Atmospheric
Measurement
Techniques

Toward autonomous surface-based infrared remote sensing of polar clouds: cloud-height retrievals

Penny M. Rowe[1,2], Christopher J. Cox[3,4], and Von P. Walden[5]

[1]NorthWest Research Associates, Redmond, WA, USA
[2]Physics Department, Universidad de Santiago de Chile, Santiago, Chile
[3]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA
[4]NOAA Earth System Research Laboratory, Physical Sciences Division, Boulder, CO, USA
[5]Department of Civil and Environmental Engineering, Washington State University, Pullman, WA, USA

Correspondence to: Penny M. Rowe (penny@nwra.com)

Abstract. Polar regions are characterized by their remoteness, making measurements challenging, but an improved knowledge of clouds and radiation is necessary to understand polar climate change. Infrared radiance spectrometers can operate continuously from the surface and have low power requirements relative to active sensors. Here we explore the feasibility of retrieving cloud height with an infrared spectrometer that would be designed for use in remote polar locations. Using a wide variety of simulated spectra of mixed-phase polar clouds at varying instrument resolutions, retrieval accuracy is explored using the CO_2 slicing/sorting and the minimum local emissivity variance (MLEV) methods. In the absence of imposed errors and for clouds with optical depths greater than ~ 0.3, cloud-height retrievals from simulated spectra using CO_2 slicing/sorting and MLEV are found to have roughly equivalent high accuracies: at an instrument resolution of $0.5\,\mathrm{cm^{-1}}$, mean biases are found to be $\sim 0.2\,\mathrm{km}$ for clouds with bases below 2 and $-0.2\,\mathrm{km}$ for higher clouds. Accuracy is found to decrease with coarsening resolution and become worse overall for MLEV than for CO_2 slicing/sorting; however, the two methods have differing sensitivity to different sources of error, suggesting an approach that combines them. For expected errors in the atmospheric state as well as both instrument noise and bias of $0.2\,\mathrm{mW/(m^2\,sr\,cm^{-1})}$, at a resolution of $4\,\mathrm{cm^{-1}}$, average retrieval errors are found to be less than $\sim 0.5\,\mathrm{km}$ for cloud bases within $1\,\mathrm{km}$ of the surface, increasing to $\sim 1.5\,\mathrm{km}$ at $4\,\mathrm{km}$. This sensitivity indicates that a portable, surface-based infrared radiance spectrometer could provide an important complement in remote locations to satellite-based measurements, for which retrievals of low-level cloud are challenging.

1 Introduction

Measurements of cloud properties are needed to improve climate and forecast models of the Arctic and Antarctic atmospheres (Hines et al., 2004; Town et al., 2007; Wesslen et al., 2014). Clouds have a strong impact on the polar regions, and recent work indicates that sensitivity to clouds may increase as polar regions warm (Cox et al., 2015b). At the same time, large errors have been found in atmospheric radiative fluxes and cloud radiative forcing in reanalysis products and climate models, which have been partially attributed to errors in cloud-base heights (Walsh et al., 2009); for ERA-Interim, Wesslen et al. (2014) find that cloud-base height is often too high.

Measurements of cloud properties at high latitudes come primarily from satellite platforms (e.g., Wang and Key, 2005; Lubin et al., 2015). Active instruments, such as lidar, can vertically profile clouds (see, e.g., Verlinden et al., 2011; Cesana et al., 2012) but have a small footprint, so that monthly or seasonal averaging is needed for global coverage. Passive instruments that measure upwelling infrared radiances have large footprints, enabling global coverage on daily timescales. These instruments have the advantage that the cloud property retrievals are derived from, and are

thus directly tied to, their radiative effect. However, passive satellite-based instruments are best suited for viewing the tops of clouds and have less sensitivity to the important region of the atmosphere that affects the surface energy budget, that is, between the surface and the base of the cloud. Thus, satellite-based measurements should be complemented by surface-based measurements.

Atmospheric observatories that are capable of surface-based remote sensing of cloud properties exist in the Arctic at a small number of coastal and interior land stations; in addition, a number of field campaigns have been conducted over the Arctic Ocean (see Uttal et al., 2015, and references therein). In the Antarctic, field stations are sparsely located, principally on the coast, and have fewer instruments for measuring cloud properties than in the Arctic. In addition to cloud measurements from existing field stations and past campaigns (Bromwich et al., 2012, and references therein), the Atmospheric Radiation Measurement (ARM) West Antarctic Radiation Experiment (AWARE) is making a broad suite of measurements from November 2015 to 2017. Nevertheless, there remains a dearth of surface-based remote sensors in the Antarctic. The lack of instrumentation at both poles is due largely to the expense and logistical challenge of deploying instruments in these remote regions. A lack of autonomous sensors prevents collection of data at locations other than established stations. New instruments are needed that address these challenges, in particular designs intended for the purposes of both climate monitoring and process studies representing a more comprehensive range of regional high-latitude climates.

Surface-based infrared spectrometers, such as the Atmospheric Emitted Radiance Interferometer (AERI) of the ARM program, are proven instruments that have been used to retrieve cloud temperature or height in the Antarctic (Mahesh et al., 2001) and Arctic (Rathke et al., 2002). There have been a limited number of cloud-height retrievals from surface-based infrared spectrometers because cloud height is more typically measured by co-located active instruments. However, a legacy of cloud-height retrievals from stand-alone passive infrared remote sensors on satellites has demonstrated the usefulness of this approach and led to refined retrieval methodologies (e.g., Smith and Platt, 1978; Minnis et al., 2001; Kahn et al., 2007). Infrared spectrometer technology can be relatively low-cost, with energy requirements that are considerably lower than active instrumentation such as lidar (e.g., Christensen et al., 2004). Thus, portable, autonomous infrared spectrometers are a viable solution for acquiring long-term, high temporal resolution, surface-based measurements of clouds and the atmospheric state from a more spatially diverse and comprehensive sample of the high latitudes, including over sea ice. Evaluating the requirements for accurate cloud-height retrievals is a first step towards development of such a system.

Here we evaluate the potential for using an autonomous infrared spectrometer capable of being deployed in remote regions for retrieving cloud-base height. In particular, noise characteristics depend on instrument resolution (which limits the instrument throughput), and hence noise decreases as resolution becomes coarser. Thus we also test the effects of instrument resolution on the accuracy of cloud-height retrievals. Since such an instrument is currently hypothetical, our analysis makes use of a simulated dataset from Cox et al. (2016). Using simulated data also affords a number of useful advantages for evaluating design aspects of an infrared spectrometer by permitting control over the sources of error and maintaining a fixed and known standard for comparison. This allows uncertainties associated with retrieval methodology and instrument characteristics to be isolated. Two established methods for retrieving cloud height using spectrally resolved infrared instruments, the minimum local emissivity variance (MLEV) technique (Huang et al., 2004) and the CO_2 slicing (e.g., Menzel et al., 1983; Mahesh et al., 2001) and sorting (Holz et al., 2006) technique, are further developed and intercompared here. Although these have been compared for satellite-based retrievals from upwelling radiances (Holz et al., 2006), they have not been compared for retrievals from downwelling radiances or with consideration of variability in noise characteristics and spectral sampling between different types of spectrometers, which are key engineering barriers to developing an autonomous surface-based system. Here we evaluate and compare these techniques to determine relative accuracies for surface-based retrievals of downwelling radiance in the Arctic and to constrain the instrument requirements for providing cloud-height information from an infrared spectrometer that is designed for autonomous deployment.

2 Simulated radiances

To serve as surrogates for measured downwelling radiances for the cloud-height retrievals, simulated cloudy-sky radiances are used. A wide variety of simulations were created, described in detail by Cox et al. (2016), and summarized here. Perfect resolution spectra were created using the Line-by-Line Radiative Transfer Model (LBLRTM) (Clough et al., 2005) and the DIScrete Ordinates Radiative Transfer (DIS-ORT) model (Stamnes et al., 1988), for a spectral range of 50 to $3000\,\mathrm{cm}^{-1}$ (spectra in the range 500 to $950\,\mathrm{cm}^{-1}$ are used in this work). This dataset is designed to be of use for feasibility studies of retrievals such as the present study. The simulated dataset allows tests of retrieval accuracy to focus on a few variables, while constraining others. To this end, the cloud and atmosphere modeling included only single-layer clouds, and a plane-parallel atmosphere was assumed. First, a base set of cases was created with ice modeled as spheres and vertically homogenous clouds. This allowed testing the accuracy of cloud retrievals for idealized cases. Second, a subset of more complex cloud simulations was created (these include spectra developed by Cox et al. (2016) as well as ad-
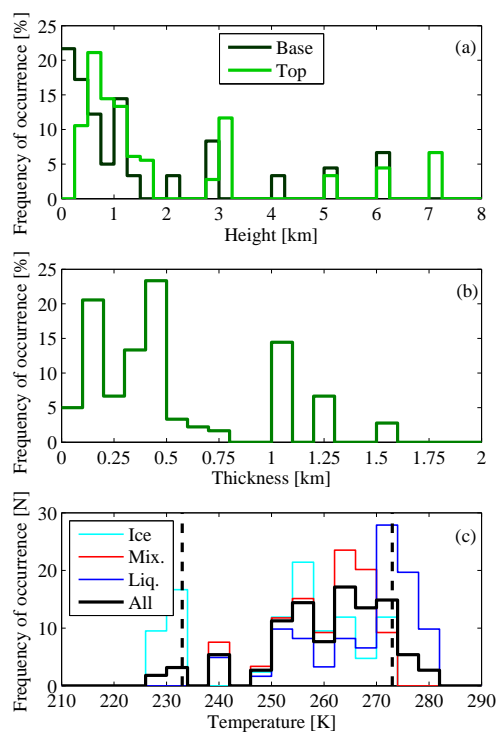
Figure 1. Reproduced from Cox et al. (2016). Distributions of macrophysical properties for 222 simulated clouds. **(a)** Cloud-base height (black) and cloud-top height (green); **(b)** physical thickness; **(c)** cloud mean temperature. The vertical lines in **(c)** represent the physical limits imposed on the cloud phase; liquid is present above the lower limit, while ice is present below the upper limit.

ditional spectra created in a similar manner). These are described below.

2.1 Base dataset

For the base dataset, a variety of typical Arctic atmospheres are represented, including conditions for all four seasons and a variety of cloud types. Because of the high incidence of mixed-phase clouds in polar regions, both single-phase and mixed-phase clouds are included. Temperature-dependent single-scattering parameters are used for liquid (see Rowe et al., 2013, and references therein), while single scattering parameters for ice spheres are from Warren and Brandt (2008). In the base dataset, mixed-phase clouds are modeled as externally mixed in a single layer. Cloud-base heights range from 0 to 7 km with temperatures ranging from 225 to 283 K. Figure 1, reproduced from Cox et al. (2016), shows the distributions of cloud height, thickness, and temperature for the base dataset. Overall, the atmospheric temperature and humidity profiles as well as the cloud optical depths, phases, effective radii, and cloud heights used in the model are intended to be realistic for the Arctic. Temperature inversions are included and cloud heights are typically low with fewer high clouds. Total cloud optical depth referenced to

the visible region (hereafter termed cloud optical depth for brevity) varies from 0 to 12. Cloud optical depth divided by cloud physical thickness (a proxy for visible extinction coefficient) varies from 0 to $0.01\,\mathrm{m}^{-1}$ near the surface and 0 to $0.001\,\mathrm{m}^{-1}$ above 4 km. To test the limits of cloud property retrievals, a few extreme and/or less likely cases were included. For example, the dataset includes a few cases of clouds with optical depths that are extremely low (< 0.2) and includes high clouds that are optically thick as well as thin. Precipitable water vapor (PWV) amounts span the range typical of the polar regions, but some cases are included that are quite high for the polar regions (mean $= 1$ cm, standard deviation $= 0.72$ cm, maximum $= 3$ cm, minimum $= 0.2$ cm). The base set is comprised of 222 clouds. Of these, 157 have bases below 2 km (hereafter referred to as "low clouds") and 65 have bases above 2 km (hereafter referred to as "high clouds"). While all simulations are for single-layer clouds, the cloud layer spans multiple model layers for 69 % of the low clouds (108 out of 157). This base set allows determination of the accuracy of the retrievals for varying atmospheric temperature and humidity profiles, precipitable water vapor amounts, cloud heights, temperatures, optical depths, ice fractions, and effective radii for clouds that are otherwise simplistic. In this work, the base dataset is used for analyses unless otherwise noted.

Unlike clouds in the base dataset, real clouds are vertically and horizontally inhomogeneous, vary temporally, and consist of a variety of ice habits. Simulations that account for these variations were created with other atmospheric and cloud properties held constant. For this purpose, subsets of the base set of clouds were selected.

2.2 Subset: cloud inhomogeneity

Cloud inhomogeneity includes vertical variation through the cloud, horizontal variation over the instrument field of view, and variation with time during the timespan of a measurement. For testing the effects of cloud inhomogeneity, cases were selected in which clouds span multiple model layers (to simulate vertically varying cloud properties), have optical depths greater than 0.5 (retrievals using the base dataset indicate that for optical depths less than about 0.5 the cloud signal is often too low for accurate retrievals), and are mixed phase (ice fractions between 0.2 and 0.8). This subset consists of 23 cases. From this subset, simulations were rerun with various attributes modified. To allow isolation of errors due to various assumptions, each new simulation was modified in only one respect; in all 92 additional spectra were created. Modifications include the following.

To simulate vertically inhomogeneous clouds, the cloud optical-depth profile was set to increase linearly with height from 0 at the cloud base to a maximum at the cloud center, then decrease to 0 at the cloud top. Based on this optical-depth profile, and accounting for the physical thicknesses of the model layers, cloud optical depths were calculated for

each model layer. Thus, for a three-model-layer cloud with layers of equal physical thickness, the top-layer optical depth would be 25 % of the total, the middle layer 50 %, and the bottom layer 25 %. The total optical depth through the cloud is the same as in the corresponding base dataset case, and ice fraction (with respect to optical depth) is kept the same.

For horizontally and temporally varying cloud simulations, the cloud measurement is expected to be a linear combination of spectra of different clouds. Such spectra were created by averaging spectra. First, an additional set of simulations for physically thin clouds was created by placing all cloud optical depth in the middle model layer. Simulations of these physically thin clouds, the set of vertically inhomogeneous clouds described above, and the base dataset were then averaged to simulate time averages of clouds that vary from physically thin and dense to thicker and more diffuse.

Because measurements indicate that Arctic clouds are often composed of an ice layer topped by a liquid layer, liquid-topped clouds were created by placing all cloud liquid in the top model layer and placing all ice in the model layers below. Total optical depths of liquid and ice were kept the same as in the corresponding cases from the base dataset.

2.3 Subset: ice habit

To create a subset of cases suitable for testing the effect of ice habit on retrievals, cases were selected for which ice optical depth was greater than 0.5 and ice fraction was greater than 0.5 %. While 79 such cases exist, 15 representative cases were selected. These include seven low clouds and eight high clouds. For both low and high clouds, winter, summer, and transition (spring/fall) seasons are represented, and optical depth varies from 0.8 to 5. The 15 cases, for five ice habits, represent 75 additional simulations. Spectra were simulated for this subset for the following ice habits: hollow bullet rosettes, smooth plates, rough plates, smooth solid columns, and rough solid columns, using the single scattering parameters of Yang et al. (2005, 2013). Further details about these simulations are provided in Cox et al. (2016).

2.4 Spectral resolution

The perfect resolution spectra were convolved with sinc functions to create sets of simulated cloudy-sky radiances at resolutions of 0.1, 0.5, 1, 2, 4, and 8 cm^{-1}. These simulated cloudy-sky radiances serve as the "observations", R_{obs}, used to test the cloud-height retrievals. Some examples are shown in Fig. 2 at resolutions of 0.5 and 4 cm^{-1}. Absorption lines are clearly evident at the finer spectral resolution but are smoothed out at the coarser resolution. At around 667 cm^{-1} the radiance depends on the surface temperature and clouds have negligible effect. Moving from 667 to 710 cm^{-1}, the effects of temperature inversions are evident: a decreasing radiance indicates temperatures decreasing with height, whereas an increasing radiance indicates tempera-
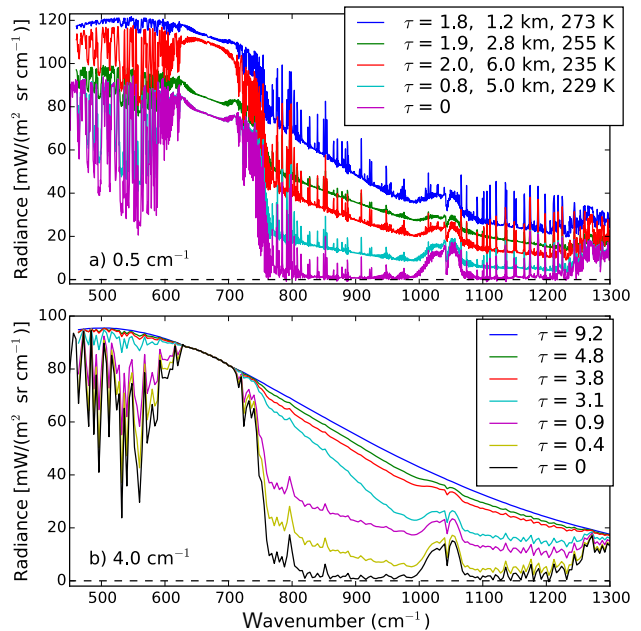
Figure 2. (a) Downwelling radiance spectra at a resolution of 0.5 cm^{-1}, for visible optical depths, cloud-base heights, and temperatures shown in legend. **(b)** Same but at a resolution of 4 cm^{-1} for a cloud-base height of 1.4 km and temperature of 249 K.

tures increasing with height. In panel (a), the uppermost three spectra have similar optical depths (≈ 2) but different cloud-base temperatures; the radiance decreases in the window region (750 to 1300 cm^{-1}) with decreasing cloud temperature. (Note that the lower two spectra have lower optical depths, with an optical depth of 0 indicating clear skies.) In panel (b), atmospheric profiles are identical and the cloud-base height is 1.4 km for all clouds, but the optical depth varies. The radiance decreases with decreasing optical depth in the window region. Note also that the shapes of the spectra differ for optical depths of 3.8 and 3.1; spectral shape also depends on thermodynamic phase, effective radius, and ice habit.

3 Cloud-height retrieval methods

In this section, we first derive the equations central to the CO_2 slicing/sorting and MLEV methods. The derivation is similar to but differs slightly in symbols and development from those of Mahesh et al. (2001), Holz et al. (2006), and Huang et al. (2004) to illustrate some key points. Next, we describe MLEV, followed by the CO_2 slicing method as applied by Mahesh et al. (2001) and the CO_2 slicing/sorting method of Holz et al. (2006). MLEV and CO_2 slicing/sorting typically use radiances from within 710 to 950 cm^{-1} (refer back to Fig. 2) and ignore scattering. Holz et al. (2006) implement the CO_2 slicing differently than Mahesh et al. and also introduce "CO_2 sorting", whereby wavenumbers are selected after sorting them roughly according to their atmo-

spheric transmittance. Finally, we describe modifications to the MLEV and CO_2 slicing/sorting methods made in this work. For MLEV, the method is modified for downwelling radiances, while for CO_2 slicing and sorting the best aspects of the methods of Holz et al. (2006) and Mahesh et al. (2001) are combined, based on experimentation with retrievals from the dataset of simulated Arctic downwelling radiances. The retrievals all assume a zenith view.

3.1 Cloud emissivity

Both MLEV and CO_2 slicing depend on approximations involving the cloud emissivity over the wavenumber range of interest: MLEV assumes it is smoothly varying with wavenumber, while CO_2 slicing traditionally assumes it is constant. Ignoring scattering, the observed downwelling radiance for a zenith view, R_{obs}, is

$$R_{obs} \approx \int_0^{TOA} B(T(z)) \frac{dt}{dz} dz, \tag{1}$$

where the parentheses represent functionality, B is the Planck function, T is temperature, z is height, t is the transmittance from the surface to z, and the integration is from the surface (height of 0) to the top of atmosphere (TOA). The integral can be broken up into contributions from the surface to the cloud base (base), from cloud base to cloud top (top), and from cloud top to the TOA. The radiance contribution from the surface to the cloud base (R_c) is unaffected by the presence of the cloud.

$$R_{obs} \approx R_c + \int_{base}^{top} B(T(z)) \frac{dt}{dz} dz + \int_{top}^{TOA} B(T(z)) \frac{dt}{dz} dz \tag{2}$$

If we assume that the cloud is in an infinitesimally thin layer ($z_{base} = z_{top}$) devoid of gases (i.e., gaseous transmittance within the cloud equals unity), a number of simplifications are possible. We let $B(T(z_{base})) = B(T(z_{top})) = B_c$, the Planck function at the cloud temperature. The first integral on the right-hand side can then be solved to give $B_c[1 - t_{cld}]t_c$, where t_c is the gaseous transmittance from the surface to the cloud base and t_{cld} is the cloud transmittance. We have

$$R_{obs} \approx R_c + B_c t_c [1 - t_{cld}] + t_{cld} \int_{top}^{TOA} B(T(z)) \frac{dt'}{dz} dz, \tag{3}$$

where t' is the cloud-free transmittance from the surface to height z. The final integral is the radiance contribution from above the cloud that makes it through the gaseous atmosphere below the cloud; it is independent of the cloud presence and is equal to $R_{clr} - R_c$. Assuming local thermodynamic equilibrium (and again ignoring scattering), the cloud

absorptivity equals the emissivity, so that $(1 - t_{cld}) = \epsilon$. We can ignore the cloud fraction when the instrument field of view is small, but the emissivity can also be thought of as an effective emissivity that takes into account any patchiness in the cloud within the field of view.

$$R_{obs} \approx R_c + \epsilon B_c t_c + [R_{clr} - R_c] t_{cld} \tag{4}$$

Substituting in $(1 - \epsilon)$ for t_{cld} and simplifying gives

$$R_{obs} \approx R_{clr} + \epsilon [B_c t_c + R_c - R_{clr}]. \tag{5}$$

The equation can be rearranged to solve for the emissivity:

$$\epsilon \approx \frac{R_{obs} - R_{clr}}{B_c t_c + R_c - R_{clr}}. \tag{6}$$

Eq. (6) is comparable to Eq. (4) of Huang et al. (2004) and Eq. (2) of Holz et al. (2006), where $R_{cld} = B_c t_c + R_c$. The equation in the form shown here demonstrates that cloud height and effective emissivity are very closely connected, making cloud-height retrievals challenging. B_c depends on cloud temperature, which in turn depends on cloud height, and t_c and R_c both depend on cloud height. Thus errors in cloud height can be largely accounted for by errors in emissivity.

The right-hand side of Eq. (6) depends on the observed radiance, R_{obs}, and quantities that can be calculated based on knowledge of the cloud-free atmospheric state. In this work, R_{clr} and R_c are calculated from atmospheric profiles of pressure, temperature, and trace gas amounts, using similar radiative transfer calculations as those performed by LBLRTM (Clough et al., 1992). R_{clr} need only be calculated once, whereas t_c and R_c are calculated for each potential cloud height. R_{clr}, t_c, and R_c all include gaseous contributions and therefore vary rapidly with frequency. By contrast, ϵ and B_c should vary slowly with frequency.

To summarize the model assumptions, they include modeling the atmosphere as a plane-parallel, layered atmosphere in local thermodynamic equilibrium, ignoring scattering, assuming the cloud is in an infinitesimally thin layer devoid of gases, and assuming that the emissivity is slowly varying or constant with frequency over ~ 710–$950\,cm^{-1}$.

3.2 MLEV

To find the MLEV, ϵ is calculated for each potential cloud height, c, for wavenumbers between limits ν_1 and ν_2. The "local" emissivity variance (LEV) is then calculated according to Eq. (5) of Huang et al. (2004):

$$LEV_c = \sum_{\nu_1}^{\nu_2} (\epsilon_{c,\nu} - \langle \epsilon_{c,\nu} \rangle)^2, \tag{7}$$

where this is defined as the local emissivity variance because the mean of ϵ is not taken over the entire spectral range ν_1 to

ν_2 but rather over a small wavenumber region ($\Delta \nu$) about ν:

$$\langle \epsilon(c, \nu) \rangle = \frac{1}{\Delta \nu} \sum_{\nu - \Delta \nu/2}^{\nu + \Delta \nu/2} \epsilon_{c, \nu}. \tag{8}$$

Huang et al. (2004) use $\nu_1 = 750$ and $\nu_2 = 950 \, \text{cm}^{-1}$ in Eq. (7) and use an interval of $\Delta \nu = 5 \, \text{cm}^{-1}$ in Eq. (8). When Eq. (7) is calculated with an incorrect height, errors in the calculated values of t_c and R_c result in errors in the calculated effective emissivity that vary rapidly with frequency due to the dependence of t_c and R_c on trace gases, causing the LEV to be large. Thus the correct cloud height is retrieved as that corresponding to the minimum LEV, or MLEV.

In this work, the MLEV method is performed similarly to that of Huang et al. (2004) but for downwelling radiances and for a variety of different spectral resolutions. Furthermore, all values are calculated for the desired instrument resolution. This is done by convolving R_{clr}, t_c, and R_c with a sinc function with the desired linewidth. As in Huang et al. (2004), we use $\nu_1 = 750$ and $\nu_2 = 950 \, \text{cm}^{-1}$. For resolutions of 0.5 and $1 \, \text{cm}^{-1}$, we use an interval of $\Delta \nu = 5 \, \text{cm}^{-1}$ in Eq. (8) (corresponding to averaging over 10 or 5 spectral points, respectively), like Huang et al. (2004). However, for a resolution of $2 \, \text{cm}^{-1}$, we use $\Delta \nu = 10 \, \text{cm}^{-1}$ (average over 5 points), and for resolutions of 4 and $8 \, \text{cm}^{-1}$ we use $\Delta \nu = 24 \, \text{cm}^{-1}$ (average over 6 or 3 points, respectively). Small variations about these values were found to give similar results.

The steps for MLEV are as follows.

1. Choose model heights (layer boundaries) for the model atmosphere. Calculate R_{clr}. Calculate R_c, B_c, and t_c for each model height for the clear-sky atmosphere based on best estimates of temperature, water vapor, and trace gas amounts.

2. Calculate the LEV (Eqs. 6–8) for each trial cloud height.

3. Find the height that corresponds to the MLEV.

3.3 CO$_2$ slicing and sorting

In Mahesh et al. (2001), CO$_2$ slicing makes use of the variation in the absorption coefficient of the CO$_2$ band from ~ 700 to $755 \, \text{cm}^{-1}$, where CO$_2$ emission dominates. (Unlike H$_2$O, CO$_2$ is a well-mixed gas and thus can be estimated fairly accurately from surface measurements.) Rearranging Eq. (5), including the wavenumber dependence explicitly, and dividing both sides by the same quantities at a reference wavenumber, ν_0, gives

$$\frac{R_{\text{obs}}(\nu) - R_{\text{clr}}(\nu)}{R_{\text{obs}}(\nu_0) - R_{\text{clr}}(\nu_0)} = \tag{9}$$
$$\frac{\epsilon(\nu)[B_c(\nu) t_c(\nu) + R_c(\nu) - R_{\text{clr}}(\nu)]}{\epsilon(\nu_0)[B_c(\nu_0) t_c(\nu_0) + R_c(\nu_0) - R_{\text{clr}}(\nu_0)]}.$$

The value of ν is varied from ~ 700 to $755 \, \text{cm}^{-1}$, while ν_0 is chosen to be a wavenumber close to this spectral region

but where the absorption coefficient of CO$_2$ is small enough (i.e., the transmittance is large enough) that the downwelling radiance is sensitive to the entire atmospheric column (Mahesh et al., 2001, chose $\sim 812 \, \text{cm}^{-1}$). It is next assumed that the emissivity is constant from 700 to $812 \, \text{cm}^{-1}$, so that the emissivity terms cancel.

$$\frac{R_{\text{obs}}(\nu) - R_{\text{clr}}(\nu)}{R_{\text{obs}}(\nu_0) - R_{\text{clr}}(\nu_0)} = \tag{10}$$
$$\frac{B_c(\nu) t_c(\nu) + R_c(\nu) - R_{\text{clr}}(\nu)}{B_c(\nu_0) t_c(\nu_0) + R_c(\nu_0) - R_{\text{clr}}(\nu_0)}$$

The left-hand side (LHS) of Eq. (10) is constant, while the right-hand side (RHS) varies with assumed cloud height. Solutions are found at each wavenumber where the RHS equals the LHS, giving a retrieved cloud height for each ν. When the RHS is not equal to the LHS at any height, the solution is found where the magnitude of the difference (RHS − LHS) is smallest. Due to model and measurement errors, retrieved cloud heights vary for different values of ν. (Note that Mahesh et al., 2001, retrieve cloud-base pressure rather than height; in this work, cloud-base height is retrieved.) Mahesh et al. take a weighted average of the results obtained, where the weights are the change in the RHS with a change in the pressure at the cloud base, determined in a 10 hPa interval centered about the retrieved cloud-base pressure. This typically provides more weight to wavenumbers with "e-folding" distances close to the cloud base.

Multiple solutions may exist at a given ν due to errors or due to the presence of near-surface temperature inversions, which are common in the polar regions. Due to a temperature inversion, the cloud temperature may exist at more than one height. Because the retrieval methodology relies to a large extent on sensitivity to cloud temperature (rather than height), choosing between heights having the same temperature can be challenging. To do this, the best result is determined for each set of solutions (e.g., the set below the inversion and the set above the inversion). Then, to choose between sets of cloud bases above and below a temperature inversion, Mahesh et al. (2001) perform a second step using "short-sighted" wavenumbers. Short-sighted wavenumbers are those with low transmittances, and are sensitive to low clouds but not high clouds. Mahesh et al. find the percentage of short-sighted wavenumbers at which a cloud is detected. When this is large, the cloud base is assumed to be within the inversion; when it is small, the cloud base is assumed to be above the inversion.

The CO$_2$ slicing/sorting method of Holz et al. (2006) refines the CO$_2$ slicing technique by selecting a subset of wavenumbers within 650 to $800 \, \text{cm}^{-1}$ to use for the retrieval. Wavenumbers are sorted roughly according to the gaseous transmittance. For downwelling radiances, the transmittance is defined from the surface up to some level in the atmosphere. As the transmittance increases with (sorted) wavenumber, the effective emitting height becomes higher.

Holz et al. (2006) use clear-sky brightness temperature as a proxy for clear-sky transmittance. (As will be discussed, this is a reasonable proxy when temperatures decrease with height in the troposphere, which is not always the case for this work.) First, clear-sky brightness temperatures are sorted. The sorted index is then applied to cloudy-sky brightness temperatures. Sorted clear-sky and cloudy-sky brightness temperatures are then compared to determine at which wavenumbers they differ. Sorted wavenumbers are only used in the retrieval when the clear-sky and cloudy-sky brightness temperatures differ. This sets the lower limit in the gaseous transmittance such that wavenumbers that have little sensitivity to the cloud are excluded from the cloud-height determination. An upper limit in gaseous transmittance is also selected, based on where the slope of the brightness temperature decreases. Finally, Holz et al. found results were improved when only wavenumbers between strong CO_2 absorption lines were used.

Once the subset of wavenumbers to be used has been determined, a unique cloud height is determined for each wavenumber in a similar manner as for CO_2 slicing, but using a different formulation, which is designed for upwelling radiances (see Eq. 1 of Holz et al., 2006).

After a unique cloud height has been found for each wavenumber, the method for determining the best overall cloud height also differs from that of Mahesh et al. (2001). Instead of weighting the cloud heights, an error function is computed for each retrieved cloud height, c.

$$\text{Err} = \sum_{v'} R_{\text{obs}}(v) - R_{\text{clr}}(v) \tag{11}$$
$$- \epsilon_c(v_0) [B_c(v) t_c(v) + R_c(v) - R_{\text{clr}}(v)]$$

The sum is over the selected wavenumbers, v'. The optimal cloud height is chosen as that which minimizes this equation.

This work uses aspects of the CO_2 slicing method of Mahesh et al. (2001) as well as the CO_2 slicing/sorting method of Holz et al. (2006) and additional adaptations for computational efficiency and for cloud retrievals specifically from downwelling radiance measurements made in the polar regions. Based on detailed sensitivity studies and trial and error, the following modifications were made.

CO_2 sorting is applied slightly differently in this work. The use of brightness temperature as a proxy for transmittance, as in Holz et al. (2006), is not a good approximation in the polar regions. While gaseous transmittance, which is defined relative to the surface, always decreases with height, clear-sky brightness temperatures do not always decrease with height in the polar regions; they can increase with height within near-surface temperature inversions, which are common in the polar regions. Thus in our method, wavenumbers are sorted by gaseous transmittance from the surface to the TOA, t_{TOA}. The gaseous transmittances are calculated for each measured radiance spectrum (at the desired resolution) based on the clear-sky atmospheric state and then sorted.

Another difference in our application of sorting involves setting the threshold for choosing the wavenumbers to use. Within the spectral range of 700 to $750 \, \text{cm}^{-1}$, at some wavenumbers CO_2 transmits so little radiance that there is little sensitivity to cloud. At these wavenumbers, $R_{\text{obs}} - R_{\text{clr}}$ is expected to be on the order of the uncertainty. Thus a threshold is needed for which there is adequate cloud signal for the retrieval. A threshold of 0.5 RU is used here (1 RU, or radiance unit, is defined to be $1 \, \text{mW}/(\text{m}^2 \, \text{sr} \, \text{cm}^{-1})$). The gaseous transmittance t_{thresh} determined as the transmittance for which the magnitude of $R_{\text{obs}} - R_{\text{clr}}$ is equal to 0.5 RU, and wavenumbers (v) are selected that correspond to $t_{\text{TOA}} \geq t_{\text{thresh}}$. A final difference is that an upper wavenumber cutoff of $755 \, \text{cm}^{-1}$ is used, rather than estimating a cutoff based on the slope of the brightness temperature; the retrieval is not sensitive to small variations in the choice of upper wavenumber.

Like Mahesh et al. (2001), we use short-sighted wavenumbers to distinguish between multiple solutions. However, whereas Mahesh et al. found that wavenumbers between 670 and $700 \, \text{cm}^{-1}$ are sensitive to clouds within the inversion, this wavenumber range was found to have negligible sensitivity to clouds at any height for the atmospheric profiles used here. Instead, the best wavenumber range for the atmospheric profiles used here is found to be 705 to $715 \, \text{cm}^{-1}$. In addition, sensitivity studies indicate that a method that gives more accurate results than the method employed by Mahesh et al. is to once more find the solution that minimizes the error function given in Eq. (11), this time summing over the short-sighted wavenumbers (step 10 below). These short-sighted wavenumbers are where the transmittance is low and thus generally represent wavenumbers that were excluded in calculating the error function previously. The steps of the CO_2 slicing/sorting method used in this work are summarized as follows.

1. Choose model heights (layer boundaries) for the model atmosphere. Calculate R_{clr}. Calculate R_c, B_c, and t_c for each model height for the clear-sky atmosphere based on best estimates of temperature, water vapor, and trace gas amounts.

2. Calculate the LHS of Eq. (10).

3. Calculate the RHS of Eq (10) for each model height and for each wavenumber.

4. Use CO_2 sorting to choose the best set of wavenumbers for the retrieval; these are typically between 720 and $755 \, \text{cm}^{-1}$.

5. Find the height(s) at which the LHS and RHS agree best (interpolate to find where they cross or, if they never cross, determine where the difference is a minimum) for each wavenumber selected by CO_2 sorting.

6. Calculate the emissivity at the reference wavenumber, $\epsilon_c(v_0)$, using Eq. (6), for each cloud height (c) retrieved.

This yields sets of cloud heights retrieved, comprised of one height for each selected wavenumber within each set, with corresponding reference emissivities. For example, there might be a set of cloud heights retrieved ($c_{\text{lower set}}(\nu)$) corresponding to heights below the inversion and a set ($c_{\text{higher set}}(\nu)$) above the inversion.

7. Calculate the error function as in Eq. (11) for the height retrieved at the first wavenumber, $c = c_{\text{lower set}}(\nu = \nu_1)$, using the corresponding $\epsilon_c(\nu_0)$ and $B_c(\nu)$, $t_c(\nu)$, and $R_c(\nu)$.

8. Repeat step 7 for each of the remaining selected wavenumbers ($\nu = \nu_2$, etc). Find $z_{c,\text{lower set}}$ that corresponds to the minimum error. This yields a single cloud-height retrieval ($c_{\text{ret, lower set}}$).

9. Repeat 7 and 8 for the higher set of retrieved heights, yielding a single cloud-height retrieval ($c_{\text{ret, higher set}}$).

10. To choose between $c_{\text{ret, lower set}}$ and $c_{\text{ret, higher set}}$, calculate the error function again for each of them. However, this time use the short-sighted wavenumbers; these are typically between 705 and 715 cm^{-1}.

4 Results

In this section we demonstrate cloud-height retrieval accuracy for the simulated spectra, including comparison of the results of the MLEV and CO_2 slicing/sorting methods as adapted for this work against the true cloud-base heights, and characterize the effects of ice habit, cloud inhomogeneity, temporal averaging of measured spectra, and sources of error. To understand how different hypothetical instrument specifications and varying amounts of ancillary information affect the results, the comparisons are made with and without imposed errors (e.g., instrument noise and bias and uncertainty in the water vapor and temperature profiles) and as functions of instrument resolution.

4.1 Cloud mask and retrieval capability

An important aspect of a cloud-height retrieval algorithm is that it must be able to determine whether there is a cloud present. Figure 3a shows a scatter plot of cloud height retrieved using CO_2 slicing/sorting vs. true cloud-base height, for the base dataset. For these retrievals, no errors were imposed, so the only error present is model error. The spectral resolution is 0.5 cm^{-1}. The points are color-coded according to the cloud optical depth (in a real experiment, the optical depth will not be known). Cases with high PWV(> 2.9 cm) are indicated in red boxes; these points will be discussed later. The true cloud bases are offset by a small random factor so that the points are spread out slightly for better visibility; the discrete cloud-base heights are evident for bases between 2 and 7 km.
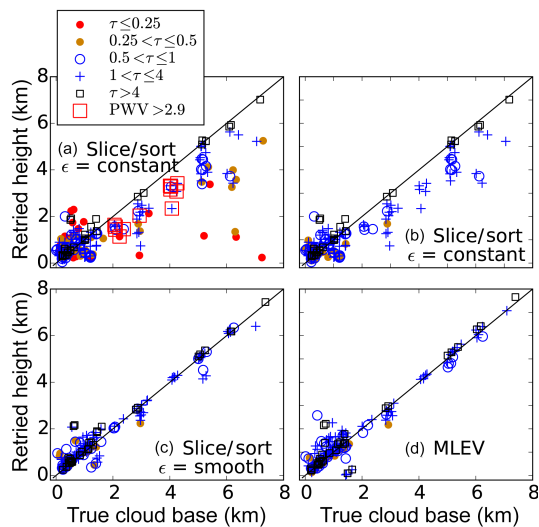
Figure 3. Retrieved cloud height vs. true cloud base for cases with no error and a simulated instrument resolution of 0.5 cm^{-1}. The CO_2 slicing/sorting method ("Slice/Sort" in the figure) and MLEV methods are described in the text. In the figure, a small positive random number (mean of 0.2 km) was added to heights above 2 km to separate points for clarity.

Retrieved cloud-base heights for clouds with very low optical depths (less than 0.5; red and orange points) stand out as having larger retrieval errors. These points constitute clouds that are below the radiance detection threshold and therefore need to be removed from the analysis. The cloud mask is set according to a threshold for a difference between measured and simulated radiance; for the wavenumbers selected using CO_2 sorting, a requirement that the root mean square (RMS) difference between observed and clear-sky radiances differ by at least 2.2 RU is found to remove most low-accuracy points, as shown in Fig. 3b. All cases with cloud optical depths below 0.25 were removed and many of the clouds with optical depths below 0.5 were removed.

Examining Fig. 3b, we see that retrieved cloud base is biased low for clouds above about 2 km (the mean bias is -0.93 km). The bias gets worse, roughly, as clouds get thinner. This bias occurs in large part because the emissivity is assumed to be constant with wavenumber, but actually varies slightly. For thinner clouds, the emissivity is typically larger at wavenumbers in the numerator of Eq. (9) than at the reference wavenumber (chosen here to be ~ 811 cm^{-1}), in the denominator. Thus rather than canceling out, the factor $\epsilon(\nu)/\epsilon(\nu_0)$ is typically about 1.05 for high, thin clouds, causing a bias of about -0.9 km on average. To remove this bias, Eq. (10), used in step 3, is replaced with

$$\frac{R_{\mathrm{obs}}(\nu) - R_{\mathrm{clr}}(\nu)}{R_{\mathrm{obs}}(\nu_0) - R_{\mathrm{clr}}(\nu_0)} = \quad (12)$$

$$\epsilon_{c,\mathrm{rat}}(\nu) \frac{B_c(\nu)t_c(\nu) + R_c(\nu) - R_{\mathrm{clr}}(\nu)}{B_c(\nu_0)t_c(\nu_0) + R_c(\nu_0) - R_{\mathrm{clr}}(\nu_0)},$$

where $\epsilon_{c,\mathrm{rat}}(\nu)$ is determined at each trial cloud height (c) as an estimate of the ratio of the emissivity at ν to the emissivity at ν_0. For each trial height $\epsilon_{c,\mathrm{rat}}(\nu)$ is determined by first calculating the emissivity, $\epsilon_c(\nu)$, according to Eq. (6). While $\epsilon_c(\nu)$ should be smooth, the observed value is highly variable due to errors. Errors are expected to be lowest where the signal is strongest. Thus the next step is to select the wavenumbers where the signal is the strongest; for this a subset of the wavenumbers selected by CO_2 sorting is used. When fewer than 16 wavenumbers are selected by CO_2 sorting, then no emissivity smoothing is attempted; $\epsilon_{c,\mathrm{rat}}(\nu)$ is set to one; that is, Eq. (12) is abandoned in favor of Eq. (10). When at least 16 wavenumbers are selected by CO_2 sorting, then the 16 to 30 points with the highest signal are used. A straight line is fitted to the emissivity at the selected wavenumbers, and its value is divided by $\epsilon_c(\nu_0)$ to get an equation for $\epsilon_{c,\mathrm{rat}}$ for the selected wavenumbers. This equation is used for all wavenumbers within the range of the first and last of the selected wavenumbers. However, outside this range, $\epsilon_{c,\mathrm{rat}}$ is set to 1 because the weakness of the signal prohibits obtaining an estimate of the emissivity that is better than the assumption $\epsilon_c(\nu) = \epsilon_c(\nu_0)$, and examination of the true emissivity indicates that it may not continue to fall on the straight line determined by the fit at the selected wavenumbers.

Using a smooth, rather than constant, emissivity removes much of the low bias observed in Fig. 3b for clouds with bases above 2 km, as shown in Fig. 3c.

Returning to cases with high PWV (red boxes in panel a), note that these occur for cloud-base heights near 2 and 4 km. We see in panels (c) and (d) that these clouds are retrieved quite accurately. It is generally true that these higher-PWV cases can be retrieved accurately even when errors are imposed, except for when large errors exist in PWV itself.

Figure 3d shows the scatter plot for cloud heights retrieved using MLEV. Comparing Fig. 3c and d, we see that both CO_2 slicing/sorting and MLEV are quite accurate for single-layer clouds in the absence of imposed errors.

fig:resultsScatterErrspt5fig:resultsScatterErrs4

In a real experiment, there are errors in the observed radiance (noise or bias) and in knowledge of the atmospheric state, most notably temperature and humidity. To probe the effects of these sources of error, cloud heights are retrieved with errors imposed on the temperature or water vapor profiles used in the retrieval or on the simulation of "observed" radiance (noise or radiation bias). Detailed studies of a variety of errors are summarized in Fig. 4 for a resolution of $0.5\,\mathrm{cm}^{-1}$ and Fig. 5 for a resolution of $4.0\,\mathrm{cm}^{-1}$. The panels of these figures are scatter plots similar to Fig. 3c and d, but both CO_2 slicing/sorting (blue pluses) and MLEV (red
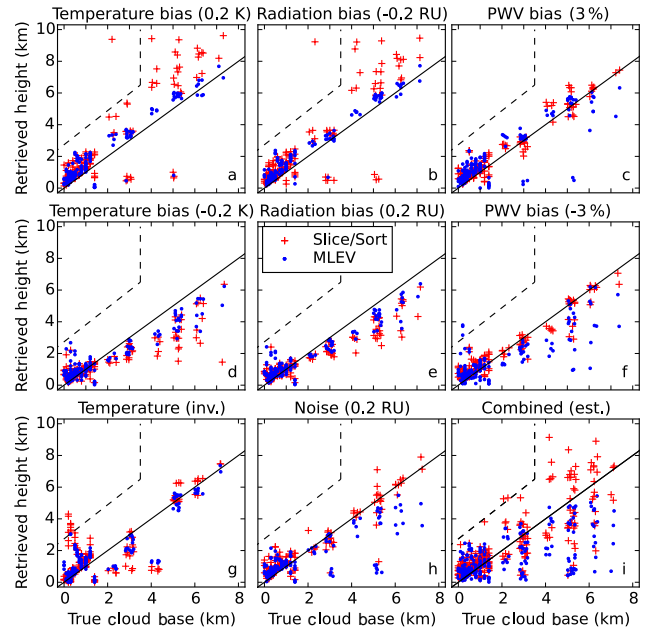
Figure 4. Retrieved vs. true cloud base for cases with the errors shown in the titles for the CO_2 slicing/sorting (Slice/Sort) and MLEV methods (at a resolution of $0.5\,\mathrm{cm}^{-1}$; see text for description of errors). The dashed lines indicate the upper left region where points rarely lie.
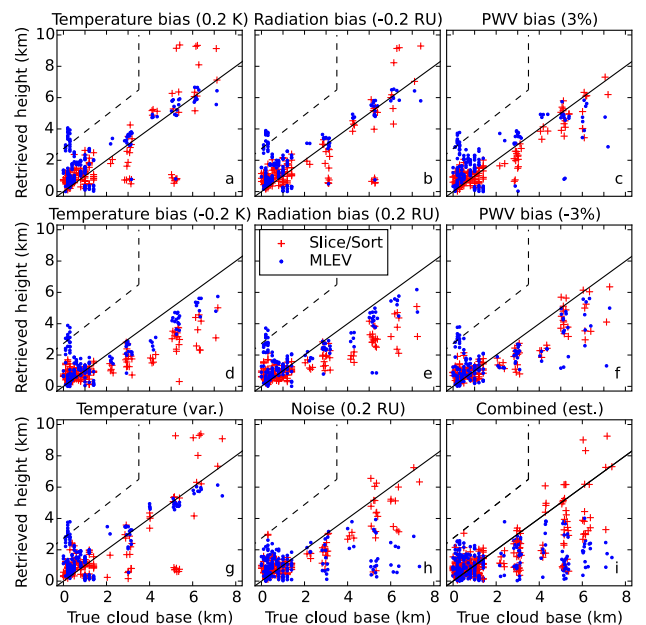
Figure 5. Retrieved vs. true cloud base for cases with the errors shown in the titles for the CO_2 slicing/sorting (Slice/Sort) and MLEV methods (at a resolution of $4\,\mathrm{cm}^{-1}$; see text for description of errors). The dashed lines indicate the upper left region where points rarely lie.

Table 1. Errors in retrieved cloud height for clouds with bases below 2 km using the CO_2 slicing/sorting and MLEV retrieval methods at a resolution of $0.5\,\mathrm{cm}^{-1}$. Errors were determined by imposing a source of error (source) on either the cloudy-sky radiance – noise or radiation bias (bias) – or on the simulated radiances used in the retrieval. For profiles, biases were imposed at all heights, except for variable temperature errors (var.; see text) and errors in the temperature inversion (inv.; see text). The mean error (mean) and the standard deviation (SD) of the errors in retrieved cloud height are given. There were 157 cases, of which some were omitted based on screening (omit). The final two rows show estimates of the combined error for realistic sources of errors, calculated as described in the text.

Source	Value	CO₂ slicing/sorting Mean (km)	SD (km)	MLEV Mean (km)	SD (km)	Omit (no.)
None	–	0.16	0.34	0.14	0.48	30
Noise	0.2	0.13	0.38	0.07	0.40	30
Bias (RU)	0.2	−0.02	0.34	0.03	0.48	29
Bias (RU)	−0.2	0.27	0.37	0.27	0.38	33
Temp. (K)	0.2	0.30	0.35	0.21	0.48	33
Temp. (K)	−0.2	−0.01	0.34	0.02	0.49	29
Temp. (K)	var.	0.28	0.34	0.22	0.45	30
Temp. (K)	inv.	0.39	1.00	0.14	0.62	30
PWV (%)	3	0.16	0.34	0.19	0.49	31
PWV (%)	−3	0.14	0.35	0.05	0.61	29
PWV (%)	10	0.16	0.34	0.21	0.60	34
PWV (%)	−10	0.13	0.36	−0.12	0.67	26
Combined	–	0.23	0.38	0.18	0.49	34
Combined	–	0.01	0.33	−0.05	0.45	27

Table 2. Errors in retrieved cloud height for clouds with bases $\geq 2\,\mathrm{km}$ using the CO_2 slicing/sorting and MLEV retrieval methods at a resolution of $0.5\,\mathrm{cm}^{-1}$. Errors were determined by imposing a source of error (source) on either the cloudy-sky radiance – noise or radiation bias (bias) – or on the simulated radiances used in the retrieval. For profiles, biases were imposed at all heights, except for variable temperature errors (var.) and errors in the temperature inversion (inv.; see text). The mean error (mean) and the standard deviation (SD) of the errors in retrieved cloud height are given. There were 65 cases, of which some were omitted based on screening (omit). The final two rows show estimates of the combined error for realistic sources of errors, calculated as described in the text.

Source	Value	CO₂ slicing/sorting Mean (km)	SD (km)	MLEV Mean (km)	SD (km)	Omit (no.)
None	–	−0.10	0.33	0.01	0.19	32
Noise	0.2	−0.16	0.95	−1.28	1.27	32
Bias (RU)	0.2	−1.10	0.58	−0.69	0.35	29
Bias (RU)	−0.2	0.60	1.63	0.38	0.20	32
Temp. (K)	0.2	0.87	1.80	0.44	0.24	32
Temp. (K)	−0.2	−1.36	1.10	−0.86	0.46	29
Temp. (K)	var.	1.02	1.74	−0.16	0.41	32
Temp. (K)	inv.	−0.58	1.23	−0.55	0.91	32
PWV (%)	3	0.11	0.47	0.01	0.70	33
PWV (%)	−3	−0.30	0.50	−1.02	0.76	27
PWV (%)	10	0.54	0.87	−2.24	1.92	35
PWV (%)	−10	−0.84	0.74	−2.74	1.66	22
Combined	–	0.67	1.73	−1.28	1.60	33
Combined	–	−1.04	0.67	−1.99	1.28	27

dots) are plotted in each panel. Error sources and amounts are given in the titles (the black dashed line is discussed in Sect. 5.2), and mean errors and the standard deviations of errors are given in Tables 1 and 2. Errors are calculated as retrieved cloud height minus true cloud base. Panels a–f show the results of biases in temperature, measured radiance, and PWV. Positive temperature biases cause biases in simulated radiances that are fairly smooth spectrally and thus have a very similar effect as negative radiation biases, and likewise for negative temperature biases and positive radiation biases. However, for water vapor, the effect is complicated by spectrally varying line strengths, and PWV biases affect retrievals differently, particularly positive PWV biases. Random errors in temperature (see Tables 1 and 2) and PWV (not shown) were also tested but were found to have a smaller effect than bias errors due to partial cancellation. The effect of noise in measured radiation is shown in panel (h). Larger PWV biases were also tested (10 %; see Tables 1 and 2, not shown in figure). In addition, errors due to failing to capture the temperature inversion were calculated, as well as the effects of estimated temperature and PWV errors based on errors found in reanalysis data. Failing to capture temperature inversions can have a large effect on low clouds (Fig. 4g). Expected errors in reanalysis data cause retrieval errors of similar magnitude as for biases in temperature (Fig. 5g) and PWV (not

shown). To determine combined errors, two sets of retrievals were performed with multiple errors imposed. First, noise of 0.2 RU (random error in measured radiance with a standard deviation of 0.2 RU), radiation bias of 0.15 K, and negative PWV bias of −3 % were imposed. This was then repeated for the same noise but for negative temperature biases and positive PWV biases. Both sets of points are shown in panel (i) of each figure, so that these panels have twice as many points as other panels. As the figure and tables show, errors (except for failing to capture the temperature inversion) typically affect high clouds much more than low clouds, for which retrievals remain quite accurate. This is not surprising given that high clouds generally have a weaker signal due to the larger atmospheric column below, and the greater sensitivity to the atmospheric column means they are more strongly affected by errors in knowledge of the atmospheric state. These errors are discussed in detail in the discussion section.

4.2 Comparison of CO₂ slicing/sorting and MLEV

In the absence of imposed errors, cloud-height retrievals are slightly more accurate for CO_2 slicing/sorting than for MLEV for low clouds, while MLEV is slightly more accurate for high clouds (Tables 1 and 2). Imposed errors are found to have differing effects on CO_2 slicing/sorting and MLEV (Figs. 4 and 5). Overall MLEV is found to be more accurate

Table 3. Errors in retrieved cloud height for macroscopically varying clouds (see text), using the CO_2 slicing/sorting and MLEV retrieval methods at a resolution of $0.5\,cm^{-1}$. For the upper set of cases (error = n), no errors were imposed on the retrieval, while for the lower set of cases (error = y) noise of $0.1\,mW/(m^2\,sr\,cm^{-1})$ and temperature bias of $0.1\,K$ were imposed. The mean error (mean) and the standard deviation (SD) of the errors in retrieved cloud height are given.

Cloud type	Error	CO_2 slicing/sorting		MLEV	
		Mean (km)	SD (km)	Mean (km)	SD (km)
Dense	n	0.20	0.35	0.11	0.39
Diffuse	n	0.25	0.33	0.35	0.55
Inhomogeneous	n	0.36	0.42	0.43	0.57
Temporally varying	n	0.24	0.27	0.23	0.41
Liquid topped	n	0.25	0.34	0.35	0.56
Dense	y	0.11	0.47	0.12	0.49
Diffuse	y	0.15	0.37	0.21	0.50
Inhomogeneous	y	0.21	0.36	0.31	0.44
Temporally varying	y	0.19	0.45	0.19	0.49
Liquid topped	y	0.23	0.57	0.24	0.49

Table 4. Errors in retrieved cloud height for clouds with a variety of ice habits, using the CO_2 slicing/sorting and MLEV retrieval methods at a resolution of $0.5\,cm^{-1}$. For the upper set of cases (error = n), no errors were imposed on the retrieval, while for the lower set of cases (error = y) noise of $0.1\,mW/(m^2\,sr\,cm^{-1})$ and temperature bias of $0.1\,K$ were imposed. The mean error (mean) and the standard deviation (SD) of the errors in retrieved cloud height are given.

Habit	Error	CO_2 slicing/sorting		MLEV	
		Mean (km)	SD (km)	Mean (km)	SD (km)
Sphere	n	0.013	0.23	−0.14	0.45
Hollow bullet rosette	n	0.013	0.25	−0.14	0.54
Smooth solid column	n	−0.007	0.28	−0.14	0.45
Rough solid column	n	−0.055	0.44	−0.14	0.45
Smooth plate	n	−0.033	0.32	−0.13	0.46
Rough plate	n	−0.072	0.40	−0.12	0.45
Sphere	y	−0.35	0.55	−0.44	0.54
Hollow bullet rosette	y	−0.40	0.59	−0.49	0.59
Smooth solid column	y	−0.32	0.43	−0.45	0.54
Rough solid column	y	−0.38	0.64	−0.41	0.60
Smooth plate	y	−0.43	0.73	−0.43	0.56
Rough plate	y	−0.35	0.44	−0.45	0.57

in the presence of biases in the observed radiance and biases in temperature, while CO_2 slicing/sorting is more accurate in the presence of noise in the observed radiance and biases in water vapor.

Factors that complicate how errors affect retrievals using CO_2 slicing/sorting include errors in the fitting of the emissivity to a smooth function and changes in the strength of the apparent cloud signal, which can affect screening-out due to low signal. For example, positive biases can make the cloud signal look stronger (fewer cases screened out), while negative biases can make it look weaker (more cases screened out). For MLEV, the consequences of errors are not as clear as for CO_2 slicing/sorting; indeed, both positive and negative biases in the water vapor profile (expressed as PWV in the figure) result in negative biases in retrieved cloud height.

4.3 Dependence of cloud-height retrievals on cloud inhomogeneity and ice habit

Sensitivity studies of the effects of cloud vertical, horizontal, and temporal inhomogeneity were performed for the subset of cases described in Sect. 2.2, for $0.5\,cm^{-1}$ in the absence of imposed errors and for imposed noise (0.1 RU) and temperature error (0.1 K). Error statistics are compared in Table 3. "Dense" clouds (physically thin) are found to have the smallest mean bias, "diffuse" clouds (equivalent to cases in the base dataset) have slightly larger mean biases, and "inhomogeneous" clouds (with optically thinner upper and lower boundaries) have the largest mean biases. However, the standard deviations of the errors do not follow this trend. Temporally varying clouds (or equivalently, horizontally varying clouds) are averages of the dense, diffuse, and vertically in-

homogeneous clouds in the first three rows. Error statistics for temporally varying clouds are typically intermediate between those for the clouds that make them up. In the absence of errors, liquid-topped clouds have nearly identical statistics as the base dataset counterparts (i.e., diffuse clouds), while errors are slightly larger for CO_2 slicing/sorting when errors are imposed.

Sensitivity studies were also performed for simulations of various ice habits. Error statistics are compared in Table 4. For CO_2 slicing/sorting in the absence of errors, mean biases vary in sign and errors are slightly larger for non-spherical ice habits. However, statistics for MLEV are nearly identical. Furthermore, in the presence of even small errors, trends in error statistics with ice habit disappear.

4.4 Dependence of cloud-height retrievals on resolution

Figure 6 shows the magnitude of the mean biases (solid lines) and the standard deviations (dashed lines) of errors in the retrieved cloud heights for MLEV (blue) and CO_2 slicing/sorting (red) as functions of instrument resolution. Upper panels are for high clouds and lower panels are for low clouds. The left panels show retrieval errors in the absence of imposed error, and the right panels are for combined errors (cases with imposed errors of 0.2 RU noise, radiation bias of $-0.15\,K$, and water vapor bias of 3 % and cases with imposed errors of 0.2 RU noise, a temperature bias of 0.15 K, and a water vapor bias of -3 %). (A single outlier each for low clouds at 0.5 cm and $1.0\,cm^{-1}$ were omitted.) Note that mean biases are positive for low clouds and negative for high
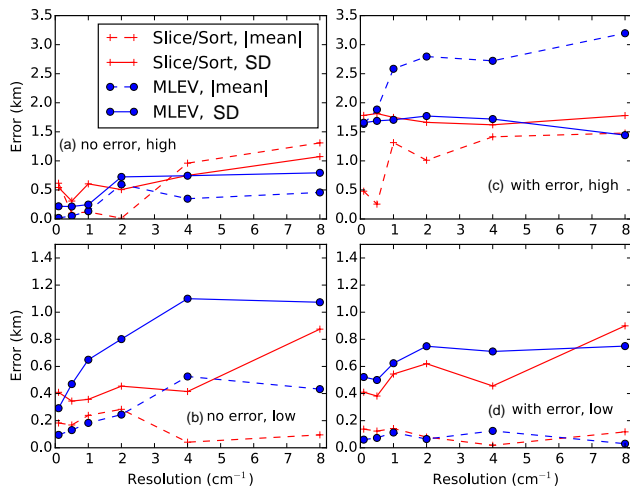
Figure 6. Absolute value of mean error in retrieved cloud heights and standard deviation in retrieved cloud heights as a function of instrument resolution for **(a)** low clouds (cloud bases below 2 km) with no imposed error, **(b)** high clouds (bases of 2 km and above) with no imposed error, **(c)** low clouds with imposed error, and **(d)** high clouds with imposed error. The errors imposed are 0.1 mW/(m^2 sr cm^{-1}) noise in the cloudy-sky radiance and a bias of -0.1 K in the temperature profile used for the retrieval.

clouds (or close to 0); the figure shows the absolute values of the biases.

In the absence of error (left panels), retrieval errors increase gradually overall as resolution becomes coarser, from 0.1 to 8 cm^{-1}. Furthermore, in the absence of error, overall MLEV is more accurate for high clouds, while CO_2 slicing/sorting is more accurate for low clouds. In the presence of imposed errors, behavior with resolution changes. For high clouds, magnitudes of mean biases increase rapidly with resolution for both methods between 0.5 and 1 cm^{-1}, while standard deviations of errors remain constant. For low clouds, by contrast, mean biases remain fairly constant with resolution, while standard deviations of errors increase. Both represent increasing errors with coarsening resolution: for high clouds this is due to increasingly negative biases, while for low clouds this is due to increasingly variable errors. Overall, errors are larger for MLEV than for CO_2 slicing/sorting when errors are imposed.

5 Discussion

5.1 Context with past studies

Mahesh et al. (2001) assume a variation of 3 % in the ratio of emissivities $\epsilon(\nu)/\epsilon(\nu_0)$ (i.e., error due to the assumption that emissivity is constant with wavenumber over the spectral region used). In their analysis, this source of uncertainty leads to uncertainty in retrieved cloud-base pressure of 5 to 13 mb (for a zenith angle of 45°). Converting these to errors

in cloud-base height gives error estimates of 0.03 to 0.11 km for low clouds (bases of 0.1 to 1 km) and 0.14 km for a single high cloud at 2.1 km. In this work, we find that for high, thin clouds the variation in $\epsilon(\nu)/\epsilon(\nu_0)$ is closer to 5 %, resulting in biases of ~ -0.9 km for high clouds (above 2 km, for a zenith view). Furthermore, we estimate that errors in retrieved height due to other sources of model error are approximately 0.2 ± 0.3 km for low clouds and -0.2 ± 0.4 km for high clouds. Thus this work expands on the error analysis of Mahesh et al. (2001) and indicates that retrieval errors for the CO_2 slicing method applied to downwelling radiances are larger than previously predicted. However, errors in actual cases will depend on the specific set of clouds sampled.

Holz et al. (2006) describe retrieval errors for CO_2 slicing/sorting, CO_2 slicing (without sorting; not included here as a separate category), and MLEV. Note, however, that Holz et al. (2006) compare cloud-top height retrieved from upwelling (aircraft-based) infrared radiances (nadir view, 0.5 cm^{-1} instrument resolution) to cloud heights from lidar measurements, whereas we compare cloud heights retrieved from simulated downwelling radiances at the surface to known model cloud-base heights. Thus our results are not suited for detailed comparisons. However, some general observations can be made. The results of Holz et al. (2006) suggest that CO_2 slicing/sorting is more accurate than MLEV for retrievals of optically thin clouds ($\tau < 1.0$) from measurements of upwelling radiance. This study indicates that, for downwelling radiances, the two are roughly equivalent and highly accurate, in the absence of errors, while in the presence of errors accuracy is highly dependent on the source of error. As an example, this work shows that humidity biases cause smaller errors for CO_2 slicing/sorting than for MLEV; thus one explanation for the higher accuracy Holz et al. found for CO_2 slicing/sorting could be errors in the humidity profiles they used. In addition, this work suggests that retrievals from upwelling radiance would benefit from a combined CO_2 slicing/sorting and MLEV method and can suggest implementation strategies based on expected error magnitudes. Finally, Holz et al. (2006) state that retrievals are challenging for clouds below 3 km using upwelling radiances. Since low clouds are retrieved most accurately using downwelling radiances, retrievals from surface-based infrared spectrometers provide an important complement to retrievals based on satellite measurements.

5.2 Dependence of cloud-height retrievals on cloud inhomogeneity and ice habit

The cloud-height retrieval is based on the assumption that the cloud is in an infinitesimally thin atmospheric layer, which is characterized by a temperature and emissivity. Thus for real clouds, which have a finite thickness, variations in temperature and emissivity through the cloud are important and the retrieved cloud height corresponds to an effective emitting height. Sensitivity studies here bear out this expectation, with

physically thicker clouds having larger retrieval errors than physically thinner counterparts. Furthermore, for optically thinner clouds, the effective emitting height will be closer to the cloud middle, while for optically thick clouds, it will be closer to the cloud bottom. In keeping with this, clouds with optically thinner boundaries were found here to have larger retrieval errors compared to true cloud base. However, standard deviations of errors do not follow these trends when errors are imposed; this suggests that for real retrievals, the effects of cloud vertical inhomogeneity will be less important than other sources of error.

Varying the vertical distribution of cloud phase by placing liquid at the cloud top is also expected to move the cloud effective emitting height upward, resulting in larger retrieval errors. However, this is only borne out here for CO_2 slicing/sorting in the presence of imposed errors, for which errors are slightly larger (mean biases are 0.1 km higher and standard deviations of errors are 35 % larger). Converting from a uniformly mixed cloud to a liquid-topped cloud is expected to have a similar effect on retrieval errors as imposing an optical depth that increases moving up through the cloud. Differences in statistics result because retrieved cloud heights are typically higher than for homogeneous mixed-phase clouds. (Note that these cases are all for clouds with bases below 2 km; for higher clouds this positive bias will work to counteract negative biases due to model errors.)

Error statistics for temporally varying clouds were found to be generally intermediate between those for the clouds that make them up, as expected. Thus we can expect that retrieved cloud heights for temporally varying clouds, or for clouds that vary horizontally within the instrument field of view, will be similar to the average cloud height to within expected retrieval error. Furthermore, an instrument such as the one proposed here can be used to measure temporal cloud homogeneity and, using multi-angle measurements, cloud horizontal inhomogeneity (Rathke et al., 2002; Neshyba and Rathke, 2003). Because of their considerably smaller fields of view, knowledge of cloud inhomogeneity from surface-based measurements would be a useful complement to satellite measurements.

Sensitivity studies were also performed for simulations of various ice habits. Differences are likely due to differences in the shape of the emissivity spectra for different habits. Recall that the retrieval does not require any a priori knowledge or assumptions about ice habit but rather relies on the assumption that the emissivity is constant or varies slowly with wavenumber; thus ice habit affects cloud property retrievals only inasmuch as it alters the frequency dependence of the cloud emissivity. Ice habits that result in spectrally flatter (i.e., closer to constant) emissivities should give more accurate results, while ice habits that result in more spectrally varying emissivities are expected to give less accurate results. When a smoothly varying emissivity is fitted, details about the variation of the emissivity, in combination with errors, will determine relative accuracy of the retrievals in a man-

ner that is difficult to predict. Here, error statistics for MLEV are found to be nearly identical for all ice habits. While errors are found to be slightly larger for non-spherical habits for CO_2 slicing/sorting, in the presence of errors, trends in error statistics with ice habit disappear. Thus differences in statistics due to ice habit are likely to be negligible compared to sources of error.

5.3 Effects of errors on cloud-height retrievals

Figure 4 and Tables 1 and 2, presented in the results section, summarize errors for a resolution of $0.5\,\mathrm{cm}^{-1}$ for a variety of sources of error. Table 1 summarizes error statistics for low clouds and Table 2 for high clouds. After screening out cases with cloud signal below 2.2 RU (see the columns indicated by "omit"), errors in retrieved cloud height (retrieved–true cloud-base height) are calculated for each remaining case. The mean error, representing the mean bias in retrieved cloud heights, and the standard deviation in the error are given in the table. The effect of model error, which is present even when no errors are imposed, is shown in the top row. Because model error is present for all retrievals, the value in the table for each source of error is an overestimate. Cases were omitted when the RMS radiance difference for cloudy/clear-sky conditions ($R_{obs} - R_{clr}$) was greater than a chosen threshold (2.2 RU). The threshold was chosen that eliminated all clouds with optical depths less than 0.25 and most with optical depths less than 0.5 in the absence of imposed error (referring back to Fig. 3a and b). In the presence of imposed error, more clouds are screened out for errors that reduce the cloudy-sky radiance or increase the clear-sky radiance, and vice versa. This typically eliminated less than about 20 low-cloud cases. For high clouds, about half (24–36 out of 67) of the clouds were screened out. High clouds emit less because they are colder and typically optically thinner than low clouds; furthermore they have a longer transmission path length through the atmosphere. Thus it seems likely that applying such a threshold to real measurements will also screen out a greater proportion of high clouds (this was true in our dataset despite the fact there is no statistical difference between the optical depths of high and low clouds). Tuning the threshold to a higher value will remove more low-signal cases, particularly high clouds.

For cloud retrievals in the polar regions using an autonomous infrared spectrometer, noise and radiation bias represent instrument characteristics, while errors in the atmospheric profiles will depend on the accuracy of knowledge of the atmospheric state. In remote locations, this will in turn depend on the accuracy of reanalysis data, such as from the European Centre for Medium-range Forecasting (ECMWF) Interim reanalysis dataset (ERA-Interim; Dee et al., 2011). Wesslen et al. (2014) measured temperature and humidity profiles and compared them to ERA-Interim. The measured profiles were not assimilated into the reanalyses, and the location of the measurements was distant from radiosonde as-

simulation sources. Thus, we assume the errors they found in ERA-Interim temperature and humidity profiles are similar to what an autonomous spectrometer would experience in remote locations. Based on their temperature errors, we also performed retrievals for varying temperature errors: imposed errors were 1 K at 10 km, decreasing to -0.5 K at 2 km, and then increasing back to 1 K at 0.2 km. Because the temperature at the surface will be measured and thus known very accurately, the imposed temperature error was reduced to 0 K at the surface. As shown in Tables 1 and 2, the effect of the varying temperature error based on Wesslen et al. (2014) was found to be roughly equivalent to the effect of a positive temperature bias of 0.2 K.

In addition to variable temperature errors, the effects of using temperature profiles in the retrieval that fail to capture temperature inversions were determined. Steep temperature inversions are common in polar regions that can be difficult to capture accurately from satellite. Such temperature inversions are included in the atmospheric profiles used here (see Cox et al., 2016). However, because measurements of surface temperature would accompany a surface-based instrument, extreme cases of error in profiling surface-based temperature inversions would be apparent by comparing the temperature measured near the instrument to the surface temperature in the assumed profile. In addition to surface-based inversions, aloft inversions are common, particularly in the presence of a cloud. To address the effects of poorly profiled temperature inversions, a set of retrievals was performed with temperature inversions removed from temperature profiles used in the retrieval. Because the surface temperature would be known, the true value was replaced in the erroneous profile, and errors were allowed to increase over several layers to provide reasonable temperature differentials across the lower layers (the lowest model layers were set such that temperature differentials would not be more than 1 K for the lowest 1 km and not more than 5 K for the lowest 3 km). Resulting errors, shown previously in (Fig. 4g), affect CO_2 slicing/sorting more than MLEV, particularly for low clouds, and thus MLEV might be preferred for such cases. While surface-based temperature inversions that were not captured by reanalysis data could be identified and screened out from cloud-height retrievals, a better use would be to perform the retrievals to provide an important check on satellite and reanalysis data, correcting for the surface inversion to the extent possible given the known surface temperature and keeping in mind the elevated uncertainties. (In fact the instrument proposed here could also be used to improve temperature profiles, particularly for the lower troposphere, as similar instruments have been in use for such a purposes both from satellite and from the surface. Such improvements would thus improve the reanalysis results and therefore the input temperature to the cloud retrievals.)

For water vapor, Wesslen et al. (2014) find mean errors to be typically positive and below 2 % for the first 3 km, after which they increase to 5 to 10 % from 4 to 8 km. For this

work we assume a relative bias of 3 % throughout the atmosphere. As for temperature, we assume the relative humidity will be measured to high accuracy at the surface and correct the surface error to 0 %. This represents an underestimate of error in the upper atmosphere; however, most of the water vapor is in the lower atmosphere, where this represents an overestimate of error compared to Wesslen et al. (2014). Water vapor biases of ± 3 % at all heights were assumed to be roughly equivalent to the errors found by Wesslen et al. (2014) for ERA-Interim. For comparison, water vapor errors of 10 % were also calculated.

CO_2 errors are expected to be on the order of 0.5 %, based on the work of Alkhaled et al. (2008). Such CO_2 errors were found to produce negligible errors (not shown).

The final two rows of each table give estimates of combined sources of error, estimated as follows. First, we note that the effects of radiation bias and temperature bias are very similar (compare Fig. 4a, b, d, and e), so only one needs to be included; here we include radiation bias. We assume some cancellation of errors between radiation bias and temperature and thus reduce the radiation bias to 0.15 RU, but we assume no cancellation between radiation bias and water vapor bias, pairing positive radiation bias with negative water vapor bias. Thus we simulate the combined error budget by imposing 0.2 RU for noise, 0.15 RU for radiation bias, and -3 % for water vapor. This is expected to be roughly equivalent to what is attained by combining errors in quadrature and is referred to in this work as combined error. In Fig. 4i, the combined errors are shown for both sets of calculations; there are approximately twice as many points on panel (i) as the other panels. For high clouds, errors for CO_2 slicing/sorting are highly variable and both positive and negative, while for MLEV they are strongly negatively biased. This occurs because the sources of error tested do not cause strong positive biases in MLEV (only strong negative biases) regardless of the sign of the error.

As Table 1 shows, for low clouds, mean biases are almost always positive and errors (as mean bias + standard deviation of error) are ≤ 0.65 km for all sources of error tested (except for when temperature inversions are absent from the temperature profiles), illustrating the accuracy with which low clouds can be retrieved. For high clouds (Table 2), the situation is quite different; error magnitudes can be larger than 2 km for a single source of error. It is thus important to distinguish low clouds from high clouds. Referring back to Figs. 4 and 5, note the dashed line in each panel. For all errors shown, only a few points fall in the region to the upper left of this line, and these are mainly in the presence of strong temperature inversions. When a strong temperature inversion is present, only clouds retrieved above 4 km by MLEV can be assumed to be high clouds. Further, in extreme cases of misrepresenting a steep temperature inversion, the same may be true for CO_2 slicing/sorting. However, if the surface temperature suggests the absence of a strong temperature inversion, the analysis here indicates that large positive biases will not generally be

found for retrievals of low clouds (for similar error levels). Instead, positive biases will generally be limited to the value shown by the line. For example, clouds with bases at 1 km will not generally be retrieved above 3.5 km. This means that when a high height is retrieved, the true cloud base is probably high. Furthermore, when MLEV and CO_2 slicing/sorting disagree by more than 2 km, the true cloud-base is probably high. This allows accurate categorization of most clouds into low or high in the absence of strong temperature inversions and for well-characterized temperature inversions. As will be shown in more detail for a resolution of 4 cm^{-1}, errors are strongly height dependent.

5.4 Hybrid methods

The fact that MLEV and CO_2 slicing/sorting show different susceptibilities to different sources of errors suggests that the best method is to use them in combination. This is reasonable as they use overlapping but distinct frequency regions. The exact details of how this is done will depend on the relative magnitudes of errors for a given case, which in turn depends on knowledge of the atmospheric state. However, a few details are worth pointing out here.

Methods for combining MLEV and CO_2 slicing/sorting are worth pursuing but are beyond the scope of this work. They could include combining them at the algorithmic level: for example, in a Bayesian analysis that determines the optimal solution based on the intersection of the mean ±1 standard deviation probabilities for CO_2 slicing/sorting and MLEV. Otherwise, they could be combined post-retrieval by calculating a weighted combination of retrieved cloud heights for the two methods, where the weights depend on the uncertainty levels of the radiance, knowledge of the atmospheric state, and retrieved heights. Regardless, how best to combine CO_2 slicing/sorting and MLEV will depend on resolution and the magnitudes of sources of error and will require estimates of how errors are propagated into errors in retrieved cloud heights. The extra computational time taken for running both CO_2 slicing/sorting and MLEV is minimal because the most time-consuming computations are the calculations of B_c, t_c, and R_c for each model layer, and this set of calculations is identical for the two methods (see step 1 for each method in Sects. 3.2 and 3.3).

In addition to the methods discussed here, an additional candidate for a hybrid cloud-height retrieval is one that relies on multi-angle sky views. Rathke et al. (2002) used a geometric method to retrieved cloud temperature from downwelling infrared radiance spectra measured using the University of Puget Sound infrared spectrometer during the Surface Heat Budget of the Arctic (SHEBA) campaign (see Rathke et al., 2002, and references therein). This method was not compared with the MLEV and CO_2 slicing/sorting methods here for three reasons. First, they found RMS errors in cloud temperatures to be 5.1 K, whereas errors for a spectral method were found to be only 2.9 K (errors were deter-

mined by comparison to radiosonde temperature at the height determined to correspond to the cloud base by co-located lidar). Second, the multi-angle method is only appropriate for horizontally homogeneous clouds. Third, the method retrieved cloud temperature and thus cannot distinguish between heights above and below an inversion. However, such a method could help improve cloud-height determination for homogeneous clouds by incorporation into a hybrid method that makes primary use of the CO_2 slicing/sorting method. For example, a multi-angle method could be used to improve knowledge of the spectral dependence of the emissivity. Homogeneous clouds can be identified using a simple test; cases in which $\ln[1 - R_{obs}/B_c]$ is found to be inversely proportional to the cosine of the zenith viewing angle are identified as homogeneous.

## 5.5 Effect of resolution on retrieval and choice of instrument characteristics

For CO_2 slicing/sorting in the absence of errors, the magnitude of the mean bias in retrievals for high clouds increases with resolution, with the bias changing from ~ 0 km at 2 cm^{-1} to -1 km at 4 cm^{-1} (refer back to dashed red line in Fig. 6b, but note that the figures shows the absolute value of the mean bias). This is primarily due to the assumption that the emissivity is constant. As described previously, the variation of the emissivity with wavenumber results in a mean bias of -0.9 at a resolution of 0.5 cm^{-1}, but this bias is removed to a large extent by fitting the emissivity at high-signal points to a best-fit line and using the best-fit line to calculate a smoothly varying (linear rather than constant) emissivity. At resolutions coarser than ~ 2 cm^{-1}, this correction is hindered by an insufficient number of points, and the bias reappears; note that a bias of -0.9 explains much of the magnitude of the bias shown for CO_2 slicing/sorting in Fig. 6b at 4 and 8 cm^{-1}. A similar increase is evident when errors are imposed (dashed red line in Fig. 6c), but it occurs at a lower resolution (1 cm^{-1}) and is overall larger due to the additional effect of the imposed errors.

For low clouds in the absence of imposed error, the positive mean biases at fine resolution (lower left of panel b) are due to an effective emitting height that is slightly above cloud base. The fact that mean biases are smaller at coarser resolution and when error is imposed is due to fortuitous cancellation of errors. Overall, for low clouds mean biases are fairly small (less than 0.2 km). Thus for low clouds, it is the variation of the error, rather than bias error, that best demonstrates expected retrieval errors. The variation in the error (expressed by the standard deviation) is larger when errors are imposed and grows with coarsening resolution. The opposite is true for high clouds: variations in errors do not change much with coarsening resolution, so it is the increasingly negative bias with imposed errors and with coarsening resolution that needs to be taken into account. However, note that overall there is less dependence on resolution in the pres-

ence of errors, particularly for CO_2 slicing/sorting. Thus in the presence of errors, there is less benefit in using measurements at finer resolution.

On the basis of this resolution dependence, we explore the error budget at $4\,cm^{-1}$. Comparing Figs. 4 and 5, we see the effect of reducing the resolution from 0.5 to $4\,cm^{-1}$ for high clouds is generally to lower the retrieved cloud base, in keeping with the enhanced mean bias shown in Fig. 6. For MLEV, at the coarser resolution, noise has a very large impact, resulting in retrieved heights that are below $\sim 4\,km$ regardless of true base height. Furthermore, for most error sources when MLEV is used, there are a number of cases for which errors are more than $2\,km$ for clouds with bases near the surface. This occurs in part because MLEV has lost the ability to differentiate between clouds with bases above and below an inversion. (It is not clear why these errors are not apparent for combined errors, but note that the differences in local emissivity variances for choices above and below inversions can be extremely small in the presence of errors for coarse resolution, resulting in high sensitivity to minor differences in errors.) Thus this analysis suggests that at a resolution of $4\,cm^{-1}$ and for noise $\geq 0.2\,RU$, MLEV is of limited utility as a stand-alone method, although MLEV could provide information in a hybrid CO_2 slicing/sorting and MLEV method. It is of interest to know whether these biases are born out in retrievals from real clouds; however, we are unaware of any measurements of downwelling radiance currently made at $4\,cm^{-1}$ resolution. Reducing the resolution of instruments currently deployed near active cloud profilers (lidar and ceilometer) and comparing retrieved cloud heights is an interesting topic for future work. For CO_2 slicing/sorting, errors are large but the retrieval can still provide information. For example, most clouds with retrieved heights above $2.5\,km$ can be reliably classified as high clouds. Overall, for CO_2 slicing/sorting (requiring a cloud signal of $2.2\,RU$), mean biases are $-0.08\,km$ for low clouds, with a standard deviation in the error of $0.43\,km$, and mean biases are $-1.3\,km$ for high clouds, with standard deviations in the error of $1.5\,km$.

If the instrument characteristics assumed here (noise of $0.2\,RU$ and bias of $0.2\,RU$) are difficult to achieve, it is also possible to screen out more optically thin cloud cases by increasing the cloud–signal threshold. The threshold of $2.2\,RU$ (corresponding to optical depths below 0.25 to 0.5) removed many of the clouds with bases of $2\,km$ and above; a larger threshold would exclude even more. Thus a stricter threshold would result in greater retrieval accuracy, at the cost of eliminating more thin clouds from the analysis. In other words, retrieval accuracy is directly dependent on cloud signal and will be greater for thicker clouds. To better understand how the magnitude of retrieval errors depends on the threshold used for cloud detection, Fig. 7 shows cloud retrieval error at $4\,cm^{-1}$ for combined errors as functions of both cloud signal and cloud height. The x axis for panels (b) and (c) is cloud signal, computed as the RMS difference between cloudy and
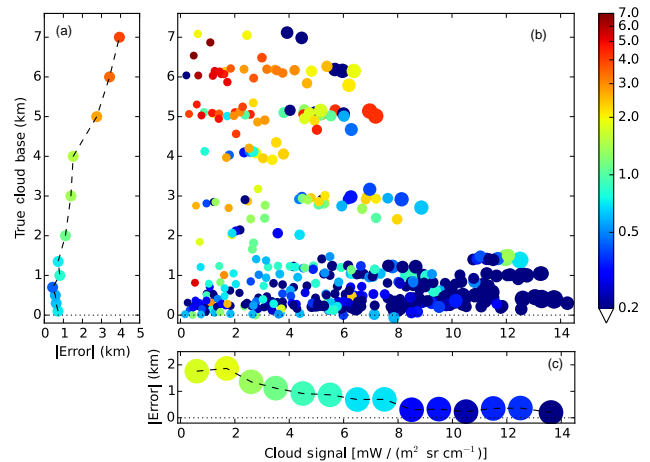
Figure 7. (a) Binned means of the absolute values of errors in retrieved cloud heights (x axis) as a function of cloud-base height (y axis), for estimated combined error budget (see text). **(b)** Absolute value of cloud-height retrieval error (given in color bar in km) as a function of cloud signal (root mean square of cloudy/clear-sky radiance) and true cloud-base height. A small random number is added to cloud-base heights in this panel to make them more easily distinguishable. **(c)** Binned means of the absolute values of errors as a function of cloud signal. Cloud heights were retrieved using CO_2 slicing/sorting from downwelling radiances at a resolution of $4\,cm^{-1}$.

clear-sky radiance at the wavenumbers used in the retrieval, and the y axis for panels (a) and (b) is true cloud base. In all panels, colors indicate the absolute value of cloud error; cloud errors are binned in panels (a) and (c). In panel (b), symbol size depends on cloud optical depth, illustrating the strong correlation between cloud signal and optical depth. Average absolute cloud-height errors decrease from $1.8\,km$ at low cloud signal, to $0.2\,km$ at high signal (panel c). The absolute values of cloud-height errors also generally decrease with decreasing cloud height (panel a). This is partly because higher clouds tend to have lower signal but also occurs independent of cloud signal.

The statistics of cloud-base heights used in this work are similar to those measured in the Arctic (see Cox et al., 2016, and references therein), with 79 % of the clouds below $2\,km$. Thus a surface-based spectrometer would be of the greatest benefit for retrieving the heights of low clouds, which are common in both the Antarctic (Bromwich et al., 2012; Mahesh et al., 2005) and the Arctic (Intrieri et al., 2002).

6   Conclusions

Two established methods for retrieving cloud height from upwelling infrared radiances are modified for retrievals from downwelling infrared radiances: the MLEV and the CO_2 slicing/sorting method. Modifications to CO_2 slicing/sorting make use of the method of Mahesh et al. (2001a) for CO_2

slicing of downwelling radiances. For CO_2 slicing/sorting, a low bias (~ -0.9 km) is found in retrievals of clouds with bases of 2 km or higher; a correction to this bias is presented that assumes a smooth, rather than constant, emissivity at wavenumbers selected by CO_2 sorting (~ 720–811 cm^{-1}). However, it is found that this correction can only be applied when other errors are low.

Working towards the goal of assessing the feasibility of cloud-height retrievals from an infrared spectrometer designed to be used in remote polar locations, and the instrumental considerations (noise and bias in measured radiances) permitting useful retrievals, errors in cloud-height retrievals using the two methods are assessed for simulated radiances. Simulated radiances include single-layer, mixed-phase clouds for a variety of cloud and atmospheric conditions characteristic of the Arctic, at resolutions of 0.1 to 8 cm^{-1}. Retrieval errors are estimated for instrumental sources of error and errors in the atmospheric state that are likely to be experienced in measurements. Retrieval errors are found to vary slightly for different ice habits; because ice habit is not an assumption of the retrieval model, this variation is attributed to differences in the spectral variation of the emissivity spectra. The effects of vertical, horizontal, and temporal variations in the cloud are investigated, including the effects of varying optical depth and phase partitioning (uniformly mixed-phase clouds vs. liquid-topped ice clouds). For clouds with bases below 2 km, retrieved cloud height for physically thinner, denser clouds and clouds with uniformly mixed ice and liquid are found to correspond more closely to true cloud-base height than more diffuse clouds and liquid-topped clouds. This is attributed to lower effective emitting heights. Clouds of intermediate effective emitting heights (diffuse but homogeneous) form the base dataset for the remaining analysis.

In the absence of imposed errors, cloud-height retrievals from simulated spectra using CO_2 slicing/sorting and MLEV are found to have roughly equivalent, high accuracies at resolutions of 0.5 cm^{-1} or finer, with retrieval errors typically < 0.5 km for clouds with visible optical depths greater than 0.3 to 0.5. As resolution becomes coarser, retrieval errors increase. However, in the presence of errors, the dependence on resolution is weakened. Overall, CO_2 slicing/sorting is found to be more accurate than MLEV for low clouds and in the presence of errors, but the two methods are found to have differing sensitivities to different sources of error: CO_2 slicing/sorting is more sensitive to bias in observed radiation and errors in the temperature profile, while MLEV is more sensitive to noise and humidity errors. This complementarity suggests that an approach that combines the two methods is ideal. In particular (for expected error magnitudes), it can be assumed that the cloud base is high when either method retrieves a high cloud. This can be helpful to improve or screen out cases where one method fails completely, e.g., when a near-surface height is retrieved for a cloud base above 2 km for only one method. Thus, a hybrid method combining CO_2

slicing and MLEV could provide greater accuracy. Other possible hybrid methods include geometric retrievals based on multi-angle measurements (Rathke et al., 2002), which would allow characterization of cloud horizontal homogeneity and could improve retrieval accuracy for cases identified as horizontally homogenous.

Retrieval accuracy is found to decrease with decreasing cloud signal, where cloud signal is defined to be the RMS difference, at the selected wavenumbers, between observed and clear-sky radiances. A cloud–signal threshold of 2.2 RU is found to screen out most cases with cloud optical depths below 0.25 and many cases with cloud optical depths below 0.5. Proportionally more high clouds are screened out than low clouds because high clouds typically have lower signal (because they emit less, due to lower cloud temperature, and less cloud emission reaches the surface). However, retrievals for high clouds are also found to be less accurate independent of cloud signal. For real clouds, high clouds are also typically thinner optically than low clouds, thus applying the threshold to real observations is expected to remove even more high-cloud cases, proportionally to low-cloud cases.

At a resolution of 4 cm^{-1}, for expected errors in the atmospheric state and instrument noise level and bias of 0.2 mW/(m^2 sr cm^{-1}), average retrieval accuracies are found to be better than ~ 0.5 km for cloud bases within 1 km of the surface, increasing to ~ 1.5 km at 4 km. The coarser resolution will allow greater instrument throughput and thus greater flexibility in instrumental characteristics such as choice of detector. To further improve the signal-to-noise level, the studies here suggest that temporal averaging of spectra will permit cloud-height retrievals that correspond to time-averaged cloud properties to within similar uncertainty. If these instrument characteristics are not feasible, retrievals may be performed for a smaller subset of thicker clouds by increasing the cloud–signal threshold, which would exclude a large portion of high clouds. Future work should include characterization of errors for cloud-height retrievals from real-world measurements by comparison to other measurements (e.g., from a collocated active instrument). The detailed analysis presented here can help optimize instrument characteristics.

The sensitivity demonstrated here for a surface-based infrared spectrometer to low clouds, which are most common in polar regions, is an important complement to satellite-based measurements, particularly infrared instruments, for which retrievals of low-level cloud are challenging.

7 Data availability

The dataset of perfect resolution downwelling radiance spectra characteristic of the Arctic (Cox et al., 2015a) is available on the Arctic Observing Network (AON) Arctic data repository (at https://arcticdata.io/catalog/#view/doi:10.5065/D61J97TT) and is described by Cox et al. (2016). Com-

puter codes for creating cloudy spectra (runDisort_mat and runDisort_py), which use DISORT (Stamnes et al., 1988), are available at https://github.com/prowe12; see also Rowe et al. (2013).

References

Alkhaled, A. A., Michalak, A. M., Kawa, S. R., Olsen, S. C., and Wang, J.-W.: A global evaluation of the regional spatial variability of column integrated CO_2 distributions, J. Geophys. Res., 113, 303–317, doi:10.1029/2007JD009693, 2008.

Bromwich, D. H., Nicolas, J. P., Hines, K. M., Kay, J. E., Key, E. L., Lazzara, M. A., Lubin, D., McFarquhar, G. M., Gorodetskaya, I. V., Grosvenor, D. P., Lachlan-Cope, T., and van Lipzig, N. P.: Tropospheric clouds in Antarctica, Rev. Geophys., 50, RG1004, doi:10.1029/2011RG000363, 2012.

Cesana, G., Kay, J. E., Chepfer, H., English, J. M., and de Boer, G.: Ubiquitous low-level liquid-containing Arctic clouds: New observations and climate model constraints from CALIPSO-GOCCP, Geophys. Res. Lett., 39, L20804, doi:10.1029/2012GL053385, 2012.

Christensen, P. R., Jakosky, B. M., Mehall, G. L., Kieffer, H. H., Ferry, S., Malin, M. C., McSween Jr., H. Y., Nealson, K., Silverman, S. H., Ferry, S., Caplinger, M., and Ravine, M.: The Thermal Emission Imaging System (THEMIS) for the Mars 2001 Odyssey Mission, Space Sci. Rev., 110, 85–130, doi:10.1023/B:SPAC.0000021008.16305.94, 2004.

Clough, S., Iacono, M. J., and Moncet, J. L.: Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor, J. Geophys. Res. Atmos., 97, 15761–15785, 1992.

Clough, S., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., Boukabara, S., and Brown, P. D.: Atmospheric radiative transfer modeling: a summary of the AER codes, J. Quant. Spectrosc. Ra., 91, 233–244, doi:10.1016/j.jqsrt.2004.05.058, 2005.

Cox, C., Rowe, P., and Walden, V. P.: Simulated Spectra for Autonomous Arctic Infrared Observer, Dataset, doi:10.5065/D61J97TT, 2015a.

Cox, C., Walden, V. P., Rowe, P. M., and Shupe, M.: Humidity trends imply increased sensitivity to clouds in a warming Arctic,

Nature Communications, 6, 10117, doi:10.1038/ncomms10117, 2015b.

Cox, C. J., Rowe, P. M., Neshyba, S. P., and Walden, V. P.: A synthetic data set of high-spectral-resolution infrared spectra for the Arctic atmosphere, Earth Syst. Sci. Data, 8, 199–211, doi:10.5194/essd-8-199-2016, 2016.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Hines, K. M., Bromwich, D. H., Rasch, P., and Iacono, M. J.: Antarctic Clouds and Radiation within the NCAR Climate Models, J. Climate, 17, 1198–1212, 2004.

Holz, R. E., Ackerman, S. A., Antonelli, P., Nagle, F., Knuteson, R. O., McGill, M., Hlavka, D. L., and Hart, W. D.: An Improvement to the High-Spectral-Resolution CO_2-Slicing Cloud-Top Altitude Retrieval, J. Atmos. Oceanic Technol., 23, 653–670, doi:10.1175/JTECH1877.1, 2006.

Huang, H.-L., Smith, W. L., Li, J., Antonelli, P., Wu, X., Knuteson, R. O., Huang, B., and Osborne, B. J.: Minimum local emissivity variance retrieval of cloud altitude and effective spectral emissivity-simulation and initial verification, J. Appl. Meteor., 43, 795–809, 2004.

Intrieri, J., Shupe, M., Uttal, T., and McCarty, B.: An annual cycle of Arctic cloud characteristics observed by radar and lidar at SHEBA, J. Geophys. Res., 107, 8030, doi:10.1029/2000JC000423, 2002.

Kahn, B. H., Eldering, A., Braverman, A. J., Fetzer, E. J., Jiang, J. H., Fishbein, E. F., and Wu, D. L.: Toward the characterization of upper tropospheric clouds using Atmospheric Infrared Sounder and Microwave Limb Sounder observations, J. Geophys. Res., 112, D05202, doi:10.1029/2006JD007336, 2007.

Lubin, D., Kahn, B. H., Lazzara, M. A., Rowe, P. M., and Walden, V. P.: Variability in AIRS-retrieved cloud amount and thermodynamic phase over west versus east Antarctica influenced by the SAM, Geophys. Res. Lett., 42, 1259–1267, doi:10.1002/2014GL062285, 2015.

Mahesh, A., Walden, V. P., and Warren, S. G.: Ground-based infrared remote sensing of cloud properties over the Antarctic Plateau – Part I: Cloud-base heights, J. Appl. Meteor., 40, 1265–1278, 2001.

Mahesh, A., Campbell, J. R., and Spinhirne, J. D.: Multi-year measurements of cloud base heights at South Pole by lidar, Geophys. Res. Lett., 32, L09812, doi:10.1029/2004GL021983, 2005.

Menzel, W. P., Smith, W. L., and Stewart, T. R.: Improved cloud motion wind vector and altitude assignment using VAS, J. Climate Appl. Meteor., 22, 377–384, doi:10.1175/1520-0450(1983)022<0377:ICMWVA>2.0.CO;2, 1983.

Minnis, P., Chakrapani, V., Doelling, D. R., Nguyen, L., Palikonda, R., Spangenberg, D. A., Uttal, T., Arduini, R. F., and Shupe, M.: Cloud coverage and height during FIRE ACE de-

rived from AVHRR data, J. Geophys. Res., 106, 15215–15232, doi:10.1029/2000JD900437, 2001.

Neshyba, S. and Rathke, C.: Statistics of Arctic cloud downwelling infrared emissivity, J. Geophys. Res., 108, 4468, doi:10.1029/2002JD003157, 2003.

Rathke, C., Neshyba, S., Shupe, M., Rowe, P., and Rivers, A.: Radiative and microphysical properties of Arctic stratus clouds from multiangle downwelling infrared radiances, J. Climate, 107, 4703, doi:10.1029/2001JD001545, 2002.

Rowe, P. M., Neshyba, S., and Walden, V. P.: Radiative consequences of low-temperature infrared refractive indices for supercooled water clouds, Atmos. Chem. Phys., 13, 11925–11933, doi:10.5194/acp-13-11925-2013, 2013.

Smith, W. L. and Platt, C.: Comparison of satellite-deduced cloud heights with indications from radiosonde and ground-based laser measurements, J. Appl. Meteor., 17, 1796–1802, doi:10.1175/1520-0450(1978)017<1796:COSDCH>2.0.CO;2, 1978.

Stamnes, K., Tsay, S.-C., Wiscombe, W. J., and Jayaweera, K.: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, Appl. Opt., 27, 2502–2509, 1988.

Town, M. S., Walden, V. P., and Warren, S. G.: Cloud cover over the South Pole from visual observations, satellite retrievals, and surface-based infrared radiation measurements, J. Climate, 20, 544–559, 2007.

Uttal, T., Starkweather, S., Drummond, J. R., et al.: International Arctic Systems for Observing the Atmosphere (IASOA): An International Polar Year Legacy Consortium, B. Am. Meteorol. Soc., doi:10.1175/BAMS-D-14-00145.1, 2015.

Verlinden, K. L., Thompson, D. W. J., and Stephens, G. L.: The Three-Dimensional Distribution of Clouds over the Southern Hemisphere High Latitudes, J. Climate, 24, 5799–5811, doi:10.1175/2011JCLI3922.1, 2011.

Walsh, J. E., Chapman, W. L., and Portis, D. H.: Arctic cloud fraction and radiative fluxes in atmospheric reanalyses, J. Climate, 22, 2316–2334, doi:10.1175/2008JCLI2213.1, 2009.

Wang, X. and Key, J. R.: Arctic surface, cloud, and radiation properties based on the AVHRR Polar Pathfinder dataset – Part I: Spatial and temporal characteristics, J. Climate, 18, 2558–2574, 2005.

Warren, S. G. and Brandt, R. E.: Optical constants of ice from the ultraviolet to the microwave: A revised compilation, J. Geophys. Res., 113, doi:10.1029/2007JD009744, 2008.

Wesslén, C., Tjernström, M., Bromwich, D. H., de Boer, G., Ekman, A. M. L., Bai, L.-S., and Wang, S.-H.: The Arctic summer atmosphere: an evaluation of reanalyses using ASCOS data, Atmos. Chem. Phys., 14, 2605–2624, doi:10.5194/acp-14-2605-2014, 2014.

Yang, P., Wei, H., Huang, H.-L., Baum, B. A., Hu, Y. X., Kattawar, G. W., Mishchenko, M. I., and Fu, Q.: Scattering and absorption property database for nonspherical ice particles in the near- through far-infrared spectral region, Appl. Optics, 44, 5512–5523, doi:10.1364/AO.44.005512, 2005.

Yang, P., Bi, L., Baum, B. A., Liou, K. N., Kattawar, G. W., Mishchenko, M. I., and Cole, B.: Spectrally Consistent Scattering, Absorption, and Polarization Properties of Atmospheric Ice Crystals at Wavelengths from 0.2 to 100 µm, J. Atmos. Sci., 70, 330–347, doi:10.1175/JAS-D-12-039.1, 2013.