Passive microwave radiometer channel selection based on cloud and precipitation information content

By SABATINO DI MICHELE* and PETER BAUER
European Centre for Medium-Range Weather Forecasts, Reading, UK

SUMMARY

The information content of microwave frequencies between 5 and 200 GHz for rain, snow and cloud water retrievals over ocean and land surfaces was evaluated using optimal estimation theory. The study was based on large datasets representative of summer and winter meteorological conditions over North America, Europe, Central Africa, South America and the Atlantic obtained from short-range forecasts with the operational ECMWF model. The information content of rain, snow and cloud water was traded off against the uncertainties due to the natural variability of other variables that microwave observations are sensitive to. These are surface emissivity, land surface skin temperature, atmospheric temperature and moisture. The estimation of the underlying error statistics was based on ECMWF model forecast error statistics.

The results suggest that a number of frequency bands are most suited for the retrieval of

(i) rain over oceans: 15–18, 35–40, 80, 145, 118.75 ± 10–14 GHz; rain over land: 85–100, 135–140 GHz,
(ii) snow over land and oceans: 95–100, 140–150, 187 GHz,
(iii) clouds over oceans: 40, 80–85 GHz; clouds over land: 90–100, 135–140 GHz.

For radiometers designed for global and multi-season applications, several channels in all of the above frequency ranges would be desirable for optimizing channel usage in hydrometeor retrievals depending on the observed situation.

KEYWORDS: Precipitation remote sensing

1. INTRODUCTION

Optimal estimation theory is widely used for the inversion of radiometric observations from space to retrieve physical state variables related to the earth's atmosphere and surface. With the advent of infrared radiometer technology applying spectroscopic principles, instruments have been launched that provide a large number of channels over a wide spectral range. Several studies were carried out to optimize the usage of these data by reducing the number of channels and by maintaining the maximum possible information content at the same time (Rabier *et al.* 2002). Iterative 'selection' methods, first introduced by Rodgers (1996), turn out to be both practicable and efficient. Fourrié and Thépaut (2003) compared these methods against empirical techniques in the context of the Atmospheric Infrared Sounder (AIRS), concluding that a significant gain can be obtained when the optimum channel selection is employed.

Microwave frequencies are commonly used to remotely sense atmospheric temperature and moisture but also clouds and precipitation as well as surface properties. However, only a few studies are available that assess optimal channel configuration based on sufficient statistics or that account for the simultaneous variability of the atmosphere, clouds, precipitation and surface emission. Lipton (2003) applied a framework similar to that presented here to clear-sky applications, i.e. temperature and water vapour sounding. He also compared the optimal channels with those selected for present and future satellite instruments such as the Special Sensor Microwave Imager Sounder (SSMIS) and the Atmospheric Technology Microwave Sounder (ATMS). English (1999) extended the information content estimation to land surfaces and based on datasets and forecast error statistics from numerical weather prediction (NWP) model output.

---

* Corresponding author: ECMWF, Shinfield Park, Reading RG2 9AX, UK.
e-mail: sabatino.di.michele@ecmwf.int

English also investigated the degradation of potentially undetected cloud liquid water on sounder performance. Prigent and Rossow (1999) included cloud water in a variational retrieval approach of integrated moisture over land. The two latter studies clearly identified the strong impact of surface emissivity uncertainties on information content and retrieval errors.

The present study further develops and applies the methodology for the purpose of cloud and precipitation remote sensing over various surfaces and assuming that the entire microwave spectrum between 10 and 200 GHz is available. This application is motivated by the potential active use of cloud and precipitation microwave observations in data assimilation systems (Moreau *et al.* 2003; Bauer *et al.* 2005, 2006a). Data assimilation requires an optimized use of observations for maximizing information but reducing computational cost.

Two quantities are usually used as figure of merit for the information content, namely the degree of freedom for signal (DFS) and the entropy reduction (ER), both reflecting the improvement given by each channel to the retrieval error. The first determines the number of independent observations given an overdetermined inversion problem (Purser and Huang 1993), while the latter measures the reduction of entropy in a posteriori probability distribution functions (p.d.f.s; Shannon and Weaver 1949).

Iterative methods loop over channels and sort them by decreasing information content, given a priori information (from the model background) and the ones previously selected. As alternatives to 'selection' methods (which work top-down from highest to lowest information content), 'deselection' methods (that work bottom-up) also exist (Sofieva and Kyrölä 2003). The latter seem to be less vulnerable to the non-uniqueness of solutions but are computationally more costly because the iteration must always be performed for the entire set of channels while selection methods may be truncated earlier.

In this paper, these concepts are applied to define the optimized usage of channels in the microwave portion of the spectrum. The basics of optimal estimation and the information content methodology employed in this study are outlined in section 2. The definition of the candidate microwave channels and the description of the atmospheric profiles dataset used to perform the channels selection are given in section 3. The results from this new application for sensor optimization are presented in section 4. Finally, the main conclusions are outlined in section 5.

2. METHODOLOGY

The methodology employed for evaluating the information content of spaceborne passive microwave observations of clouds and precipitation is based on the optimal estimation theory (e.g. Rodgers 2000). In this section, the basis of the forward and inverse problems will be briefly outlined, then the formulation used for the information content evaluation will be described.

(*a*) *Optimal estimation*

The problem of the inversion of observations from space is not fully constrained, in particular in the presence of clouds and precipitation. The application of statistical principles is therefore fundamental for solving the inverse problem. In the following, the state of the atmosphere (here hydrometeor content profiles to be retrieved) is denoted as a vector $\mathbf{x}$ and the multiple-channel observations from a theoretical microwave radiometer are contained in a vector $\mathbf{y}$. The statistical link between $\mathbf{y}$ and $\mathbf{x}$ is expressed

as conditional p.d.f.s, through the application of Bayes' theorem:

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}, \tag{1}$$

where $P(\mathbf{x}|\mathbf{y})$ is the posteriori probability of $\mathbf{x}$ when $\mathbf{y}$ is observed. $P(\mathbf{y}|\mathbf{x})$ is the probability of making observation $\mathbf{y}$ when $\mathbf{x}$ is present, while $P(\mathbf{x})$ and $P(\mathbf{y})$ are the a priori probabilities of $\mathbf{x}$ and $\mathbf{y}$, respectively. $P(\mathbf{x})$ is assumed to fully describe the a priori distribution of $\mathbf{x}$.

The physical link between $\mathbf{x}$ and $\mathbf{y}$ is described by the observation operator, H, that may be nonlinear. In our case, the observation operator only consists of a radiative transfer model. The modelled observations, $\mathbf{y}$, are:

$$\mathbf{y} = H(\mathbf{x}) + \epsilon \tag{2}$$

and ϵ summarizes observation errors (e.g. radiometer noise) and forward modelling errors (e.g. radiative transfer model uncertainties).

The estimation of $P(\mathbf{x}|\mathbf{y})$ is determined by the observations and an a priori estimate of the state, $\mathbf{x}_b$, as well as the error covariance matrices of that a priori state, $\mathbf{B}$, and the errors originating from observation and modelling, $\mathbf{R} = \mathbf{E} + \mathbf{F}$:

$$P(\mathbf{x}|\mathbf{y}) = \exp[-\tfrac{1}{2}\{\mathbf{y} - H(\mathbf{x})\}^{\mathrm{T}}\mathbf{R}^{-1}\{\mathbf{y} - H(\mathbf{x})\} - \tfrac{1}{2}\{\mathbf{x} - \mathbf{x}_b\}^{\mathrm{T}}\mathbf{B}^{-1}\{\mathbf{x} - \mathbf{x}_b\}]. \tag{3}$$

Superscripts '−1' and 'T' denote inverse and transpose matrices, respectively. This only holds if the distributions $\mathbf{y} - H(\mathbf{x})$ and $\mathbf{x} - \mathbf{x}_b$ are uncorrelated and have zero means as well as if $\mathbf{B}$ and $\mathbf{R}$ have Gaussian characteristics.

The information content estimation relies on the estimation of the additional contribution to retrieval accuracy provided by the observation. This means that the retrieval accuracy or the retrieval error have to be calculated. In general, this calculation assumes that the observation operators are linear. This is because the estimation of $\mathbf{x}$ requires finding the minimum of a cost function that contains the right-hand side of Eq. (3). The minimum is found using the first derivative of this term that contains the first derivative of the observation operator. This is rather simple because the linearity implies a quadratic cost function.

For a nonlinear H, the second derivative of H is needed whose actual computation may become difficult. Nonlinearity introduces additional errors but the calculated information content is only used as a measure of relative information content between channels rather than for absolute quantification. Therefore, we will assume that H is only weakly nonlinear and therefore that $\mathbf{x}_b$ is rather close to $\mathbf{x}$. Since we are not carrying out an actual retrieval this assumption is not critical. But even for retrieval studies, this assumption may be valid if, for example, a statistically consistent perturbation approach is applied (Bauer *et al.* 2005) that is based on small perturbations of the same order of magnitude as $\mathbf{x} - \mathbf{x}_b$ differences.

It can be shown (Rodgers 2000) that in the linear case the optimal analysis state $\mathbf{x}_a$ has the following expression:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\{\mathbf{y} - H(\mathbf{x}_b)\}, \tag{4}$$

with $\mathbf{H}$ and $\mathbf{H}^{\mathrm{T}}$ being the tangent-linear and adjoint of the observation operator, and $\mathbf{A}$ the analysis-error covariance matrix:

$$\mathbf{A} = \mathbf{B} - \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{H}\mathbf{B}. \tag{5}$$

(b)   *Channel selection*

The information content of an observation is a quantitative measure of the reduction of the estimation error produced by the observation. The improvement of estimation error depends on the sensitivity of the observation to the state and on the accuracy of the measurement and modelling, i.e. on $\mathbf{H}$ and $\mathbf{R}$, their respective magnitudes, and finally on the accuracy of the a priori information (through $\mathbf{B}$). This makes information content similar to the signal-to-noise ratio. Historically, there have been several formulations of information content (Rodgers 2000 and references therein). As previously mentioned, in our application the DFS and the ER are possible choices. The former estimates the number of independent pieces of information in a measurement vector that are related to the signal versus those that are related to the uncertainty and sensor noise. The latter calculates the probability of solutions in terms of entropy that has a maximum when all states have equal probability and that has a minimum if all states except one have zero probabilities. DFS is defined as the expectation value of the normalized difference between analysis state, $\mathbf{x}_a$, and a priori state:

$$\text{DFS} = \varepsilon\{(\mathbf{x}_a - \mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_a - \mathbf{x}_b)\} = \text{Tr}(\mathbf{I} - \mathbf{A}\mathbf{B}^{-1}). \tag{6}$$

The entropy reduction is defined as the difference between the entropy, S, of $P(\mathbf{x})$ and that of the posteriori probability $P(\mathbf{x}|\mathbf{y})$:

$$\text{ER} = S\{P(\mathbf{x})\} - S\{P(\mathbf{x}|\mathbf{y})\} = \frac{1}{2}\log_2\left(\frac{|\mathbf{B}|}{|\mathbf{A}|}\right), \tag{7}$$

where $|\ |$ indicates the determinant operator. In our framework of Gaussian error distributions, the right-hand sides of Eqs. (6) and (7) hold. The log with base 2 is usually chosen for expressing ER in units of bits. The actual channel selection is an iterative procedure proposed by Rodgers (1996), in which the contribution of each channel is sequentially quantified based on the hypothesis of error uncorrelated among channels, i.e. that $\mathbf{R}$ is diagonal.

Given a set of candidate channels, the iterative selection method starts with no channels selected and sequentially chooses the channel with the highest information content taking into account the information provided by previously selected channels. Equations (6) and (7) then require an update of the error covariance matrix $\mathbf{A}$ for each iteration (initially, $\mathbf{A}_0 = \mathbf{B}$). It is convenient to derive an expression for the decrease of DFS or ER for each iteration step, respectively. $\mathbf{H}$ is independent of iteration. For unification of geophysical units, the Jacobian matrix $\mathbf{H}$ can be normalized by the observation-error covariance matrix $\mathbf{R}$ before the procedure:

$$\mathbf{H}' = \mathbf{R}^{-1/2}\mathbf{H}. \tag{8}$$

The updated analysis error covariance matrix at each iteration step, i, can be calculated from the previous step, $i - 1$, using Eq. (5) by:

$$\mathbf{A}_i = \mathbf{A}_{i-1} - \frac{(\mathbf{A}_{i-1}\mathbf{h}')(\mathbf{A}_{i-1}\mathbf{h}')^{\mathrm{T}}}{1 + (\mathbf{A}_{i-1}\mathbf{h}')^{\mathrm{T}}\mathbf{h}'} \tag{9}$$

looping over all remaining channels. Column vector $\mathbf{h}'$ is equal to the row of $\mathbf{H}'$ that corresponds to the channel that is investigated. The difference of $\mathbf{A}$ between two iteration steps (Eq. (9)) determines the information gain or reduction of DFS and ER:

$$\delta\text{DFS} = \text{Tr}\{(\mathbf{A}_{i-1} - \mathbf{A}_i)\mathbf{B}^{-1}\} = \frac{(\mathbf{A}_{i-1}\mathbf{h}')^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{A}_{i-1}\mathbf{h}')}{1 + (\mathbf{A}_{i-1}\mathbf{h}')^{\mathrm{T}}\mathbf{h}'} \tag{10}$$

and

$$\delta\text{ER} = \frac{1}{2}\log_2\left(\frac{|\mathbf{A}_{i-1}|}{|\mathbf{A}_i|}\right) = \frac{1}{2}\log_2(1 + \mathbf{h}'^{\text{T}}\mathbf{A}_{i-1}\mathbf{h}'). \tag{11}$$

Generally, both ER and DFS produce very similar results. However, DFS requires the normalization by $\mathbf{B}^{-1}$ in every iteration step which makes the computational procedure less efficient. We therefore decided to use only ER in the following analysis. The iteration procedure may be terminated when the information content of additional channels reaches a certain threshold. This cut-off reduces computational cost but also affects the resulting distributions of ER when integrated over large datasets. For this purpose, a threshold value of 0.025 was chosen. This figure (which is usually reached after 30 iterations) roughly corresponds to a 5% decrease of the determinant of $\mathbf{A}$ when a new channel is added (Eq. (11)). This may be a sensible threshold given the assumptions in this study.

Depending on the application the radiometer is designed for, the different geophysical variables contributing to the measured radiation can be considered either sources of information or of uncertainty. In our case, the information content is analysed for clouds and precipitation. Therefore, atmospheric temperature $\mathbf{T}$, specific humidity $\mathbf{q}$, as well as surface emissivity ϵ, and skin temperature T_0 represent a source of 'geophysical uncertainty' from which the 'geophysical information' on clouds and precipitation has to be separated. Although the control vector includes all geophysical variables that contribute to the signal, ER is only expressed with respect to the hydrometeor contents. Radiometer receiver noise is accounted for as well, but is expected to be small compared to the geophysical uncertainty.

In our formulation, the control vector, $\mathbf{x}$, contains:

$$\left.\begin{aligned}
\mathbf{x} &= (\mathbf{x}_1, \mathbf{x}_2) \\
\mathbf{x}_1 &= (\mathbf{w}_R, \mathbf{w}_S, \mathbf{w}_L, \mathbf{w}_I) \\
\mathbf{x}_2 &= (\epsilon, T_0, \mathbf{T}, \mathbf{q}),
\end{aligned}\right\} \tag{12}$$

with profiles of hydrometeor contents of rain ($\mathbf{w}_R$), snow ($\mathbf{w}_S$), cloud liquid water ($\mathbf{w}_L$), and cloud ice ($\mathbf{w}_I$). The latter has been included but the sensitivity of channels between 5 and 200 GHz to cloud ice is rather weak so that cloud ice is not further analysed in this study. The corresponding block forms of $\mathbf{B}$, $\mathbf{H}$, and $\mathbf{A}$ respectively are:

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{22} \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} \mathbf{H}_1 & \mathbf{H}_2 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}, \tag{13}$$

with

$$\mathbf{B}_{11} = \mathbf{B}_{\mathbf{w}}, \quad \mathbf{B}_{22} = \begin{pmatrix} \sigma_\epsilon^2 & 0 & 0 & 0 \\ 0 & \sigma_{T_0}^2 & 0 & 0 \\ 0 & 0 & \mathbf{B}_{\mathbf{T}} & 0 \\ 0 & 0 & 0 & \mathbf{B}_{\mathbf{q}} \end{pmatrix}. \tag{14}$$

Error correlations between $\mathbf{x}_1$ and $\mathbf{x}_2$ in the background error $\mathbf{B}$ have been neglected. This maximizes the effect of the variance. Using the above block matrix notation in Eq. (5), the analysis-error covariances for $\mathbf{x}_1$ and $\mathbf{x}_2$ have the following expressions:

$$\left.\begin{aligned}
\mathbf{A}_{11} &= \mathbf{B}_{11} - \mathbf{B}_{11}\mathbf{H}_1^{\text{T}}(\mathbf{H}_1\mathbf{B}_{11}\mathbf{H}_1^{\text{T}} + \mathbf{H}_2\mathbf{B}_{22}\mathbf{H}_2^{\text{T}} + \mathbf{R})^{-1}\mathbf{H}_1\mathbf{B}_{11} \\
\mathbf{A}_{22} &= \mathbf{B}_{22} - \mathbf{B}_{22}\mathbf{H}_2^{\text{T}}(\mathbf{H}_2\mathbf{B}_{22}\mathbf{H}_2^{\text{T}} + \mathbf{H}_1\mathbf{B}_{11}\mathbf{H}_1^{\text{T}} + \mathbf{R})^{-1}\mathbf{H}_2\mathbf{B}_{22}.
\end{aligned}\right\} \tag{15}$$

This leads to the following recursive formulae, similar to Eq. (9):

$$
\left.\begin{aligned}
\mathbf{A}_{i,11} &= \mathbf{A}_{i-1,11} - \frac{(\mathbf{A}_{i-1,11}\mathbf{h}_1')(\mathbf{A}_{i-1,11}\mathbf{h}_1')^T}{1 + (\mathbf{A}_{i-1,11}\mathbf{h}_1')^T\mathbf{h}_1' + (\mathbf{A}_{i-1,22}\mathbf{h}_2')^T\mathbf{h}_2'} \\
\mathbf{A}_{i,22} &= \mathbf{A}_{i-1,22} - \frac{(\mathbf{A}_{i-1,22}\mathbf{h}_2')(\mathbf{A}_{i-1,22}\mathbf{h}_2')^T}{1 + (\mathbf{A}_{i-1,22}\mathbf{h}_2')^T\mathbf{h}_2' + (\mathbf{A}_{i-1,11}\mathbf{h}_1')^T\mathbf{h}_1'}.
\end{aligned}\right\}
\tag{16}
$$

As before, $\mathbf{h}_1'$ and $\mathbf{h}_2'$ refer to a row of the normalized Jacobian matrix containing the sensitivity of a channel to elements in $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively. δER can be evaluated from:

$$
\delta\mathrm{ER}_{i,11} = \frac{1}{2}\log_2\left(\frac{|\mathbf{A}_{i-1,11}|}{|\mathbf{A}_{i,11}|}\right) = \frac{1}{2}\log_2\left\{\frac{1 + (\mathbf{A}_{i-1,11}\mathbf{h}_1')^T\mathbf{h}_1' + (\mathbf{A}_{i-1,22}\mathbf{h}_2')^T\mathbf{h}_2'}{1 + (\mathbf{A}_{i-1,22}\mathbf{h}_2')^T\mathbf{h}_2'}\right\}.
\tag{17}
$$

The channel selection can be performed according to the information content of rain, snow, or cloud if in the previous expression $\mathbf{A}_{i-1,11}$ and $\mathbf{A}_{i,11}$ refer to the portion of the error covariance matrix that corresponds to the hydrometeor type of interest.

3. MODELS AND DATA

(a) Observation operator

The observation operator, H, only consists of a radiative transfer model that accounts for multiple scattering at microwave frequencies in clouds and precipitation (Bauer *et al.* 2006b). The model has been part of the Rapid Transmittance TIROS Operational Vertical Sounder (RTTOV; e.g. Saunders *et al.* 2001) since RTTOV version 8. Since RTTOV aims at applications in NWP, the code is designed for optimal computational efficiency and makes available the forward, tangent-linear, adjoint and k-versions of the model. The k-version refers to the code that can be used for calculating the model Jacobian matrices that contain the partial derivatives of brightness temperature due to perturbations in input parameters. Compared to a multiple-stream radiative transfer model, standard deviations of less than 0.5–1 K were obtained for channel frequencies of the SSMIS (Bauer *et al.* 2006b) which cover similar frequency bands as exploited here.

While the relative contribution of temperature and water vapour profiles to the total signal in cloud and precipitation is small, surface emissivity largely affects the observations in thin and semi-transparent clouds. For an ocean surface, emissivity is modelled according to the fast emissivity model FASTEM-2 (Ellison *et al.* 2003) that is part of RTTOV. Land surface emissivity is particularly difficult to simulate due to the complex interaction of electromagnetic radiation with soil, vegetation and snow cover as a function of a large number of unknown state variables. Therefore, emissivity climatologies produced from Special Sensor Microwave/Imager (SSM/I) observations and integrated NWP and satellite products (Prigent *et al.* 1997) were employed. The climatologies also contain information on temporal emissivity variability over one-month averaging periods. These were used for the error standard deviations, σ_ϵ, required in Eq. (14). Data between July 1992 and June 1993 were matched with the corresponding dates of the atmospheric model fields from the European Centre for Medium-Range Weather Forecasts (ECMWF) to ensure realistic surface conditions. The interpolation from SSM/I frequencies to the channel frequencies that are investigated are carried out with a parametric fit. Details of the emissivity calculation are given

in Bauer *et al.* (2005). Besides emissivity, the contribution of background to radiation also depends on surface temperature. For its uncertainty, σ_{T_0}, a reference value of 5 K was assumed. The radiometer noise was taken to be uncorrelated between channels with a constant value of 1.5 K and is accounted for in the normalization in Eq. (8). With a conically scanning radiometer in mind, a constant viewing angle of 53.1° has been chosen. Jacobian calculations have been performed considering only the vertically polarized component of the radiation.

(b)   Candidate channels

This study aims at a general evaluation of the suitability of frequencies between 5 and 200 GHz for hydrometeor profile retrievals. Therefore, the problem of frequency protection from intrusion by mostly commercial emitters is not included in the analysis. This will be targeted in a separate effort. Ideally, as many channels as possible could be chosen as candidates for the selection based on information content.

Microwave imager and sounder channels have finite channel bandwidths that originate from the best compromise between noise reduction and spectral resolution. The latter is mainly an issue for sounding channels so that spectral window channels usually have larger bandwidths. For example, the Advanced Microwave Sounding Unit (AMSU-A) channels between 50 and 55 GHz have bandwidths between 0.15 and 0.4 GHz while the SSM/I window channels' bandwidths range from 0.25 to 1.5 GHz. Assuming a channel bandwidth of 0.25 GHz that is constant throughout the spectrum, 780 channels are obtained. This number can be further reduced in those spectral regions where the gradient of atmospheric absorption is small. For simplicity, an analytical formula was applied that derives sampling density, $\delta\nu$, as a function of the gradient, δk, of clear-sky atmospheric absorption coefficient, k, relative to a reference profile:

$$\delta\nu = \delta\nu_{\mathrm{min}} + \frac{\delta\nu_{\mathrm{max}} - \delta\nu_{\mathrm{min}}}{\exp(\delta k/0.005)}, \qquad (18)$$

with $\delta\nu_{\mathrm{min}} = 0.25$ GHz and $\delta\nu_{\mathrm{max}} = 3.0$ GHz. This formula scales the spectral sampling between 0.25 and 3 GHz. The gradients of k were derived from finite differences over $\delta\nu = 0.02$ GHz windows. This reduces the total number of channels to 300 with the highest spectral resolution of 0.25 GHz in the area of strong absorption (22, 50–70, 114–123 and 168–195 GHz). Figure 1 shows the candidate channels as dots superimposed on the reference absorption spectrum. Note that all radiative transfer calculations have been carried out for monochromatic channels. Once the optimal frequency bands are identified, more focused computations can be carried out taking into account technical specifications that are proposed for specific instruments.

(c)   Model profiles

Profiles of temperature and specific humidity were extracted from short-range (6-hour) forecasts of the operational ECMWF model in January and July 2004. The atmospheric fields are representative for a $\approx$39 km resolution (truncation at wave number 511, $T_L 511$). The hydrometeor content profiles were obtained from linearized moist physics parametrizations of large-scale condensation (Tompkins and Janisková 2004) and convection (Lopez and Moreau 2005). The required input fields besides T and q are their tendencies, latent and sensible heat surface fluxes as well as surface wind stress on the local pressure grid that is set up on 60 levels as a function of surface pressure. The moist physics parametrizations produce profiles of cloud water and ice mixing ratio (in kg kg^{-1}) and rain and snow flux (in kg m^{-2}s^{-1}).
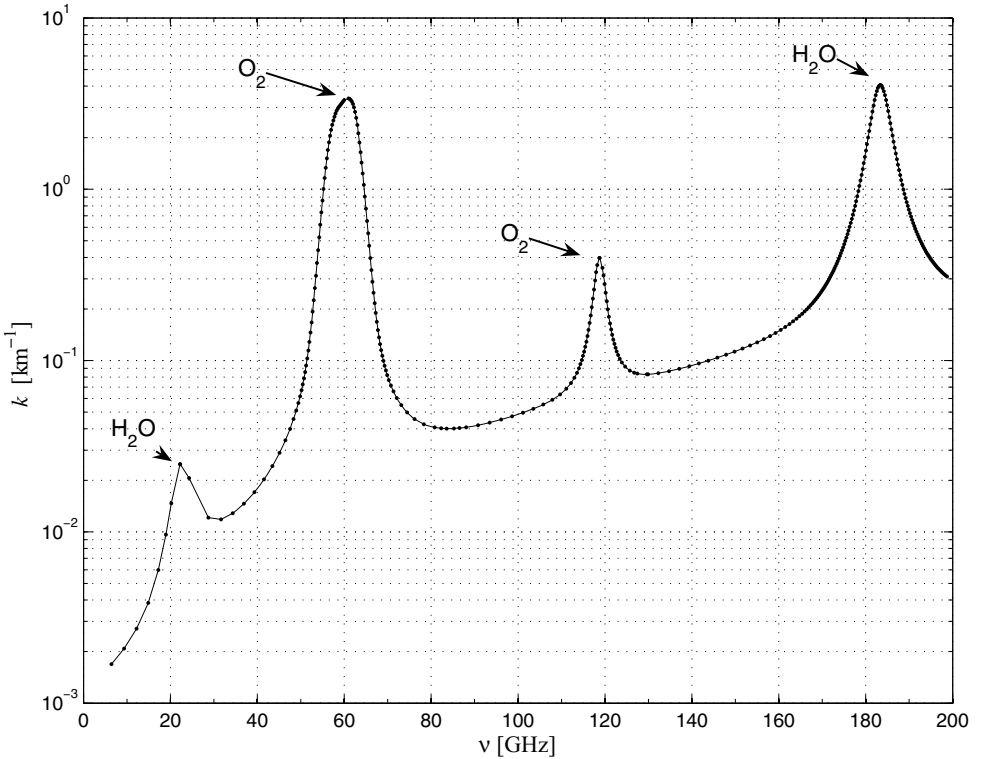
Figure 1.   Atmospheric absorption coefficient, k, as a function of frequency, ν, for the reference profile. Spectral locations of 300 candidate channels are superimposed as dots.

From the forecasts, only profiles were chosen for which rain or snow flux at any level exceeded 10^{-5} kg m^{-2}s^{-1}. For convenience and to comply with the input requirements for the radiative transfer model, all hydrometeor profiles were converted to hydrometeor densities (in g m^{-3}). This conversion employs assumptions on particle size distributions and particle density that were chosen according to the radiative transfer model set-up (Bauer 2001).

For the calculation of information content, the background-error covariance matrices for T, q and hydrometeor contents are required. For T and q the operational forecast-error fields were extracted that serve as background errors in the subsequent analyses ($\mathbf{B}_{T,q}$; Andersson *et al.* 2005). The corresponding errors of the hydrometeor profiles, $\mathbf{B_w}$, were calculated from:

$$\mathbf{B_w} = \mathbf{H_w}\mathbf{B}_{T,q}\mathbf{H_w^T}, \tag{19}$$

where $\mathbf{H_w}$ denotes the linearized moist physics parametrizations for which tangent-linear and adjoint versions have been implemented at ECMWF (Lopez and Moreau 2005).

Since this study aims at a globally valid channel selection, the sampled profiles have to be representative of both ocean and land surfaces as well as different seasons. Therefore the profiles collected from the model have been separated into the following four main categories according to their season and location:

(i)   WINTER–OCEAN (North Atlantic in January 2004),
(ii)   SUMMER–OCEAN (North Atlantic in July 2004),

TABLE 1.   SAMPLE SIZES OF PROFILE DATASETS

	Original	Redistributed	Representative
WINTER–OCEAN	133 829	46 766	10 000
SUMMER–OCEAN	72 992	31 420	25 000
WINTER–LAND	85 888	25 317	20 000
SUMMER–LAND	156 412	63 367	45 000

(iii) WINTER–LAND (North America, Canada and Europe in January 2004),
(iv) SUMMER–LAND (South America and Africa in July 2004).

Figure 2 shows the spatial occurrence of the profiles, while Table 1 presents the size of each sample.

   To avoid excessive computation, an optimum between sample size and representativeness was targeted. For large enough datasets, hydrometeor contents approach lognormal distributions which means lower amounts occur much more often than large amounts. The original samples were reduced in a first step by a random resampling into uniform distributions in order not to bias the results of channel selection towards profiles with less intense clouds and precipitation. The numbers of resulting profiles after this operation are shown in Table 1. Note that the data amount has been reduced to ≈40% of the original size. This reduction is more or less identical for each dataset. The datasets are further reduced in a second step to remove redundant profiles and maintain the statistical properties of the entire set at once. Profiles are deleted randomly until the variance reaches 95% of the original one. The variance has been calculated from the eigenvalues of the hydrometeor covariance matrix, thus all hydrometeor types are accounted for.

   In this step the reduction depends more strongly on the dataset than in the first processing step. This difference reflects the characteristics that are driven by the represented meteorological situations. WINTER–OCEAN is reduced most because less active and variable precipitation is observed while SUMMER–LAND is reduced least because more active convection is present and cloud morphology becomes more heterogeneous. The final number of profiles for each dataset is shown in Table 1. The resulting means and standard deviations of rain, snow, cloud water and ice contents per dataset are shown in Fig. 3. Standard deviations are usually twice as high as the means.

4.   RESULTS

(a)   *Example profiles*

   Figure 4 shows example profiles from different meteorological conditions over land and ocean surfaces. The land profile exhibits rather cold and dry conditions with minor amounts of rain and cloud water but quite significant snowfall. The ocean profile has a near-surface temperature of 300 K and substantial moisture producing convection. The snow amounts above the freezing level are very high and the near-surface rainrate is about 3.5 mm h^{-1}.

   Figure 5 shows the resulting δER calculations for rain, snow and cloud water relative to the land profile in Fig. 4. Since the surface emission contribution is rather high at spectral window frequencies and with less atmospheric opacity, δER increases with frequency for all parameters. This trend show gaps where the clear-sky atmospheric absorption becomes too strong near the centres of absorption lines. The best compromise for all parameters is obtained in the wings of the absorption lines near 118 GHz. This is because (i) $\mathbf{B_T}$ produces less signal than $\mathbf{B_q}$ and (ii) cloud and precipitation
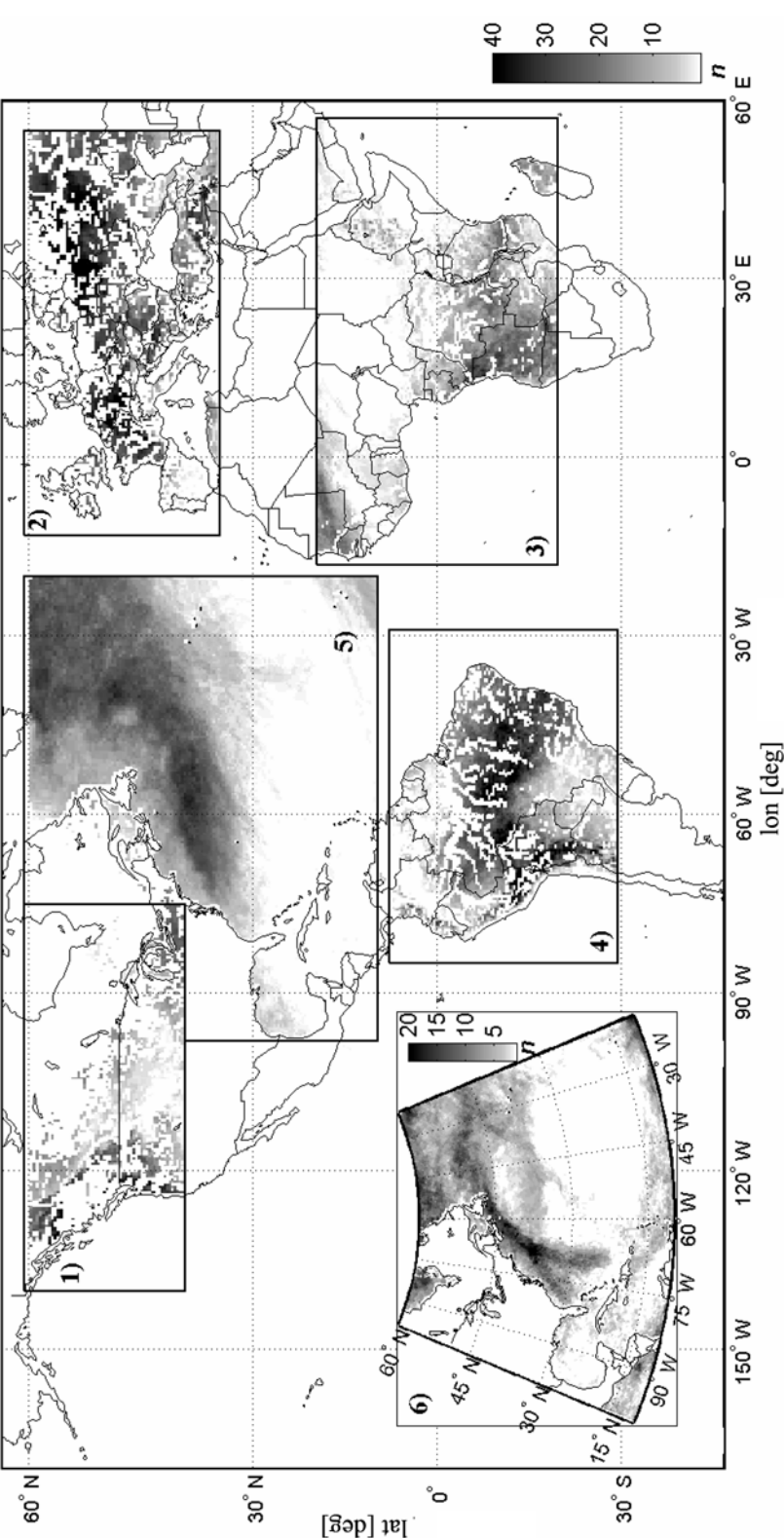
Figure 2.   Location and frequency distribution, n, of profile datasets chosen for the channel selection statistics. (1)+(2):WINTER–LAND, (3)+(4):SUMMER–LAND, (5):WINTER–OCEAN, (6):SUMMER–OCEAN.
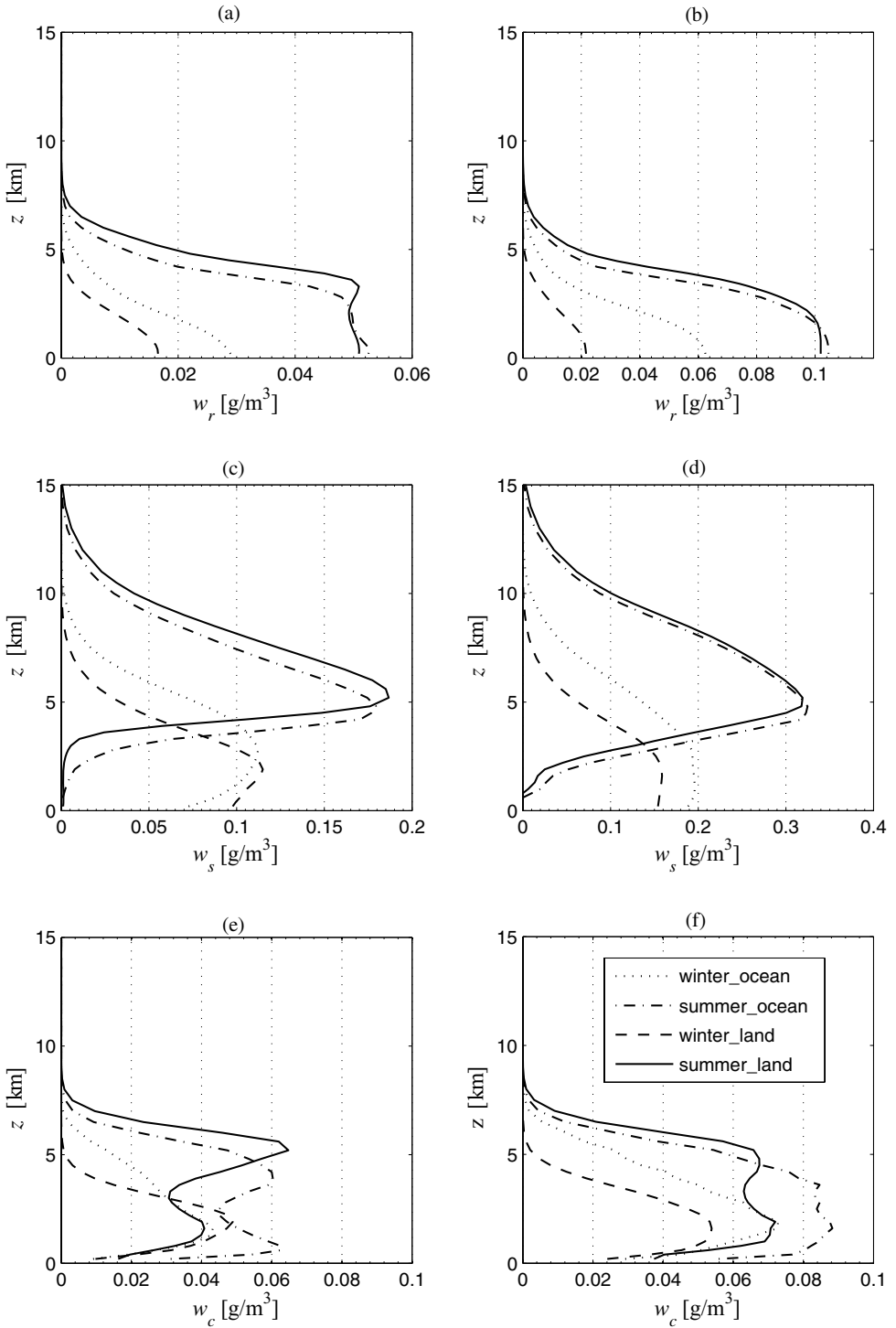
Figure 3. (a) Mean and (b) standard deviation profiles from representative datasets for rain. (c, d) and (e, f) are as (a, b) but for snow and cloud liquid water, respectively.
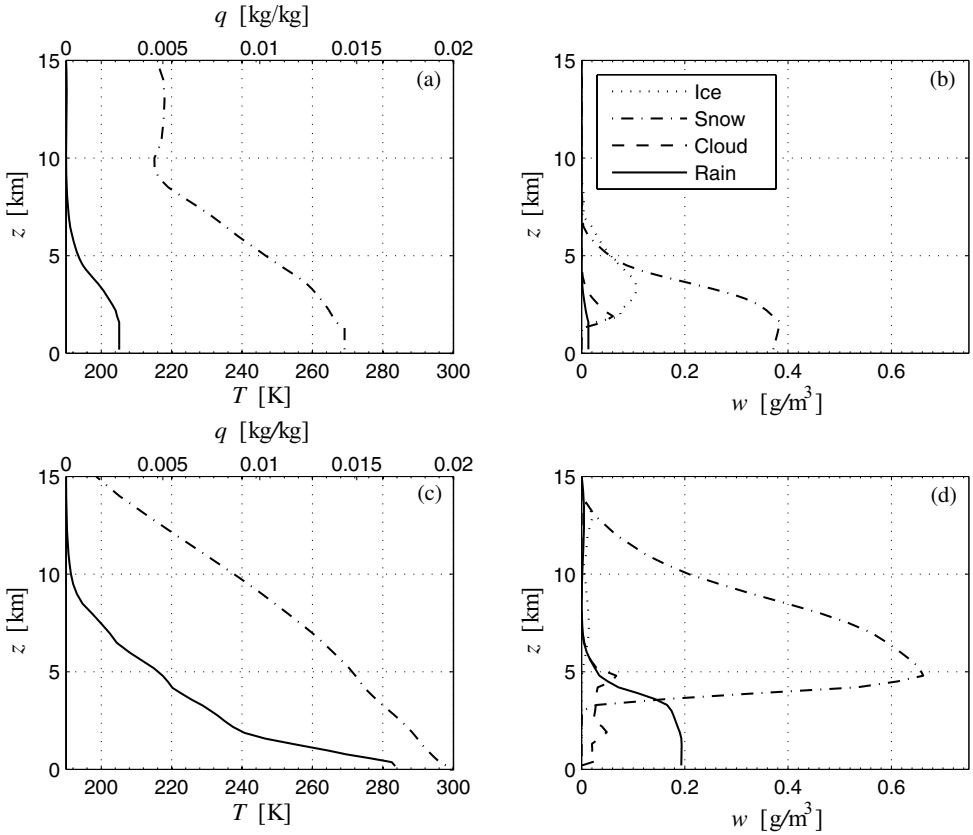
Figure 4.   Profiles from WINTER–LAND datasets of (a) temperature (dash-dotted) and specific humidity (solid), and of (b) corresponding rain, cloud, snow and ice. (c, d) are as (a, b), but from SUMMER–OCEAN datasets.

emission/scattering effects are stronger at 118 GHz than near 50–60 GHz. In this situation, both rain and cloud water δER values are rather low compared to those of snow. Another indication for the lower amount of information contributed by microwave observations is the smaller relative reduction of δER between iterations. For snow, δER values are much higher and the gradient of δER with iteration is larger.

Over the ocean (Fig. 6), δER decreases with frequency for rain because σ_ϵ is rather small and σ_{T_0} of little significance. This and the large signal dynamic range due to low surface emissivities produce the highest δER for rain near 15 GHz and for cloud water near 110 GHz. For snow, the spectrum is rather flat; at lower frequencies this is because the surface and clear-atmosphere contributions are weak, as is the sensitivity to snow, while at higher frequencies the contribution from geophysical uncertainties increases as does the sensitivity to snow itself. Finally, note the different magnitude between δER values of rain and snow compared to cloud water.

(b)   Full datasets

The mean δER as well as the 25%, median and 50% levels of the occurrence distributions for rain, snow and cloud water are displayed in Figs. 7, 8, and 9, respectively. For each frequency, values are evaluated only on those profiles having δER above the cut-off value.
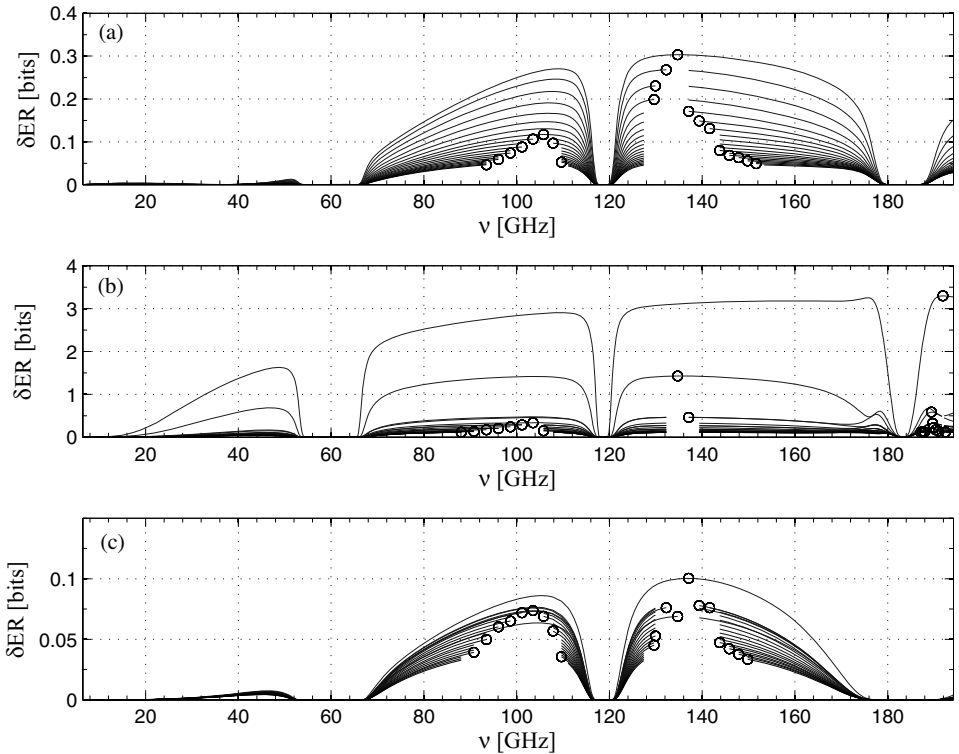
Figure 5. Entropy reduction for the single WINTER–LAND profile as a function of microwave frequency for (a) rain, (b) snow, and (c) cloud water. Individual lines refer to a single iteration, maximum values are highlighted as circles.

Figure 7(a) suggests that for rain over oceans in winter the largest information is provided by frequencies near 40, 140, 80 GHz (in decreasing order). The more intense precipitation in summer amplifies channel contributions near 10–15 GHz and shifts the peak from 40 to 37 GHz. As for the higher frequencies, channels near the oxygen absorption complex at 118.75 ± 10–14 GHz are preferred to the ones at 80 and 140 GHz (Fig. 7(b)). Over land, the channel selection for rain is also sensitive to season. In winter, one main peak for the median δER is located near 90 GHz, while wide portions of the spectrum between 30 and 40 GHz and between 130 and 160 GHz exhibit lower (similar) values. This result can be explained by the typical structure of rain profiles which have their maximum below 3 km (cf. Fig. 3), thus having their scattering signal hidden by the strong water vapour absorption. Also to be noted is a small response by sounding channels near 118.75 ± 4 GHz. The distribution in summer shows maxima near 118.75 ± 14 GHz, similarly to the ocean dataset. However, unlike the ocean case, the peaks around 52, 67, and 118.75 ± 4 GHz show how sounding channels can provide relevant information content in both oxygen absorption bands. The latter is explained by the reduced sensitivity to surface emission due to the more opaque atmosphere traded off against sufficient sensitivity to precipitation. Sounding channels near 183.31 GHz are less favourable because of the strong variability of background water vapour that is considered as geophysical uncertainty in our case.

For snow, sounding channels in the oxygen absorption bands are dominant in each of the four cases (Fig. 8). In winter, over both land and ocean, sounding frequencies near
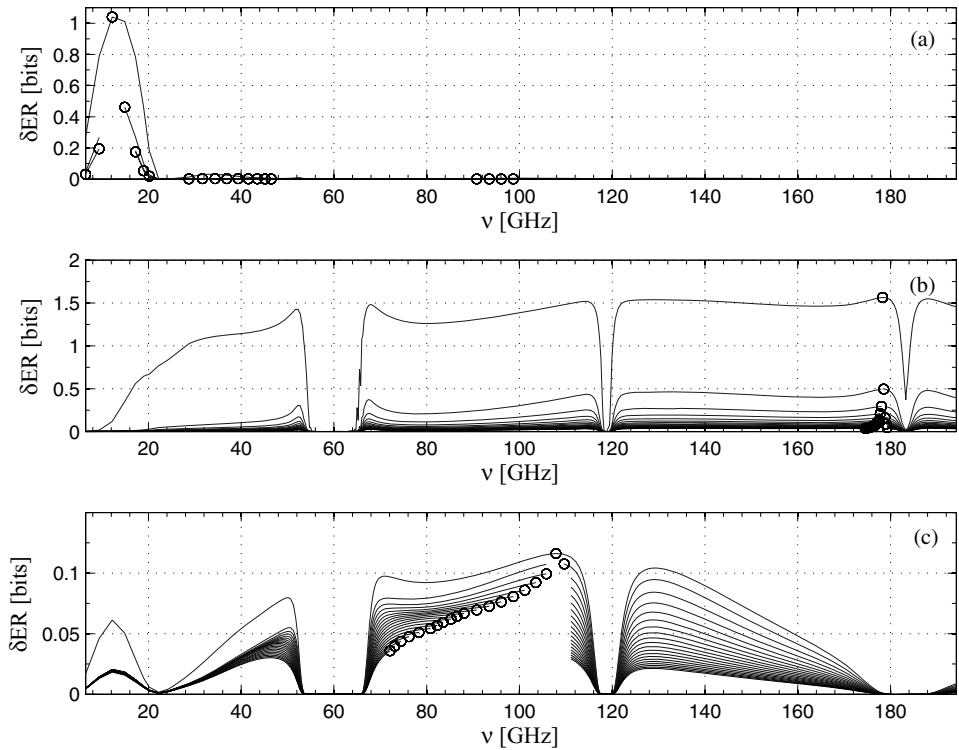
Figure 6.   As Fig. 5, but for the single SUMMER–OCEAN profile.

118.75 ± 5 GHz emerge, while window channels (around 90 and 145 GHz over ocean, near 100 and 140 GHz over land) and water vapour sounding (near 183.31 ± 4 GHz) exhibit much weaker secondary maxima. In summer (particularly over ocean surface), the peaks around 118.75 ± 5 GHz are weaker than for the winter cases. Frequencies near the 60 GHz absorption lines appear important with a maximum near 67 GHz. Secondary maxima observed in window bands for winter cases are also seen in summer, but in this case new ones around 15–19 GHz and 35–45 GHz appear.

The results for cloud water show very distinct peaks in the δER distributions near two or three frequency bands (Fig. 9). Over ocean, this is independent of season and the bands are located near 42 and 82–83 GHz while the secondary maxima are rather weak (Figs. 9(a), (b)). Over land, the lower frequency band δER is highly reduced due to the surface emission contribution. In winter (Fig. 9(c)), frequencies near 100 and 135 GHz show better information contributions. In summer (Fig. 9(d)), maxima are between 80 and 100 and between 135 and 150 GHz, but they are much weaker due to the saturation of channels by precipitation contributions.

(c)   *Channel recommendation*

The above results have to be analysed in conjunction with the number of profiles for which they provided significant information. Otherwise, calculating mean δER distributions may become misleading if only a few situations contribute to the mean with very high values of δER. This is quantified in Fig. 10, which shows the relative
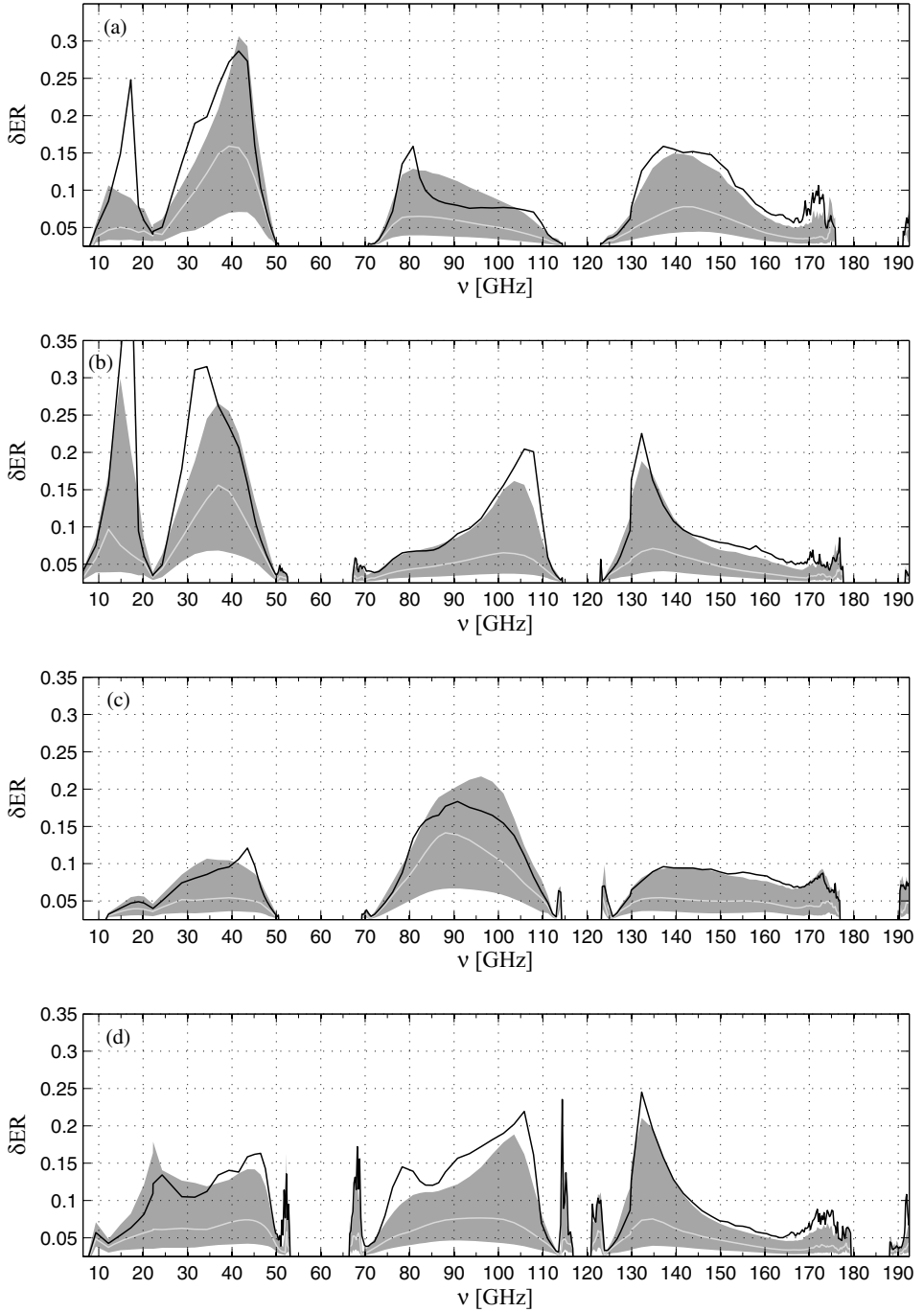
Figure 7. δER frequency distributions for rain from datasets (a) WINTER–OCEAN, (b) SUMMER–OCEAN, (c) WINTER–LAND and (d) SUMMER–LAND. Black lines denote the mean, white lines denote the median and grey-shaded areas are the 25% and 75% limits of the distribution.
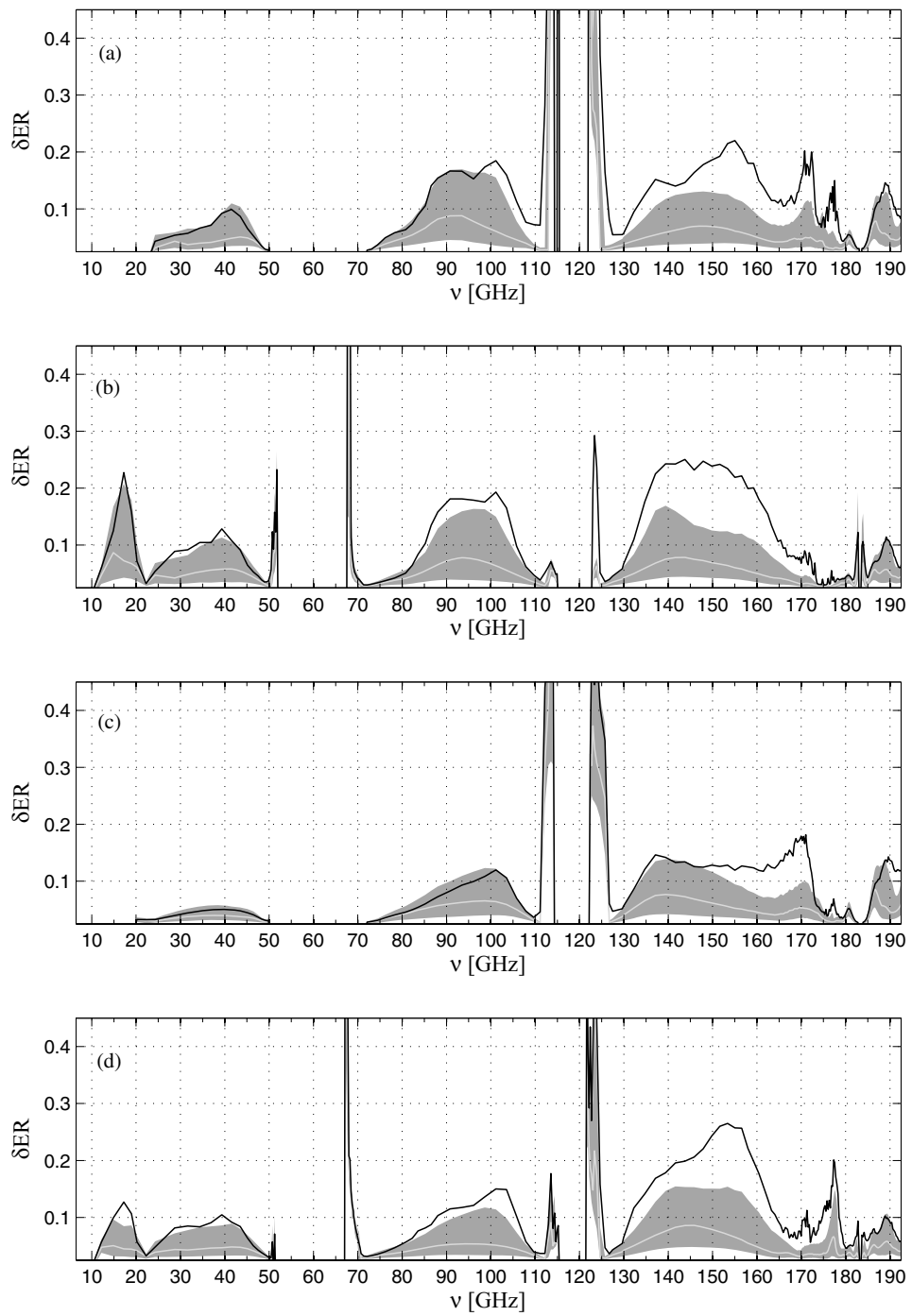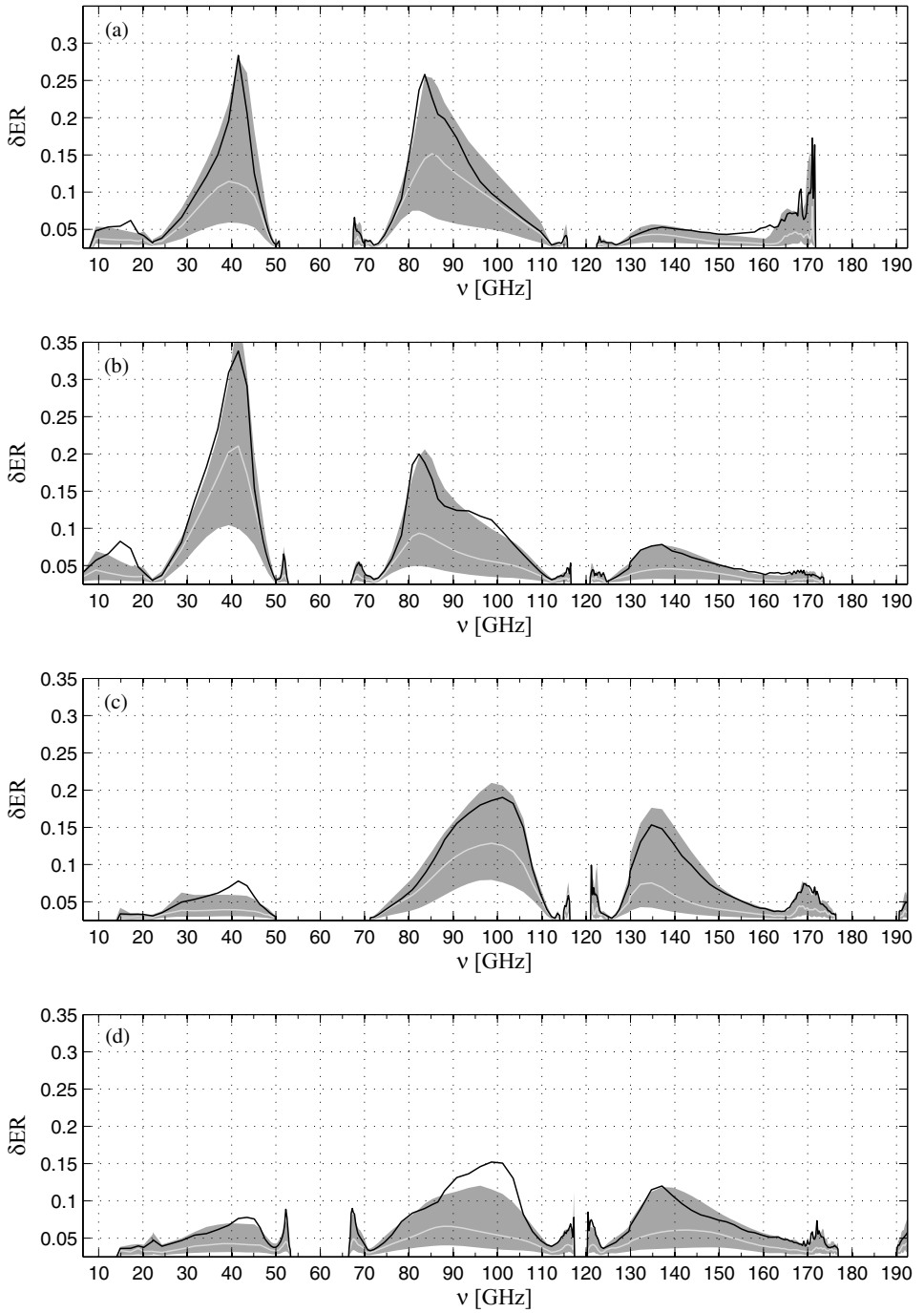
Figure 8.   As Figure 7, but for snow.

Figure 9. As Figure 7, but for cloud.

fraction of profiles, f, for which a channel is useful, i.e. δER remains above the cut-off threshold for rain, snow and cloud water. In most situations maxima correspond to window frequencies, while the fraction rapidly decreases approaching sounding bands.

For rain over oceans, frequencies between 30 and 40 GHz provide sufficient information in 80% of the cases (Figs. 10(a) and (b)). Other frequencies provide sufficient information only in a smaller fraction of the situations; in winter they are near 80 GHz (50%) and 145–155 GHz (45%), while in summer they lie near 110 GHz (45%) and 135 GHz (50%). This shift corresponds to a preference of window channels in winter and channels near the 118.75 GHz oxygen absorption line in summer. This can be related to saturation effects at window frequencies that are more sensitive to the integrated water column than sounding channels, which sense only the upper layers. Over land, this behaviour is less evident because sensitivity to changes in rain water is generally weaker (Figs. 10(c) and (d)). There, the lower frequencies are strongly affected by surface emissivity and lose their relative contribution of information content to the sample. Channels near 100 GHz dominate in 80% of the cases followed by channels near 135 GHz.

For cloud water over oceans, the relative number of cases is higher than for rain and reaches 90% at 80 GHz and 85% at 40 GHz in winter, as well as 90% at 40 GHz and 70% at 80 GHz in summer. The reduction at 80 GHz between seasons originates from signal saturation due to the higher occurrence of deep clouds. Over land, lower frequencies are penalized; in winter maxima are localized around 90–100 GHz (95%) and 130 GHz (80%). In summer the absolute maximum is at 85 GHz, while the secondary peak shifts to 135 GHz.

The results for snow are less dependent on season and surface because they are related to particle scattering rather than emission. In all four datasets, frequencies between 140 and 155 GHz dominate with significant information content contributions of 80% followed by the ones around 105 GHz with 40–60%.

Note that the frequencies that emerge in Fig. 10 represent the obvious choices that will contribute significant information in most cases. These figures cannot imply that (for example), if for 80% of the profiles the information content is sufficient, a retrieval is not possible in 20% of the cases. The relative contributions are complementary and therefore channel selection should account for situation-specific performances. In particular, sounding channels require a case-dependent analysis because for smaller data samples they were found to produce large information contributions.

The mean statistics of δER weighted with the number of cases for which it is above the minimum threshold are shown in Figs. 11–13. Besides a shifting of the maxima, in all cases the main change is the removal of peaks near the absorption lines. As mentioned before, this only means that these frequencies require a case-dependent analysis because they only provide information in fewer cases but may nonetheless be very important for specific applications. Except for rain over the ocean, another effect of the convolution is the reduction in importance of frequencies below 20 GHz.

Table 2 summarizes the most important frequencies obtained from selecting the maxima of median curves in Figs. 11–13. The frequency bands listed in Table 2 show a general agreement with the frequencies that are already in use for retrieving hydrometeors from radiometric observations, however the available channels do not necessarily agree with the optimal ones.

Rain over oceans is best observed with channels at 40, 80, 145 GHz in winter and 37, 15–18 and 118.75 ± 14 GHz in summer. This choice is represented by SSMIS channels at 19.35, 37.0, 85 and 150 GHz even though these channels are not optimal but comply with frequency protection. Over land, the lower frequencies become more
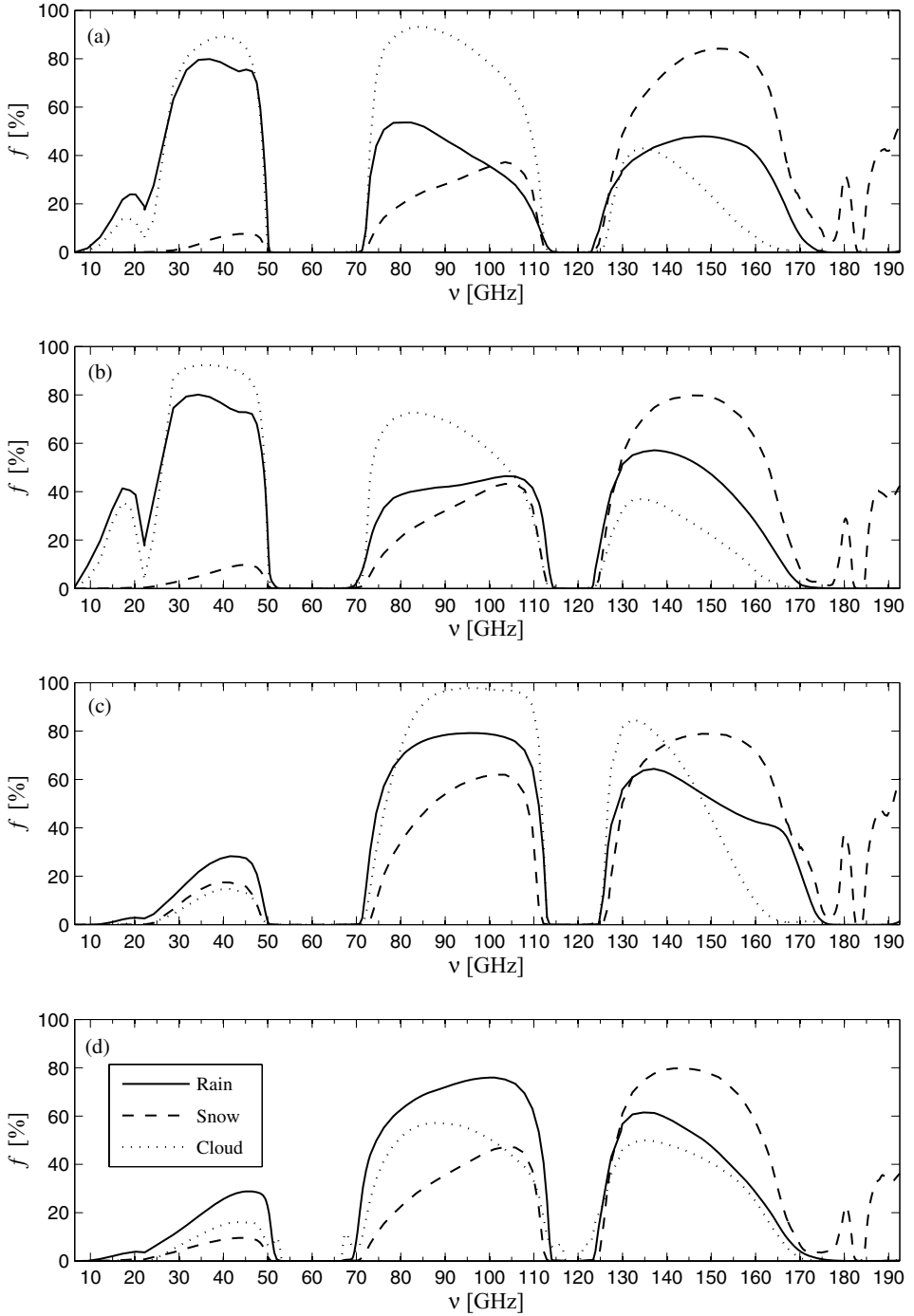
Figure 10. Fraction of profiles for which $\delta ER > 0.025$ per frequency for rain (solid), snow (dashed) and cloud water (dotted) for datasets (a) WINTER–OCEAN, (b) SUMMER–OCEAN, (c) WINTER–LAND and (d) SUMMER–LAND.
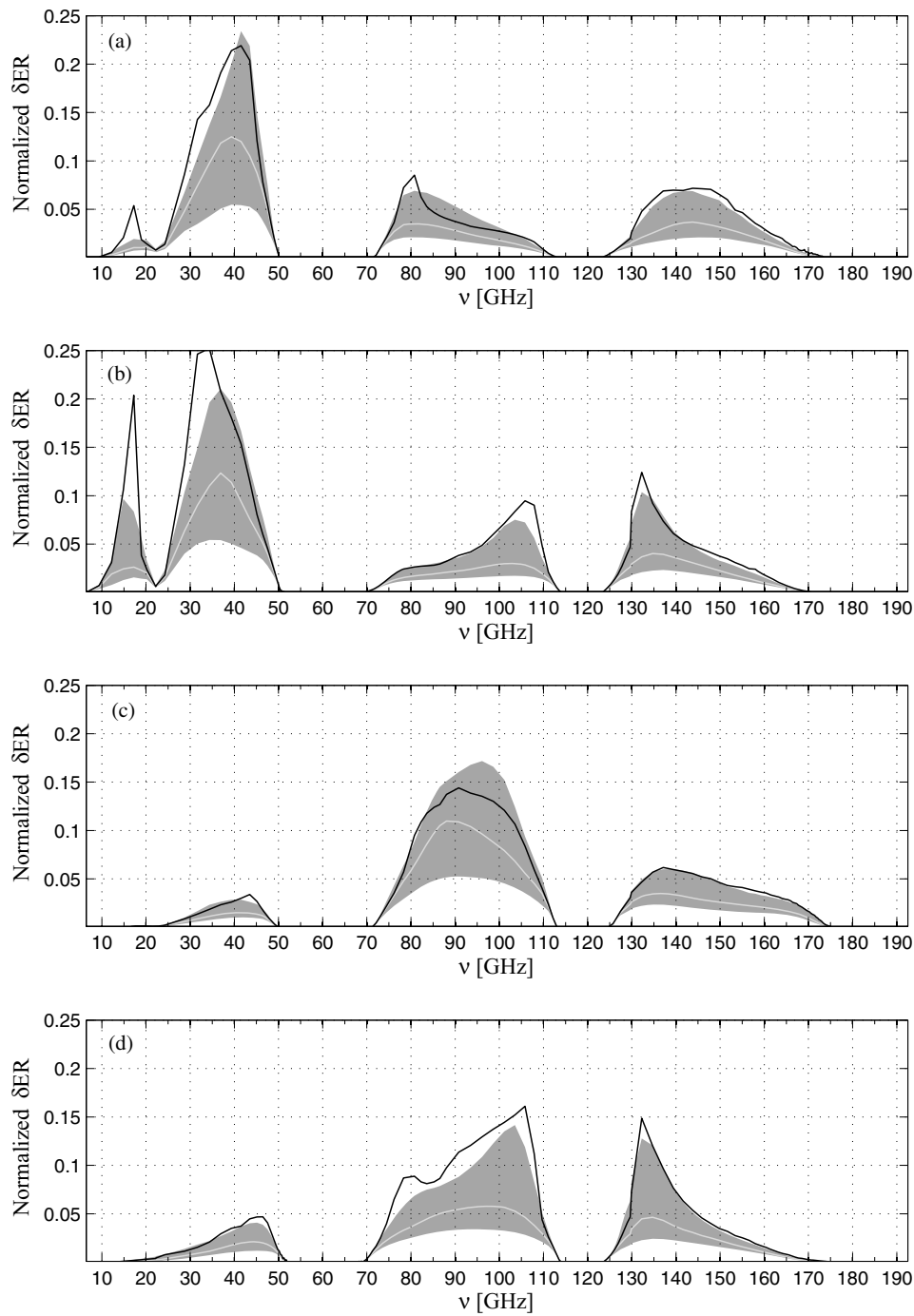
Figure 11. Frequency distributions of normalized $\delta ER'$ for rain from datasets (a) WINTER–OCEAN, (b) SUMMER–OCEAN, (c) WINTER–LAND and (d) SUMMER–LAND. Black lines denote the mean, white lines denote the median and grey-shaded areas are the 25% and 75% limits of the distribution.
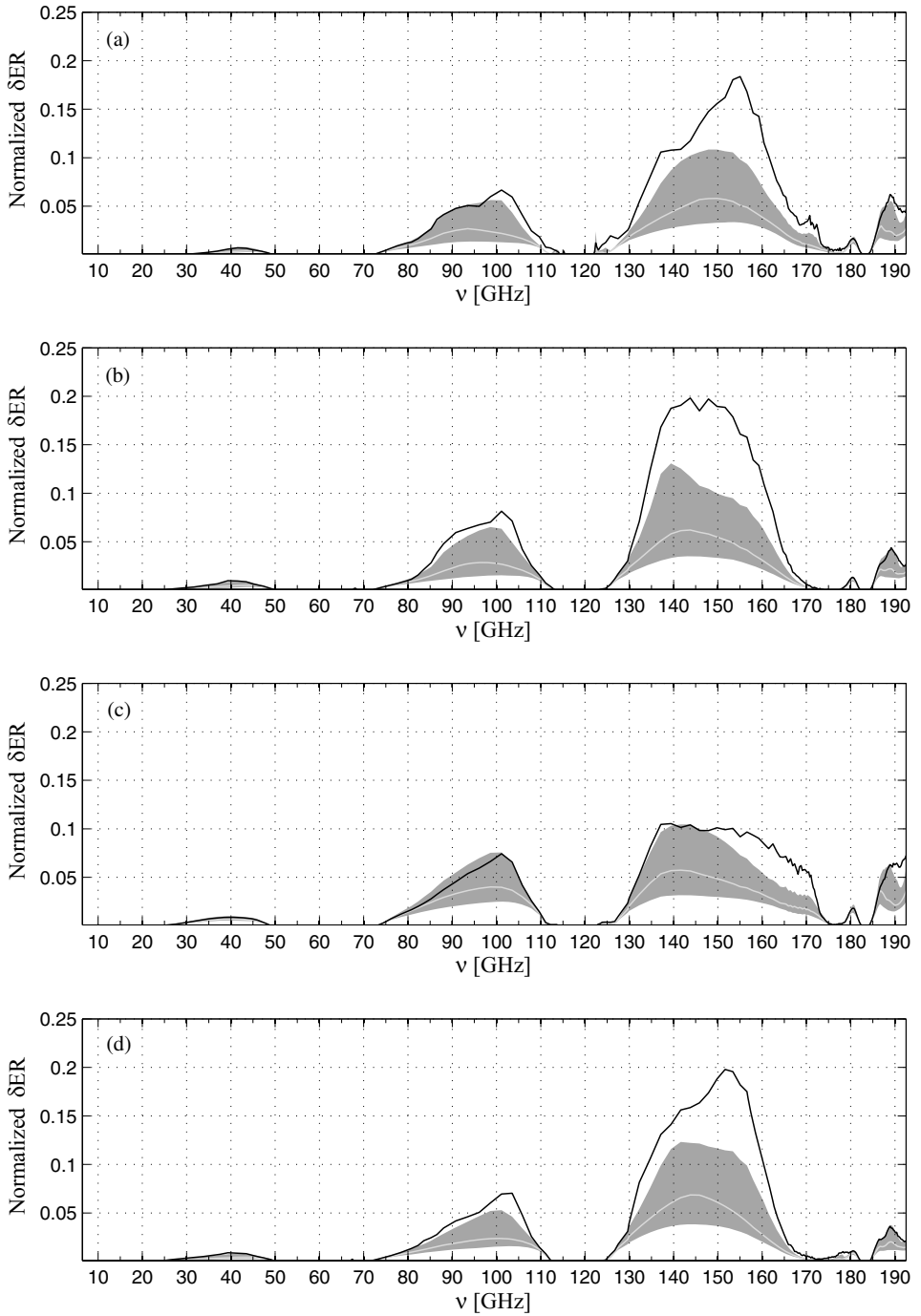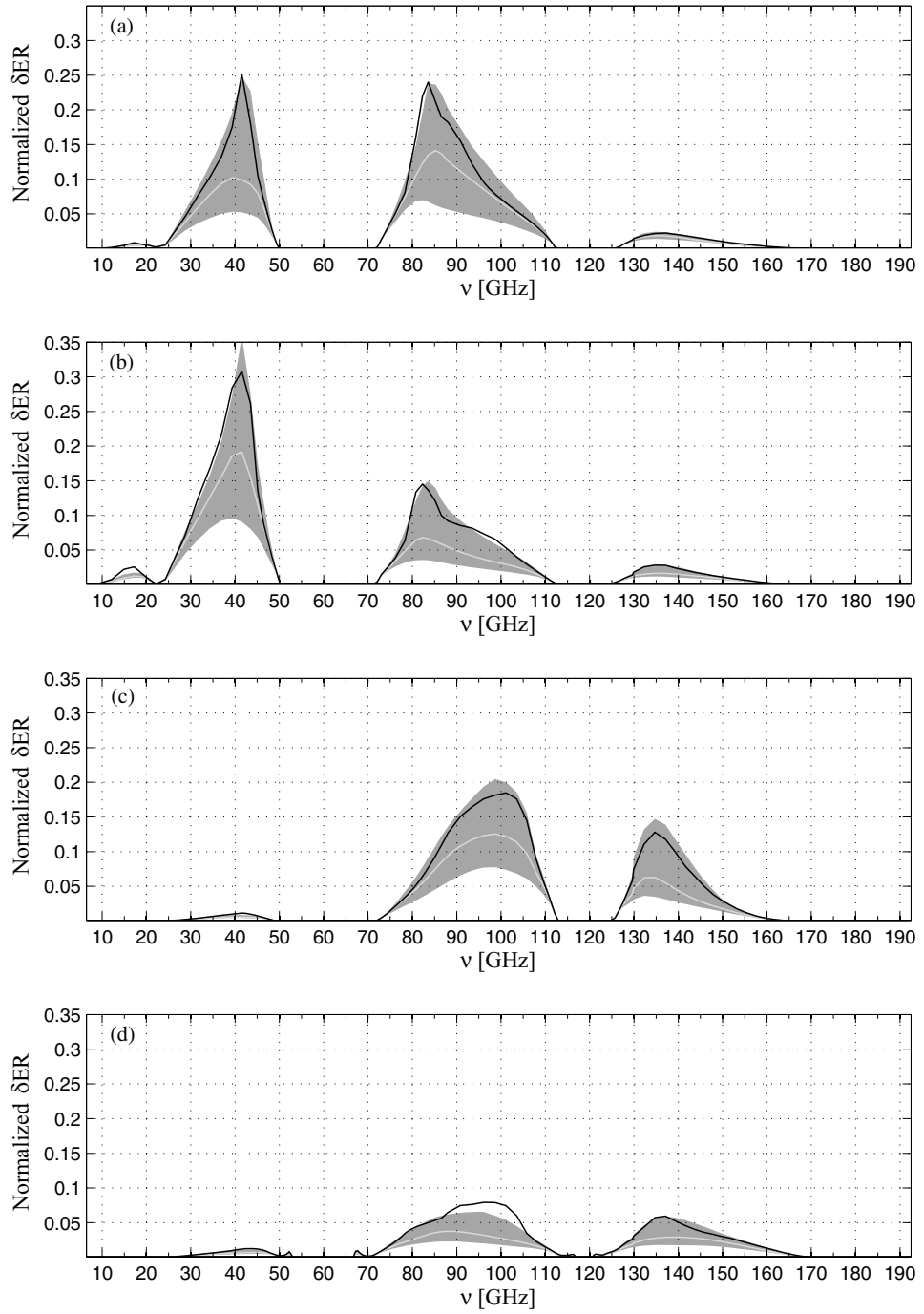
Figure 12.   As Figure 11, but for snow.

Figure 13.   As Figure 11, but for cloud.

TABLE 2. CHANNEL RECOMMENDATION BY PRIORITY AS A FUNCTION OF
DATASET AND RETRIEVAL PARAMETER (UNITS ARE GHz)

		Rain	Snow	Cloud water
WINTER–OCEAN	1	40	150	85
	2	80	90–100	40
	3	145	187	
SUMMER–OCEAN	1	37	140–145	40
	2	15–18	95–100	80–85
	3	118.75 ± 10–14		
WINTER–LAND	1	85–90	140	100
	2	135–140	100	130–135
	3		187	
SUMMER–LAND	1	90–100	145	85–90
	2	135	95–105	135–145

affected by uncertain surface emissivity, and channels between 85 and 100 (as first choice) and between 135 and 140 GHz (as second choice) dominate.

For *snow*, the picture is less dependent on surface and season. Distinct maxima are at 140–150, but channels at 95–100 GHz are also identified. In winter additional secondary maxima emerge near 187 GHz.

Cloud water is best observed at 40 and 85 GHz over oceans with no alternative channels at higher frequencies. Over land, frequencies at 90–100 and 135–140 GHz take over.

5. CONCLUSIONS

This is the first study investigating the information on hydrometeor contents contained in passive microwave measurements from satellites on a global and representative data basis and in a objective way.

An iterative method was developed for calculating the information content contributed by individual channels to an optimal analysis estimate of the atmospheric state. This method distinguishes between information on hydrometeor profiles that are to be retrieved and uncertainties contributed by unknown moisture and temperature profiles as well as land surface temperature and emissivity. Three hundred synthetic-channel radiometer channels were defined to sample the microwave spectrum between 5 and 200 GHz. ECMWF model output over several continental areas and from January and July 2004 provided sets of temperature, moisture and hydrometeor profiles as well as the associated background-error covariance statistics. For each profile, the potential channels have been ranked according to their information contribution. From the application of this method to the entire datasets, average performance statistics were calculated to identify those channels that are best suited for hydrometeor profile retrievals from microwave radiometer data. To investigate the general suitability for global application for the purpose of cloud and precipitation remote sensing, the number of profiles for which none of the potential radiometer channels produced significant contributions to potential hydrometeor retrievals has also been taken into account.

The results show that some of the 'classical' window channel frequencies (19, 37, 85 GHz) that were defined for existing and planned microwave radiometers appear very useful. However from our study the exact channel locations often differ depending on season and surface. An intelligent retrieval scheme would aim at recognizing the optimal channel combination for a specific situation because the less optimal channels may degrade the retrieval (Bauer *et al.* 2005).

For rain observations over oceans in winter, channels near 40 GHz seem stronger than channels at 36.5–37 GHz; for summer the optimum is shifted towards 35 GHz. In intense precipitation another choice is 18 GHz which is close to the available range of 18.7–19.35 GHz, and also channels that are located in the wings of the oxygen absorption line at 118.75 GHz. Over land, frequencies are obtained that are dominated by scattering of microwave radiation at hydrometeors, namely 85–100 and 135–140 GHz.

For snow, channels between 140 and 150 GHz show strong information contributions but temperature sounding channels (near 118.75 ± 5) are likely to perform better in individual situations (Bauer *et al.* 2005).

For cloud water, very strong information contributions were obtained for channels at 40 and 80–85 GHz over oceans as well as 90–100 and 135–140 GHz over land.

These results have to be interpreted as guidelines for the most sensitive wavelength bands rather than selection of individual channels. The determination of exact channel locations can be obtained taking into account detailed technical specifications of, for example, protected frequencies, channel bandwidths and detector noise. These issues will be targeted in a forthcoming study together with a quantitative estimation of hydrometeor retrieval accuracy.

The realism of the study can be enhanced with a better definition of modelling errors. For example, so far, identical model errors were assumed for all channels, but sounding-channel modelling errors are generally smaller than window-channel modelling errors due to the missing surface emissivity component. Also, a better specification of optimum polarization may introduce additional information at wavelengths where radiation exhibits significant sensitivity to this. For example, window-channel information content over oceans and some stronger polarizing land surfaces may increase if both polarizations are available.

REFERENCES

Andersson, E., Bauer, P., Beljaars, A., Chevallier, F., Hólm, E., Janiskovà, M., Kållberg, P., Kelly, G., Lopez, P., McNally, A., Moreau, E., Simmons, A. and Thépaut, J.-N.   2005   Assimilation and modelling of the hydrological cycle. *Bull. Am. Meteorol. Soc.*, **86,** 387–402

Bauer, P.   2001   Including a melting layer in microwave radiative transfer simulation for clouds. *Atmos. Res.*, **57,** 9–30

Bauer, P., Moreau, E. and Di Michele, S.   2005   Hydrometeor retrieval accuracy using microwave window and sounding channel observations. *J. Appl. Meteorol.*, **44,** 1016–1032

Bauer, P., Lopez, P., Benedetti, A., Salmond, D. and Moreau, E.   2006a   Implementation of 1D+4D-Var assimilation of microwave radiances in precipitation at ECMWF, Part I: 1D-Var. *Q. J. R. Meteorol. Soc.*, in press

Bauer, P., Moreau, E., Chevallier, F. and O'Keefe, U.   2006b   Multiple-scattering microwave radiative transfer for data assimilation applications. *Q. J. R. Meteorol. Soc.*, **132,** 1259–1281

Ellison, W. J., English, S. J.,
    Lamkaouchi, K., Balana, A.,
    Obligis, E., Deblonde, G.,
    Hewison, T. J., Bauer, P.,
    Kelly, G. and Eymard, L.
2003 A comparison of new permittivity data for sea water with AMSU, SSM/I and airborne radiometers observations. *J. Geophys. Res.*, **108**(D21)**,** 4663, doi: 10.1029/2002JD003213

English, S. J. 1999 Estimation of temperature and humidity profile information from microwave radiances over different surface types. *J. Appl. Meteorol.*, **38,** 1526–1541

Fourrié, N. and Thépaut, J.-N. 2003 Evaluation of the AIRS near-real-time channel selection for application to numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **128,** 2425–2439

Lipton, A. 2003 Satellite sounding channel optimization in the microwave spectrum. *IEEE Trans. Geosci. Remote Sens.*, **41,** 761–781

Lopez, P. and Moreau, E. 2005 A convection scheme for data assimilation: Description and initial tests. *Q. J. R. Meteorol. Soc.*, **131,** 409–436

Moreau, E., Bauer, P. and
    Chevallier, F.
2003 Variational retrieval of rain profiles from spaceborne passive microwave radiance observations. *J. Geophys. Res.*, **108**(D16)**,** 4521, doi: 10.1029/2002JD003315

Prigent, C. and Rossow, W. R. 1999 Retrieval of surface and atmospheric parameters over land from SSM/I: Potential and limitations. *Q. J. R. Meteorol. Soc.*, **125,** 2379–2400

Prigent, C., Rossow, W. B. and
    Matthews, E.
1997 Microwave land surface emissivities estimated from SSM/I observations. *J. Geophys. Res.*, **102,** 867–890

Purser, R. J. and Huang, H.-L. 1993 Estimating effective data density in a satellite retrieval or an objective analysis. *J. Appl. Meteorol.*, **32,** 1092–1107

Rabier, F., Fourrié, N., Chafaï, D.
    and Prunet, P.
2002 Channel selection methods for the Infrared Atmospheric Sounding Interferometer. *Q. J. R. Meteorol. Soc.*, **128,** 1011–1027

Rodgers, C. D. 1996 Information content and optimisation of high spectral resolution measurements. *Optimal spectroscopic techniques and instrumentation for atmospheric and space research II*, SPIE Proceedings, **2830,** Eds. P. B. Hays and J. Wang, 136–147

2000 *Inverse methods for atmospheric sounding. Theory and practice.* Series on Atmospheric, oceanic and planetary physics. World Scientific, Singapore

Saunders, R., Brunel, P.,
    Chevallier, F., Deblonde, G.,
    English, S. J., Matricardi, M.
    and Rayer, P.
2001 'RTTOV-7 science and validation report'. Forecasting and Research Technical Report No. 387, Met Office, Exeter, UK

Shannon, C. E. and Weaver, W. 1949 *The Mathematical Theory of Communications.* (repub. 1963) University of Illinois Press, Urbana, Illinois, USA

Sofieva, V. F. and Kyrölä, E. 2003 Information approach to optimal selection of spectral channels. *J. Geophys. Res.*, **108**(D16)**,** 4513, doi: 10.1029/2002JD002980

Tompkins, A. and Janisková, M. 2004 A cloud scheme for data assimilation: Description and initial tests. *Q. J. R. Meteorol. Soc.*, **130,** 2495–2518