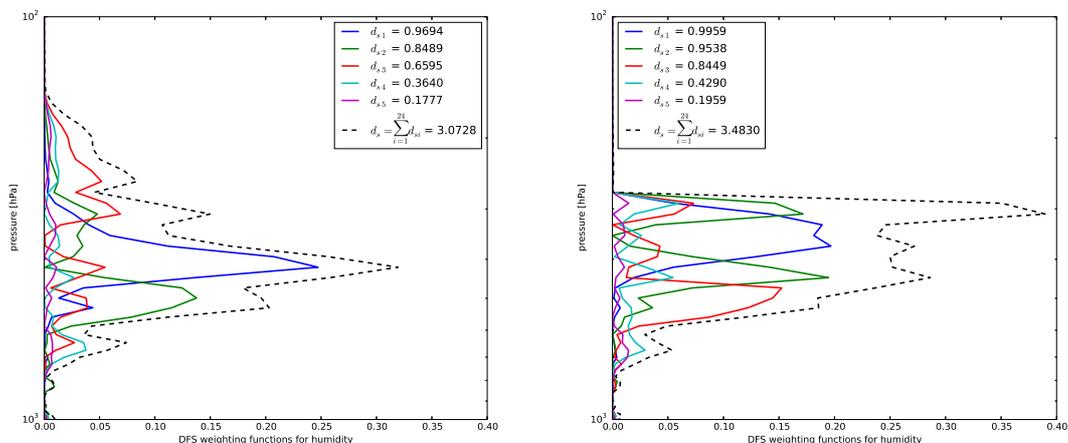


Figure 13: DFS weighting functions (see text for their definitions) for humidity and corresponding DFS values (for the whole state vector) achieved by the 24 selected channels listed in Table 1 at a clear sky location for the non-calibrated (left panel) and calibrated (right panel) forecast error standard deviation case. The black line shows the humidity portion of the cumulative DFS weighting function.

about 150 hPa and 800 hPa) coincides with that of the uncalibrated cumulative DFS weighting function. The calibrated cumulative DFS weighting function has also a similar dependence with height, although it becomes negligible above about 270 hPa due to both the rapid decline of the water vapour forecast error standard deviation with height and the effects of its calibration, which reduce the magnitude of the standard water vapour mixing ratio by more than 80% in the region above 290 hPa.

6 Summary and conclusions

In this study the iterative channel selection method, which is in standard use at operational meteorological centres to select an optimal subset of all available channels from advanced infrared sounding instruments for assimilation, was revisited in order to select channels with correlated errors (due to both apodization and interference from contaminant species) using a flow-dependent forecast error uncertainty both in clear-sky and overcast conditions. Also, the standard channel selection figure of merit, defined by the number of DFS expressed by the channels already selected plus that of an additional candidate channel, was modified so as to be able to be optionally used for selecting an optimal set of channels for estimation of a portion of the state space, e.g., tropospheric temperatures. To this end, the new concept of (cumulative) DFS weighting function was introduced, which can also be used to provide a synthetic, nondimensional and normalized picture of the region of the state space from which is possible to extract the (cumulative) contributions to the DFS expressed by a given set of channels. Note, however, that the “traditional” jacobians provide a measure of sensitivity of the radiation emerging from the atmosphere in a given channel to infinitesimal variations of the state, in a way that depends only on the characteristics of the instrument and on radiative transfer processes and not on those of the estimation system (i.e., on the observation and forecast error covariance matrices used for estimation).

The observation-error-correlation-aware channel selection method discussed in this paper was then used – in its standard figure of merit formulation – to select a set of 100 temperature-sensitive (below 900 cm^{-1}) and 50 humidity-sensitive (between 900 and 2200 cm^{-1} to avoid solar and non-LTE contamina-

tion) IASI channels, at a number of clear-sky and overcast locations for a case study in July 2012. Care was taken to select a final short-list of 24 humidity-sensitive channels from the set of humidity-sensitive channels that were selected both in clear-sky and overcast conditions over at least 40% of all considered locations. Finally, a ranking of the shortlisted channels was provided, based on their selection frequency and average selection iteration step. Future work will investigate the potential of an all-sky assimilation of (a subset of) the selected humidity-sensitive IASI channels on improving the ECMWF forecast skill scores over suitable case studies.

Acknowledgements

The author is partly funded by the NERC National Centre for Earth Observation. The author would like to thank F. Baordo, M. Bonavita, N. Bormann, S. English, R. Eresmaa, A. Geer, M. Hamrud, E. Holm, C. Lupu, M. Matricardi, T. McNally for their help, suggestions and comments.

References

- [1] E Anderson and M Fisher. Developments in 4D-Var and Kalman Filtering. *ECMWF Tech. Memo*, 247, 2001.
- [2] Peter Bauer, Alan J Geer, Philippe Lopez, and Deborah Salmond. Direct 4d-var assimilation of all-sky radiances. part i: Implementation. *Q. J. R. Meteorol. Soc.*, 136(652):1868–1885, 2010.
- [3] M. Bonavita, L. Isaksen, and E. Hólm. On the use of eda background error variances in the ECMWF 4D-Var. *Q. J. R. Meteorol. Soc.*, 138:1540–1559, 2012.
- [4] N. Bormann, A. Collard, and P. Bauer. Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. ii: Application to airs and IASI data. *Q. J. R. Meteorol. Soc.*, 136(649):1051–1063, 2010.
- [5] AD Collard. Selection of IASI channels for use in numerical weather prediction. *Q. J. R. Meteorol. Soc.*, 133:1977–1991, 2007.
- [6] AD Collard and AP McNally. The assimilation of infrared atmospheric sounding interferometer radiances at ECMWF. *Q. J. R. Meteorol. Soc.*, 135(641):1044–1058, 2009.
- [7] A. Dudhia, V.L. Jay, and C.D. Rodgers. Microwindow selection for high-spectral-resolution sounders. *Applied optics*, 41(18):3665–3673, 2002.
- [8] N. Fourrié and J-N Thépaut. Evaluation of the airs near-real-time channel selection for application to numerical weather prediction. *Q. J. R. Meteorol. Soc.*, 129(592):2425–2439, 2003.
- [9] G. Gaspari and S. E. Cohn. Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.*, 125:723–757, January 1999.
- [10] G.H. Golub and C.F. van Loan. *Matrix Computations*. The John Hopkins University Press, third edition, 1996.
- [11] T.M. Hamill, J.S. Whitaker, and C. Snyder. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, 129:2776–2790, 2001.

- [12] J. Hocking, P. Rayer, D. Rundle, R. Saunders, M. Matricardi, A. Geer, P. Brunel, and J. Vidot. RTTOV v11 users guide. *NWPSAF-MO-UD-028 v1.2*, 2013.
- [13] E.V. Holm and T Kral. Flow-dependent, geographically varying background error covariances for 1D-VAR applications in MTG-IRS L2 processing. *ECMWF Tech. Memo*, 680, 2012.
- [14] P.L. Houtekamer and H.L. Mitchell. Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, 126:796–811, 1998.
- [15] P.L. Houtekamer and H.L. Mitchell. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, 129:123–137, 2001.
- [16] P Martinet, N Fourrié, V Guidard, F Rabier, T Montmerle, and P Brunel. Towards the use of microphysical variables for the assimilation of cloud-affected infrared radiances. *Q. J. R. Meteorol. Soc.*, 2012.
- [17] M. Matricardi. The inclusion of aerosols and clouds in RTIASI, the ecmwf fast radiative transfer model for the infrared atmospheric sounding interferometer. *ECMWF Tech. Memo.*, 474, 2005.
- [18] M. Matricardi. A principal component based version of the RTTOV fast radiative transfer model. *Q. J. R. Meteorol. Soc.*, 136(652):1823–1835, 2010.
- [19] S. Migliorini. Information-based data selection for ensemble data assimilation. *Q. J. R. Meteorol. Soc.*, 2013.
- [20] P. Prunet, J-N Thépaut, and V. Cassé. The information content of clear sky IASI radiances and their potential for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, 124(545):211–241, 1998.
- [21] F. Rabier, N. Fourrié, D. Chafäi, and P. Prunet. Channel selection methods for infrared atmospheric sounding interferometer radiances. *Q. J. R. Meteorol. Soc.*, 128(581):1011–1027, 2002.
- [22] P. Räisänen. Effective longwave cloud fraction and maximum-random overlap of clouds: A problem and a solution. *Mon. Wea. Rev.*, 126(12):3336–3340, 1998.
- [23] C.D. Rodgers. Information content and optimization of high-spectral-resolution measurements. *SPIE's 1996 International Symposium on Optical Science, Engineering, and Instrumentation*, pages 136–147, 1996.
- [24] C.D. Rodgers. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific Publishing Co. Pte. Ltd., 2000.
- [25] L. Ventress and A. Dudhia. Improving the selection of IASI channels for use in numerical weather prediction. *Q. J. R. Meteorol. Soc.*, 2013.