TECHNICAL MEMORANDUM

475

Passive Microwave Radiometer Channel Selection Based on Cloud and Precipitation Information Content Estimation

Sabatino Di Michele and Peter Bauer

Research Department

Submitted to Q. J. Royal Meteor. Soc.

July 2005

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

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

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

Abstract

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 was traded off against noise that is mainly produced by geophysical variables such as surface emissivity, land surface skin temperature, atmospheric temperature and moisture. The estimation of the required error statistics was based on ECMWF model forecast error statistics. The results suggest a number of frequency bands most suited for the retrieval of (1) rain over oceans: 17-18, 32-45, 80, 135-150, 118.75±10-14 GHz; rain over land: 90, 135-140, 118.75±10-14 GHz; (2) snow over land and oceans: 135-155, 118.75±10-14, 195 GHz, (3) clouds over oceans: 42, 82-83 GHz; clouds over land: 90-100, 135 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 impact on hydrometeor retrievals depending on the observed situation.

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 this 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 they are employed. Microwave frequencies are commonly used to remotely sense atmospheric temperature and moisture but also clouds and precipitation as well as surface properties. With increasing computer power and improvement in data assimilation techniques, cloud and precipitation microwave observations have been applied for Numerical Weather Prediction (NWP) showing encouraging results (Moreau et al. 2003, Bauer et al. 2005a). Iterative methods loop over channels and sort them by decreasing information content given a priori information (from the model background) and that from the ones previously selected. Alternatively 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 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. Two quantities are usually used as figure of merit for the information content, namely the degree of freedom for signal and the entropy reduction, 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 (pdf, Shannon and Weaver 1949). 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 1. 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 conclusions from this new application for sensor optimization are finally presented in Section 4.

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, first the basis of the forward and inverse problems will be briefly outlined, then the formulation used for the information content evaluation will be exposed.

2.1 Optimal estimation

The problem of the inversion of observation 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 (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 $\mathbf{y}$. Their statistical link is expressed as conditional pdf's, through the application of Bayes' theorem:

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

$P(\mathbf{x}|\mathbf{y})$ 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 non-linear. In our case, the observation operator only consists of a radiative transfer model. The modelled observation, $\mathbf{y}$, is:

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

and ε summarizes observation errors (e.g. radiometer noise) and forward modeling 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, $\mathbf{R} = \mathbf{E} + \mathbf{F}$:

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

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 as well as $\mathbf{B}$ and $\mathbf{R}$ have Gaussian characteristics and zero means.

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 is greatly facilitated by having linear observation operators. This is because the estimation of $\mathbf{x}$ requires finding the minimum of a cost-function that contains the right-hand side of Equation (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 non-linear H, the second derivative of H is needed whose actual computation may become difficult.

Therefore, we will assume that H is only weakly non-linear 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. 2005b) 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:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{A}\mathbf{H}^{\mathbf{T}}\mathbf{R}^{-\mathbf{1}}[\mathbf{y} - H(\mathbf{x}_\mathbf{b})] \tag{4}$$

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

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

2.2 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 'degrees of freedom for signal', DFS, and the 'entropy reduction', ER, are possible choices. The former estimates the number of independent pieces of information in a measurement vector that are related to the signal vs. those that are related to the 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:

$$DFS = \varepsilon\{(\mathbf{x}_a - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}_a - \mathbf{x}_b)\} = tr[\mathbf{I} - \mathbf{A}\mathbf{B}^{-1}] \tag{6}$$

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

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

where $|\ |$ indicates the determinant operator. In our framework of Gaussian error distributions, the right-hand sides of Equations (6) and (7) hold. The log with basis '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 Equation (5) by:

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

looping over all remaining channels, respectively. 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 (Equation 8) determines the information gain or reduction of DFS and ER, called respectively:

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

and

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

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 and was thus chosen for reducing computational cost. In principal, depending on the application the radiometer is designed for, the different geophysical sources contributing to the measured radiation can be considered either information or noise. When the information content is analyzed for clouds and precipitation, the variability of atmospheric temperature $\mathbf{T}$, specific humidity $\mathbf{q}$, as well as surface emissivity ε, and skin temperature T_o represent the 'geophysical noise' from which the 'geophysical information' on clouds and precipitation has to be separated. Therefore for our purpose, although the control vector includes all geophysical variables that contribute to the signal, ER is only expressed with respect to the components relative to hydrometeor contents. Let's express the control vector, $\mathbf{x}$ as:

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

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 5-200 GHz to cloud ice is rather weak so that cloud ice is not further analyzed in this study. The corresponding block forms of $\mathbf{B}$, $\mathbf{H}$, and $\mathbf{A}$ respectively are:

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

with:

$$\mathbf{B}_{11} = \mathbf{B}_w \qquad \mathbf{B}_{22} = \begin{pmatrix} \sigma_\varepsilon^2 & 0 & 0 & 0 \\ 0 & \sigma_{T_o}^2 & 0 & 0 \\ 0 & 0 & \mathbf{B}_T & 0 \\ 0 & 0 & 0 & \mathbf{B}_q \end{pmatrix} \tag{14}$$

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 Equation (5) the analysis error covariances for $\mathbf{x}_1$ and $\mathbf{x}_2$ have the following espressions:

$$\begin{aligned} \mathbf{A_{11}} &= \mathbf{B_{11}} - \mathbf{B_{11}}\mathbf{H_1^T}(\mathbf{H_1}\mathbf{B_{11}}\mathbf{H_1^T} + \mathbf{H_2}\mathbf{B_{22}}\mathbf{H_2^T} + \mathbf{R})^{-1}\mathbf{H_1}\mathbf{B_{11}} \\ \mathbf{A_{22}} &= \mathbf{B_{22}} - \mathbf{B_{22}}\mathbf{H_2^T}(\mathbf{H_2}\mathbf{B_{22}}\mathbf{H_2^T} + \mathbf{H_1}\mathbf{B_{11}}\mathbf{H_1^T} + \mathbf{R})^{-1}\mathbf{H_2}\mathbf{B_{22}} \end{aligned} \tag{15}$$

The previous leads to the following recursive formulae, similar to Equation (9):

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

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 finally evaluated as:

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

3 Models and data

3.1 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. 2005c). The model is 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 temperatures 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 Special Sensor Microwave Imager Sounder (SSMIS; Bauer et al. 2005c) which cover similar frequency bands as exploited here.

While the relative contribution of temperature and water vapor profiles to the total signal in clouds and precipitations is small, surface emissivity largely affects the observations in thin and semi-transparent clouds. For an ocean surface, emissivity is modeled 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 electro-magnetic radiation with soil, vegetation and snowcover as a function of a large number of unknown state variables. Therefore, emissivity climatologies produced from 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 Equation (14). Data between July 1992 and June 1993 was matched with the corresponding dates of the atmospheric ECMWF model fields 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. (2005b). Beside emissivity, the contribution of background to radiation also depends on surface temperature. For its uncertainty (σ_{T_o}) a reference value of 5 K was fixed. The radiometer noise was assumed to be uncorrelated between channels with a constant value of 1.5 K and is accounted for in the normalization in Equation (8).

3.2 Candidate channels

This study consciously aims at a general evaluation of the suitability of frequencies between 5-200 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 done 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 50-55 GHz have bandwidths between 0.15 and 0.4 GHz while the Special Sensor Microwave / Imager (SSM/I) window channels' bandwidths range from 0.25-1.5 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 clear-sky atmospheric absorption coefficient k relative to a reference profile:

$$\Delta\nu = \Delta\nu_{min} + \frac{\Delta\nu_{max} - \Delta\nu_{min}}{exp\left(\frac{\Delta k}{0.005}\right)} \tag{18}$$

with $\Delta\nu_{min} = 0.25$ GHz and $\Delta\nu_{max} = 3.0$ 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.

3.3 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 ≈39 km resolution (wavenumber truncation at 511, T_L511). The hydrometeor content profiles were obtained from linearized moist physics parameterizations 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 parameterizations produce profiles of cloud water and ice mixing ratio (in $kg\,kg^{-1}$) and rain and snow flux (in $kg\,m^{-2}s^{-1}$). 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 ($\mathbf{B}_{T,q}$; Holm et al. 2002). 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 parameterizations 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 samples profiles have to be representative of both ocean and land surfaces as well as different seasons. The profiles collected from the model have been therefore separated into the following four main categories according to their season and location:

1. WINTER-LAND (North America, Canada and Europe in January 2004),
2. SUMMER-LAND (South America and Africa in July 2004),
3. WINTER-OCEAN (North Atlantic in January 2004),

4. SUMMER-OCEAN (North Atlantic 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 log-normal 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 toward the weakest profiles. The number of profiles resulting after this operation are shown in Table 1. We note that the data amount has been reduced to ≈40% of the original size. This reduction is more or less identical for each data set. The datasets so obtained are further reduced in a second step in order to remove redundant profiles because not enough different from the others. The method consists in randomly deleting profiles until the total variance reaches 95% of the original one. The variance has been calculated from the eigenvalues of the hydrometeor covariance matrix, respectively, thus including all hydrometeor types. In this step the reduction depends stronger on the dataset than in the first one where profiles have been redistribuited. This is interpreted as due to the profile variability which is generally larger over land and generally larger in summer. Therefore, WINTER-OCEAN is reduced most and SUMMER-LAND is reduced least in the second step. The final number of profiles for each dataset are reported in Table 1. Figure 3 shows the resulting mean profiles of all hydrometeors per dataset. The profiles indicate the wide range of cloud depths and precipitation intensities between ocean and land as well as summer and winter situations.

4 Results

4.1 Example profiles

Figure 4 shows two example profiles from different meteorological conditions over land and ocean surfaces, respectively. 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 freezing level are very high and the near surface rainrate is about 3.5 mm/h.

Figure 5 shows the resulting ER calculations for rain, snow and cloud water relative to the land profile. Since the effect of geophysical noise from the surface is rather strong at spectral window frequencies and with less atmospheric opacity, ER increases with frequency for all parameters. This trend show gaps where the clear-sky atmospheric absorption becomes too strong near the centers of absorption lines. The best compromise for all parameters is obtained in the wings of the absorption lines near 118 GHz. This is because (1) $\mathbf{B}_T$ produces less signal than $\mathbf{B}_q$ and (2) cloud and precipitation emission/scattering effects are stronger at 118 GHz than near 50-60 GHz. In this situation, both rain and cloud water ER-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 ER between iterations. For snow, ER-values are much higher and the gradient of ER with iteration is larger.

Over ocean (Figure 6), ER decreases with frequency for rain because σ_ε is rather small and σ_{T_o} of little significance. This and the large signal dynamic range due to low surface emissivities produces the highest ER for rain near 15 GHz and for cloud water near 110 GHz. For snow, the spectrum is rather flat because at lower frequencies the surface and clear-atmosphere contributions are weak but the sensitivity to snow as well while at higher frequencies the contributions from geophysical noise increase as does the sensitivity to snow itself. Note the different magnitude between ER-values of rain and snow compared to cloud water.

4.2 Full datasets

If all profiles are considered, the probability distribution functions of ER have to be compared to the total number of profiles for which each channel provided significant information. This allows the estimation of the relative fraction of situations where a channel is not useful. Otherwise, calculating mean ER distributions may become misleading if only a few situations contribute to the mean with very high ER's. For this purpose, a cut-off value for ER of 0.001 was chosen which is usually reached after 150 iterations. This threshold also reduces the computational burden significantly.

Figure 7 shows the resulting relative fraction of profiles for which ER remained above the threshold for rain, snow and cloud water, respectively. Over oceans and (Figure 7a, b) nearly 100% are obtained between 30 and 50 GHz for rain, between 70-116 GHz and 122-172 GHz for snow. Observations in all window channel spectral regions above 20 GHz are suited for cloud observations over oceans. Over land, the situation becomes worse for frequencies below 65-70 GHz due to the stronger contribution of surface emission. Over both surface types, water vapor absorption reduces the observational sensitivity to all hydrometeor types above 170 GHz. In this region about 10% of the situations will make hydrometeor content retrievals very difficult because the geophysical noise is too strong compared to the targeted signal.

These results have to be viewed in conjunction with the distributions of ER that were calculated for the fraction of profiles for which each channel's information contribution sufficed. The mean ER as well as the 25%, median and 50% levels of the frequency distributions for rain, snow and cloud water are displayed in Figures 8, 9, and 10, respectively. The recommendation for optimal channel selection must trade-off the fraction of active profiles and the resulting ER. This is because a channel should have large information contributions in almost all situations.

Figure 8a suggests that this is the case for rain at frequencies near 40-45, 80, 135-150 GHz over oceans (in decreasing order) in winter. The more intense precipitation in summer amplifies channel contributions near 17-18 GHz, shifts the peak from 40-45 to 32-35 GHz and also exhibits channels near the oxygen absorption complex at 118.75$\pm$10-14 GHz as a potential choice (Figure 8b). The latter is defined rather widely because, ideally, several sounding channels with different offsets from the line center at 118.75 GHz would be chosen to adjust to different situations. The shape of all distributions is skewed, with peaks at large values of ER and a long tail towards low values of ER. Over land, the channel selection for rain is also sensitive to season. The peaks of the mean ER values are located near 90 and 135-140 GHz in winter and near 118.75$\pm$10-14 GHz in summer. The latter distribution is very similar to the ocean data set in summer.

For snow, frequencies below 80 GHz show smaller information contribution in all seasons and over land and ocean (Figure 9). The strongest frequency bands are located in the window region between 120 and 170 GHz. Over oceans, the mean ER distribution peaks at 155 GHz in winter and between 140-150 GHz in summer. Over land there is a slight shift to lower frequencies (135-145 GHz) in winter but a peak near 150-155 GHz is obtained for the summer dataset. The strong gradients of ER with frequency suggest several channels between 135 and 155 GHz to cover the entire range. The secondary peaks are located near sounding channel bands at 118.75$\pm$10-14 GHz with a very high fraction of valid cases and near 195 GHz with slightly less valid cases (see Figure 7). The sounding channels at 118.75$\pm$10-14 GHz have a very persistent performance in all situations for snow.

The results for cloud water show very distinct peaks of the ER distributions near 2-3 frequency bands (Figure 10). Over ocean, this is independent of season and the choice would be at 42 and 82-83 GHz while the secondary maxima are rather weak (Figure 10a, b). Over land, the lower frequency band is again suppressed due to the surface emission contribution and frequencies near 100 and 135 GHz show strong information contributions. In summer, besides 90-95 GHz, also channels near 45 GHz present an option. This has to be viewed with

caution because the *ER*-values are generally much lower in this situation due to the saturation of channels with precipitation contributions (Figure 10c, d).

In summary, *rain* is best observed with channels near 17-18, 32-45, 80-90, 135-150, and 118.75±10-14 GHz. The exact channel locations and their hirarchy depend on season and surface. Besides the classical window channels, sounding channels located in the wings of the 118.75 oxygen absorption line show a strong information contribution, in particular in the summer datasets over both land and ocean. Note that the 17-18 GHz channels provide significant information in less than 90% of the cases as does the 80 GHz channels in winter over ocean. For *snow*, less channels are required. These should be mainly located between 135-155 GHz and 118.75±14 GHz. Again, the exact location depends on season and surface. For *cloud water*, channels near 42 and 82-83 GHz are very strong over oceans as are channels near 90-100 GHz over land. Table 2 summarizes these results.

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. The objective is to provide guidelines to future satellite radiometer developments for optimal channel selection as well as to optimize the usage of radiometric observations in data assimilation systems employed by NWP centres. Methodologies and datasets are based on the ECMWF model output and variational assimilation principles because they provide a self-consistent framework and globally representative view. The information content of microwave observations over land surfaces and for mid-to-high latitude weather regimes with less intense and mixed phase and frozen precipitation represented an important part of this study. This is because most of the currently available sensors are not optimized for these situations and therefore large terrestrial areas are underrepresented in global precipitation datasets.

ECMWF model output over several continental areas and from January and July 2004 provided sets of temperature, moisture and hydrometeor contents as well as the associated background error covariance statistics. The dataset's representativeness was ensured by covariance analyses from all meteorological parameters. Surface emissivity was obtained from model calculations over sea and from interpolated emissivity climatologies over land. Three hundred synthetic channels radiometer channels were defined to sample the microwave spectrum between 5 and 200 GHz. The spectral sampling was adjusted to be finer in regions where large absorption gradients occur in the vicinity of oxygen and water vapor absorption lines.

An iterative method for calculating the information content contributed by individual channels to an optimal analysis estimate of the atmospheric state was developed. This method distinguishes between information on hydrometeor profiles that are to be retrieved and noise from moisture and temperature profiles as well as land surface temperature and emissivity. For each profile, the potential channels can be ranked according to their information contribution. From the application of this method to the entire datasets, average performance statistics can be calculated and those frequency bands can be identified which are best suited for hydrometeor profile retrievals from microwave radiometer data. Another important information is the number of profiles for which none of the potential radiometer channels produced significant contributions to potential hydrometeor retrievals. This must be taken into account when drawing conclusions on the general suitability of microwave radiometer measurements for the purpose of cloud and precipitation remote sensing.

The results show that some of the 'classical' window channel frequencies that were defined for existing and planned microwave radiometers (AMSR-E, SSM/I, SSMIS, CMIS) appear very useful. However from our study, the exact channel locations often differ depending on season and surface. For rain observations over oceans in winter, channels near 40-45 GHz seem stronger than channels at 36.5-37 GHz; for summer the

optimum is shifted towards 32-35 GHz. In intense precipitation another choice is 17-18 GHz which is close to the range of 18.7-19.35 GHz.

Over land, frequencies are obtained that are dominated by scattering of microwave radiation at hydrometeors, namely 80, 135-155 GHz and channels that are located in the wings of the oxygen absorption line at 118.75 GHz. The latter is explained by the reduced sensitivity to surface emission due ot the more opaque atmosphere traded off against sufficient sensitivity to precipitation. This is effective over both ocean and land surfaces and most seasons. Sounding channels near 183.31 GHz are less favorable because of the strong variability of background water vapor that is considered as geophysical noise in our case. The far-wing sounding channels (118.75$\pm$10-14 GHz) also show strong contributions for snow retrievals as do channels between 140 and 150 GHz. For cloud water, very strong information contributions were obtained for channels at 42 and 82-83 GHz over oceans as well as 90-100 GHz over land.

The exact location of channels is difficult to determine at this stage because the variability of the included meteorological situations is very large. A study about the hydrometeor profile retrieval accuracy is needed (e.g. Bauer et al. 2005b) that also incorporates constraints from the satellite mission objectives such as higher vs. lower latitude or land vs. ocean observations as well as technical constraints (number of channels, polarization, bandwidths, radiometer noise etc.).

Acknowledgements

The authors are grateful to Jean-Nöel Thépaut for his valuable comments. This study was funded by the European Space Agency (ESA) through ESTEC Contract No.18101/04/NL/GSNG.

References

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

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, and E. Moreau, 2005a: Implementation of 1D+4D-Var Assimilation of Microwave Radiances in Precipitation at ECMWF, Part I: 1D-Var. *Q. J. Roy. Meteor. Soc.*, submitted.

Bauer, P., E. Moreau, and S. DiMichele, 2005b: Hydrometeor Retrieval Accuracy using Microwave Window and Sounding Channel Observations. *J. Appl. Meteor.*, **44**, 1016–1032.

Bauer, P., E. Moreau, F. Chevallier, and U. O'Keefe, 2005c: Multiple-scattering microwave radiative transfer for data assimilation applications. *Q. J. Roy. Meteor. Soc.*, submitted.

Ellison, W.J., S.J. English, K. Lamkaouchi, A. Balana, E. Obligis, G. Deblonde, T.J. Hewison, P. Bauer, G. Kelly, and L. Eymard, 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.

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

Holm, E., E. Andersson, A. Beljaars, P. Lopez, J-F. Mahfouf, A.J. Simmons, and J-N. Thépaut, 2002: Assimilation and modelling of the hydrological cycle: ECMWF's status and plans. *ECMWF Technical Memorandum 383*.

Lopez, P., and E. Moreau, 2005: A convection scheme for data assimilation: Description and initial tests. *Q. J.*

Roy. Meteor. Soc., **131**, 409–436.

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

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

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

Rabier, F., N. Fourrié, D. Chafaï̈, and P. Prunet, 2002: Channel selection methods for Infrared Atmospheric Sounding Interferometer. *Q. J. Roy. Meteor. 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, Vol. 2830, Paul B. Hays and Jinxue Wang eds., 136–147.

Rodgers, C.D., 2000: Inverse methods for atmospheric sounding. Theory and practice. *Series on Atmospheric, oceanic and planetary physics*, Vol. 2. World Scientific, Singapore, New Jersey, London, Hong Kong, pp. 238.

Saunders, R., P. Brunel, F. Chevallier, G. Deblonde, S. J. English, M. Matricardi, and P. Rayer, 2001: RTTOV-7 Science and validation report. *Met Office Forecasting and Research Technical Report*, No. 387, pp. 51.

Shannon, C.E., and W. Weaver, 1949: The Mathematical Theory of Communications. *University of Illinois Press, Urbana* (1962).

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

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

Table 1: Sample size of profile datasets.

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

Table 2: Channel recommendation by priority as a function of dataset and retrieval parameter.

	Rain	Snow	Cloud water
WINTER-LAND:			
1	90	135-145	100
2	135-140	195	135
3		118.75±10-14 GHz	
SUMMER-LAND:			
1	118.75±10-14 GHz	150-155	90-95
2		118.75±10-14 GHz	135
WINTER-OCEAN:			
1	40-45	155	42
2	80	118.75±10-14 GHz	83
3	135-150		
SUMMER-OCEAN:			
1	32-35	140-150	42
2	17-18	118.75±10-14 GHz	82
3	118.75±10-14 GHz		

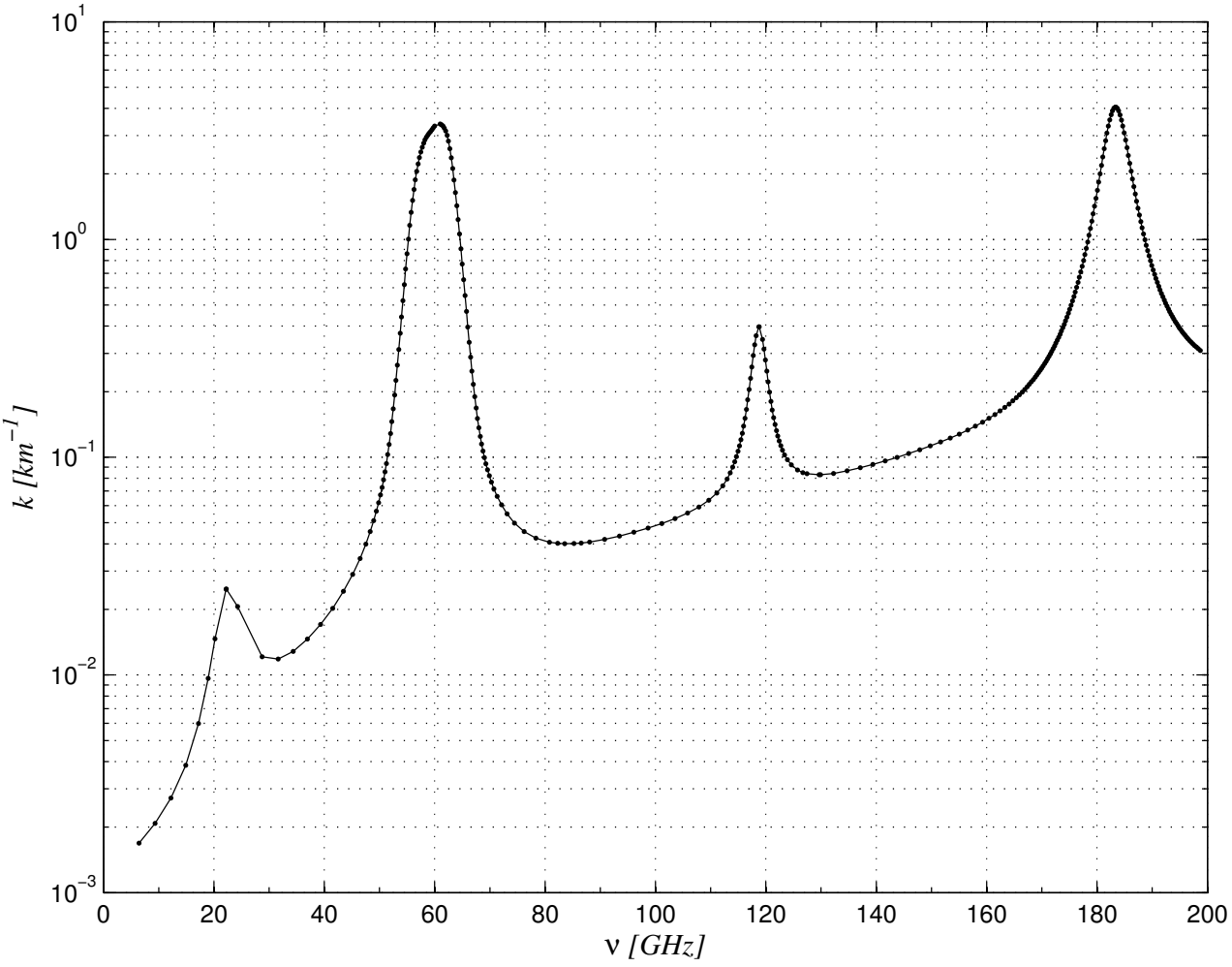

Figure 1: Atmospheric absorption coefficient, k, as a function of frequency, ν, for reference profile. Spectral location of 300 candidate channels is superimposed as dots.

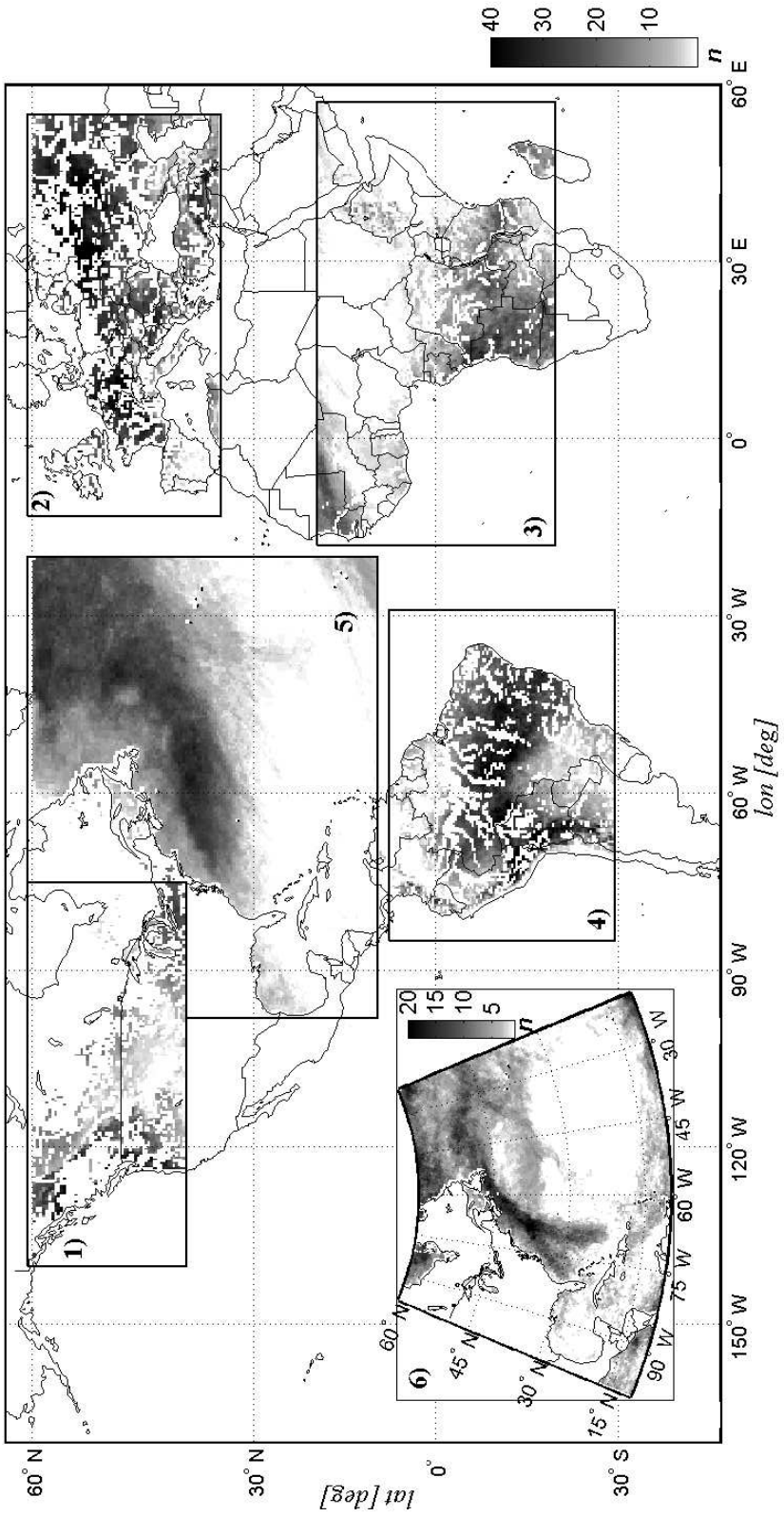

Figure 2: Location and frequency distribution (n) of profile datasets chosen for the channel selection statistics: WINTER-LAND = (1) + (2), SUMMER-LAND: (3) + (4), WINTER-OCEAN (5), SUMMER-OCEAN: (6).

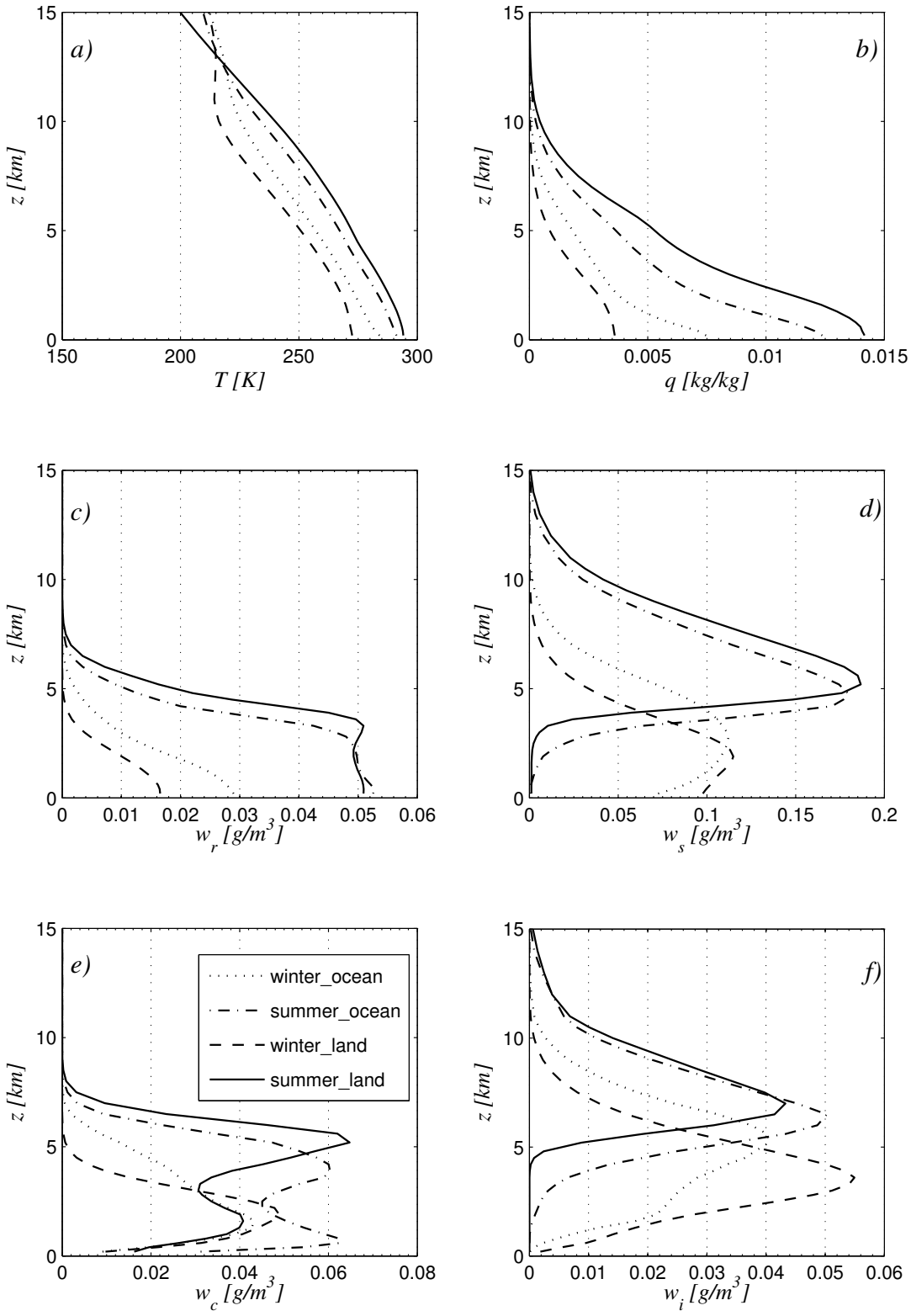

Figure 3: Mean temperature (a), specific humidity (b), rain (c), snow (d), cloud liquid water (e) and cloud ice (f) profiles from representative datasets WINTER-OCEAN (dotted), SUMMER-OCEAN (dash-dotted), WINTER-LAND (dashed) and SUMMER-LAND (solid).

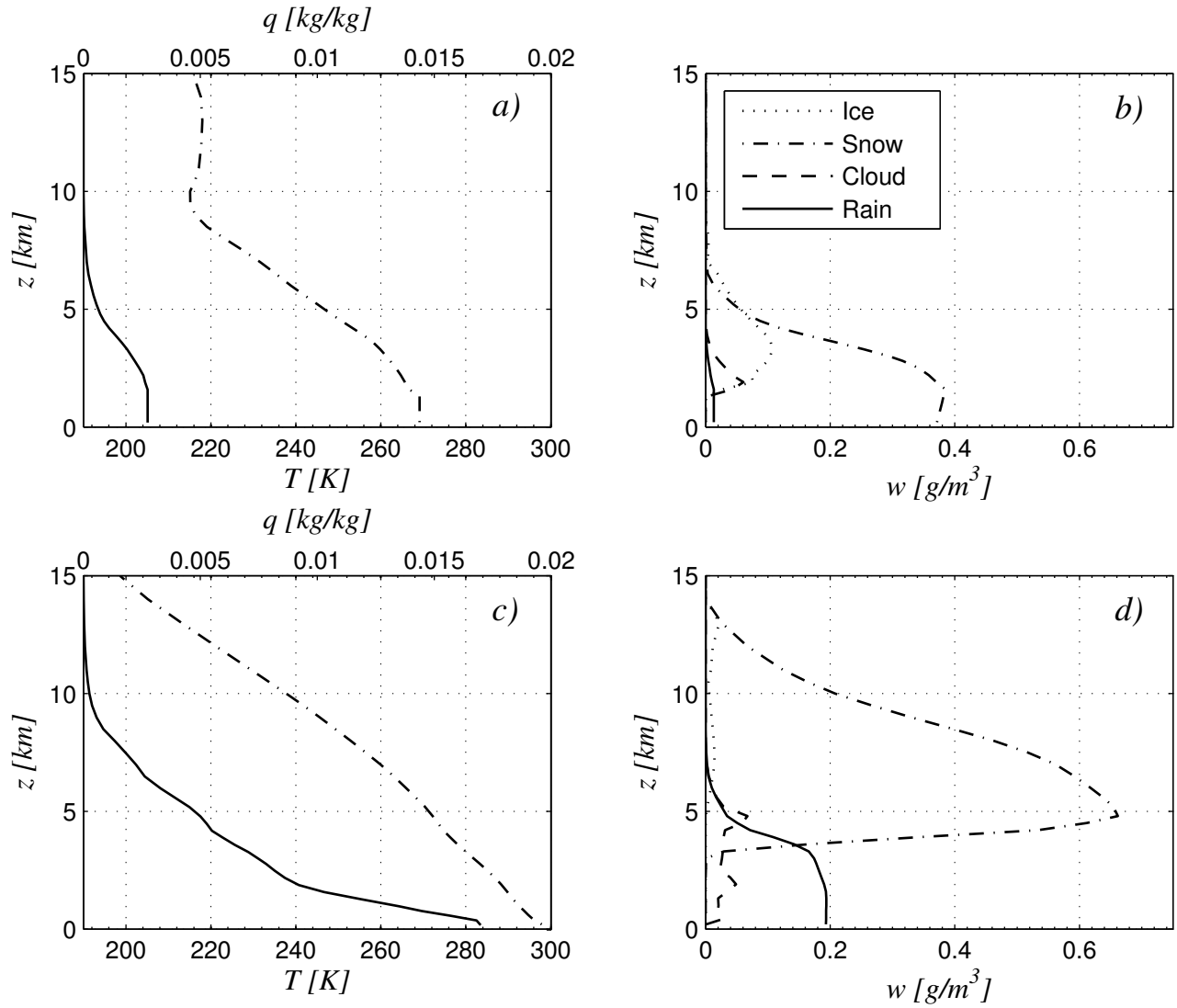

Figure 4: Profiles of temperature (dash-dotted) and specific humidity (solid) from WINTER-LAND (a) and SUMMER-OCEAN (c) datasets. Corresponding rain (solid), cloud (dashed), snow (dash-dotted), and cloud water (dotted) profiles are shown in (b) and (d).

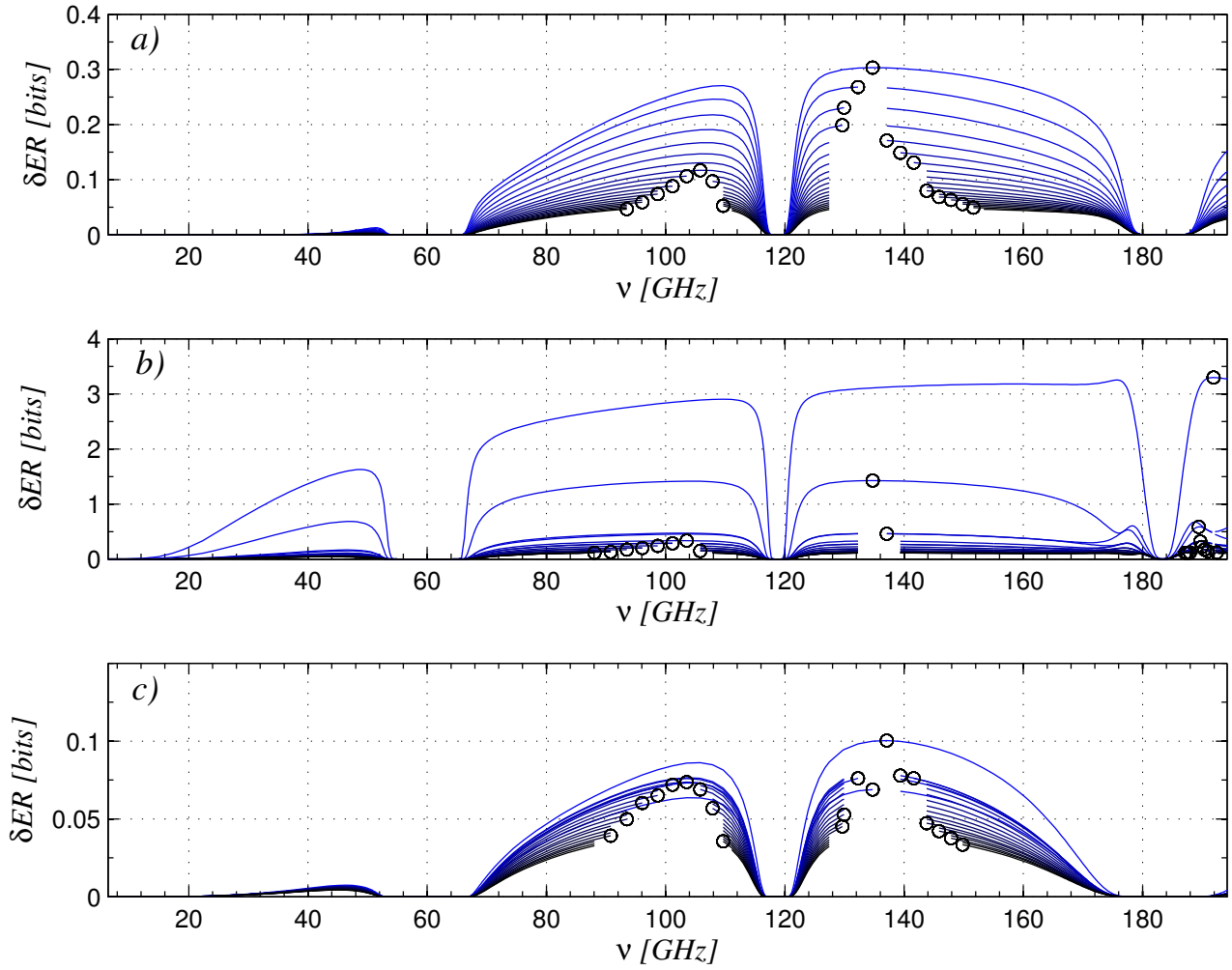

Figure 5: Entropy reduction for single WINTER-LAND profile as a function of microwave frequency for rain (a), snow (b), and cloud water (c). Individual lines refer to a single iteration, maximum values are highlighted as circles.

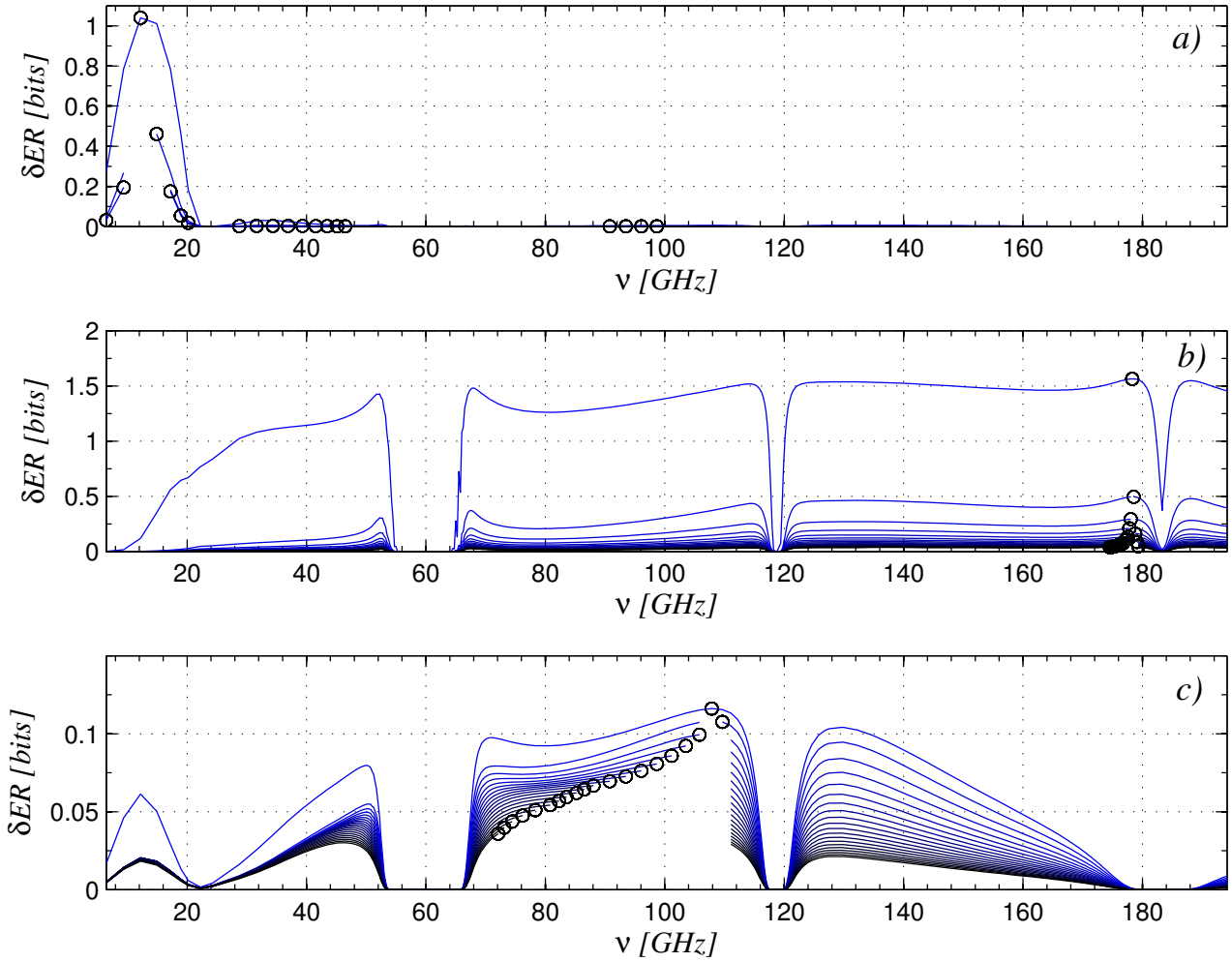

Figure 6: See Figure 5 for single SUMMER-OCEAN profile.

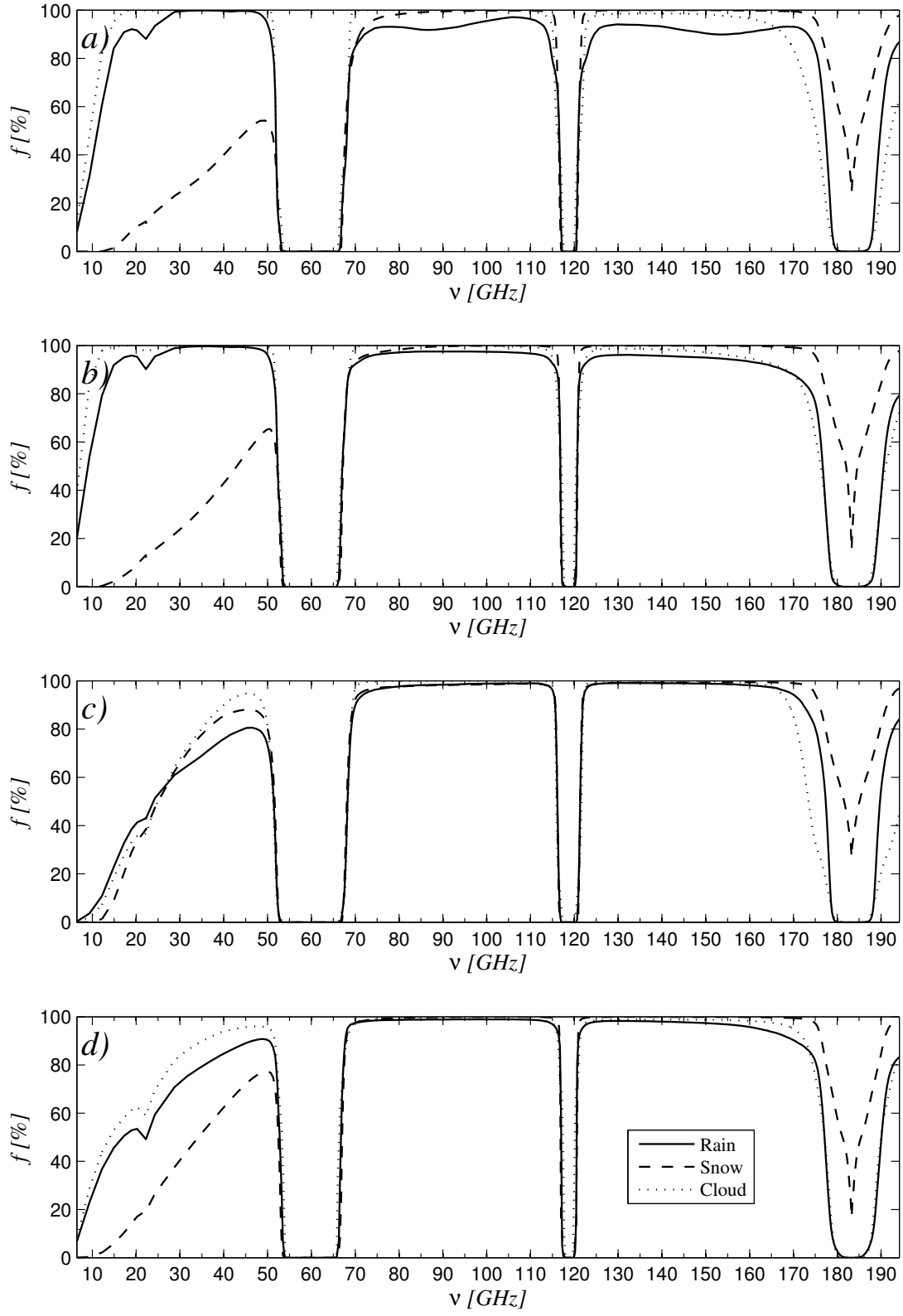

Figure 7: Fraction of profiles for which $\delta ER > 0.001$ per frequency for rain (solid), snow (dashed) and cloud water (dotted) for datasets WINTER-OCEAN (a), SUMMER-OCEAN (b), WINTER-LAND (c) and SUMMER-LAND (d).

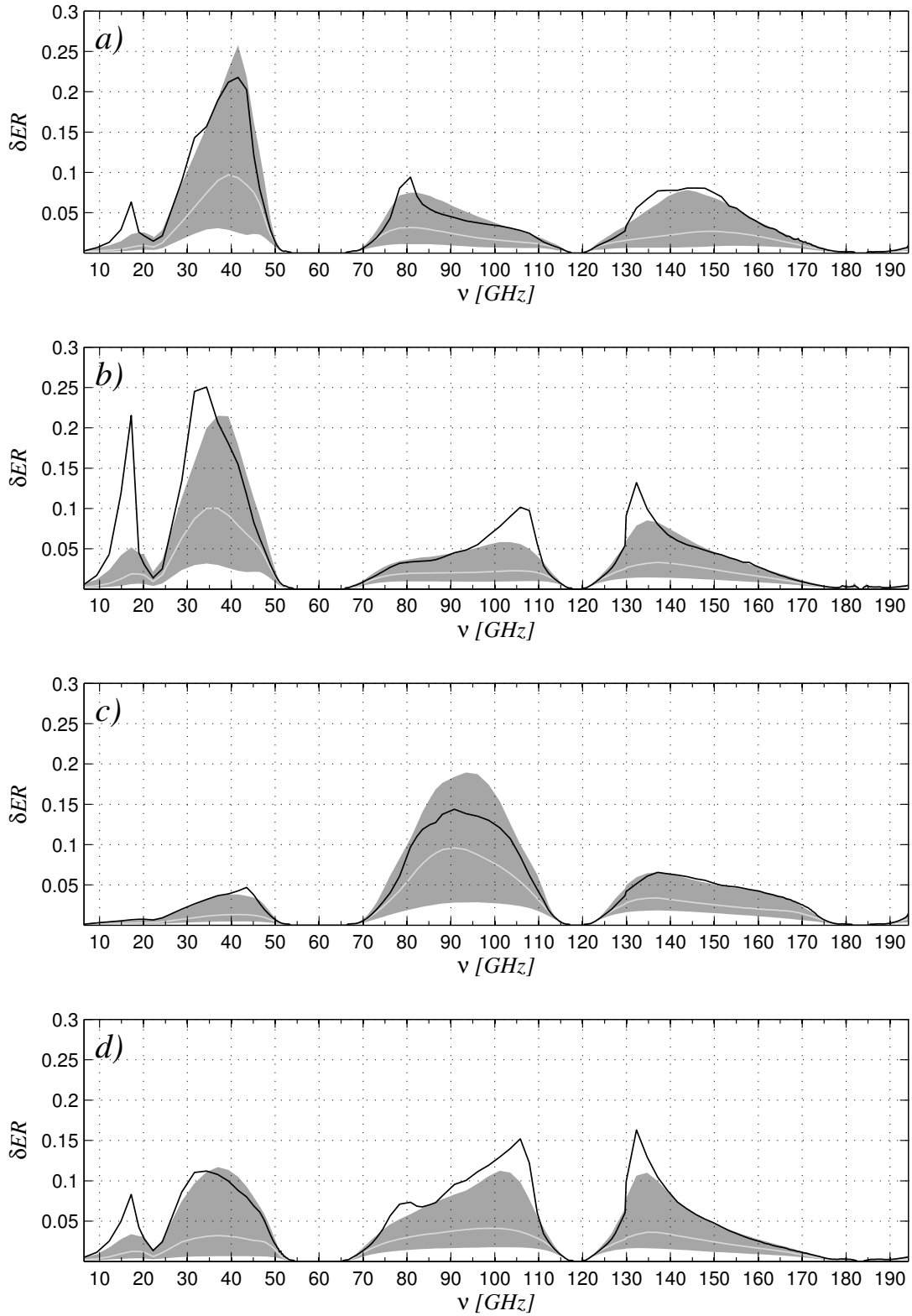

Figure 8: δER frequency distributions for rain from datasets WINTER-OCEAN (a), SUMMER-OCEAN (b), WINTER-LAND (c) and SUMMER-LAND (d). Black lines denote mean, white lines denote median and grey-shaded areas are the 25% and 75% limits of the distribution.

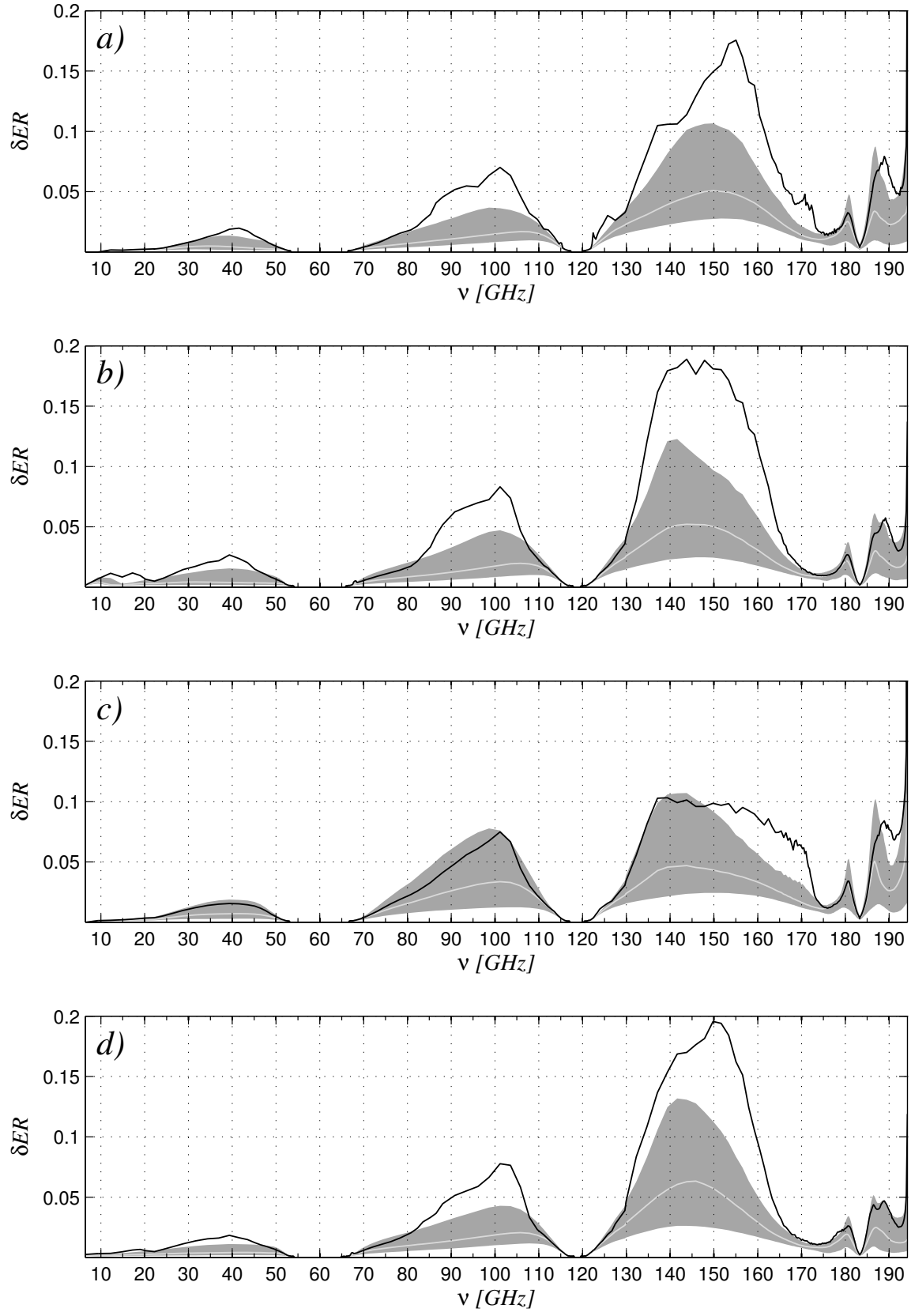

Figure 9: See Figure 8 for snow.

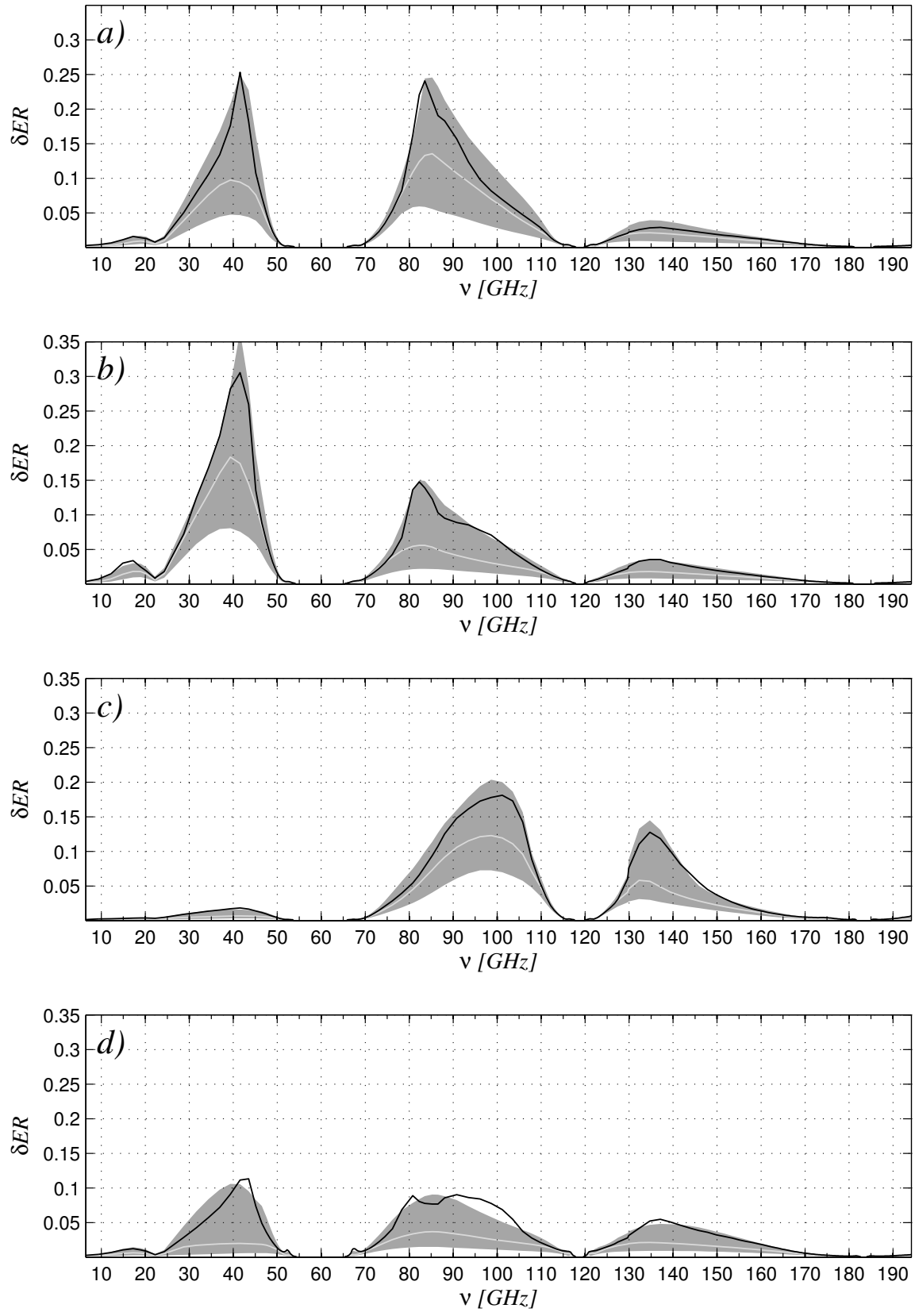

Figure 10: See Figure 8 for cloud.