

Performance simulations for a submillimetre-wave satellite instrument to measure cloud ice

C. Jiménez,^{a,d,*} S. A. Buehler,^a B. Rydberg,^b P. Eriksson^b and K. F. Evans^c

^a Department of Space Science, Luleå University, Luleå, Sweden

^b Department of Radio and Space Science, Chalmers University of Technology, Gothenburg, Sweden

^c Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, USA

^d Laboratoire d'Etudes de Rayonnement et de la Matière en Astro-physique, Centre National de la Recherche Scientifique, Observatoire de Paris, Paris, France.

ABSTRACT: The performance of a conically scanning satellite instrument for the measurement of cloud ice was studied. The instrument measures radiances in 12 channels placed around the 183, 325 and 448 GHz water vapour lines and the 243, 664 and 874 GHz window channels, and is designed to provide estimations of ice water path (IWP), the equivalent sphere diameter (DME), and the median ice mass height (ZME). Overall median relative errors of around 20% for IWP, 33 μm for DME, and 240 m for ZME for a midlatitude winter scenario, and 17% for IWP, 30 μm for DME, and 310 m for ZME for a tropical scenario were found. Detection limits (relative retrieval error reaching 100%) of around 2 gm^{-2} were estimated for both scenarios. The performance of a five-receiver instrument, where either the 664 or 874 GHz channel is dropped, was close, but with increased errors for very thin and high clouds. A trade-off between having the 874 GHz receiver or two infrared channels at 10.7 and 12 μm emerged, as very similar performance was found between the six-receiver instrument and the five-receiver instrument with the infrared channels. Another trade-off between receiver selection and noise was also apparent, with some of the four-receiver selections operating at half noise levels being able to compete with the standard six-receiver instrument. Dual-polarized measurements were also tested, but they did not significantly improve the retrievals of IWP or DME. Copyright © 2007 Royal Meteorological Society

KEY WORDS ice water path; microwaves; neural networks

Received 25 February 2007; Revised 16 June 2007; Accepted 21 June 2007

1. Introduction

Clouds are an essential component of the climate system with a strong impact on the Earth's radiative balance (e.g. Stocker, 2001). Our interest here is for clouds composed of ice crystals. They reflect solar radiation, which tends to cool the surface, but they also absorb outgoing infrared radiation, which tends to warm it. The ice crystal shapes and sizes, cloud altitude, and optical thickness are the main parameters determining the relative strength of these effects, but their global distributions are still poorly known. This has an impact on our present models predicting weather and climate. The models represent cloud ice by fields of ice water content (IWC), but poor observations of this parameter or its vertically integrated parameter, the ice water path (IWP), lead the models to poorly constrain the IWC fields. Mean values of IWP in present global climate models have been shown to differ by even an order of magnitude (Del Genio, 2001). The problem is that the IWC fields can vary considerably depending on the assumptions made regarding ice sedimentation velocities and the treatment of advection from

one time step to the next, as shown by recent sensitivity studies (e.g. Wilson, 2000; Reinhardt *et al.*, 2004). Better observations of IWC and IWP will thus be very useful to better constrain the model fields, reducing the related uncertainties in the weather and climate predictions.

Existing sensors, mainly in the visible, infrared and millimetre-wave regions, have advanced our knowledge about cloud ice, but measurements still have a large degree of uncertainty. Thermal infrared methods (e.g. Ou *et al.*, 1993) saturate for clouds with moderate optical depth, millimetre-wave methods (e.g. Evans and Stephens, 1995b) have low sensitivity for optically thinner clouds, and visible and near-infrared methods (e.g. Rolland *et al.*, 2000) are most sensitive to the particles at the cloud top. Further, their observing wavelengths makes them mostly sensitive to one of the ends of the typical particle size distributions. In order to estimate the ice mass from their observations, assumptions about the particle size distribution are needed, which can result in large errors. To overcome these limitations, the use of the submillimetre range for cloud ice sensing has been proposed in the last years (e.g. Evans *et al.*, 1998). A down-looking submillimetre cloud ice sensor works by measuring brightness temperature depressions produced by the scattering of the up-welling radiation from the lower atmosphere. The radiation mainly comes from the

* Correspondence to: C. Jiménez, LERMA, Observatoire de Paris, 61 avenue de l'Observatoire, 75014 Paris, France.
E-mail: carlos.jimenez@obspm.fr

strong water vapour emission at these frequencies, and it is scattered by the ice particles present in the clouds. The size of the depressions will depend on the amount of ice and the particle size, with the added advantage that the observing wavelengths have the highest sensitivity to the particles responsible for most of the ice mass. For a given ice mass and particle size, different frequencies will have different responses, and this allows a multi-frequency sensor to determine independently the ice mass and the particle size.

A number of aircraft instruments have been built to prove the concept (e.g. Evans *et al.*, 2002, 2005) and first cloud ice retrievals with the existing satellite limb-sounding submillimetre sensors have been published (Li *et al.*, 2005; Eriksson *et al.*, 2007). However, aircraft instruments are useful for specific campaigns but cannot provide global coverage, while the limb-sounding geometry results in poor horizontal resolution. A satellite down-looking millimetre and submillimetre sensor can provide both global coverage and good horizontal resolution, and a mission concept for this type of instrument has been proposed by Buehler *et al.* (2007). The suggested instrument is a conical scanning instrument placed in a sun-synchronous orbit at between 830 and 870 km, with an incident angle of around 53 degrees providing an instrument swath of around 1800 km, a footprint resolution of around 10 km, and nearly global coverage in about a day. The instrument has been designed to provide the integrated ice water path (IWP), the equivalent sphere diameter (DME), and the median ice mass height (ZME) as main cloud parameters. For brevity, we will refer to this instrument type as a submillimetre instrument, even though the lowest frequency channels typically have wavelengths longer than a millimetre.

This paper studies the performance of the proposed instrument. This is achieved by means of a system model composed of three main blocks: a dataset containing a large number of atmospheric states and corresponding radiances, a sensor model representing the instrument to be studied, and a retrieval algorithm, to estimate the retrieval performance of the instrument. Most of the simulations use a dataset stochastically generated from cloud geometry statistics obtained from radar data. However, for completeness some retrieval tests are also carried out with another independent dataset. This second dataset was produced with a different radiative transfer code and used radar reflectivity measurements to generate clouds with ice profiles matching the measured radar backscatter. In both datasets, only one-dimensional (1D) atmospheric states are included, i.e. only vertical variations in the atmospheric fields are considered. This means that retrieval aspects related to the inhomogeneity of the clouds fields are not addressed, and possible associated errors are not discussed. This is a simplification. For instance, Davis *et al.* (2006) have shown that in some cases the 3D structure of clouds can have a significant impact on simulated radiances. The dominant effect was found to be that of sub-pixel structure (beam-filling effect), rather than 3D radiative transfer effects.

Despite these known shortcomings of the 1D approach, it was judged to be adequate for a basic mission performance study. This is planned to be refined in later studies, which should yield a quantitative estimate of the error introduced by the 1D assumption and related beam-filling issues.

The retrieval method is based on a neural network algorithm that numerically estimates the mean state of the posterior probability density distribution of the state vector. Other possibilities were also considered, namely the optimal estimation method (e.g. Rodgers, 2000) and a Monte Carlo integration of the posterior distribution (e.g. Evans *et al.*, 2002). In principle, these three methods can be set up to provide the mean state of the posterior distribution. Nevertheless, limitations related to the practical implementation of these algorithms can have a large impact on the solutions. The optimal estimation method was quickly discarded as the nonlinearity of the inversion problem implies the very costly recalculation of the weighting functions. For the number of cases to be inverted here, this cost was prohibitive. The Monte Carlo integration was considered further, but it was found out that for the size of the available datasets the reported retrieval errors were larger than for the equivalent neural network inversions. This is because there is a source of retrieval error related to the finite size of the dataset for the less-represented regions of state space (e.g. Evans *et al.*, 2005). This error also exists in the neural network inversions, but simultaneous inversions clearly showed smaller errors for the neural network inversions.

The sensor model completes the system model. For this work it has simply been simulated with the estimates of the receivers' system temperature, and this means that the instrument performance is judged by an error characterization based on simulating only main random uncertainties. Systematic errors can also be part of the error budget. For instance, imperfect knowledge of the spectroscopic parameters or the temperature of the calibration loads can result in retrieval biases. They are not considered in this study under the assumption that, for a properly characterized forward model and sensor, they are smaller than the investigated stochastic errors, but this assumption needs to be verified in further work once a more defined observing system exists.

The paper is structured as follows. First the set-up and methodology used to simulate the observing system is described. The concept of a neural network is introduced, followed by the description of a specific neural network algorithm modelling the mapping between radiances and retrieval parameters, and the two datasets used to specify this mapping. This is followed by the sensor model and a description of the error characterization adopted in the paper. Then this system model is used to analyse specific aspects of the instrument performance. The impact of receiver noise, channel selection and measuring polarization are studied in detail. The synergy of adding infrared data to the instrument radiances is also considered. Finally the instrument performance is

summarized and the main conclusions of the study are given.

2. Set-up and methodology

This section describes the retrieval method and simulation set-up used to assess the retrieval errors of the instrument. The retrieval errors are characterized by taking the difference between retrieved and true state vectors and then estimating the median of the absolute values of the retrieval errors for all the realizations in the validation dataset. If relative errors are given, they are calculated as the difference between the retrieved and true value divided by the true value. The median is reported instead of the standard deviation, as the latter is more influenced by the presence of outliers. DME and ZME errors are discussed in absolute values while IWP, due to its large dynamical range, is presented in relative values for most of the dynamical range and absolute values for very low IWP values, where the relative errors are not useful. Note that the error analysis only focuses on the cloudy cases, so clear-sky cases are removed from the datasets. This implies that an algorithm to detect the existence of clouds must exist and should be used before running the inversions, as the retrievals only deal with cloudy situations. Note also that the use of two completely independent datasets has resulted in two slightly different definitions for DME. The exact definitions are given when describing the datasets. The descriptions of the neural network retrieval algorithm, the sensor model, and the datasets used for the error analysis now follow.

2.1. Neural network technique

A neural network (NN) is an interconnected assembly of processing units called neurons. Each neuron processes its inputs by first weighting them, adding a bias, and then transforming the resulting value with a certain function. The structure of a neuron is sketched in Figure 1. If \mathbf{w} is a vector with the weights, \mathbf{i} is a vector with the inputs to the neuron, b is a bias value, and f is the transforming function, usually called the activation function, the output of a neuron o is given by

$$o = f(\mathbf{w}^T \mathbf{i} + b). \tag{1}$$

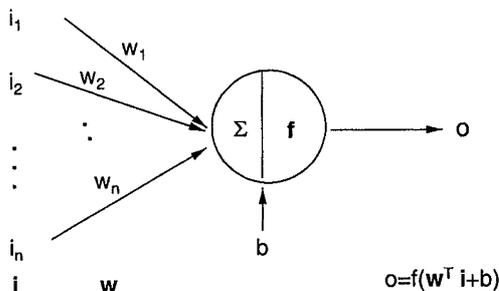


Figure 1. Schematic of a neural network neuron. The input vector is weighted, the values are summed together, a bias term is added, and the result is used to feed an activation function giving the output of the neuron.

Typical activation functions include linear functions and different types of sigmoid functions.

Different types of NN can be set up by organizing the neurons in different layers and allowing the input to the NN to propagate in different ways. For instance, if the input signal is allowed to propagate only in the forward direction, there is more than one layer, and at least one of the layers has differentiable activation functions, the NN is called a feed-forward multilayer perceptron (MLP). Such NNs are broadly used because they can represent any continuous functional mapping from one finite dimensional space to another. Other methods to represent nonlinear mappings between multidimensional spaces exist, but NNs and similar techniques using superpositions of nonlinear functions have the advantage of requiring a smaller number of free parameters when the dimension of the input or output spaces starts to increase (Bishop, 1995).

The topology of a MLP is sketched in Figure 2. If w_{lt}^j is the weight of input t for the neuron l in layer j , the weighting matrix \mathbf{W}^j takes the form

$$\mathbf{W}^j = \begin{pmatrix} w_{11}^j & \dots & w_{1R}^j \\ \vdots & \vdots & \vdots \\ w_{S1}^j & \dots & w_{SR}^j \end{pmatrix}, \tag{2}$$

where S is the number of inputs and R is the number of neurons for layer j . If the input to the layer j is expressed as a vector \mathbf{i}^j , the bias of the neurons is expressed as a vector \mathbf{b}^j , the output from the neurons is grouped as a vector \mathbf{o}^j , and the activation function of all neurons in layer j is f_j , the output of layer j is given by

$$\mathbf{o}^j = f_j(\mathbf{W}^j \mathbf{i}^j + \mathbf{b}^j). \tag{3}$$

If the input vector of a MLP of M layers is a vector \mathbf{p} ($\mathbf{p} = \mathbf{i}^1$) and the output of the MLP is \mathbf{u} ($\mathbf{u} = \mathbf{o}^M$), the way the input signal propagates through the MLP is given by

$$\begin{aligned} \mathbf{u} &= \mathbf{f}_M(\mathbf{W}^M \mathbf{i}^M + \mathbf{b}^M) \\ &= \mathbf{f}_M[\mathbf{W}^M \mathbf{f}_{M-1}(\mathbf{W}^{M-1} \mathbf{i}^{M-1} + \mathbf{b}^{M-1}) + \mathbf{b}^M] \\ &= \dots \end{aligned} \tag{4}$$

The number of layers, number of neurons for each layer, and type of activation functions are variables to be determined for each specific application of the MLP.

The processing ability of a NN lies in the modification of its weights and biases when the NN is subject to adaptation to a set of examples, given as L pairs of vectors $\{\mathbf{x}^l, \mathbf{t}^l\}_{l=1, \dots, L}$. Training consists in finding the weights and biases minimizing the cost function that determines the learning process. A typical cost function is the mean sum of squares of the difference between targets, \mathbf{t}^l , and current outputs of the MLP to the corresponding input vectors, \mathbf{x}^l , i.e.

$$\sum_{l=1}^L \|\mathbf{u}(\mathbf{x}^l, \mathbf{w}) - \mathbf{t}^l\|^2, \tag{5}$$

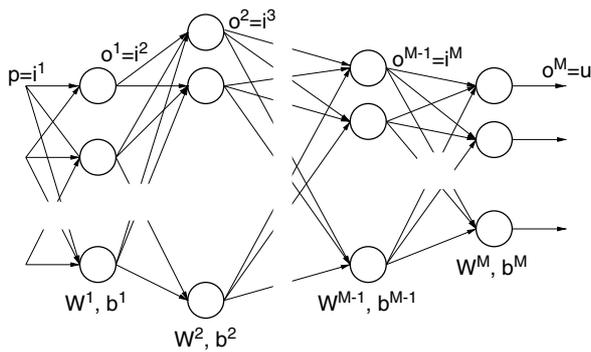


Figure 2. Schematic of a feed-forward multilayer perceptron. Neurons are organized into an input layer, a number of hidden layers, and an output layer. The signals propagate only in the forward direction from the input layer to the output neurons.

where L is the number of neurons of the output layer, $\|\cdot\|$ is the standard 2-norm, and $\mathbf{u}(\mathbf{x}^l, \mathbf{w})$ is the output vector of the MLP for the corresponding spectra \mathbf{x}^l and weights \mathbf{w} . Different numerical methods can be used to minimize the cost function; those widely used are quasi-Newton methods (e.g. Gill *et al.*, 1981), and methods based on the Marquardt–Levenberg algorithm (e.g. Hagan and Mehraj, 1994). All of them need the gradient of the cost function with respect to the different weights and biases, which can be calculated by propagating the error backwards through the MLP (Rumelhart *et al.*, 1986).

When training the NN, it is important to avoid overfitting situations. In this situation a very small error is obtained for the training set, but when new data are presented, the error is much larger; i.e. the NN has learned the specific features of the training set but not the general features of the mapping presented. A way to prevent this would be to use a NN just large enough to provide the right fit; the NN would then not be complex enough to learn the specific mapping of the training set and would only learn its general features. However, determining this optimal NN is difficult, and other possibilities exist. For instance, regularization by weight decay, where the MLP is trained with the cost function

$$\alpha \sum_{l=1}^L \|\mathbf{u}(\mathbf{x}^l, \mathbf{w}) - \mathbf{t}^l\|^2 + \beta \|\mathbf{w}\|^2, \quad (6)$$

where α and β are parameters setting a trade-off between the term minimizing the difference between targets and current outputs of the MLP, and the term minimizing the sum of the MLP weights (e.g. Bishop, 1995). This cost function favours small weights, which encourages smoother mappings, less likely to over-fit to the training set.

2.2. Retrieval algorithm

The retrieval algorithm is based on training a set of MLPs to provide a model of the distribution underlined by the training data. Here, this is the posterior distribution of

a parameter of the state vector once a measurement is taken, and the MLP is trained to provide the mean state of this posterior distribution.

The algorithm employs one MLP per parameter of the state vector, that is, one MLP to retrieve IWP, a second one for DME, and a third one for ZME. The number of input nodes is the number of inputs to the MLP, in this case the number of receiver channels used in the inversion. There is one hidden layer consisting of a number of neurons with hyperbolic tangent activation functions, followed by an output neuron with a linear activation function. The weights of the MLP are initialized following the method by Nguyen and Widrow (1990). Before training, inputs and outputs to the MLP are linearly transformed into the range $[-1, 1]$ to make the initialization of the weights more effective.

Once the topology is decided, the MLP needs to be trained. This requires a suitable dataset with an ensemble of atmospheric states and associated radiances. These are described in Sections 2.4 and 2.5. The cost function given in Equation (6) is suitable for the training, as it can be proven that a Bayesian estimation of the mean state with a prior distribution for the weights favouring small values corresponds to minimizing that function (e.g. Bishop, 1995). This means that weight decay is used to improve the generalization capacity of the MLP, as discussed in 2.2.

The practical implementation of the training algorithms follows the work of Foresee and Hagan (1997), where the parameters α and β of Equation (6) are statistically estimated to produce a MLP with good generalization and the right model complexity. This makes the choice of the number of nodes in the hidden layer not too critical, though a reasonable number has to be selected by taking into account the number of MLP inputs. Tests were conducted with a different number of hidden nodes, and MLPs with 12 nodes in the hidden layer were found adequate.

2.3. Sensor model

The sensor is modelled as an instrument having a combination of receivers measuring at different frequencies. The observing frequencies were selected after a systematic study where 26 different channel sets consisting of different combinations of 44 submillimetre frequencies were tested. The frequencies considered were either window channels or frequencies placed around water vapour and oxygen lines. The major water vapour lines between 183 and 916 GHz were included in the analyses, apart from the lines at 557 and 752 GHz, which are too strong for tropospheric use. Atmospheric opacity at frequencies beyond 1 THz is also too high for tropospheric observations, so that region was excluded from the analysis. Regarding the placement of the channels around the lines, the range in offset frequencies was made as wide as possible, limited by the amplifier bandpass of the receivers, to give a large range in weighting function altitudes. The oxygen lines at 715 and 774 GHz were included in some

of the selections, but were not selected in the best combination. This was attributed to their narrowness, which leads to smaller bandwidth and thus noisier channels that penalize the retrieval error. The window channels included frequencies used in previous proposals, or used on existing aircraft radiometers, or placed at the geometric mean of two absorption line frequencies, but Evans (2004) showed that the exact placement of the channels did not matter much.

The best selection from this analysis is adopted as a baseline for our sensor. It consists of twelve channels grouped in six receivers, three of them placed at the 183, 325 and 448 GHz water vapour lines, and the remaining three placed at the 243, 664 and 874 GHz window channels. The exact placement of the channels is given in Table I.

A sensor with the number of receivers discussed here is necessarily a complex instrument, but the analysis at this stage of the mission development was kept simple and only a characterization of the sensor based on the effective system temperature of its receivers was considered. The derived noise-equivalent temperature sensitivity for each channel is given in Table I and is used as the main measurement uncertainty in the simulations. Note that when studying some specific issues, such as polarization or the impact of the receiver noise, the noise levels can be different from those given in the table. Also

note that the noise figures in Table I for some channels deviate slightly from those given in Buehler *et al.* (2007), which are based on a later performance estimate of the engineering team. The differences are small and do not affect the results significantly.

2.4. Colorado training dataset

The training dataset was created by running an Eddington second-approximation radiative transfer code on a large number of randomly generated atmospheric profiles containing simulated clouds. The simulations have been run for two different atmospheric scenarios representing tropical anvil clouds and midlatitude winter synoptic cirrus. A summary of the statistics used for the dataset generation is given in Table II, and a summary of the randomly generated clouds can be found in Table III. A detailed description of the dataset generation follows.

The atmospheric temperature and relative humidity profile statistics were obtained from soundings from the Atmospheric Radiation Measurement (ARM) programme. The midlatitude winter simulation uses 156 Oklahoma soundings from December through to March for two years (1998 and 2000). The tropical simulation uses 442 Manus and Nauru soundings selected throughout the year over about four years, with morning and evening soundings selected with about equal frequency. Stochastic profiles of temperature and relative humidity were generated using Gaussian random amplitudes as coefficients of the empirical orthogonal functions (EOFs) obtained from the atmospheric profile statistics. To simulate auxiliary temperature profile information during the retrievals, the temperature profile variability was reduced to achieve 1.0 K standard deviation in the middle to upper troposphere. This was achieved by multiplying the midlatitude winter temperature variability by a factor of 0.25, and the tropical temperature variability by a factor of 0.75. As the window channel near 243 GHz sees the surface, surface emissivity variations were included in the dataset. The surface emissivity in the dataset was Gaussian distributed with mean 0.93 and standard deviation 0.03, independent for each receiver.

The clouds were generated in a stochastic manner, with the cloud geometry statistics obtained from ARM cloud radar data. Simulated clouds consist of one single-layer ice cloud and an optional liquid boundary-layer

Table I. Frequencies selected for the instrument grouped in six receivers with their corresponding radiometric sensitivities.

Receiver	Channel	Frequency (GHz)	NE ΔT (K)
R1	1	183.31 \pm 1.5	0.6
	2	183.31 \pm 3.5	0.5
	3	183.31 \pm 7.0	0.4
R2	4	243.2 \pm 2.5	0.5
R3	5	325.15 \pm 1.5	0.8
	6	325.15 \pm 3.5	0.7
	7	325.15 \pm 9.5	0.6
R4	8	448.00 \pm 1.4	1.3
	9	448.00 \pm 3.0	1.0
	10	448.00 \pm 7.2	0.7
R5	11	664.0 \pm 4.2	1.4
R6	12	874.4 \pm 4.5	2.3

Table II. Parameters of the probability distributions describing the Colorado ice cloud and liquid boundary layer cloud geometry.

Parameter	Tropical	Midlatitude winter
Mean cloud-top height (km)	8.5	12.0
Standard deviation of cloud-top height (km)	2.5	2.5
Mean ice cloud thickness (km)	3.0	2.5
Probability of boundary-layer (BL) cloud	0.25	0.20
Mean of log-normal BL cloud-base height (km)	0.8	1.8
Standard deviation of BL cloud-base height (km)	0.5	1.4
Mean of exponential BL cloud thickness (km)	0.9	0.6
Mean of BL cloud LWP distribution (gm^{-2})	120	90

Table III. Statistics of the randomly generated ice clouds.

Parameter	Mean	Median	Std. dev.	Minimum	Maximum
Colorado: original					
IWP (g m ⁻²)	74 (290)	27 (43)	141 (781)	0.01 (0.01)	6944 (9987)
DME (μm)	246 (211)	254 (191)	120 (120)	21 (17)	629 (630)
ZME (km)	7.32 (10.68)	7.16 (10.57)	2.39 (2.56)	0.91 (4.14)	19.66 (22.45)
Colorado: uniform					
IWP (g m ⁻²)	342 (415)	263 (366)	279 (297)	0.10 (0.10)	999 (999)
DME (μm)	370 (294)	382 (295)	101 (98)	51 (20)	599 (502)
ZME (km)	5.77 (9.07)	5.61 (8.95)	1.79 (1.77)	2.01 (5.01)	11.9 (13.9)
Chalmers: non-polarized					
IWP (g m ⁻²)	88	12	177	0.10	998
DME (μm)	232	211	78	106	598
ZME (km)	6.48	6.45	2.11	2.25	11.85
Chalmers: polarized					
IWP (g m ⁻²)	489	286	496	1	2000
DME (μm)	259	231	84	107	492
ZME (km)	6.14	6.40	2.32	0.6	11.40

For the Colorado dataset, the figures are given for both the original dataset and the modified more uniform version.

The numbers in brackets correspond to the tropical scenario, and those without brackets to the midlatitude winter scenario.

The Chalmers datasets are generated only for a midlatitude scenario, and no uniform procedure was applied due to a smaller number of cases. The statistics are given for the non-polarized and polarized versions.

cloud. The ice cloud may have mixed phase or even pure liquid parts, with a Gaussian distribution of top height and an exponential distribution of thickness. The mean cloud-top height is obtained from the current temperature profile by specifying a mean cloud-top temperature. The mean ice cloud thickness is fixed to 3 km for the tropical dataset (Jay Mace, personal communication), and 2.5 km for the midlatitude winter dataset (Zhien Wang, personal communication), taking into account that the thickness represents the integrated effect of multilayer clouds in nature. The temperatures of the top and bottom of the ice clouds are used to choose random IWC and DME at top and bottom from a tri-variate Gaussian distribution of temperature, $\ln(\text{IWC})$, and $\ln(\text{DME})$. This same IWC is used for the mixed-phase cloud total mass content. The DME is calculated as a mean mass-weighted equivalent sphere diameter. For each particle, an equivalent sphere diameter D is calculated as the diameter that the particle would have if it was an ice sphere with the density of solid ice ρ and equal mass m as the particle, i.e.

$$D = \left(\frac{6m}{\rho\pi} \right)^{\frac{1}{3}}, \quad (7)$$

followed by a mass-weighted integration over all the particles in the cloud

$$\text{DME} = \frac{\int_0^{\infty} D m(D) n(D) dD}{\int_0^{\infty} m(D) n(D) dD}, \quad (8)$$

where $n(D)$ is the particle size distribution and $m(D)$ the distribution of mass with particle size for the given cloud.

The boundary-layer cloud occurs with a certain probability, and is overcast and plane-parallel. The cloud-base height is distributed log-normally with a specified mean and standard deviation, the thickness is exponentially distributed, and the mean liquid water path (LWP) is specified. The cloud LWP is quadratically related to the thickness. All the distributions are independent. The cloud phase (liquid, ice, or mixed) and the LWC mixing fraction are chosen randomly for each independent mixed-phase layer with specified thickness. The probability of a layer being pure ice, p_{ice} , is a linear function of temperature:

$$p_{\text{ice}} = (T_{\text{noice}} - T)/(T_{\text{noice}} - T_{\text{allice}}), \quad 0 \leq p_{\text{ice}} \leq 1. \quad (9)$$

The probability of a layer being pure liquid, p_{liq} , is an exponential function of temperature:

$$p_{\text{liq}} = \exp\{(T - T_{\text{all liq}})/T_{\text{scale liq}}\}, \quad 0 \leq p_{\text{liq}} \leq 1. \quad (10)$$

Homogeneous freezing of droplets is enforced: if ($T < T_{\text{hom. freeze}}$) then ($p_{\text{liq}} = 0$; $p_{\text{ice}} = 1$). The liquid phase has priority: if ($p_{\text{liq}} + p_{\text{ice}} > 1.0$) then ($p_{\text{ice}} = 1 - p_{\text{liq}}$). Typical values for the parameters are: $T_{\text{all ice}} = -34^\circ\text{C}$, $T_{\text{noice}} = +6^\circ\text{C}$, $T_{\text{hom. freeze}} = -39^\circ\text{C}$, $T_{\text{all liq}} = +4^\circ\text{C}$, and $T_{\text{scale liq}} = +8^\circ\text{C}$. If the layer is not pure ice or liquid, then it is mixed phase. The liquid mass fraction is distributed according to a beta distribution (range 0 to 1) with the mean fraction fixed (0.5) and the standard deviation fixed (typically 0.35).

The tropical microphysical input statistics are an even combination of Central Equatorial Pacific Experiment (CEPEX) and Cirrus Regional Study of Tropical

Anvils and Cirrus Layers–Florida Area Cirrus Experiment (CRYSTAL–FACE) statistics. The CEPEX microphysics are based on the 2DC probe (Evans *et al.*, 2002), while the CRYSTAL–FACE microphysics are based on FSSP, 2DC, and HVPS probes (Evans *et al.*, 2005). The midlatitude winter statistics are based on 2DC and 2DP probes from three flights during the ARM 2000 Cloud Intensive Operation Period (IOP) (Heysmfield *et al.*, 2002). Three randomly oriented particle shapes are simulated with the Discrete Dipole Approximation (Evans and Stephens, 1995a; Evans *et al.*, 1998) in the microwave and volume/area equivalent spheres in the infrared: solid columns, hexagonal plates, and four-bullet rosettes. Three gamma size distribution widths are used ($\alpha = 0, 1, 2$). A truncated distribution is used as there is an upper limit on the particle diameter due to limitations of the radiative transfer code. The distribution follows

$$N(\text{DME}) \propto D + \exp(a) \exp \left\{ -(a + 3.67) \frac{\text{DME}}{b} \right\},$$

where a and b are parameters adjusted so that the truncated distribution has the moments of the desired untruncated, continuous gamma distribution defined by α and DME. A single cloud is purely one particle shape and one alpha, which emphasizes the retrieval error due to particle shape. Due to computational limitations, the largest DME ranges from 447 to 632 μm (depending on particle shape and alpha), and this limits the highest simulated cloud DME.

Running the radiative transfer code on each of the cloud realizations produces a corresponding set of radiances. Associating each radiance with their corresponding IWP, DME and ZME forms a training dataset with a prior distribution consistent with the original statistical assumptions. From a Bayesian point of view, this is the right dataset to train a NN as it follows the true statistics of the state space. However, using the prior distribution might result in regions of state space where the NN fails to model the distribution due to the very small number of data points present in that region. This is the case, for instance, for IWP, where following the prior distribution results in very few cases with high IWP. It can then be argued that less realistic but more uniform datasets with all regions more equally represented can help to reduce the errors related to how well the different regions of the state space are represented. It should be noted that this difficulty is not specific for a NN algorithm, as discussed in Section 1. A possible drawback is that a retrieval algorithm using a more uniform dataset is not strictly Bayesian when observations following the original prior distribution are inverted, as the true prior information is not used in the retrievals. A characterization of the retrieval errors based on a Bayesian propagation of the observation uncertainties should then be approached with caution, as the real prior information is not perfectly known. However, for simulations where the true fields are known, as is the case here, the retrieval error is characterized by taking differences

between true and retrieved values. As the Bayesian propagation of uncertainties is not needed, the reduction of these representation errors outweighs the departure from a strictly Bayesian inversion.

Following this argument, a more uniform dataset was extracted from the original dataset. For the midlatitude scenario, this dataset was built by arbitrarily dividing the region of state space with IWP between 0.1 and 1000 g m^{-2} , DME between 100 and 600 μm , and ZME between 2 and 12 km in $20 \times 20 \times 20$ equally sized regions, and then filling them with 200 data points if more than 200 points were available in that region, otherwise leaving them empty. The same was done for the tropical scenario but dividing the region of state space between 0.1 and 1000 g m^{-2} for IWP, 50 and 500 μm for DME, and 5 and 15 km for ZME. Due to the large dynamical range for IWP, a second possibility would have been to select the bins in a log-space, i.e. giving more relevance to the small values of IWP. This would have filled the regions with low IWP values with a higher number of cases, but at the expense of very few cases left in the very broad bins for intermediate and high IWP values. The first possibility was preferred. This means that a conservative approach is taken regarding reported retrieval errors for low IWP values or detection limits, as the smaller number of cases means a larger possibility for higher representation errors even when using the more uniform dataset. This is discussed further in Section 3.

The original dataset of 1×10^6 cases is downsized to nearly one third after this procedure. This does not present a problem, as retrieval tests showed that a dataset of 5×10^4 cases was valid to train the MLPs. A validation dataset of 2×10^5 cases, chosen from the remaining part of the dataset not used for the training, is inverted by the trained MLP to characterize the retrieval error. The same sizes are taken when the original dataset is used for training. Note that the retrieval focuses only on the cloudy cases, so clear-sky cases (IWP = 0) are removed from the datasets before the inversions. The minimum IWP contained by a cloud is 0.01 g m^{-2} for the original datasets, and 0.1 g m^{-2} for the uniform version of the datasets.

2.5. Chalmers training dataset

This alternative training dataset was created by running the radiative transfer code ARTS (Buehler *et al.*, 2004) on a set of simulated atmospheres containing a large number of clouds generated from radar reflectivity measurements from Cloudnet radar data. Only a short description of the dataset generation follows; Rydberg *et al.* (2007) provide a complete description of the dataset and its generation. A summary of the randomly generated clouds can be found in Table III.

Three midlatitude radar stations located at Chilbolton (UK), Pailaseau (France), and Cabauw (The Netherlands) are used for the radar data. This means that the dataset is only valid for midlatitude conditions. Single profiles of radar backscatter are first averaged, based on wind data

from the European Centre for Medium-range Weather Forecasts (ECMWF), to better represent mean properties matching the instrument footprint size. Then vertical profiles of temperature, humidity, and liquid water content are computed based on temperature and water vapour from ECMWF for the time and location of the radar measurements, and liquid water content from Cloudnet products. This is followed by the generation of vertical profiles of ice cloud microphysical parameters.

The particle size DME is calculated as a median mass equivalent sphere diameter, i.e.

$$\int_0^{\text{DME}} m(D)n(D) dD = \frac{\text{IWP}}{2},$$

where D , $n(D)$ and $m(D)$ are defined as in Equations (7) and (8). Note that this definition of DME is different from the definition of DME in the Colorado dataset given in Equation (8), though their values are very close. For instance, in the Colorado dataset for a gamma particle-size distribution with $\alpha = 1$, the ratio of both particle definitions is 1.07. Note that in the text both equivalent sphere diameters are referred to as DME, since the context (i.e. the specific dataset discussed) makes clear the definition in use.

The particle size distribution parameters are taken from results from *in situ* measurements reported by Heymsfield (2003) and Ivanova *et al.* (2001), and the shape distributions are represented by a mixture of three randomly orientated particle types (spheres, ellipsoids and columns), with the three particle types present in each cloud. The parameters of the distributions are randomly varied to simulate clouds covering a large range of microphysical conditions. Aspect ratios of 1 to 2.5 are assumed for the particles. Aspect ratios for individual particles can be larger than this, but for a particle ensemble population the aspect ratios are generally well below 2 (e.g. Korolev and Isaac, 2003). The aspect ratio is fixed for each particle shape for a given cloud, and is selected with a uniform probability. Finally IWC profiles are generated consistent with the previous calculations and the radar data, i.e. each ice cloud in the dataset produces the same backscatter as observed in its corresponding radar measurement.

ARTS is then run on each of the cloud cases to generate the corresponding radiances. Different scattering algorithms can be used within ARTS; an iterative discrete ordinate algorithm (Emde *et al.*, 2004) is used here to handle the scattering calculations. The surface is assumed to be a black body at the lowermost atmospheric temperature. This means that, unlike the Colorado dataset, the Chalmers dataset does not include surface emissivity variations. A second version of this dataset simulating polarized measurements has also been produced. In this case the prior IWP distribution did not follow the original distribution, as data cases were selected in order to have a more uniform dataset in IWP, and the clouds were populated by two azimuthally randomly orientated particle types (ellipsoids and columns, only one type in

each cloud), as azimuthally randomly orientated particles are more likely to give rise to polarization effects in the radiances (e.g. Davis *et al.*, 2005).

The non-polarized dataset has close to 2×10^5 cases. However, as they mimic the radar data, a large proportion of the cases do not contain ice clouds. Once the clear-sky cases are removed, the dataset is downsized to around 5×10^4 cases. The polarized dataset has a smaller number of cases, 1×10^4 , as the simulations are much more expensive to run, but it is still large enough for a reasonable first estimation of errors. Due to the smaller number of cases, no procedure to make these datasets more uniform has been applied. Half of the cases are used for the training dataset and half for the validation. The minimum IWP contained by a cloud is 0.1 gm^{-2} for the non-polarized dataset, and 1.0 gm^{-2} for the polarized version.

3. Results and discussion

This section presents the retrieval performance of the cloud ice observing system described in the previous section. Most of the simulations were conducted with the Colorado dataset, as it allowed study of the performance of the instrument for two different atmospheric scenarios.

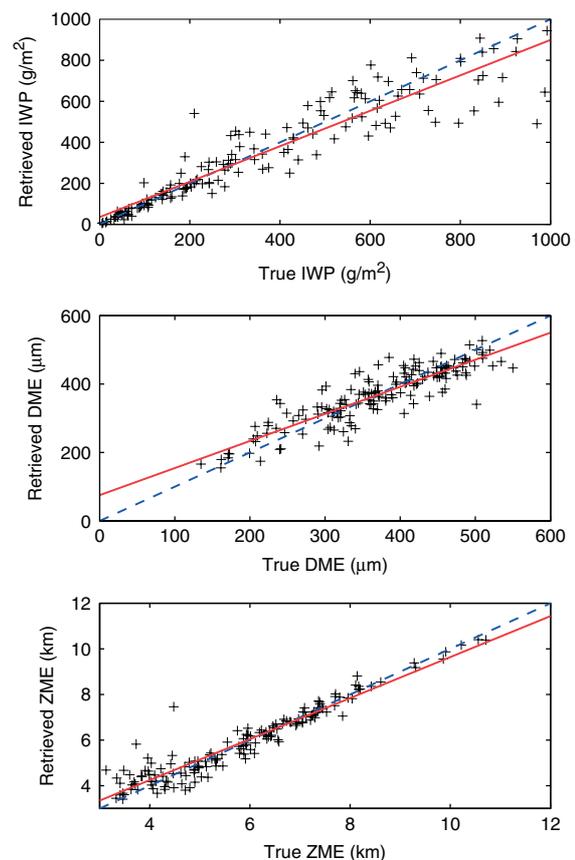


Figure 3. Noise-free all-receiver retrieval performance scatter plots for IWP (top), DME (middle), and ZME (bottom). The red line shows the best linear fit, and the 1:1 line is shown as blue dashed. The fit is calculated for all the cases in the validation dataset, but for clarity only a fraction of the cases is plotted. The simulations are for the midlatitude winter scenario of the uniform Colorado dataset.

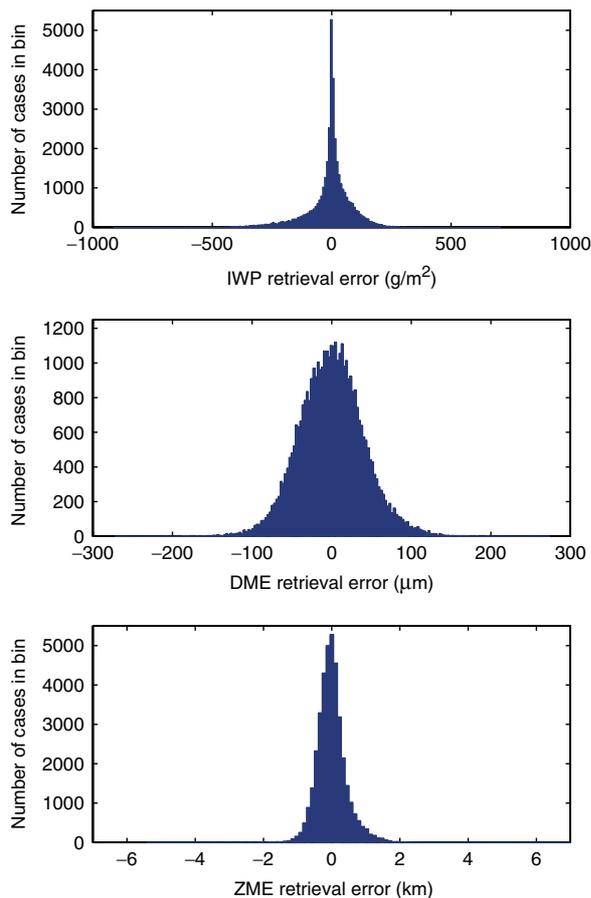


Figure 4. Histograms of the retrieval error for the noise-free all-receiver simulations of Figure 3 for IWP (top), DME (middle), and ZME (bottom). The retrieval errors are separated into 200 equally spaced bins, and the vertical axis shows the number of retrieval cases in each bin. This figure is available in colour online at www.interscience.wiley.com/qj

The retrieval errors for different channel selections and radiometric noise levels were investigated with this dataset. The synergy of adding infrared channels to the instrument is also studied with this dataset. The Chalmers dataset is used for consistency checks and to study a dual-polarized instrument.

3.1. Performance of a noise-free instrument

We start by inverting noise-free radiances from the uniform version of the midlatitude winter Colorado dataset. This is not very realistic, as radiances cannot be measured without a receiver, but it can be used to determine a lower limit of the retrieval error, as a real instrument will never be able to perform better than in the noise-free case. The retrieval error arises from different contributions. The ill-posedness of the inversion problem means that some components of state space cannot be determined from the measurements and have to be estimated from the prior information. The propagation of the uncertainties in the knowledge of some of the forward model parameters, such as the temperature or surface emissivity, results in another contribution. Also,

as discussed in Section 2, the neural networks approximate the posterior distribution of the retrieval variables, and errors related to this approximation, especially in the less-represented areas of the state space, are possible. For these simulations, the overall errors with all channels included in the inversions take values of around 17% for IWP, 27 μm for DME, and 240 m for ZME. Figure 3 shows typical scatter plots for these inversions, while Figure 4 shows the corresponding histograms of the retrieval error.

The performance of noise-free single channels was also considered. One channel from each receiver was selected. For receivers with multiple channels, the channel furthest away from the line centre was considered. The exact placement of the channels can be seen in Table I. The channels are identified below by the receiver number, i.e. R1 means the simulations for channel 183.31 ± 7.0 GHz, R2 for 243.2 ± 2.5 GHz, R3 for 325.15 ± 9.5 GHz, and so on. Their retrieval errors for IWP are plotted in Figure 5 for the midlatitude winter simulations. A similar curve for the all-receiver instrument is also displayed. The plot shows that the performance for the high-frequency channels R4-R5-R6 is poorer than for the low-frequency channels. This is at first sight confusing, but it can be explained if the atmospheric opacity is taken into consideration. The cloud-scattering signal is generally stronger at high frequencies but, on the other hand, high-frequency channels are limited to relatively high clouds, because the high opacity at these frequencies hides the lower clouds. Low-frequency channels do not suffer from this and can see clouds at different altitudes by placing the channels properly. As the high-frequency channels get opaque at higher altitudes, they miss a larger part of the ice mass column compared with the lower-frequency channels. As the true IWP is a vertical integration of the ice mass, the IWP retrieval errors get larger for the high-frequency channels. Figure 6 shows the same error curves as Figure 5, but only for clouds with ZME larger than 7 km included in the dataset. The IWP errors are now more similar for all the channels, though nearly identical curves cannot be expected. R4, R5 and R6 have smaller errors than R1, R2 and R3 at low IWP, due to the stronger cloud signal of the higher-frequency channels, but as they saturate at lower IWP they have larger errors for large IWP.

If a detection limit is defined as the IWP where the relative error becomes larger than 100%, detection limits as a function of channel selection can be drawn from the left panel in Figure 6 by looking at the interception of the absolute error curves with the 1:1 straight line. It can be seen in the figure that there is a direct relation between observed frequency and detection limit, with better detection limits the higher the channel frequencies. As discussed in the previous paragraph, this is related to the general stronger cloud signal of the higher frequencies.

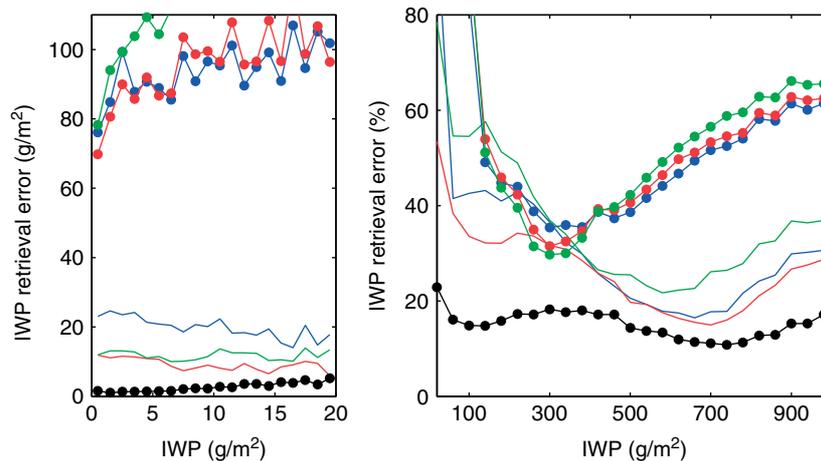


Figure 5. Retrieval performance for noise-free single channels. The simulations are for the midlatitude winter scenario of the uniform Colorado dataset. Clouds at all altitudes are present in the dataset. The lines without circles plot the error for the receivers R1 at 183.3 GHz (blue), R2 at 243.2 GHz (red) and R3 at 325.1 GHz (green), the lines with circles for R4 at 448 GHz (blue), R5 at 664.0 GHz (red), and R6 at 874.4 GHz (green). Only one channel per receiver is used; for the receivers with three channels the furthest away from the line centre is used. The errors for the all-receiver instrument are also plotted as a black line with circles.

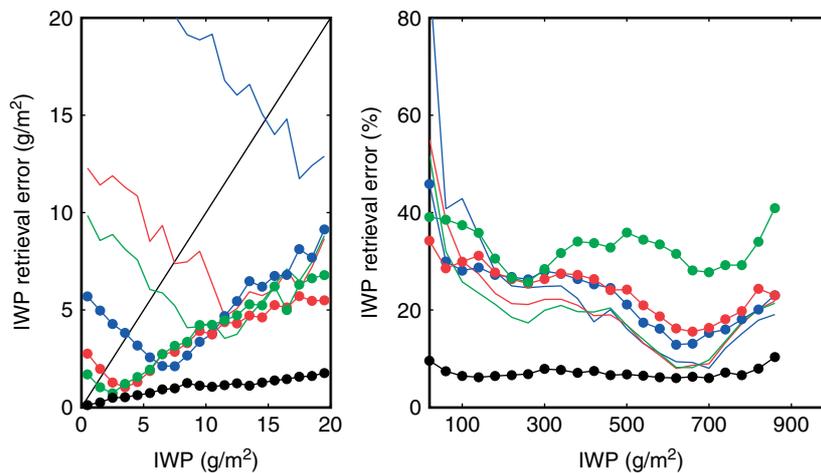


Figure 6. Retrieval performance for noise-free single channels as in Figure 5, but only for clouds with ZME larger than 7 km present in the dataset.

3.2. Performance for different receiver selections

We continue the discussion of the instrument performance by studying the inversion of the uniform Colorado dataset by different receiver selections. Now the inversions assume receivers with the noise figures given in Table I. The placement of the receiver centre frequencies and channel offsets was done in order to have sensitivity to a large range of ice mass, particle size and cloud altitude, and instruments with different receiver selections will have different sensitivities.

We start by assuming a simplified instrument with only three receivers. Figure 7 shows the retrieval performance for such an instrument with different receiver combinations. The simulations shown are for the midlatitude winter scenario. It can be seen that the receiver selections having the lower frequencies lower the DME error for large particles and the ZME error for low clouds, while the combinations having the high-frequency

channels lower the DME error for small particles and the ZME error for high clouds. Regarding IWP, the receivers having the high-frequency channels have larger errors than the other selections. As discussed for the single-channel instruments, these receivers are more sensitive to IWP than the low-frequency channels, but they also saturate at smaller IWP values, so the selections having mostly high-frequency channels have the larger IWP error for very thick clouds. The effect of the sounding altitude can also be seen here by comparing for instance the errors for R134 and R345. Having R1 instead of R5 lowers considerably the IWP error for the whole IWP range. The errors for an all-receiver instrument are also plotted for comparison, and it can be seen that no 3-receiver instrument can match the retrieval performance of the all-receiver instrument.

Figure 7 also shows that the retrieval performance tends to deteriorate at the extremes of the retrieval

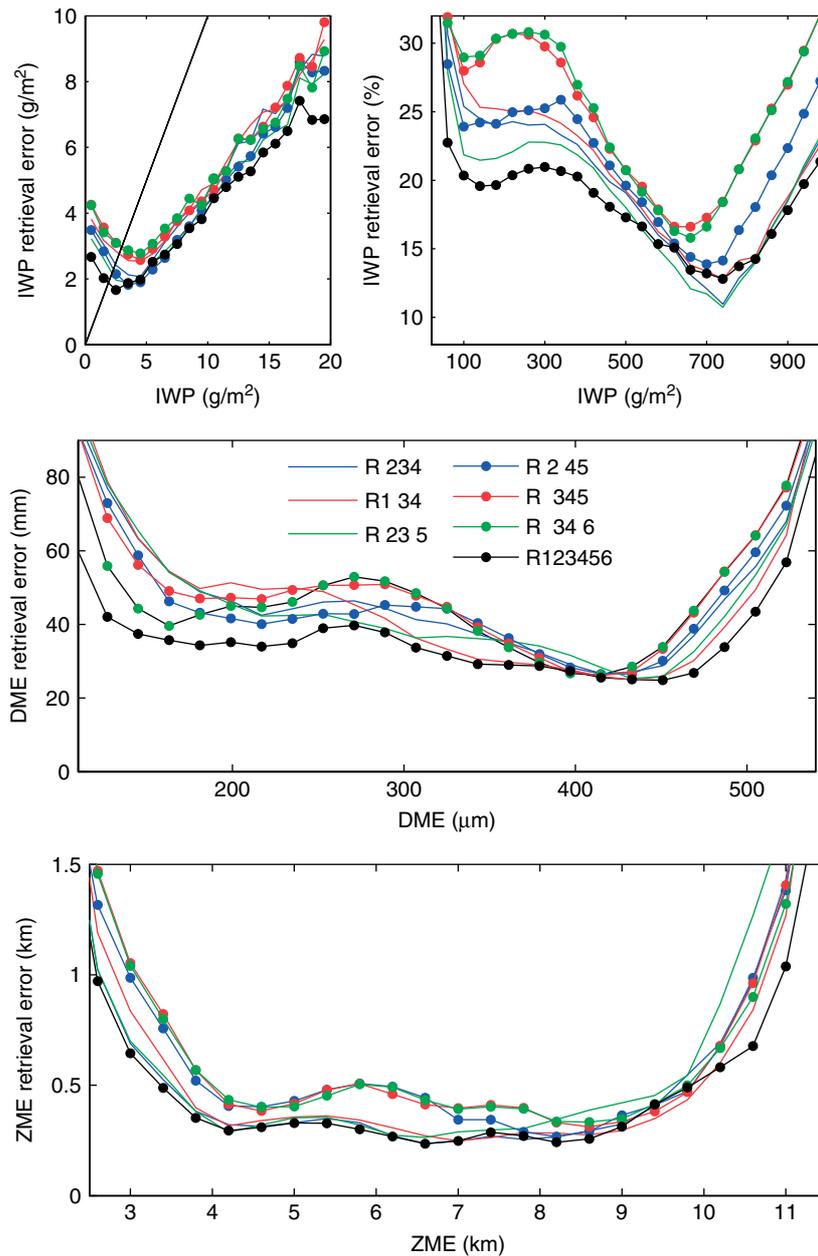


Figure 7. Retrieval performance for different 3-receiver instruments. The simulations are for the midlatitude winter scenario of the uniform Colorado dataset, and the receivers are simulated with the noise levels of Table I. The solid lines without circles plot the error for the selections R234 (blue), R134 (red) and R235 (green), the lines with circles for R245 (blue), R345 (red), and R346 (green). The errors for the all-receiver instrument are plotted as a black line with circles.

ranges. In some cases there are physical reasons for such high errors. For instance, the high errors for very low IWP are expected as the cloud signature in the radiances is very weak for very low IWP. Larger ZME errors for very high clouds are also possible as the brightness temperature depression for a given IWP and DME becomes constant above a certain altitude, and consequently there is no more altitude information in the cloud signal. However, these errors are also accentuated in the present simulations by the representation errors discussed in Section 2. Even the more uniform dataset used for the simulations has a small number of data cases at the extremes of the state space, and the posterior

distribution cannot be modelled correctly. For instance, although the wavelengths of the channels used impose a limit in the retrievable particle size, and large retrieval errors for very small or very large particles can be expected, with the present range of frequencies the sensitivity to 500 or 550 μm should not be as different as the simulations show, and a large part of this error increase seems related to the limited number of data cases at the end of the retrieval range. Another example is the apparent minimum in the IWP error around 700 g m^{-2} for all the receiver combinations, connected to an increase of the representation error associated with a decrease in the number of data cases for very large IWP values.

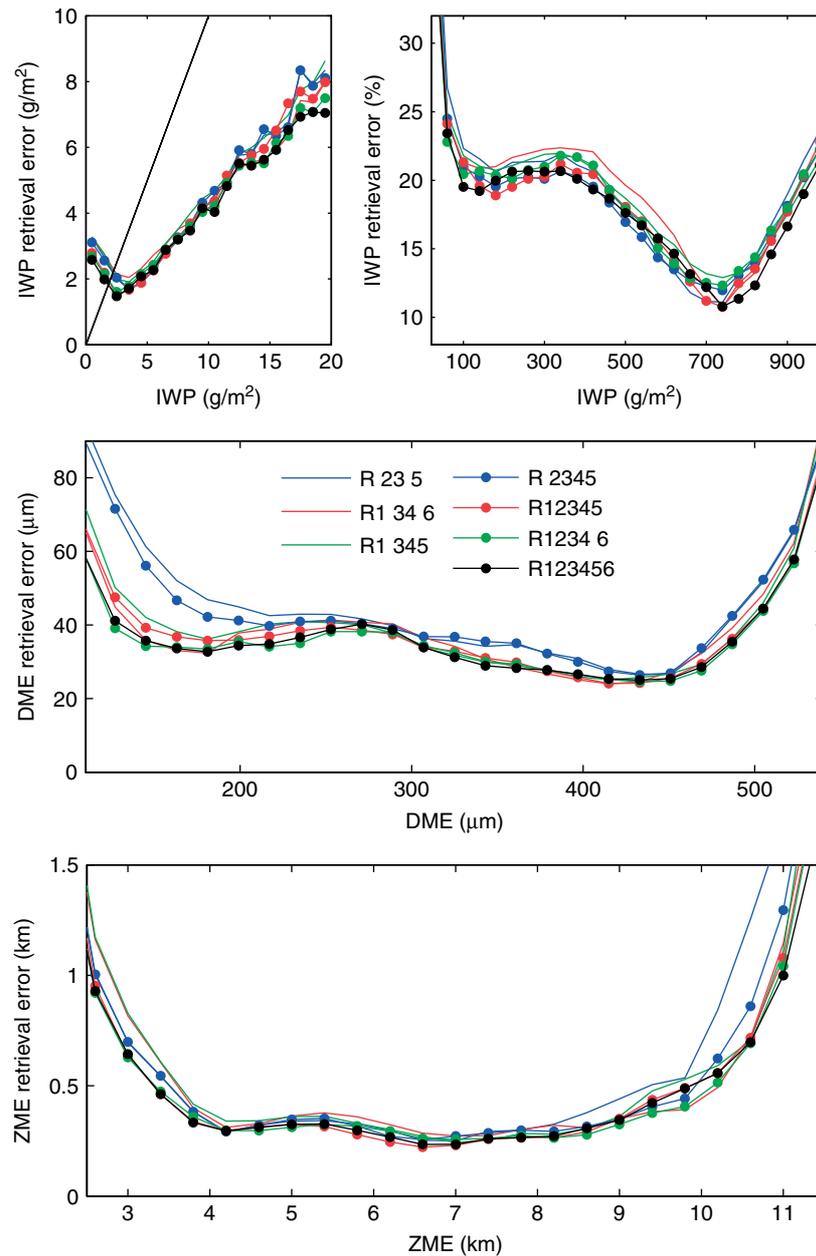


Figure 8. Retrieval performance as in Figure 7, but for instruments with 3 to 6 receivers. The lines without circles plot the error for the selections R235 (blue), R1346 (red) and R1345 (green), the lines with circles for R2345 (blue), R12345 (red), and R12346 (green). The all-receiver errors are plotted as a black line with circles.

When a larger number of receivers is available in the instrument, the error curves for the different receiver selections start to get closer. The error curves for the midlatitude winter scenario are plotted in Figure 8. For IWP, the two 4-receiver instruments not having R2 have larger errors for thin clouds compared with R2345. Selection R2345 performs nearly as well as the 5-receiver selections. R12345 and R12346 have also errors very close to the all-receiver instrument. Depending on IWP range, R12345 performs slightly better or slightly worse, but the differences are relatively small. A detection limit of around 2 gm^{-2} can be estimated from the absolute error curves. For DME, the 4-receiver instrument R1346 has smaller errors than the R2345 selection. This is

related to a larger sensitivity to the smallest and largest particles, as this selection has the lowest and highest frequency channels. R1346 performs nearly as well as the 5-receiver selections R12345 and R12346. Selections R12345 and R12346 have again very similar errors, though R12346 seems to perform a little better for very small particles due to the smaller wavelength of R6. For ZME, the curves for the different 4-5-6 receiver selections are very close, with only R2345 doing a little worse than the others for very high clouds.

The same analysis was repeated with the tropical uniform version of the Colorado dataset. Figure 9 displays the retrieval performance. The relative importance of the different selections is not very different from the

midlatitude winter analysis, but the differences between dropping R5 or R6 are now larger for thin high-altitude clouds with small particles. This can be explained by the presence of smaller ice particles and higher clouds in the tropical scenario, making the relevance of the highest frequency channel different between the two datasets. As the plots show, having R6 contributes to the lowering of the retrieval error for these cases. The plots also show a detection limit comparable to the value of 2 gm^{-2} obtained for the midlatitude winter scenario.

A summary of the retrieval errors in three different bins for each of the three retrieval variables and the different receiver selections is given in Table IV. For completeness, Table V shows the retrieval errors for similar inversions but with the original Colorado dataset.

The relative importance of the different selections is similar for both datasets. The errors are for most of the cases smaller for the simulations with the more uniform dataset, apart from the low bins of IWP and DME. This pattern can also be observed in Figure 10, where the retrieval errors for the all-receiver instrument are plotted for both original and modified datasets. As anticipated in Section 2, due to the procedure to extract the uniform dataset, there are very few data cases in the low bins compared with the original datasets, and this results in a larger possibility for higher representation errors there.

Figure 10 also shows the Chalmers dataset retrieval errors for the all-receiver instrument. As the datasets differ in several ways, an identical error analysis with

Table IV. Retrieval errors for the different receiver selections and scenarios presented in Section 3, for the more uniform Colorado dataset.

Receiver	IWP (%)			DME (μm)			ZME (km)		
	low	medium	high	low	medium	high	low	medium	high
134	51(61)	24(26)	18(18)	61(43)	37(33)	42(49)	0.60(0.46)	0.31(0.27)	0.47(0.41)
234	48(60)	24(27)	19(18)	61(53)	42(36)	49(56)	0.53(0.38)	0.29(0.25)	0.53(0.43)
235	46(55)	28(23)	19(17)	59(44)	38(36)	46(54)	0.53(0.38)	0.31(0.28)	0.66(0.82)
245	46(51)	25(23)	22(19)	52(38)	44(38)	52(60)	0.76(0.60)	0.43(0.35)	0.53(0.45)
345	52(53)	29(26)	27(21)	54(35)	46(38)	57(64)	0.60(0.55)	0.31(0.32)	0.53(0.43)
346	51(48)	30(25)	27(21)	47(28)	46(38)	7(63)	0.80(0.55)	0.45(0.33)	0.53(0.43)
1345	44(48)	22(21)	17(16)	41(26)	34(32)	40(47)	0.60(0.41)	0.31(0.25)	0.53(0.38)
2345	43(48)	20(20)	18(16)	52(34)	38(35)	46(51)	0.53(0.40)	0.30(0.24)	0.50(0.44)
1346	43(43)	22(21)	18(17)	38(24)	36(32)	42(45)	0.48(0.42)	0.30(0.24)	0.45(0.40)
12345	41(44)	20(19)	18(16)	40(27)	35(32)	39(45)	0.47(0.00)	0.26(0.00)	0.51(0.00)
12346	40(42)	21(20)	18(16)	36(23)	34(32)	38(45)	0.48(0.32)	0.30(0.20)	0.45(0.37)
123456	39(39)	21(19)	17(15)	35(24)	34(32)	39(44)	0.47(0.32)	0.28(0.22)	0.49(0.41)

The numbers in brackets denote the tropical scenario, and those without brackets the midlatitude winter scenario.

The errors are estimated in three different bins for each retrieval variable.

For IWP and both scenarios, the low bin corresponds to $0\text{--}50 \text{ gm}^{-2}$, the medium to $200\text{--}400 \text{ g m}^{-2}$, and the high to $800\text{--}1000 \text{ g m}^{-2}$.

For DME and the midlatitude winter scenario, the bins correspond to $100\text{--}200 \mu\text{m}$, $250\text{--}350 \mu\text{m}$, and $450\text{--}550 \mu\text{m}$, while for the tropical scenario the bins are $50\text{--}150 \mu\text{m}$, $225\text{--}325 \mu\text{m}$, and $400\text{--}500 \mu\text{m}$, respectively.

For ZME and the midlatitude winter scenario, the bins correspond to $2\text{--}4 \text{ km}$, $5\text{--}7 \text{ km}$, and $9\text{--}11 \text{ km}$, while for the tropical scenario the bins are $5\text{--}7 \text{ km}$, $8\text{--}10 \text{ km}$, and $12\text{--}14 \text{ km}$, respectively.

Table V. Retrieval errors as in Table IV but for the original Colorado dataset.

Receiver	IWP (%)			DME (μm)			ZME (km)		
	low	medium	high	low	medium	high	low	medium	high
134	42(56)	25(27)	30(31)	33(28)	39(39)	86(66)	1.35(0.80)	0.41(0.37)	0.70(0.53)
234	47(60)	22(28)	31(29)	35(31)	43(41)	81(68)	1.13(0.65)	0.38(0.41)	0.73(0.53)
235	41(51)	21(25)	27(28)	34(31)	43(43)	77(66)	1.03(0.66)	0.38(0.42)	0.73(0.65)
245	42(51)	22(26)	34(32)	36(31)	44(45)	83(66)	1.36(0.90)	0.55(0.54)	0.77(0.57)
345	41(51)	26(27)	36(37)	34(29)	45(44)	93(73)	1.50(0.97)	0.50(0.51)	0.72(0.55)
346	41(47)	27(28)	39(34)	32(28)	45(43)	93(73)	1.50(0.98)	0.50(0.51)	0.70(0.55)
1345	37(48)	22(22)	30(28)	30(25)	39(37)	83(59)	1.25(0.70)	0.39(0.36)	0.67(0.54)
2345	39(49)	21(22)	30(27)	33(29)	43(41)	77(62)	1.02(0.71)	0.38(0.39)	0.71(0.57)
1346	36(44)	24(22)	36(27)	28(23)	38(35)	83(57)	1.29(0.71)	0.39(0.37)	0.64(0.53)
12345	36(47)	21(21)	31(27)	29(24)	39(36)	73(54)	0.94(0.60)	0.35(0.36)	0.62(0.52)
12346	35(43)	23(21)	30(27)	27(23)	37(36)	71(55)	0.98(0.54)	0.36(0.36)	0.62(0.53)
123456	34(41)	19(20)	27(25)	27(23)	37(35)	72(54)	0.99(0.61)	0.35(0.37)	0.62(0.52)

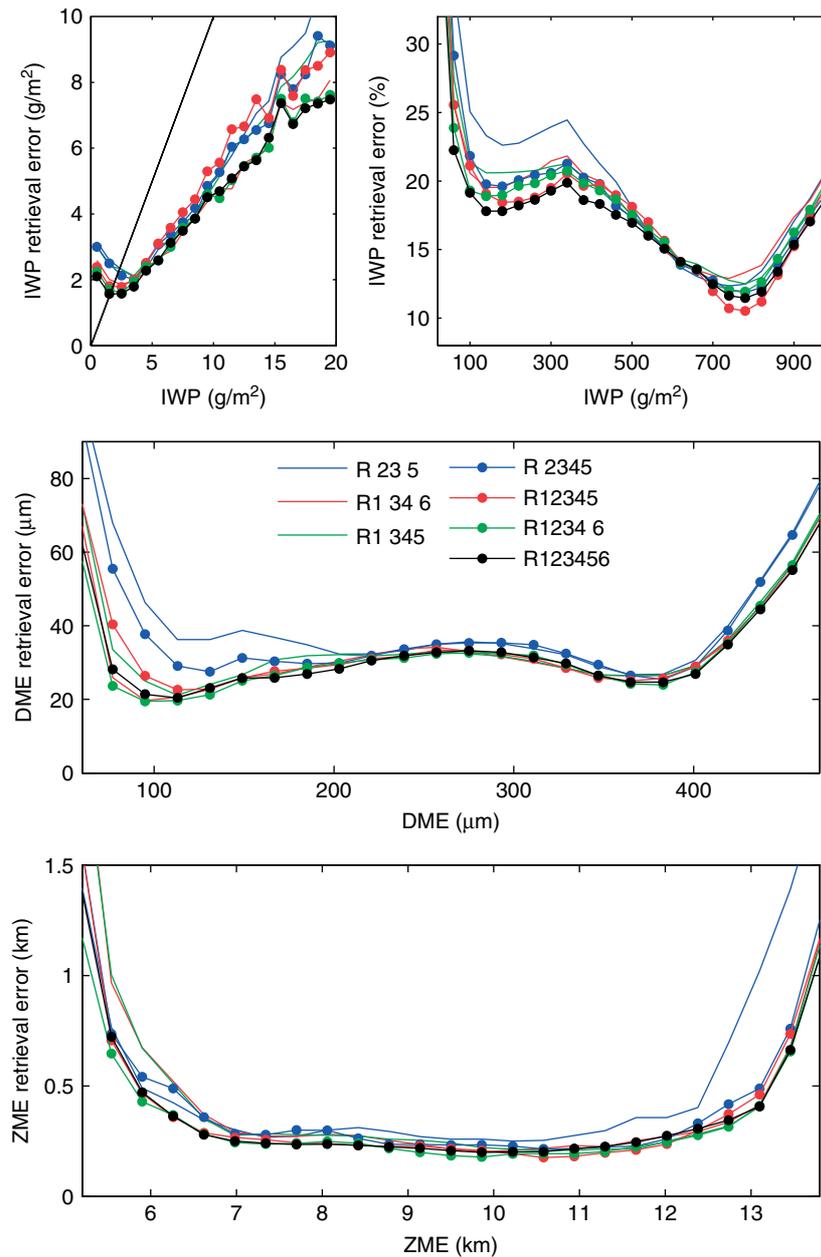


Figure 9. Retrieval performance as in Figure 8, but for the tropical scenario of the uniform Colorado dataset.

both datasets cannot be expected. However, some degree of consistency is desirable as both datasets share the same objective, to simulate a cloud ice observing system. This seems to be the case here, although some differences in the error curves can clearly be seen. For IWP values smaller than 5 gm^{-2} , Chalmers retrievals show smaller absolute errors and a lower detection limit, and this seems related to the dataset prior distributions. The Chalmers dataset was not made more uniform, resulting in a distribution with a much larger number of cases in the range with very low IWP, as in the original Colorado dataset. Differences can also be observed for intermediate IWP values. In this case it is more difficult to find out the exact reason for the disagreement, but it could be related to the different assumptions on particle shape and size distributions in the two datasets, resulting in a more

complicated retrieval situation for the Colorado dataset. For the Colorado dataset a single cloud is purely one particle shape, which emphasizes the retrieval error due to particle size, while for the Chalmers dataset there are three different particle types in each cloud. Also, the particle size distribution inside the cloud has a more detailed structure in the Chalmers dataset than in the Colorado dataset. But this detailed structure tends to be averaged out when estimating the integrated parameters, and the variations from cloud to cloud are not as large as the variations in the Colorado dataset. This should result in smaller errors for the Chalmers dataset, reflecting an easier retrieval situation. Regarding DME, Chalmers retrievals are better between 150 and 300 μm , possibly for the same reasons, but noticeably worse outside this range, especially for DME values larger than 300 μm .

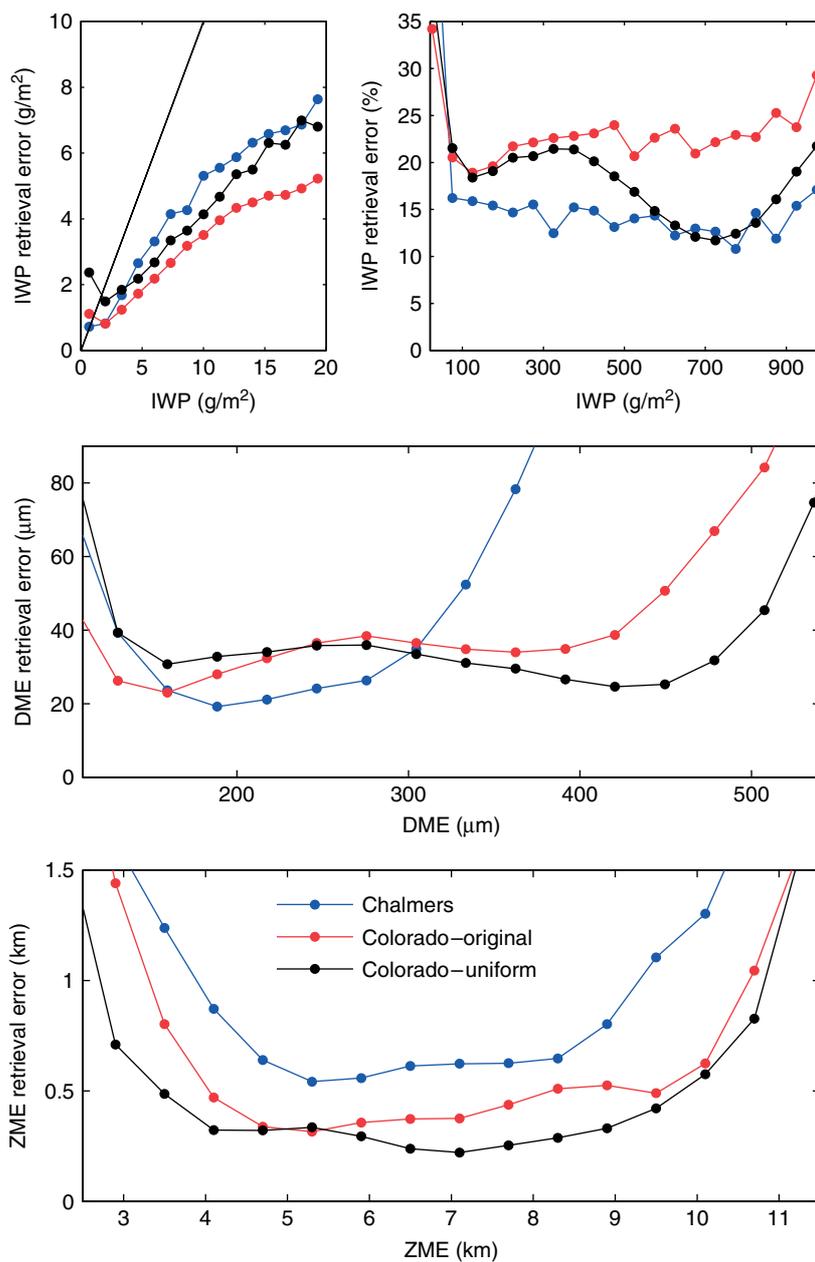


Figure 10. Retrieval performance for the all-receiver instrument. The simulations are for the midlatitude scenario of the original Colorado dataset (red), the modified more uniform Colorado dataset (black), and Chalmers datasets (blue), and the instrument is assumed to have the noise levels of Table I.

This seems related to the prior distribution of the dataset, rather than a physical limitation in the retrievals, as the number of data cases with DME larger than $300\ \mu\text{m}$ in the Chalmers dataset is small. The largest difference in the retrievals occurs for ZME, where the errors for the Chalmers dataset are nearly double the errors of the Colorado dataset for the whole range. This is possibly due to the larger variability in temperature and water vapour in the Chalmers dataset. On one hand, the Colorado dataset contains only winter variability, while the Chalmers dataset reflects a whole year variability. On the other hand, the temperature variability in the Colorado dataset was reduced in order to simulate the existence of auxiliary temperature information in the

retrieval. The result is a larger ZME error when the associated uncertainties propagate through the retrieval.

3.3. Performance for different noise levels

It is of interest to see the robustness of the retrieval performance with respect to possible changes in the receiver noise levels. Not only possible changes in the radiometric noise, but other sources of error not modelled here might contribute with a random component that could be added in quadrature to the estimated receiver noise. To test this, the noise levels assumed in the previous simulations (see Table I) were scaled by factors of 0.5, 1.0, 1.5, 2.0, 2.5 and 3.0. Figure 11 gives the

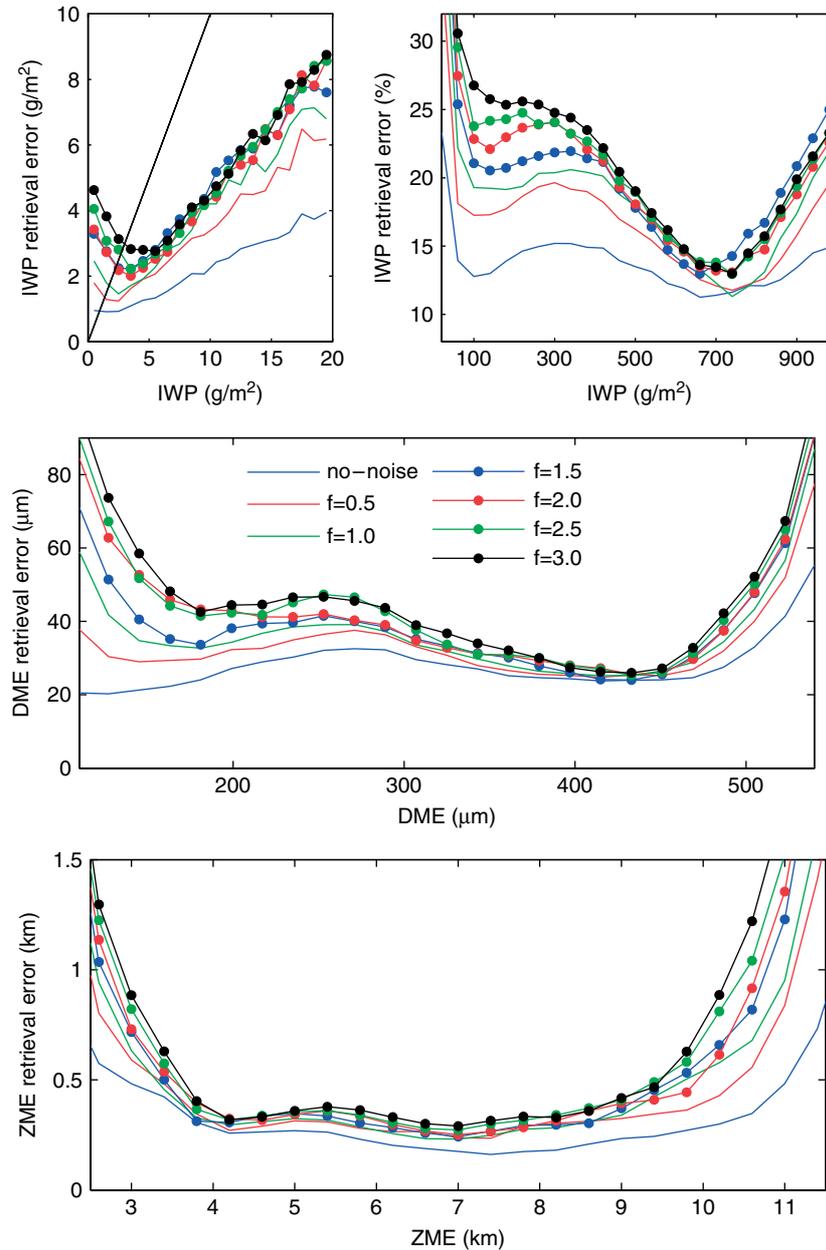


Figure 11. Retrieval performance for the all-receiver instrument with different noise levels. The simulations are for the midlatitude scenario of the uniform Colorado dataset. The instrument is assumed to have the noise levels of Table 1 scaled by different factors ($f = 0.5, 1.0, \dots, 3.0$). The lines without circles plot the error for the noise-free case (blue), noise scaled by $f = 0.5$ (red) and by $f = 1.0$ (green), the lines with circles for noise scaled by $f = 1.5$ (blue), $f = 2.0$ (red), $f = 2.5$ (green), and $f = 3.0$ (black).

error distribution for each of the retrieval variables for the midlatitude winter version of the uniform Colorado dataset. The plots show that there is a scaling of the errors when the noise levels are increased. For instance, the total error figures for the noise-free case were around 17% for IWP, 27 μm for DME, and 240 m for ZME, while for the nominal noise level (scaling factor of 1.0) they increase to values of around 23% for IWP, 33 μm for DME and 310 m for ZME. The difference between these figures is then the contribution of the sensor to the error budget. The plot also suggests that the instrument is robust with respect to variations around the estimated noise figures. A 50% increase in the estimated noise levels does not degrade the retrieval performance by

much; the error curves for the 1.5 scaling factor are only slightly higher than the curves for a scaling factor of 1.0. The detection limit is also affected by the noise levels; from a detection limit of around 1 gm^{-2} for the noise-free case, we move up to around 3 gm^{-2} for the largest noise levels. As expected, the higher the noise level, the more difficult it is to extract a weak cloud signal from the radiances and the poorer the detection limit.

The comparison of Figure 5 with Figure 11 suggests that there is some trade-off between receiver selection and noise levels. By looking at the figures, one might expect that an instrument with a smaller number of less noisy receivers could compete with an instrument with a larger number of noisier receivers. Retrievals with different

receiver selections and half and double the estimated noise levels of Table I were added to the previous retrievals of Figure 8. The retrieval errors for IWP are summarized in Figure 12, and the approximated detection limits in Figure 13. The retrieval error for each retrieval variable, receiver selection and noise level was calculated for three different bins representing low, middle and high values for each variable. If a retrieval error was given, the figures could be used to select a range of receiver-noise selections with an acceptable error. For instance, it seems that the selections with four receivers included in the analysis could have retrieval errors similar to the selections with larger number of receivers if the noise levels could be reduced by a factor of 2.

3.4. Performance of a dual-polarized instrument

The Chalmers dataset containing oriented particles was used to check the usefulness of having dual-polarized receivers. The results are evaluated in the context of the state vector to be retrieved, which includes an estimation of an integrated particle size but not its shape and orientation. Nevertheless, particle aspect ratios were also retrieved and results are shown.

Three instrument concepts were tested. The two immediate options were an instrument with total power receivers and an instrument with all receivers measuring vertical and horizontal polarization. A third concept was an instrument with channels 183.31 ± 7.0 GHz, 325.15 ± 9.5 GHz and 664.0 ± 4.2 GHz measuring vertical and horizontal polarization, while the remaining channels measure total power. The polarized channels were selected after studying the polarization signal in the radiances, i.e. the difference between the vertical and horizontal components. The strongest polarization

effects occur for R2 and R3. The polarization signal of R2 can be problematic to analyse due to the surface effect, so one channel for R1 was selected instead. For the high-frequency receivers, R5 was chosen as it had slightly higher polarization signals than R6. For all channels, the polarization signal was below 5 K for particles with aspect ratios below 1.3, though for larger aspect ratios polarizations up to 16 K were observed. Taking into account that the polarization signal comes from the differences of two measurements, such small signals can be hard to distinguish from the receiver noise, and these simulations were started with lower noise levels than the standard levels given in Table I. A noise-equivalent system temperature of 0.5 K for each polarized channel has been assumed, and correspondingly $0.5/\sqrt{2}$ K for the total power receivers, to have equivalent noise levels.

The results are presented in Table VI. Some decrease in the IWP and DME errors is observed for the noise-free simulations, but even adding the modest noise levels used here degrades the performance of the instrument having polarized channels to the levels of the non-polarized instrument. The retrieval of aspect ratios benefits from the polarized instruments, with nearly half the errors of the total power instrument for the noisy simulations. However, this result should be taken with some care, as polarization may depend strongly on the inhomogeneity of the 3D atmospheric cloud field and these simulations assumed a 1D atmosphere. Although the result is of interest, it cannot on its own justify the inclusion of polarized channels in the discussed instrument as aspect ratio is not one of the primary retrieval parameters.

3.5. Synergy with infrared data

Inversions of combined submillimetre and infrared radiances were performed in order to test the synergy of

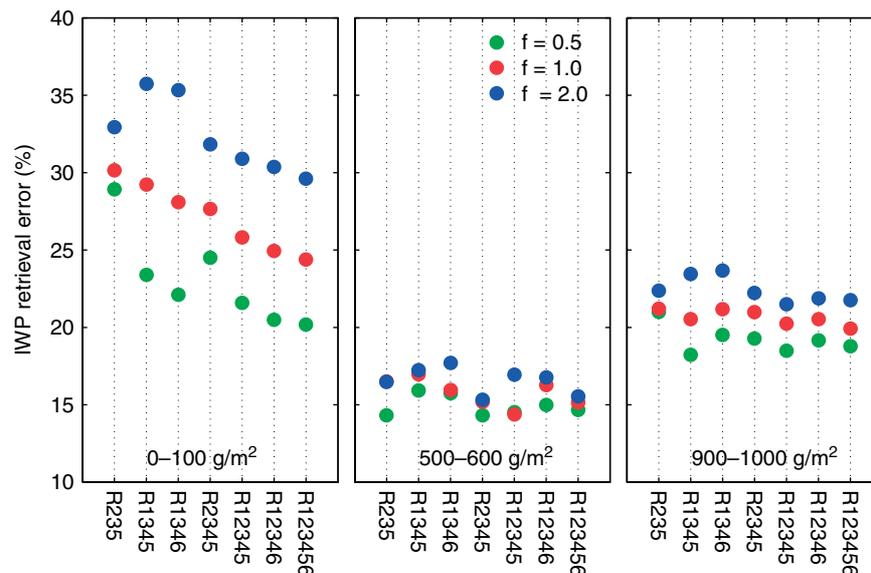


Figure 12. IWP retrieval error for different combinations of receivers and noise scenarios. The simulations are for the midlatitude winter scenario of the uniform Colorado dataset. The selections of receivers are given on the horizontal axis, and the scalings by a factor $f = 0.5, 1.0, 2.0$ of the nominal noise levels given in Table I are plotted with different colours. The retrieval error is calculated in three different bins for each retrieval variable, with the span of the bin given at the bottom of each plot.

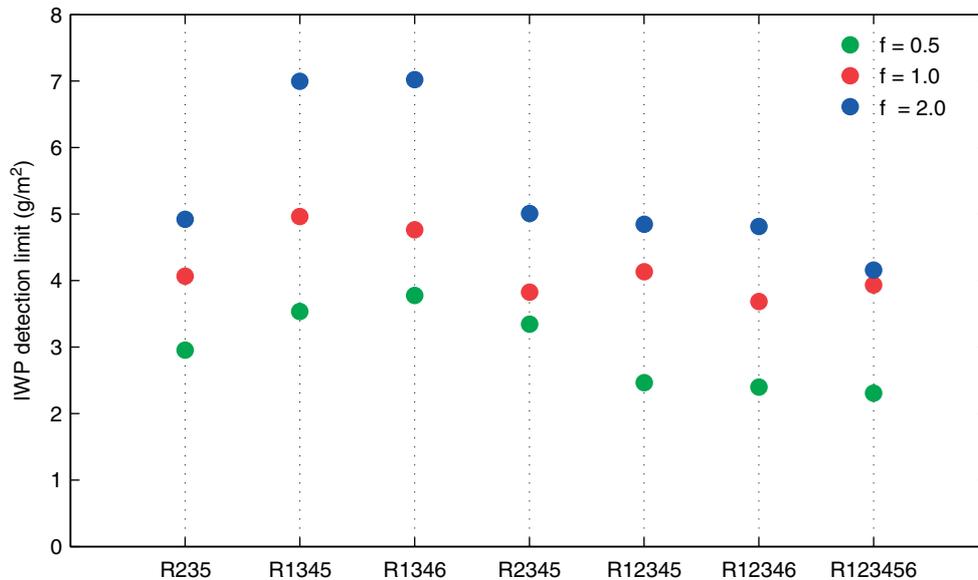


Figure 13. IWP detection limit for different combinations of receivers and noise figures for the midlatitude winter scenario of the uniform Colorado dataset. The selections of receivers are given on the horizontal axis, and the scalings by a factor $f = 0.5, 1.0, 2.0$ of the nominal noise levels given in Table I are plotted with different colours.

having submillimetre and infrared data in the cloud ice retrievals. The infrared channels were assumed to sample identical atmospheres as the submillimetre radiances, i.e. collocation between the submillimetre and infrared measurements was assumed to be perfect and the propagation path to be identical. The submillimetre instrument was assumed to have either receivers R12345, i.e. all channels but the highest frequency channel, or R123456, i.e. all the channels. The simulated infrared channels were placed at the old Geostationary Operational Environmental Satellite (GOES) window channels 10.7 and 12.0 μm and the training datasets extended by running the forward model on the different clouds and atmospheric realizations. Note that no optimization of the infrared channel placement was performed. The surface emissivity in the dataset was Gaussian distributed, as for the submillimetre channels, but with a mean of 0.97 and standard deviation of 0.01. A noise level of 0.6 (mW m^{-2})(cm sr^{-1}), corresponding to about 0.5 K at 260 K in the Planck function, was added to both infrared channels. For the submillimetre radiances the standard noise levels of Table I were added. The inversions are presented for the tropical version of the uniform Colorado dataset, as this is the scenario where adding the infrared channel has the larger effect due to the presence of smaller particles.

Figure 14 shows the retrieval performance for the submillimetre and infrared instruments and when the instruments are combined. The first thing to note is that the performance of the pure infrared instrument is very poor. This error curve does not really do justice to such a type of instrument, since no attempt has been made to optimize its retrieval performance. The retrieval simulations for the infrared instrument use exactly the same algorithm as for the submillimetre instrument, and it is quite likely that meaningful retrievals from a pure infrared

Table VI. Retrieval errors (given as relative values) for the three instrument concepts described in the text.

	A	B	C
IWP	17.0 (15.0)	15.7 (13.4)	15.6 (11.0)
DME	5.9 (5.4)	5.7 (4.5)	6.0 (4.4)
AR	10.2 (6.7)	6.3 (4.6)	6.4 (3.9)

A: an instrument with total power receivers.

B: a mixture of polarized and unpolarized channels.

C: with all receivers measuring vertical and horizontal polarizations.

Retrieval errors for the particle aspect ratio (AR) are also given.

The figures in brackets denote the noise-free case, and those without brackets the situation when noise is added to the radiances.

instrument would have required stronger prior assumptions, for instance, on temperature. Error curves for the case where the instrument would not add any information to the measurements, i.e. the retrieval algorithm outputs the mean state of the prior distribution, are added to the plots to help understand the pure infrared performance. The similarity between these curves and the infrared errors shows that not much information can be extracted from the retrieval algorithms when only the infrared channels are considered. Nevertheless, adding the infrared channels to the submillimetre instrument has a noticeable effect on the retrievals for the higher and thinner clouds with the smallest particles. The information provided by the infrared channels can now be effectively used once the information from the submillimetre channels helps reduce the dependency of the infrared channels on the prior information. For IWP the infrared channels lower the detection limit from the submillimetre value of around 2 gm^{-2} to a value of around 1 g m^{-2} . Reductions close to 10% can be seen in retrieval error in the lowest IWP range. The reduction in the IWP error for particles

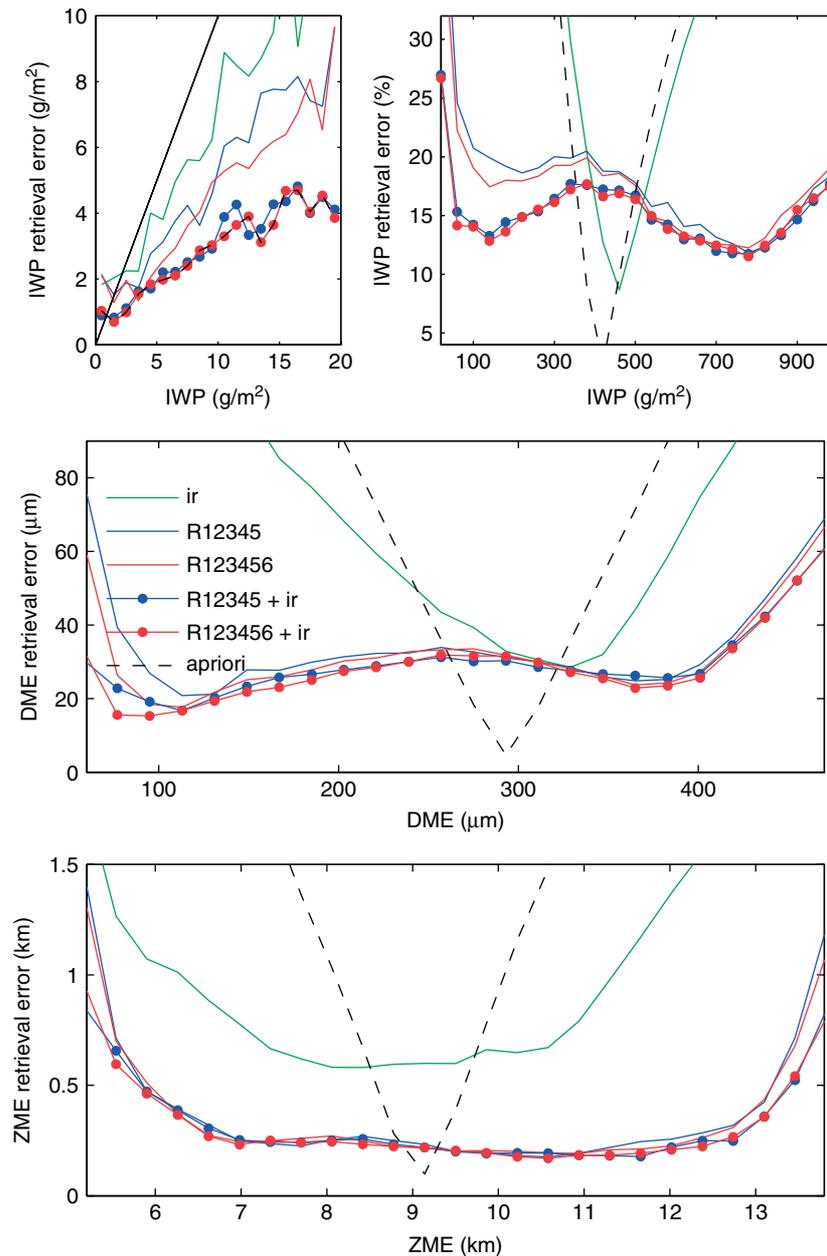


Figure 14. Retrieval performance for different selections of submillimetre receivers and infrared channels for the tropical scenario of the uniform Colorado dataset. The receiver noise levels for the submillimetre channels are given in Table I, and for the infrared channels in the text. The lines without circles plot the error for the infrared channels (green), the submillimetre selection R12345 (blue) and R123456 (red), the lines with circles when the infrared channels are added to the submillimetre selections R12345 (blue) and R123456 (red). A dashed black line displays the error if there was no information in the measurements, i.e. the retrieval solution is the mean state of the prior distribution.

smaller than $100\ \mu\text{m}$ is very remarkable as well. It can also be seen that adding the infrared instrument to the full submillimetre instrument, or the instrument missing the highest frequency channel, results in a very similar IWP retrieval performance. The loss of sensitivity to very small particles that occurs when dropping the R6 channel appears to be compensated by the large sensitivity of the infrared channels to very small particles. For DME, the infrared channels also lower the error for very small particles, and it can be seen that the selection without R6, but with the infrared channels, can match the performance of the full submillimetre instrument.

4. Summary

This paper summarizes the performance simulations for a submillimetre-wave cloud ice satellite instrument. The proposed instrument is a conical scanner measuring the brightness temperature depressions caused by the scattering of up-welling radiation by the ice particles. The size of the depressions at different frequencies is used for an estimation of the cloud integrated ice water path, the cloud equivalent sphere diameter, and the cloud median ice mass height by an inversion algorithm based on neural networks.

The neural network algorithm proved adequate to study the retrieval performance of the instrument. The neural network retrieval models the posterior distribution of the state vector and provides as a solution its mean state. Another retrieval algorithm based on a Monte Carlo integration of the posterior distribution was also tried. However, for the size of the datasets available for this study, the neural networks were able to do retrievals with smaller errors, making the neural networks the preferred choice for the inversions.

Two independent datasets were available to study the retrieval performance of the instrument. Most of the retrieval simulations were carried out with the Colorado dataset, as it provided a larger number of realizations for two different atmospheric scenarios. A number of retrievals were also performed with the Chalmers dataset, to add robustness to the reported results and to study an instrument with polarized receivers. Inversions of IWP from both datasets showed retrieval errors of the same order. The fact that two datasets generated completely independently of each other and with two different radiative transfer codes gave consistent retrieval errors suggests some general validity of the results presented here.

The prior distribution of the datasets was found to affect the retrieval errors reported. The number of data cases in some regions of the state space was very limited, resulting in poor modelling of the posterior distribution. To avoid this, a more uniform dataset where all the regions were more equally represented was extracted from the original Colorado dataset and used for most of the retrieval simulations. Comparisons with retrievals from the original dataset proved that this approach was useful to reduce the representation error, but it did not completely eliminate it, specially at the ends of the retrieval ranges where the data cases were still very sparse.

The specific mapping given by the datasets determines the details of the error analysis, so the limitations of the datasets ought to be considered when evaluating the study findings. Only 1D atmospheric fields were included in the datasets, i.e. the impact of cloud inhomogeneity in the retrieval errors has not been addressed; only two possible atmospheric scenarios were included—tropical anvil clouds and midlatitude synoptic cirrus; and only three particles shapes per dataset were considered. Nevertheless, the datasets generated are consistent with our present capability to simulate a cloud ice observing system and were judged sufficient for a detailed analysis of the instrument retrieval performance.

A six-receiver instrument was considered as baseline for the sensor model. Three receivers were placed around the 183, 325 and 448 GHz water vapour lines, with three channels per receiver, while the other three were placed in the 243, 664 and 874 GHz windows. The placement of the centre frequencies and offset of the channels was done in order to have sensitivity to a large range of ice mass, particle size and sounding altitude. The sensor model was kept simple, simulated by the estimated system

temperature for each channel. However, simulations with different noise levels were also conducted to see the impact of an increase in the sensor noise levels or the possible impact of other unmodelled uncertainties resulting in a new source of random errors. A scaling of the originally estimated noise levels resulted in a scaling of the retrieval errors and affected the usefulness of the different channels, but the performance was quite robust against changes in the noise levels, as even a 50% increase in the estimated noise levels did not degrade the retrieval performance much.

Simulations with different receiver selections were run to see the relative merits of the different channels and to find an optimal selection in terms of retrieval performance. For the all-receiver baseline sensor, overall median relative errors of around 20% for IWP, 33 μm for DME, and 240 m for ZME for a midlatitude winter scenario, and 17% for IWP, 30 μm for DME, and 310 m for ZME for a tropical scenario were found. If a detection limit is defined as the value where the retrieval error reaches a 100% relative error, values of around 2 gm^{-2} were found. For the midlatitude winter scenario, a five-receiver instrument, where one of the two receivers at the highest frequencies was dropped, gave errors very close to the all-receiver instrument. However, for the tropical scenario where higher clouds and smaller ice particles are present, the highest-frequency receiver proved useful to reduce the errors for very thin and high clouds. Some selections with four receivers also had errors not very different from the all-receiver instrument, but could not completely match its performance for all retrieval ranges. The same combinations of receivers but with different noise levels were also investigated. A possible trade-off between receiver selections and receiver noise was apparent. For instance, some of the four-receiver selections could compete with the all-receiver instrument if the four-receiver instrument could operate at half noise levels.

Dual-polarized measurements were also simulated by using a dataset containing oriented particles. Although the polarized channels notably improved the retrieval of the particles' aspect ratio, they did not prove useful to improve the retrievals of IWP and DME if realistic receiver noise was added to the instrument. As the main retrieval parameters of the proposed sensor are not the particle shape or orientation, the inclusion of polarized channels in the instrument could not be justified.

Finally the usefulness of adding information from two infrared channels to the inversions of the submillimetre instrument was considered. The infrared channels lowered the retrieval errors for the regions of state space with the smallest particles due to their larger sensitivity to small particles. Some trade-off between having the highest-frequency (submillimetre) receiver or the infrared channels emerged, as comparable performance was found between the submillimetre six-receiver instrument and the selection combining all the submillimetre channels, except the highest in frequency, with the infrared channels.

Acknowledgements

This work was carried out within an ESA-funded project: Establishment of mission instrument requirements to observe cirrus clouds at submillimetre wavelengths, ESTEC contract No. 19053/05/NL/AR. It is also a contribution to COST Action 723 'Data exploitation and modelling for the upper troposphere and lower stratosphere' (<http://www.cost.esf.org>). Thanks go to the ARTS radiative transfer community, many of whom have indirectly contributed by implementing features to the ARTS model used for some of this work.

References

- Bishop M. 1995. *Neural Networks for Pattern Recognition*. Oxford University Press: New York.
- Buehler SA, Eriksson P, Kuhn T, von Engeln A, Verdes C. 2004. ARTS, the Atmospheric Radiative Transfer Simulator. *J. Quant. Spectrosc. Radiat. Transfer* **91**: 65–93. DOI:10.1016/j.jqsrt.2004.05.051.
- Buehler SA, Jiménez C, Evans KF, Eriksson P, Rydberg B, Heymsfield AJ, Stubenrauch C, Lohmann U, Emde C, John VO, Streerekhya TR, Davis C. 2007. A concept for a satellite mission to measure cloud ice water path and ice particle size. *Q.J.R. Meteorol. Soc.* 109–128.
- Davis CP, Wu DL, Emde C, Jiang JH, Cofield RE, Harwood RS. 2005. Cirrus-induced polarization in 122 GHz Aura Microwave Limb Sounder radiances. *Geophys. Res. Lett.* **32**: DOI:10.1029/2005GL022681.
- Davis CP, Evans KF, Buehler SA, Wu DL, Pumphrey HC. 2006. 3-D polarised simulations of space-borne passive mm/sub-mm midlatitude cirrus observations: a case study. *Atmos. Chem. Phys. Discuss.* **6**: 12701–12728.
- Del Genio AD. 2001. CGM simulations of cirrus for climate studies. Chapter 15 in *Cirrus*. Oxford University Press: Oxford, UK.
- Emde C, Buehler SA, Davis C, Eriksson P, Sreerekha TR, Teichmann C. 2004. A polarized discrete ordinate scattering model for simulations of limb and nadir long-wave measurements in 1-D/3-D spherical atmospheres. *J. Geophys. Res.* **109**(D24): D24207 DOI:10.1029/2004JD005140.
- Evans KF. 2004. 'Submillimeter-wave ice cloud radiometry channel selection study'. University of Colorado: Boulder. (<http://nit.colorado.edu/>).
- Evans KF, Stephens GL. 1995a. Microwave radiative transfer through clouds composed of realistically shaped ice crystals. Part I: Single scattering properties. *J. Atmos. Sci.* **52**: 2041–2057.
- Evans KF, Stephens GL. 1995b. Microwave radiative transfer through clouds composed of realistically shaped ice crystals. Part II: Remote sensing of ice clouds. *J. Atmos. Sci.* **52**: 2058–2072.
- Evans KF, Walter SJ, Heymsfield AJ, Deeter MN. 1998. Modeling of submillimeter passive remote sensing of cirrus clouds. *J. Appl. Meteorol.* **37**: 184–205.
- Evans KF, Walter SJ, Heymsfield AJ, McFarquhar GM. 2002. Submillimeter-wave cloud ice radiometer: Simulations of retrieval algorithm performance. *J. Geophys. Res.* **107**: 2.1–2.21.
- Evans KF, Wang JR, Racette PE, Heymsfield G, Li L. 2005. Ice cloud retrievals and analysis with the Compact Scanning Submillimeter Imaging Radiometer and the Cloud Radar System during CRYSTAL FACE. *J. Appl. Meteorol.* **44**: 839–859.
- Eriksson P, Ekström M, Rydberg B, Murtagh PD. 2007. First Odin sub-mm retrievals in the tropical upper troposphere: Ice cloud properties. *Atmos. Chem. Phys.* **7**: 471–483.
- Foresee FD, Hagan MT. 1997. 'Gauss–Newton approximation to Bayesian regularization'. Pp 1930–1935 in Proceedings of the 1997 International Joint Conference on Neural Networks, Houston, Texas.
- Gill PE, Murray W, Wright MH. 1981. *Practical Optimization*. Academic Press: San Diego.
- Hagan MT, Menhaj M. 1994. Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Networks* **5**: 989–993.
- Heymsfield AJ. 2003. Properties of tropical and midlatitude ice cloud particle ensembles, Part II: Applications for mesoscale and climate models. *J. Atmos. Sci.* **60**: 2592–2611.
- Heymsfield AJ, Lewis S, Bansemer A, Jaquinta J, Miloshevich LM, Kajikawa M, Twohy C, Poellot MR. 2002. A general approach for deriving the properties of cirrus and stratiform ice cloud particles. *J. Atmos. Sci.* **59**: 3–29.
- Ivanova D, Mitchell DL, Arnott WP, Poellot M. 2001. A GCM parameterization for bimodal size spectra and ice mass removal rates in midlatitude cirrus clouds. *Atmos. Res.* **59–60**: 89–113.
- Korolev A, Isaac G. 2003. Roundness and aspect ratio of particles in ice clouds. *J. Atmos. Sci.* **60**: 1795–1808.
- Li JL, Waliser DE, Jiang J, Wu DL, Read D, Read W, Waters JW, Tompkins AM, Donner LJ, Chern JD, Tao JD, Tao WK, Atlas R, Gu Y, Liou KN, Genio AD, Khairoutdinov M, Gettelman A. 2005. Comparisons of EOS MLS cloud ice measurements with ECMWF analyses and GCM simulations. Initial results. *Geophys. Res. Lett.* **32**: DOI: 10.1029/2005GL023788.
- Nguyen D, Widrow B. 1990. 'Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights'. Pp 21–26 of Proceedings of the 1990 International Joint Conference on Neural Networks, San Diego.
- Ou SC, Liou KN, Goch WM, Takano Y. 1993. Remote sensing of cirrus cloud parameters using advanced very-high-resolution radiometer 3.7 and 10.9 μm channels. *Appl. Opt.* **32**: 2171–2180.
- Reinhardt T, Wacker U. 2004. Impact of ice particle habits on simulated clouds. *Geophys. Res. Lett.* **31**: L21106, DOI: 1029/2004GL021134.
- Rodgers CD. 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific Publishing: Singapore.
- Rolland P, Liou KN, King MD, Tsay SC, McFarquhar GM. 2000. Remote sensing of optical and microphysical properties of cirrus clouds using moderate-resolution imaging spectroradiometer channels: Methodology and sensitivity to physical assumptions. *J. Geophys. Res.* **105**: 11721–11738.
- Rumelhart DE, Hinton GE, Williams R. 1986. Learning representations by back-propagating errors. *Nature* **323**: 533–536.
- Rydberg B, Eriksson P, Buehler SA. 2007. Prediction of cloud ice signatures in sub-mm emission spectra by means of ground-based radar and *in situ* microphysical data. *Q. J. R. Meteorol. Soc.* **133**: 151–162.
- Stocker T. 2001. Physical climate processes and feedbacks. Chapter 7 in *Climate Change 2001: The Scientific Basis*. IPCC Third Assessment Report. Cambridge University Press: Cambridge, UK.
- Wilson D. 2000. The impact of a physically based microphysical scheme on the climate simulation of the Meteorological Office Unified Model. *Q. J. R. Meteorol. Soc.* **126**: 1281–1300.