Characterizing the near-global cloud vertical structures over land using high-resolution radiosonde measurements

Hui Xu[1], Jianping Guo[1], Bing Tong[2], Jinqiang Zhang[3], Tianmeng Chen[1], Xiaoran Guo[1], Jian Zhang[4], and Wenqing Chen[5]

[1]State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, 100081, China
[2]State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China
[3]Key Laboratory of Middle Atmosphere and Global Environment Observation, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China
[4]Hubei Subsurface Multi-Scale Imaging Key Laboratory, Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan, 430074, China
[5]Business Science and Technology Division, National Meteorological Information Centre, Beijing, 100081, China

Correspondence to: Jianping Guo (jpguocams@gmail.com) and Bing Tong (bingtong@rcees.ac.cn)

Abstract. Cloud remains one of the largest uncertainties in weather and climate research due to the lack of fine-resolution observations of cloud vertical structure (CVS) on large scale. In this research, near-global CVS is characterized by high-vertical-resolution twice daily radiosonde observations from 374 stations over land. It is found that the cloud base heights (CBHs) from the radiosondes have a higher correlation coefficient (R = 0.91) with the millimeter wavelength cloud radar than that with the ERA5 reanalysis (R = 0.49). Overall, cloudy skies occur 65.3 % (69.5 %) of the time, of which 55.4 % (53.8 %) are one-layer clouds at 0000 (1200) UTC. Most multi-layer clouds are two-layer clouds, accounting for 62.2 % (61.1 %) among multi-layer clouds for 0000 (1200) UTC. Geographically, one-layer clouds tend to occur over arid regions, whereas two-layer clouds do not show any clear spatial preference. The cloud bases and tops over arid regions are higher compared with humid regions albeit with smaller cloud thickness (CT). Clouds tend to have lower bases and thinner layer thicknesses as the number of cloud layer increases. The global mean CT, CBH, and cloud top height (CTH) are 4.89 ± 1.36 (5.37 ± 1.58), 3.15 ± 1.15 (3.07 ± 1.06), and 8.04 ± 1.60 (8.44 ± 1.52) km above ground level (AGL) at 0000 (1200) UTC, respectively. The occurrence frequency of clouds is bimodal with lower peaks between 0.5 and 3 km AGL and upper peaks between 6 and 10 km AGL. The CBH, CTH and CT undergo almost the same seasonality that their magnitudes are greater in the boreal summer than in the winter. As expected, the occurrence frequencies of clouds exhibit pronounced diurnal cycles in different seasons. In boreal summer, clouds tend to form as sun rises and the occurrence frequencies increase from morning to later afternoon, with the peak in the early afternoon at altitudes 6–12 km; while in boreal winter, clouds have peak occurrence frequencies in the morning. The relations between surface meteorological variables and moisture with CBH are investigated as well, showing that CBH are generally more significantly correlated with 2m RH (RH_{2m}) and 2m T (T_{2m}) than with surface pressure and 10m wind speed. Larger T_{2m} and smaller RH_{2m} always correspond to higher CBH. In most cases CBHs are negatively correlated

to soil water content. The near-global CVS obtained from high-vertical-resolution radiosonde in this study can provide key
35 data support for improving the accuracy of cloud radiative forcing simulation in climate models.

1 Introduction

Clouds cover nearly two-thirds of Earth's surface area and have a significant impact on the radiative budget of Earth-
atmosphere system (e.g., Ramanathan, 1989; Houghton et al., 1996; Crewell et al., 2004; Stephens et al., 2005; Trenberth et
al., 2009). Many studies have pointed out that the cloud radiative effects (CRE) are critically dependent on the height of clouds
40 (e.g., Wielicki et al. 1995; Weare, 2000; Nam et al., 2012; Stephens et al., 2012; Wild, 2012; Lee et al., 2015; George et al.,
2018). Different cloud types defined by cloud height can even cause two opposite CREs (i.e., the "greenhouse" and the
"umbrella" effects, Meehl and Washington, 1995). Low and thick clouds tend to cool the surface through reflecting solar
radiation, whereas high and thin cirrus clouds tend to warm it by preventing longwave radiation emitting outward (Liou, 1986;
Naud et al., 2003; Solomon et al., 2007). Even small variations in cloud vertical structure (CVS) can lead to significant
45 differences in the mean CRE and radiative heating/cooling rates (Costa-Surós et al., 2014). However, the vertical structures of
clouds are often not accurately represented in current climate models, leading to large uncertainties in estimating the CRE (e.g.,
Randall et al., 2003; Waliser et al., 2009; Cesana and Chepfer, 2012; Cesana and Waliser, 2016; IPCC, 2021). Therefore, it is
imperative to procure high-quality global CVS to improve the predictive capabilities of current climate models.

Satellite observations are efficient in detecting cloud properties. Passive satellite sensors like moderate resolution imaging
50 spectroradiometer (MODIS) can resolve a global coverage of cloud fraction and top height, but not the vertical information of
clouds (Platnick et al., 2003). Chang and Li (2005) retrieved near-global CVS for one-layer and overlapped clouds based on
MODIS data. However, these retrievals existed large discrepancies, especially for high cirrus overlapping lower water clouds.
Active sensors such as the cloud-aerosol lidar with orthogonal polarization (CALIOP) onboard CALIPSO (Winker et al., 2009)
and the millimeter wavelength cloud profiling radar (CPR) onboard CloudSat (Stephens et al., 2002) can provide vertical
55 structure of cloud layers on global scale (Oreopoulos et al., 2017). However, polar orbiting satellites have relatively long revisit
periods (e.g., 16-day) and nadir views (e.g., Winker et al., 2007; Kim et al., 2011; Guo et al., 2016).

Ground-based instruments (Hahn et al., 2001), such as lidars (Gouveia et al., 2017), ceilometers (Costa-Surós et al., 2013),
and cloud radar (Mace et al., 1998), have proven to be effective in providing CVS with continuous temporal coverage and
relatively high accuracy (Zhou et al., 2020). Lidars and ceilometers can pinpoint cloud bases and cloud radar can resolve multi-
60 layer clouds (Willen et al., 2005; Nowak et al., 2008; Reddy et al., 2018). However, the global coverage of these instruments
is too sparse and limited to a few locations.

Radiosonde can provide reliable measurements of the profiles of air temperature (T) and relative humidity (RH) at a great
many locations on a global scale (although underrepresented over oceans), making it possible to obtain global CVS given that
the cloud formation is highly associated with the water vapor and thermal conditions of the atmosphere (e.g., Poore et al., 1995;
65 Wang and Rossow, 1995; Chernykh and Eskridge, 1996; Wang et al., 2000; Minnis et al., 2005; Zhang et al., 2010). In order

to detect CVS from radiosonde observations, two main methods have been proposed. One is the threshold method, in which cloud layers are determined using RH thresholds. Poore et al. (1995) proposed a T-dependent dewpoint depression threshold for cloud detection. Wang and Rossow (1995) detected cloud layers using single RH threshold with the maximum and minimum RH thresholds were 87 % and 84 %, respectively. Zhang et al. (2010) improved the single threshold method by

70 using an altitude-dependent RH thresholds to characterize the base and top of cloud layers. Another method is the gradient method, in which cloud layers are obtained by examining the variations of RH and T profiles. Chernykh and Eskridge (1996) used a second derivative of the vertical profiles of RH and T to determine cloud boundaries. However, by comparing the CVS derived from above methods with ground-based remote sensing measurements, Costa-Surós et al. (2014) indicated that the performances of these mentioned CVS retrieval methods may need further improvement. The possible reasons can be

75 concluded as (1) the resolution of atmospheric profiles provided by radiosonde is low, and (2) refined RH thresholds remain lacking for cloud detection.

With the emergence of growing number of high-resolution radiosonde measurements worldwide, improved retrievals of CVS on large scale are now plausible. However, to our knowledge no studies have reported near-global CVS from high-resolution radiosonde measurements. The main objective of present study is to provide the first attempt to retrieve near-global

80 vertical structures of clouds, including the number of cloud layers, cloud base height (CBH), cloud top height (CTH), and cloud thickness (CT) of each layer, using two years' worth (2018–2019) of high-vertical-resolution (5–10 meters) radiosonde observations from 374 radiosonde stations across the world. Unless otherwise noted, CBH represents the base for the lowermost cloud layer, CTH represents the top for the uppermost cloud layer, and CT represents the total cloud thickness (CTH - CBH). To make this propose, we first develop a CVS detection method, considering both the vertical gradients of RH

85 and T and the altitude-dependent thresholds of RH. The radiosonde derived CVS is compared with that obtained by millimeter wavelength cloud radar (MMCR) observations and ERA5 reanalysis. The remainder of this paper proceeds as follows: Section 2 describes the data and methods to determine the CVS. The performance of the CVS retrieved in this study is evaluated in Section 3. We analyse the frequency of clouds with various layers, their vertical and horizontal distributions, and the diurnal variation of cloud occurrence frequency with height. To examine the potential key factors that affect the CVS we also

90 investigate the relationship between CBH and surface meteorological variables and moisture. The main conclusions are summarized in Section 4.

2. Data and Methods

2.1 Data

2.1.1 High-resolution radiosonde data

Radiosonde instruments provide in situ measurements of atmospheric environmental variables, including RH, T, pressure, and wind profiles, which are recognized to be able to derive the vertical distribution of cloud (Chernykh and Eskridge, 1996). Radiosonde measurements are usually made twice a day at 0000 and 1200 UTC. For high-resolution radiosondes, the sampling period is approximately 1–2 seconds (s), and the vertical resolution is approximately 5–10 meters (m) throughout the atmosphere (Guo et al., 2016; 2019; 2021). In this study, high-vertical-resolution radiosonde measurements obtained from 426 sites across the globe are provided by several organizations, including the China Meteorological Administration (CMA), the National Oceanic and Atmospheric Administration (NOAA) of United States, the German Deutscher Wetterdienst (Climate Data Center), the Centre for Environmental Data Analysis (CEDA) of United Kingdom, the Global Climate Observing System (GCOS) Reference Upper Air Network (GRUAN), and University of Wyoming. In total, there are more than 400 profiles at most radiosonde stations, which provide a roughly even distribution near-global and sufficient samples to characterize the climatology of the CVS. To minimize the uncertainty in results, the radiosonde stations with the profile number less than 400 are not considered, and radiosonde measurements at the remaining 374 stations are used to derive near-global CVS. The remaining 374 radiosonde stations include 120 L-band radiosonde stations from the CMA and 150 stations from NOAA. An additional 104 sites come from the German Deutscher Wetterdienst (Climate Data Center), the CEDA, the GRUAN, and the University of Wyoming. Overall, 407,688 soundings at near-global 374 high-vertical-resolution radiosonde stations for the period of 2018–2019 are used here, including 206,130 soundings at 0000 UTC and 201,558 soundings at 1200 UTC. The geographic distribution of the numbers of profiles at these radiosonde stations is displayed in Figure 1.

2.1.2 Ka-band millimeter-wave cloud radar

The Ka-band (35 GHz) MMCR (Moran et al., 1998) has the advantage of penetrating clouds and continuously detecting the vertical structure of clouds, which can reach an accuracy within 150 m for cloud boundaries (Clothiaux et al., 2000; Hollars et al., 2004). The CVS extracted from MMCR are widely used as reference for inter-comparison with radiosonde derived CVS (e.g., Naud et al., 2005; Zhang et al., 2013). In this study, we use the data measured by a Ka-band MMCR installed at Beijing Nanjiao weather observatory (BNWO, 39.81°N, 116.47°E, 32 m above sea level, ASL) from the Meteorological Observation Center of the CMA during the period of 2019. This MMCR can measure altitudes up to 15 km above ground level (AGL) with a vertical resolution of 30 m and a temporal resolution of 1 min. The CBHs and CTHs are identified based on the minimum threshold method by using the MMCR reflectivity measurements (Zhang et al., 2019).

2.1.3 ERA5 reanalysis

The ERA5 is the fifth generation of global atmospheric, land and oceanic climate reanalysis produced by the European Centre for Medium-Range Weather Forecasts. The ERA5 reanalysis dataset, covering 1979 to present, is produced using 4D-Var data assimilation and is closely associated with the excellence of the forecast products. Compared with former ERA-Interim, ERA5 assimilates more model simulation outputs and historical observations, improving the temporal resolution to 1 hour and the spatial resolution to $0.25° \times 0.25°$ (Hersbach et al., 2020). The product consists of hourly analysis fields on 137 levels, from the surface to a height of 80 km ASL. In this study, we use five meteorological variables such as 10m u-component of wind (u), 10m v-component of wind (v), surface pressure (PS), 2m air temperature (T_{2m}), and 2m dewpoint temperature (TD_{2m})), two moisture variables (soil volumetric water content of the 0–7 cm layer(θ), and mean vertically integrated moisture flux divergence (MFD)), and two cloud variables, including total cloud fraction and CBH, both of which are obtained from the hourly ERA5 reanalysis. In order to obtain the 2m relative humidity (RH_{2m}), the equation proposed by Lawrence (2005) is used, which is shown as follows:

$$RH_{2m} = 100 - 5 \times (T_{2m} - TD_{2m}) \tag{1}$$

2.2 Methods

2.2.1 Pre-processing

Previous studies suggested that pure ice exists in part of the cloud when T drops below –20 °C (Minnis et al., 2005). Since radiosonde measurements only provide the RH profile with respect to liquid ($RH_{liquid}(z)$), the portion of sounding observed $RH_{liquid}(z)$ should be converted to RH profile with respect to ice ($RH_{ice}(z)$) for those observations at altitudes with T below –20 °C (Austin et al., 2009). According to Murray (1967) and Monteith and Unsworth (2008), $RH_{ice}(z)$ is converted from $RH_{liquid}(z)$ when T < –20 °C based on the following equations:

$$e_{liquid}(z) = 6.1078\, exp\left[\frac{17.2693882 T(z)}{T(z) + 237.3}\right] \tag{2}$$

$$e_{ice}(z) = 6.1078\, exp\left[\frac{21.8745584(T(z) - 3)}{T(z) + 265.5}\right] \tag{3}$$

$$RH_{ice}(z) = RH_{liquid}(z) \times \frac{e_{liquid}(z)}{e_{ice}(z)} \tag{4}$$

where z represents altitude in units of km, T(z) represents the profile of air temperature in units of °C, $e_{liquid}(z)$ (hPa) and $e_{ice}(z)$ (hPa) denote saturation vapor pressure in the pure liquid and ice phase with respect to altitude, respectively. Note that the altitude represents the height AGL.

When T is between –20 °C and 0 °C, liquid water and ice coexist in cloud (Austin et al., 2009). To obtain the RH profile with respect to liquid-ice mixed ($RH_{mixed}(z)$), Austin et al. (2009) assumed that the solutions of ice phase and liquid phase are scaled linearly with $T(z)$ and obtained $RH_{mixed}(z)$ from $RH_{ice}(z)$ at –20°C to $RH_{liquid}(z)$ at 0 °C, shown as follows:

$$RH_{mixed}(z) = \frac{-20 - T(z)}{-20} RH_{liquid}(z) + \frac{T(z)}{-20} RH_{ice}(z) \qquad (5)$$

In our retrieval method, the final RH profile $RH(z)$ used to derive the CVS can be described as $RH_{liquid}(z)$, $RH_{ice}(z)$, and $RH_{mixed}(z)$, depending on the phase state, which are calculated using the following formula:

$$RH(z) = \begin{cases} RH_{liquid}(z), & T \geqslant 0°C \\ RH_{mixed}(z), & -20°C < T < 0°C \\ RH_{ice}(z), & T \leqslant -20°C \end{cases} \qquad (6)$$

2.2.2 Determination of CVS

The presence of clouds greatly affects the $RH(z)$ and $T(z)$, which changes sharply when a sounding balloon enters or leaves cloud layers (Pietrowicz and Schiermeir, 1978; Matveev, 1981; Lawson and Cooper, 1990). Thus, the changes in the vertical gradients of $RH(z)$ and $T(z)$ can be used to identify cloud boundaries. Chernykh and Eskridge (1996) presented a method to determine cloud boundaries by using the second derivatives of $RH(z)$ and $T(z)$ ($RH''(z)$ and $T''(z)$), respectively. However, only using $RH''(z)$ and $T''(z)$ would cause us to misidentify cloud layers, tending to detect more cloud layers than observations, especially for very thin cloud layers (Zhang et al., 2012; Costa-Surós et al., 2014). To reduce this possibility, RH threshold should be combined to determine cloud layers. Zhang et al (2010) proposed altitude-dependent thresholds of RH to detect cloud layers. However, their method tends to identify less cloud layers (Costa-Surós et al., 2014). This is due to the application of slightly high RH thresholds, resulting in a lower detection rate of clouds. In order to improve the accuracy of cloud detection, we develop a CVS retrieval method by combing the vertical gradients of $RH(z)$ and $T(z)$ and altitude-dependent thresholds of RH as well. The detailed description of this method proceeds as follows:

After preprocessing the $RH(z)$, the first step in our method is to detect the bases and tops of moist layers. Generally, $RH(z)$ increases when a radiosonde balloon enters a moist, which suggests that the first derivative of $RH(z)$ is greater than zero ($RH'(z) > 0$). At the base of a moist layer, there is a jump in $RH(z)$ (Wang and Rossow, 1995), thus considering that $RH(z)$ reaches a local maximum increase ($RH''(z) < 0$). At the same time, $T(z)$ stops decreasing (i.e., $T'(z) < 0$) near the base of a moist layer, due to the condensation of water vapor and its accompanying release of latent heat. Hence, T reaches a local minimum decrease, which means $T''(z) > 0$. Therefore, by examining the first and second derivatives of $RH(z)$ and $T(z)$ starting from the surface upward, the bases of moist layers can be detected when the following criteria are satisfied:

$$\begin{cases} RH'(z) > 0 \text{ and } RH''(z) < 0 \\ T'(z) < 0 \quad \text{and} \quad T''(z) > 0 \end{cases} \qquad (7)$$

175

where $RH'(z)$ and $RH''(z)$ denote the first and second derivatives of $RH(z)$, whereas $T'(z)$ and $T''(z)$ denote the first and second derivatives of $T(z)$.

Similarly, the tops of moist layers are detected as follows:

$$\begin{cases} RH'(z) < 0 \ \ \text{and} \ \ RH''(z) < 0 \\ T'(z) > 0 \ \ \ \text{and} \ \ \ T''(z) > 0 \end{cases} \tag{8}$$

180   To obtain $RH''(z)$ and $T''(z)$, Chernykh and Eskridge (1996) approximated radiosonde observed $RH(z)$ and $T(z)$ by cubic splines (Bartels et al., 1987), which leaded to bias in cloud detection. To avoid errors due to this approximation, we use real radiosonde observed $RH(z)$ and $T(z)$ to calculate $RH''(z)$ and $T''(z)$ as follows:

$$RH'(z) = \frac{dRH(z)}{dz} = \frac{RH(z_{i+1}) - RH(z_i)}{z_{i+1} - z_i}$$
$$RH''(z) = \frac{d}{dz}\left(\frac{dRH}{dz}\right) = \frac{RH'(z_{i+1}) - RH'(z_i)}{z_{i+1} - z_i}$$
$$T'(z) = \frac{dT}{dz} = \frac{T(z_{i+1}) - T(z_i)}{z_{i+1} - z_i} \tag{9}$$
$$RH'(z) = \frac{dT}{dz}\left(\frac{dT}{dz}\right) = \frac{T'(z_{i+1}) - T'(z_i)}{z_{i+1} - z_i}$$

Then, we identify cloud layers from moist layers determined above using height-resolving RH thresholds defined in Table
185   1, which are lower than the thresholds proposed by Zhang et al. (2010) to reduce the restriction for a moist layer being identified as a cloud layer. The moist layer is identified as a cloud layer if the following three conditions are met: (1) the base of moist layer is greater than 280 m (Zhang et al., 2010); (2) the thickness of moist layer is larger than 30.5 m and 61 m for the base of moist layer less than 2 km and larger than 2 km, respectively (Poore et al., 1995; Zhang et al., 2010); (3) the minimum RH (min-RH) within the moist layer is greater than the corresponding min-RH threshold at the base of moist layer (Table 1); (4)
190   the maximum RH (max-RH) within the moist layer is greater than the corresponding max-RH threshold at the base of moist layer (Table 1). Otherwise, the layer of moist is discarded.

2.2.3 Post processing

To obtain the more robust cloud vertical structures, the cloud layers determined above have to be further reprocessed as follows:

(1) If the distance between two contiguous cloud layers is less than 300 m, these two cloud layers are merged (Zhang et
195   al., 2010);

(2) If the min-RH between the continues cloud layers is greater than the corresponding minimum RH threshold (inter-RH) within this distance (Table 1), these two cloud layers are merged (Zhang et al., 2010).

The whole process, including the pre-processing of the RH profile, the determination of the cloud layers in vertical direction and the post-processing of detected cloud layers are schematically summarized in Figure 2.

3. Results and discussion

3.1 Comparison of CVS between high-resolution radiosonde with MMCR and ERA5

Before presenting the characteristics of near-global vertical structure of clouds, we first compare the CVS determined from radiosonde measurements by using our proposed retrieval method with those from the MMCR in Beijing. Figure 3 presents the comparisons of the cloud base and top heights of each layer at four selected cases. As seen from Figure 3a, the one-layer cloud derived from radiosonde (left) agrees well with that obtained by MMCR (right), with the CBH (CTH) being 0.28 (9.57) km and 0.48 (9.66) km for radiosonde and MMCR, respectively. For high cloud case, the radiosonde derived CBH and CTH are also consistent with those from MMCR (Figure 3b). The radiosonde classifications of two-layer clouds are also in good agreement with those from MMCR (Figure 3c-d). Above results indicate that the radiosonde has the potential to obtain accurate CVS for both low cloud and high cloud, assuming that there is no obvious change in CVS during short time (~15 min). Note that there are 90 samples used in the comparison analysis, since only 191 days MMCR measurements are available in 2019.

For the statistical analysis during the whole year of 2019 in Beijing (Figure 4), the cloud base and top heights obtained by sounding are generally consistent with those from MMCR, with correlation coefficients (R) being up to 0.91 and 0.81, respectively. For low-level clouds (CBH < 2 km), the radiosonde derived CBHs are obviously higher than those from MMCR. A possible reason is that MMCR detects cloud particles using millimeter-wavelength and is readily affected by the presence of large precipitation particles, as well as insects and bits of vegetation, which commonly suspended in the atmospheric boundary layer, resulting in lower cloud base (Zhang et al., 2013). For CTHs, the values retrieved from radiosonde are systematically higher than those from MMCR by about 0.86 km (Figure 4b). There are three possible reasons for this difference. One is related to the variations in clouds from radiosonde observations caused by the horizontal drift of radiosonde balloons (Clothiaux et al., 2000; Comstock et al., 2002; Zhang et al.2013), which may explain the large discrepancies. Another reason is that MMCR is not sensitive to small cloud particles far from the radar, thereby tends to underestimate CTH (Zhang et al., 2019). The last one is that radiosonde tends to detect higher cloud top heights than those retrieved from MMCR due to delay (time lag) after the radiosonde balloon passes through a cloud layer because of the wetness of the sensors (Zhang et al., 2013). Overall, the results of Figures 3–4 indicate that our proposed CVS retrieval method is reliable and can detect CVS correctly.

Next, we compare the radiosonde derived CVS over 2018–2019 with those from ERA5 reanalysis on a global scale, as shown in Figure 5. In general, the CBHs from ERA5 tend to be lower than those from radiosonde, and their R values are 0.24 and 0.49 at 0000 UTC (Figure 5a) and 1200 UTC (Figure 5b) for the whole data, respectively. This result is similar with that reported by Li et al (2022), which demonstrated that the differences in CBHs between ERA5 and the surface observation are much higher at 0000 UTC. The reason for which the correlation coefficient at 1200 UTC is more than twice as large as at 0000 UTC may be that the ERA5 reanalysis product assimilates satellite observations including CALIPSO and CloudSat at 1330

LST (Stephens et al., 2002; Platnick et al., 2003; Winker et al., 2009), which is closer to 1200 UTC for most regions. At a more detailed level, the CBHs from ERA5 tend to be much lower than those from radiosonde for clouds with CBHs from radiosonde > 6 km. Previous studies reported that the underestimation for high-level clouds in the ERA5 results were caused by the poor specification or parameterization of critical RH from model (Miao et al., 2019). To clearly examine the case where most data exist in Figure 5a–b, the comparison for CBHs from radiosonde < 2 km is further given at Figure 5c–d. Compared to the results of the all cloud layers, the fitting slopes of the low cloud layers are slightly smaller and the R values are smaller by about 25%. This indicates that the CBHs from ERA5 reanalysis are roughly underestimated compared with radiosonde observations at relatively low altitude. Miao et al. (2019) also found that low clouds from ERA-Interim were lower than those from CALIPSO/CloudSat observations. Note that there exist several cases with very low CBH from ERA5 (e.g., CBH < 0.1 km) which is not reasonable.

3.2 Frequency statistic of clouds with various layers

The annual and seasonal frequency distributions of clouds with the number of layers at 1200 UTC over 2018–2019 are presented in Figure 6. On the annual time scale, a total of 201,558 profiles from radiosonde are retrieved. Of all detected profiles, 30.5 % cases are clear skies, indicating the cloudy skies occur 69.5 % of the time (Figure 6a), which is in good agreement with the global mean cloud fraction (about 68.0 % ± 3.0 %) reported by Stubenrauch et al. (2013), who analysed the global cloud fraction using multiple satellite observations (e.g., MODIS, MISR, POLDER and CALIPSO) . Over cloudy skies, the cloud frequency decreases as the number of layers increases (Figure 6a), with the one-, two-, three-, four-, five-, six-, and seven-layer clouds accounting for 53.8 %, 28.0 %, 11.4 %, 4.2 %, 1.4 %, 0.6 %, and 0.4 %, respectively. Wang et al. (2000) and Subrahmanyam and Kumar (2017) found similar decreasing trends by using global radiosonde measurements (vertical spatial resolution of ~100 m) and global satellite observations (CALIOP) and they demonstrated that 58 % of clouds were one-layer clouds. However, the occurrence frequencies of multi-layer clouds in our study are obviously higher than those provided by Subrahmanyam and Kumar (2017), which reported that two-, three-, four-, and five-layer clouds occurred at frequencies of 20.00 %, 3.50 %, 0.40 %, and 0.04 %, respectively. The possible cause for this difference could be that satellite has insufficient detection capabilities for multi-layer clouds, leading to an underestimation of their occurrence frequencies (Chang and Li, 2005).

Among the multi-layer clouds, 61.1 % are two-layers clouds at 1200 UTC. This result is slightly lower than that reported by Wang et al. (2000), who suggested that about 67 % of the multi-layer clouds were two-layer clouds. The possible reasons are the discrepancies in the study periods, the spatial distributions of radiosonde stations, the vertical resolution of the radiosonde measurements, and the CVS derivation methods. At 0000 UTC, cloudy skies occur 65.3 %. Of all cloud figuration, the occurrence frequency of one-layer clouds accounts for 55.4 %. Among multi-layer clouds detected, 62.2% are two-layer clouds (Figure S2a).

The seasonal frequency distributions of clouds with the number of layers (Figure 6b–e) are similar with those at annual time scale, namely, frequency decreases as the number of layers increases. By comparison, more clouds occur in December–

January–February (DJF; 72 %) and March–April–May (MAM; 70.8 %) than do in June–July–August (JJA; 69.3 %) and September–October–November (SON; 66.3 %). These results are consistent with those from Xi et al. (2010) based on a 10–year climatology of cloud fraction from surface observations. The seasonal variations are caused by the fact that cloud fractions vary monthly, with a maximum of 83 % in February and a minimum of 65 % in September (Figure S1b). For one-layer clouds, the occurrence frequency is highest in MAM and DJF and lowest in JJA. The possible reason is that the radiosonde sites in this study are mainly located in northern hemisphere (NH) (Figure 1) and more stratus clouds occur in boreal winter (DJF) and boreal spring (MAM) than boreal summer (JJA) (Dong et al., 2005). In contrast, the occurrence frequency of multi-layer clouds is highest in boreal summer (JJA), owing to the deeper troposphere and more convective storms in summer (Zhang et al., 2010). The frequencies of clouds with various layers at 0000 UTC on seasonal scale are also similar with those at 1200 UTC to some degree (Figure S2b–e).

3.3 Near-global vertical distribution of CVS

The annual mean CVS for one-, two-, three-, four- and five-layer clouds at 1200 UTC are shown in Figure 7a. In general, the CBH (CTH) decreases slightly (increases significantly) as the number of cloud layers increases. The CBHs of one-, two-, three-, four-, and five-layer cloud are 3.58 ± 3.31, 2.61 ± 2.76, 2.07 ± 2.44, 1.92 ± 2.49, and 1.97 ± 2.71 km, and the corresponding CTHs are 6.52 ± 3.94, 9.42 ± 3.04, 10.71 ± 2.94, 11.81 ± 3.04, and 13.17 ± 3.06 km, respectively. This result is consistent with Zhang et al. (2010), who reported that one-layer clouds are roughly located at altitudes that fall somewhere among the altitudes of the multi-layer cloud configurations. It is obvious that the lowermost cloud layer in multi-layer clouds occurs below 3 km with slight variations, while the upper layer exhibits more significant variations. Similarly, Wang et al. (2000) pointed out that the lowermost layers of two- and three-layer clouds occur mostly below 3 km, and the uppermost layer of two- and three-layer clouds occurs over a wide range of heights centered in 6 to 7 km, and 7 to 8 km, respectively. Interestingly, the thickness of one-layer clouds (2.94 ± 3.00 km) is greater than that of multi-layer clouds of any number of layers (0.67 ± 0.88 ~ 2.44 ± 2.07 km), and the thickness of each cloud layer for multi-layer clouds decreases as the number of cloud layers increases, which is consistent with previous studies (e.g., Luo et al., 2009; Chi et al., 2022). For multi-layer clouds, the thickness of the lowermost layer is larger than that of the upper layer. These CVS results are mainly attributed to the combined effects of solar heating and exchange of longwave radiation between surface and cloud layers (e.g., Rogers and Koracin, 1992; Guan et al., 1997; Zhang et al., 2010). Note that for the three-, four-, and five -layer cloud configurations the maximum distance between two continues cloud layers is around 2.5 km, existing between the first and second lowest layer. For the upper layers, the distance between two continues cloud layers decreases to be about 1.0 km.

To further quantify the CVS for one- and multi-layer clouds, their boxplots are shown in Figure 7b. Overall, for multi-layer clouds, the mean CBH is lower than that for the one-layer clouds by 1.20 km, while the mean CTH (CT) is much greater than that for one-layer by 3.47 km (4.67 km). The possible reasons are that multi-layer clouds often occur at humid climate regions (e.g., southeast of China and east of USA as shown in Figure 12c–f) and the air parcel tends to reach the lift condensation level (LCL) at relatively low altitude, resulting in lower CBH compared to one-layer clouds. Most multi-layer

clouds are generated from one-layer clouds by extending the range of CBH and CTH of one-layer clouds. Thus, we expect that the CTH and CT of multi-layer clouds are larger than that of one-layer clouds. These results are consistent with those provided in Figure 7a. Note that the different layer cloud configurations occur at different frequencies as shown in Figure 5a.

The CTH and CBH are two of the most important CVS parameters that play a significant role on estimating CRE at the top of atmosphere and the surface, respectively (Wang and Rossow, 1995; Loeb et al., 2012; Xu et al., 2021c). In this study, we present the vertical distributions of annual and seasonal occurrence frequencies of CBH, CTH, and cloud at 1200 UTC in Figure 8. As references for altitude, the annual and seasonal mean planetary boundary layer height (PBLH) and tropopause height retrieved from radiosonde observations are also given in Figure 8. Generally, these results of the vertical occurrence frequencies of clouds at annual scale are close to those at seasonal scale. It is obvious that the relatively large occurrence frequencies of CBH occur within 1 km, with the highest frequency being at about 0.5 km (Figure 8a, d), which is lower than the annual mean PBLH (~0.76 km). This indicates that most of cloud bases exist in the atmospheric boundary layer as previous studies (Zhang et al., 2014a). Above the top of the boundary layer, the occurrence frequencies of CBH decrease with altitude, since most clouds are suppressed by the inverse layer at the top of boundary layer (Sugimoto et al., 2000). Over 6 km, the occurrence frequency of CBH in boreal summer (JJA) is higher than that in other seasons (Figure 8d). The reason may partly be that more cirrus clouds occur in summer (Wang et al., 2000). Below 1.5 km, the occurrence frequency of CBH is the highest during in boreal winter (DJF). This phenomenon can be explained by the combined effects of low temperature and upward-motion conditions in winter, which promotes condensation or collision/coalescence of water vapor to form cloud droplets at relatively low altitudes during the ascending motion, resulting in low cloud bases (Dong et al., 2005; Chi et al., 2022). These results that the occurrence frequency of CBH generally peak in boreal summer (JJA) and boreal spring (MAM) and reach its lowest value in boreal winter (DJF) and boreal autumn (SON) are in good agreement with previous studies (e.g., Dong et al., 2005; Xi et al., 2010; Zhang et al., 2019).

Different from CBH, the occurrence frequencies of CTH exhibit a bimodal distribution, with a lower peak between 0.5 and 3.0 km and an upper peak between 6 and 12 km (Figure 8b, e), which agrees with previous studies based on ground-based lidar and radar measurements (Comstock and Jakob, 2004). The occurrence frequency of CTH reaches a maximum at around 11 km, which is slightly lower than the corresponding global annual mean tropopause height (almost 12 km); and then it decreases rapidly with altitude. Above 9 km, the maximum occurrence frequency of CTH occurs in boreal summer (JJA), which is due to the deeper troposphere and the higher frequency of deep convective clouds in summer (Johnson et al., 1999; Zhang et al., 2019). This result is consistent with that reported by Zhang et al. (2021), which demonstrated that CTH peaks around in summer. Below 3 km, the occurrence frequency of CTH is the highest in boreal winter (DJF). The occurrence frequencies of CTH tend to increase with altitude between 3 and 9 km irrespective of seasons.

As for CTH, the vertical distribution of cloud occurrence frequency is distinctly bimodal, but in different ranges of altitude, namely, a lower peak between 0.5 and 3 km and a higher peak between 6 and 10 km (Figure 8c, f). These findings are consistent with previous studies (e.g., Mace and Benson, 2008; Zhang et al. 2014a), which reported that the vertical profile of cloud occurrence peaks both in the boundary layer and upper troposphere by analysing cloud retrievals from radiosonde, models,

330   and satellites at both regional and global scale. Above 6 km, more clouds occur in boreal summer (JJA) than other seasons, while below 3 km, there are more clouds in boreal winter (DJF).

To further study the characteristics of CVS over different regions of the world, the vertical distributions of the occurrence frequencies of CBH, CTH, and cloud over six regions of interest (ROI; Figure 1) at 1200 UTC are shown in Figures 9–11. The regional annual mean PBLH and tropopause height as functions of longitude and latitude are also shown. Similar with the
335   results of global mean CBHs (Figure 8a, d), the maximum occurrence frequencies of regional mean CBHs are also within the regional mean PBLH (Figure 9a–c, g–i). The occurrence frequencies of CBHs are nearly zero above the tropopause. The frequency of cloud bases between 1 and 3 km in East Asia is roughly larger than that over North America and Europe, which is consistent with previous studies (e.g., Zhang et al., 2013; Zhou et al., 2021; Sharma et al., 2022). The possible reason is that the air pollution issue in China is more severe. The higher aerosol loading can change the macro- and micro-physical properties
340   of clouds and may invigorate convective clouds at high altitudes at the expense of low-level clouds (Wall et al., 2014; Guo et al., 2018; and Liu et al., 2020). The variation pattern of CBH frequency with altitude in the Southern Hemisphere (SH) is similar with that in NH, but its fluctuation is roughly greater due to fewer stations available there. A close inspection of the horizontal distribution of CBHs shows that there is a west-high-east-low pattern in America and a north-high-south-low pattern in China (Figure 9e), which correspond to arid (high CBHs) and humid (low CBHs) climates there. Most (67 %) annual mean
345   CBHs range from 2 to 4 km with an average of 3.07 ± 1.06 km (Figure 9d–f).

The characteristic of CTHs over the six ROIs at 1200 UTC are shown in Figure 10. Similar with the global mean CTHs, the regional mean occurrence frequencies of CTHs above tropopause also decrease rapidly with altitude. As expected, most of cloud tops are within the tropopause height in East Asian and North America. However, this predominance is not obvious in Europe and Antarctica, where the tropopause is lower since the climates are humid. In East Asian, the CTH frequency is
350   generally highest in boreal summer (JJA) above 6 km since there is more frequent deep convective cloud during the summer monsoon (Wang et al., 2000), while highest in boreal winter (DJF) below 6 km as few solar radiations reach the surface in winter that the cloud development is weak, thus the CTH is relatively low. This result is consistent with previous cloud radar observations (Zhang et al., 2019). Relatively large CTHs occur at southern USA and southeast China which are affected by monsoon, resulting in more deep convective clouds (Figure 10e). Most (85 %) annual mean CTHs vary from 7 to 12 km with
355   an average of 8.44 ± 1.52 km (Figure 10d–f).

The annual and seasonal mean occurrence frequencies of clouds at different regions are shown in Figure 11. Bimodal distribution of the occurrence frequencies of clouds with altitude appears at North America and Europe, peaking at the top of boundary layer (~1 km) and the tropopause (10~11 km). While unimodal distribution is at East Asia, peaking at about 3 km, namely fewer clouds occur in the boundary layer compared to the result of global mean. This is attributed to more serious air
360   pollution issue in East Asia compared to North America and Europe, resulting in higher occurrence frequencies of convective clouds at high altitudes (Liu et al., 2020; Sharma et al., 2022). This result is consistent with Xu et al. (2021a). Above 6 km in East Asian (~9 km at North America; ~10 km at Europe), the cloud occurrence frequency in JJA is higher than those in other seasons, because the deeper troposphere and the more frequent occurrence of convective storms in summer (Xi et al., 2010).

Below 3 km in East Asia (~1 km in North America and Europe), the frequency of cloud occurrence in DJF is highest, since

365   stratus clouds usually occur in winter (Dong et al., 2005). In Antarctica, the peak of cloud occurs at low altitude (~1 km, a

little bit higher than the PBLH ~ 0.70 km) due to the cold climate there. Contrary to CBHs, the annual mean CTs have a west-

lower-east-higher pattern in North America, and a northwest-lower-southeast-higher pattern in East Asia. Relatively large CTs

exist in Europe since there are sufficient water vapours for cloud development under the ocean climate. The CTs at 20–30 °

zone are relatively smaller (Figure 11f), because there coincide with the mean subsidence zone of the Hadley circulation (Poore

370   et al., 1995). Most (52 %) annual mean CTs vary from 4 to 6 km with an average of 5.37 ± 1.58 km (Figure 11d–f).

The corresponding near-global and regional mean vertical distributions of occurrence frequencies of CVS at 0000 UTC

on annual and seasonal time scales (Figures S3–7) are similar with those at 1200 UTC.

3.4 Near-global spatial distribution of CVS characteristics

Figure 12 presents the horizontal distribution of the annual mean occurrence frequencies for different layer clouds. Generally,

375   more clouds occur at humid climate regions (e.g. eastern USA, southern China, and western Europe) (Figure 12a). This is due

to sufficient water vapor supply from the surface, and clouds can be generated if areodynamic and thermodynamic conditions

required for cloud decoupling are met. Similarly, the wetter the climate, the higher the occurrence frequencies of clouds for

three-, four- and five-layer clouds in cloudy skies (Figure 12d–f). On the contrary, more one-layer clouds in cloudy skies occurs

at regions of relatively arid climate compared to the humid climate regions (Figure 12b). Reddy et al. (2018) reported the

380   occurrence of multi-layer clouds was more significant under moist atmospheric conditions, while more one-layer clouds

occurred under dry atmospheric condistions. As the transiation between one- and multi-layer clouds, the spatial feature of two-

layer clouds (Figure 12c) is not apparent compared to other cloudy conditions. Spatially, more one-layer clouds occur over

Asia than in North America (Figure 12b), while there are roughly larger cloud frequencies for multi-layer clouds over North

America (Figure 12c–f). These results are consistent with Wang et al. (2000), which demostrated that the frequency of multi-

385   layer clouds was higher in North America than in Asia. Few studies provided the global spatial distribution of the occurrence

frequencies of clouds base for cloud with various number of layers (from one to five) by radiosonde measurements as shown

in Figure 12.

Figure 13 shows the horizontal distributions of the near-global CBH, CTH, and CT at 1200 UTC. In terms of seasonality,

the magnitudes of CBH follow the order of boreal summer (JJA: 3.74 ± 1.56 km) > autumn (SON: 3.08 ± 1.22 km) ≈ spring

390   (MAM: 3.05 ± 1.11 km) > winter (DJF: 2.55 ± 1.02 km) (Figure 13a–d). This result is consistent with previous satellite

observations that the CBHs reached its maximum in summer and minimum in winter (Chi et al., 2022). The reason can be

explained that more solar radiation energy is available for cloud development in summer (Zhang et al., 2018). In East Asia,

the CBHs exhibit a northwest-high-southeast-low pattern, namely, high clouds (as high as 4.0 km) tend to occur in northwest

China while low clouds (less than 2.0 km) tend to occur in southeast China and the northern part of the South China Sea (Figure

395   13a–d). This phenomenon further illustrates that humid (arid) climate tends to generate low (high) clouds. The reason may be

due to that the higher dew point at humid climate regions always cause a lower LCL, resulting in lower cloud bases (Zhang et

al., 2018). In north America, the CBHs are lower in the east and higher in the west, which is consistent with the results of An et al. (2017) based on the information from the Automated Surface Observing System Observation for a 5–year period (2008– 2012). Compared to other seasons, the CBHs in boreal winter (DJF) vary little with space.

400   For CTH, its relative magnitudes have the same seasonality with those for CBH, namely, higher CBH occurs in boreal summer (JJA:9.24 ± 1.67 km) and relatively lower value occurs in boreal winter (DJF: 7.63 ± 2.05 km) (Figure 13e–h). This is because more deep convective clouds occur in summer compared to other reasons (Johnson et al., 1999; Zhang et al., 2019). During boreal summer (JJA), the spatial distribution of CTHs exhibit an obvious southeast-high-northwest-low pattern in China, and an east-high-west-low pattern in contiguous United State. The phenomenon of the southeast-high-northwest-low 405 pattern in China is much related to the East Asian summer monsoon, which causes strong atmospheric convection in the southeastern China and makes the convective instability layer much thicker. This allows the water vapor to be transported to a higher altitude and favors the production of deep convective clouds (Sun et al., 2019; Chi et al., 2022). The east-high-west-low pattern of CTH in contiguous United State is attributed to the humid (arid) climate in the east (west) of United States.

       The CT values (Figure 13i–l) undergo almost the same seasonality with the CBHs and CTHs with the order of JJA (5.50 410 ± 1.89 km) > SON (5.23 ± 1.58 km) ≈ MAM (5.20 ± 1.59 km) > DJF (5.08 ± 1.96 km). The spatial pattern of the total CT is similar with that of CTH with larger CT at the humid climate regions and smaller CT at the arid climate regions. The reason is that higher RH not only contributes to a lower cloud base but also helps an air parcel to reach higher levels, and thus leads to larger cloud thickness (Zhang and Klein, 2013). Interestingly, the total CTs in Europe are relatively large compared to those in China and USA. As mentioned before, this is partly explained by that oceanic climates dominate in Europe where more 415 water vapor supply is available for cloud development.

       Overall, the spatial distributions of CVS characteristics in different seasons are largely affected by the climates. The arid (humid) climates favor larger (smaller) CBH, smaller (larger) CTH and total CT. The corresponding results of CVS at 0000 UTC are roughly close to those at 1200 UTC, as shown in Figures S8–9.

3.5 Diurnal variation of cloud occurrence frequency with height

420 Since near-global radiosonde sites cover almost all time zones, the diurnal cycle of cloud vertical structures can be obtained by converting UTC to local solar time (LST). Based on the radiosonde derived CVS, the diurnal variations of height-resolved occurrence frequencies of clouds in boreal summer (JJA) and boreal winter (DJF) are presented in Figure 14. The mean vertical profiles of the occurrence frequency of liquid, ice, and mixed clouds at both 0000 and 1200 UTC are also shown for the corresponding season. Ice clouds are identified when T below -20 °C, liquid clouds are detected with T above 0 °C, and mixed 425 clouds correspond to T between -20 °C and 0 °C, as described in the Section 2.2.1.

       In boreal summer, the maximum occurrence frequencies of clouds mainly appear at altitudes of 0.5–3 km and the second peak extend upward to be 6–10 km (Figure 14b). The lower peak of the cloud occurrence frequency (0.5–3 km) roughly corresponds to the peak of the liquid cloud occurrence frequency (0.5–3 km), while the upper peak of the cloud occurrence frequency (6–12 km) roughly corresponds to the combined peak of the ice (8–12 km) and mixed cloud occurrence frequencies

430 (6–8 km) as shown in Figure 14bc and Figure 8c, f. Note that the vertical distribution of liquid and ice clouds occurrence frequencies are expected, since T tends to decrease as altitude increases and thus, ice clouds tend to be generated at higher altitude compared to liquid clouds. Previous studies also reported that ice clouds were located higher than liquid clouds based on satellite data (Chang and Li, 2005). The clouds in the lower atmosphere tend to appear as sun rises and increases from morning to early afternoon (1500 LST) within 3 km altitude. The reason is that, during daytime solar radiation heats the surface

435 and drives turbulence and convection in the boundary layer, which largely affects the boundary layer clouds (Noel et al., 2018). In later afternoon (1600–1800 LST), clouds tend to form most frequently at relatively high altitudes (6–12 km), which is consistent with Chen et al. (2018). The reason may be that after solar radiation reaches peak the sun heats less and cloud top begins to cool, which increases the atmospheric instability, and fuels the development of one-layer cloud and the uppermost layer of cloud (Zhang et al., 2010). These findings are consistent with those reported by Chang and Li (2005) who analysed

440 satellite cloud retrievals on a global scale. After 1800 LST, clouds tend to occur below 3 km again.

For boreal winter, clouds mainly occur with maximum frequencies at altitudes of 0.5–3 km, which roughly corresponds to the combined peak of the liquid (0.5–3 km) and mixed cloud occurrence frequencies (0.5–6 km) as shown in Figure 14e–f and Figure 8c, f. Note that there is a relatively large occurrence frequency of clouds during 0600–0800 LST, which is consistent with Betts and Tawfik (2016), who demonstrated that clouds had a near-sunrise peak in the cold season. The reason may be

445 that the temperatures are lowest around sunrise for a stable boundary layer, which are beneficial for water vapor to accumulate and form cloud (e.g., Dai, 2001; Eastman and Warren, 2014; Gao, et al., 2019; An et al., 2017, 2020).

In boreal spring and autumn, the characteristics of the diurnal variations of cloud occurrence frequencies falls between that in boreal summer and winter. The diurnal variation of height-resolved occurrence frequencies of clouds provided by this study can be a reference for satellite-retrieved CVS products, even though the sampling at each hour is not even (Figure 14a–

450 d).

3.6 Association of CBH with meteorological variables and moisture

We further investigate the potential key factors linked to CBH, since CBH is one of the most important factors affecting CRE in numerical weather forecast and regional climate model (Kato et al., 2011; Viúdez-Mora et al., 2015; Prein et al., 2015). Previous studies suggested that CBH is affected by thermodynamic conditions (e.g., surface T, RH, and pressure) and large-

455 scale dynamic process (wind speed) at the near surface layer, and moisture conditions (Mauger and Norris, 2010; Gbobaniyi et al., 2011). In this study, we investigate the relationship of CBH with meteorological variables (T_{2m}, PS, RH_{2m}, and $WS_{10m} = \sqrt{u^2 + v^2}$), and moisture (θ and MFD).

The spatial distributions of the correlation relationships between CBH and meteorological variables are displayed in Figure 15. Each point is the R value of a given CBH and meteorological variable at 1200 UTC over 2018–2019. Also given

460 are the mean value and one standard deviation of the Rs on a global scale, which is obtained by averaging the absolute R value of each point. Generally, the magnitudes of the global mean absolute R values follow the order of R (CBH & RH_{2m}) = 0.29 ± 0.17 > R (CBH & T_{2m}) = 0.22 ± 0.16 > R (CBH & PS) = 0.17 ± 0.14 > R (CBH & WS_{10m}) = 0.15 ± 0.13. Since the CBH is

detected using the profiles of T and RH, it is reasonable to expect that the correlation between CBH and T_{2m} and RH_{2m} are relatively high. Our result is consistent with Gbobaniyi et al. (2011), which demonstrated that surface temperature and specific

465 humidity were strongly coupled with CBH. Interestingly, the CBHs are negatively related with RH_{2m} while positively related with T_{2m}, which further convince us that humid and cold climate tends to generate low clouds. The physical explanation could be that, compared to dry air, humid air tends to reach condensation at relatively low altitude, thus the corresponding CBH is low. Note that the correlation between T_{2m} (RH_{2m}) and CBH tends to be larger in northern China compared to the other regions of China. This is attributed to the combined effects of the larger T_{2m} and smaller RH_{2m}, resulting in larger CBH (as shown in

470 Figure 13a–d). Compared with T_{2m} and RH_{2m}, the correlation relationship between CBH and PS and WS_{10m} are relatively weaker (most of absolute R < 0.3). Note that the correlation of CBH and WS_{10m} is negative in southeastern China but positive in northwestern China. The reason may be that the climate of southeastern China is mainly affected by monsoon and the large WS brings more humid air, causing lower cloud bases; while the air is relatively dry in the northwestern China, a region that is largely affected by the high pressure from Siberia, and large WS brings more dry air, corresponding to higher cloud bases.

475 Figure 16 shows the spatial distributions of the correlation relationships between CBH and soil moisture (θ,left panels) and MFD (right panels). Obviously, most CBHs are negatively correlated to θ with mean absolute R (CBH & θ) = 0.22 ± 0.15. This suggests that large soil moisture corresponds to low CBH, since more surface fluxes are partitioning to latent heat flux with a large soil moisture, causing a reduced boundary layer height. Thus, clouds are generated with low bases (Betts, 2004; Cook et al., 2006; Huang and Margulis, 2013). The correlation of MFD with CBH is relatively weak, with most R values

480 ranging from – 0.1 to 0.1 (Figure 16b). This is not surprising since CBH is not significantly affected by the transportation of water vapor. Overall, the correlation between CBH and RH_{2m} is the strongest, followed by T_{2m}, θ, PS, and WS_{10m}, and MFD is the least.

To further obtain a quantitative understanding of the effects of meteorological variables and moisture on CBH, the 2D joint distribution of CBH is presented in Figure 17. At given T_{2m}, larger RH_{2m} corresponds to lower CBH; while at given

485 RH_{2m}, larger T_{2m} tends to result in higher CBH (Figure 17a), which are similar with Figure 15. Besides, the combined effect of T_{2m} and RH_{2m} on the variation of CBH shows that larger T_{2m} and smaller RH_{2m} result in higher CBH. The same as Figure 16, the CBH varies slightly as MFD changes, indicating that MFD exerts little impact on CBH. Different with MFD, larger θ attributes to lower CBH when θ > 0.2 m3 m-3. However, CBH is not sensitive to low θ. The relationship of CBH with meteorological variables and soil moisture at 0000 UTC are similar to those at 1200 UTC, as shown in Figures S10–12.

490 **4. Summary and conclusions**

Based on high-vertical-resolution radiosonde observations from 374 radiosonde stations across the world, two years' worth (2018–2019) of near-global, high-quality CVS has been determined. A novel retrieval method is developed which combines the vertical gradients of T and RH and the altitude-resolved thresholds of RH. The accuracy of radiosonde derived CVS has been assessed by comparison with MMCR measurements at Beijing site during the year of 2019. The good agreement in CBHs

495   (R = 0.91) and CTHs (R = 0.81) confirms that this retrieval method performs reasonably well. On global scale, the CBHs from ERA5 tend to be lower than those from radiosonde with correlation coefficients (R) of 0.24 and 0.49 at 0000 and 1200 UTC, respectively. The characteristics of near-global CVS are summarized as follows:

The near-global annual mean occurrence frequency of all clouds is 65.3 % (69.5 %) at 0000 (1200) UTC, which is close to the result from multiple satellite observations (68.0 % ± 3.0 %). Over cloudy skies, the cloud frequency decreases as the

500   number of cloud layers increases. The one-layer clouds are predominant accounting for 55.4 % (53.8 %) of all cloud configurations. Among the detected multi-layer clouds, 62.2% (61.1%) are two-layer clouds at 0000 (1200) UTC. More clouds occur in boreal winter (DJF) and spring (MAM) than in summer (JJA) and autumn (SON).

The CBH and CT of each cloud layer (CTH) tend to decrease (increase) as the number of cloud layers increases. For one-layer clouds, they are roughly located at altitudes that fall somewhere among the altitudes of the multi-layer cloud

505   configurations, and their thicknesses are greater than those of any layer of multi-layer clouds. For multi-layer clouds, the lowermost cloud layer almost occurs below 3 km with small variation, while the upper layer exhibits more significant variation. The maximum altitude differences between cloud layers is around 2.5 km , which exists between the first and second layer for the three- and more multi-layer cloud configurations. For the upper layers, altitude difference decreases to be about 1.0 km.

In most of cases CBH are located within 1 km with the highest frequency at 0.5 km, which lower than the global annual

510   mean PBLH (0.76 km). This result indicates that in most of the time cloud bases are in the atmospheric boundary layer. The vertical distributions of the CTH occurrence frequencies have two peaks with the upper peak at around 11 km, slightly lowering than the global annual mean tropopause height (11.76 km). Similar with CTHs, the cloud occurrence frequencies also exhibit a bimodal distribution, with a lower peak being between 0.5 and 3 km and an upper peak being between 6 and 10 km, agreeing well with previous studies based on radiosonde, models, and satellites at both regional and global scale. Regionally, the

515   occurrence frequencies of CBH in East Asian between 1 and 3 km is generally larger than those over North America and Europe, probably due to the severe air pollution issue in East Asia that invigorates convective clouds at high altitudes which suppresses the occurrence of low-level clouds.

As for the horizontal distribution of CVS, we find that there are more clouds occur at humid climate regions (e.g. eastern USA, southern China, and western Europe) since sufficient water vapor supply from the surface benefits the generation of

520   clouds. More one-layer clouds occur at arid cimate regions, while more clouds with more than two layers occur at humid climate regions. Compared to one-layer clouds and multi-layer clouds, there is no apparent spatial feature for two-layer clouds.

The global mean CBH, CTH, and total CT are 3.15 ± 1.15 (3.07 ± 1.06), 8.04 ± 1.60 (8.44 ± 1.52), and 4.89 ± 1.36 (5.37 ± 1.58) km AGL at 0000 (1200) UTC, respectively. The CBH, CTH, and total CT have almost the same seasonality, with magnitudes following the order of boreal summer > autumn ≈ spring > boreal winter. The spatial distributions of CVS are

525   largely affected by the climate, namely, the arid (humid) climates correspond to larger (smaller) CBH, smaller (larger) CTH and CT.

In terms of the diurnal variation of cloud occurrence frequency, in boreal summer the maximum cloud occurrence frequencies mainly appear at the altitudes of 0.5–3 km, corresponding to the peak of the occurrence frequencies of liquid

clouds; and the second peak extends upwards to be 6–12 km, corresponding to the combined peak of the occurrence frequencies
530   of ice and mixed clouds. During daytime, clouds are affected by solar radiation, tending to increase from morning to early
afternoon (1500 LST) within 3 km altitude, peak in the later afternoon (1600–1800 LST) at relatively high altitude (6–12 km),
and then decrease after 1800 LST. In boreal winter, clouds are relatively low with maximum frequencies appear at altitudes of
0.5–3 km, corresponding to the combined peak of liquid and mixed cloud occurrence frequencies. And the clouds have a peak
during the period 0600–0800 LST, associating with the lowest temperature around sunrise, which are beneficial for water
535   vapor to accumulate and form clouds. The seasonal diurnal variations of cloud occurrence frequency are in accordance with
those from satellite and ground measurements.

As for the correlation of CBH with meteorological variables and moisture, we find that the CBH is correlated with both
the RH_{2m} and T_{2m}, with global mean absolute R values of 0.29 ± 0.17 and 0.22 ± 0.16, respectively. That is, larger air
temperature and smaller humidity tend to result in larger CBH. Most CBH is negatively correlated to soil moisture with global
540   mean absolute R value of 0.22 ± 0.15 indicating larger soil moisture corresponds to lower cloud bases. The relationship
between CBH and moisture flux divergence is weak.

This study is the first attempt to characterize the near-global CVS from high-vertical-resolution radiosonde data. The
results of CVS provide a valuable addition to the study of cloud-radiation-dynamics interactions and have the potential to
improve the performance of current climate model. The formation mechanism of cloud is complex, studying features affecting
545   the CVS appears to be a promising avenue for future work.

Acknowledgments. This work was jointly supported by the National Natural Science Foundation of China (NSFC) under
grants 42205089 and U2142209, the Youth Cross Team Scientific Research Project of the Chinese Academy of Sciences under
grant JCTD-2021-10, and the NSFC under grants 41975041 and 41875183. We also thank the National Meteorological
550   Information Centre (NMIC) of CMA, NOAA, German Deutscher Wetterdienst (Climate Data Center), UK Centre for
Environmental Data Analysis (CEDA), GRUAN, and the University of Wyoming for providing the high-resolution sounding
data. We would like to thank the ECWMF for ERA5 data National Meteorological Information Center, China Meteorological
Administration for generously providing the radiosonde datasets used here (https://data.cma.cn/en).

References

555   An, N, Wang, K. C., Zhou, C. L., and Pinker, R. T.: Observed variability of cloud frequency and cloud-base height within
3600m above the surface over the contiguous United States, J. Clim., 30, 3725–3742, https://doi.org/10.1175/JCLI-D-16-
0559.1, 2017.

An, N., Pinker, R. T., Wang, K. C., Rogers, E., and Zuo, Z. Y.: Evaluation of cloud base height in the North American Regional
Reanalysis using ceilometer observations, Int. J. Climatol., 40, 3161–3178, https://doi.org/10.1002/joc.6389, 2020.

560   Austin, R. T., Heymsfield, A. J., and Stephens, G. L.: Retrieval of ice cloud microphysical parameters using the CloudSat millimeter-wave radar and temperature, J. Geophys. Res.-atmos., 114, D00A23, https://doi.org/10.1029/2008JD010049, 2009.

Betts, A. K.: Understanding hydrometeorology using global models, Bull. Am. Meteorol. Soc., 85, 1673–1688, https://doi.org/10.1175/BAMS-85-11-1673, 2004.

565   Betts. A. K., and Tawfik, A. B.: Annual climatology of the diurnal cycle on the Canadian prairies, Font. Earth Sci., 22, https://doi.org/10.3389/feart.2016.00001, 2016.

Bosman, P. J. M., van Heerwaarden, C. C., and Teuling, A. J.: Sensible heating as a potential mechanism for enhanced cloud formation over temperate forest, Q. J. R. Meteorol. Soc., 145, 450–468, https://doi.org/10.1002/qj.3441, 2019.

Cesana, G., and Chepfer, H.: How well do climate models simulate cloud vertical structure? A comparison between CALIPSO-
570   GOCCP satellite observations and CMIP5 models, Geophys. Res. Lett., 39, L20803, https://doi.org/10.1029/2012GL053153, 2012.

Cesana, G., and Waliser, D. E.: Characterizing and understanding systematic biases in the vertical structure of clouds in CMIP5/CFMIP2 models, Geophys. Res. Lett., 43, 10538–10546. https://doi.org/10.1002/2016GL070515, 2016.

Chang, F. L., and Li, Z.: A near global climatology of single-layer and overlapped clouds and their optical properties retrieved
575   from TERRA/MODIS data using a new algorithm, J. Clim., 18, 4752–4771, https://doi.org/10.1175/JCLI3553.1, 2005.

Chen, D. D., Guo, J. P., Wang, H. Q., Li, J., Min, M., Zhao, W. H., and Yao, D.: The cloud top distribution and diurnal variation of clouds over East Asia: Preliminary results from advanced Himawari imager, J. Geophys. Res.-Atmos., 123, 3724–3739, https://doi.org/10.1002/2017JD028044, 2018.

Chernykh, I. and Eskridge, R.: Determination of cloud amount and level from radiosonde soundings, J. Appl. Meteorol. Clim.,
580   35, 1362–1369, https://doi.org/10.1175/1520-0450(1996)035<1362:DOCAAL>2.0.CO;2, 1996.

Chernykh, I. V., and Aldukhov, O. A.: Long-term estimates of the number of cloud layers from radiosonde data for 1964–2017 in different latitudinal zones, Russ. Meteorol. Hydro., 45,227–238, https://doi.org/10.3103/S1068373920040020, 2020.

Chi, Y. L., Zhao, C. F., Yang, Y. K., Ma, S. S., and Yang, J.: Cloud microphysical characteristics in China mainland and east
585   coast from 2006 to 2017 using satellite active remote sensing observations, Int. J. Climatol., https://doi.org/10.1002/Joc.7790, 2022.

Clothiaux, E. E., Ackermann, T. P., Mace, G. C., Moran, K. P., Marchand, R. T., Miller, M. A., and Martner, B. E.: Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites, J. Appl. Meteorol., 39, 645−665, https://doi.org/10.1175/1520-0450(2000)039<0645:ODOCHA>2.0.CO;2, 2000.

590   Comstock, J. M., and Jakob, C.: Evaluation of tropical cirrus cloud properties derived from ECMWF model output and ground based measurements over Nauru Island, Geophys. Res. Lett., 31, L10106, https://doi.org/10.1029/2004GL019539, 2004.

Cook, B. I., Bonan, G. B., and Levis, S.: Soil moisture feedbacks to precipitation in southern Africa, J. Clim., 19, 4198–4206, https://doi.org/10.1175/JCLI3856.1, 2006.

Costa-Surós, M., Calbó, J., González, J. A., and Martin-Vide, J.: Behavior of cloud base height from ceilmeter measurements, Atmos. Res., 127, 64–76, https://doi.org/10.1016/j.atmosres.2013.02.005, 2013.

Costa-Surós, M., Calbo, J., Gonzalez, J. A., and Long, C. N.: Comparing the cloud vertical structure derived from several methods based on radiosonde profiles and ground-based remote sensing measurements, Atmos. Meas. Tech., 7, 2757–2773, https://doi.org/10.5194/amt-7-2757-2014, 2014.

Crewell, S., Bloemink, H., Feijt, A., Garcia, S. G., Jolivet, D., Krasnov, O. A., van Lammeren, A., Lohnert, J., van Meijgaard, E., Meywerk, J., Quante, M., Pfeilsticker, K., Schmidt, S., Scholl, T., Simmer, C., Schroder, M., Trautmann, T., Venema, V., Wendisch, M., and Willen, U.: The BALTEX Bridge Campaign - An integrated approach for a better understanding of clouds, Bull. Am. Meteorol. Soc., 85, 1565–1584, https://doi.org/10.1175/BAMS-85-10-1565, 2004.

Dai, A.: Global precipitation and thunderstorm frequencies. Part II: Diurnal variation, J. Clim., 14, 1112–1128, https://doi.org/10.1175/1520-0422(2001)014<1112:GPATFP>2.0.CO;2, 2001.

Dong, X. P., Minnis, P., and Xi, B.: A climatology of midlatitude continental clouds from the ARM SGP central facility. Part I: low-level cloud macrophysical, microphysical, and radiative properties, J. Clim., 18, 1391–1410, https://doi.org/10.1175/JCLI13342.1, 2005.

Eastman, R., and Warren, S. G.: Diurnal cycles of cumulus, cumulonimbus, stratus, stratocumulus, and fog from surface observations over land and ocean, J. Clim., 27, 2386–2404, https://doi.org/10.1175/JCLI-D-13-00352.1, 2014.

Gao, C. C., Li, Y. Y., and Chen, H. W.: Diurnal variations of different cloud types and the relationship between the diurnal variations of clouds and precipitation in central and east China, Atmosphere, 10, https://doi.org/10.3390/atmos10060304, 2019.

Gbobaniyi, E. O., Abiodun, B. J., Tadross, M. A., Hewitson, B. C., and Gutowski, W. J.: The coupling of cloud base height and surface fluxes: a transferability intercomparison, Theor. Appl. Climatol., 106,189–210, https://doi.org/10.1007/s00704-011-0421-0, 2011.

George, G., Sarangi, C., Tripathi, S. N., Chakraborty, T., and Turner, A.: Vertical Structure and Radiative Forcing of Monsoon Clouds Over Kanpur During the 2016 INCOMPASS Field Campaign, J. Geophys. Res.-Atmos., 123, 2152–2174, https://doi.org/10.1002/2017JD027759, 2018.

Gouveia, D. A., Barja, B., Barbosa, H. M. J., Seifert, P., Baars, H., Pauliquevis, T., and Artaxo, P.: Optical and geometrical properties of cirrus clouds in Amazonia derived from 1 year of ground-based lidar measurements, Atmos. Chem. Phys., 17, 3619–3636, https://doi.org/10.5194/acp-17-3619-2017, 2017.

Guo, J. P., Miao, Y. C., Zhang, Y., Liu, H., Li, Z. Q., Zhang, W. C., He, J., Lou, M. Y., Yan, Y., Bian, L. G., and Zhai, P. M.: The climatology of planetary boundary layer height in China derived from radiosonde and reanalysis data, Atmos. Chem. Phys., 16, 13,309–13,319, https://doi.org/10.5194/acp-16-13309-2016, 2016.

Guo, J. P., Liu, H., Li, Z. Q., Rosenfeld, D., Jiang, M. J., Xu, W. X., Jiang, J. H., He, J., Chen, D. D., Min, M., and Zhai, P. M.: Aerosol-induced changes in the vertical structure of precipitation: a perspective of TRMM precipitation radar, Atmos. Chem. Phys., 18, 13329-13343, https://doi.org/10.5194/acp-18-13329-2018, 2018.

Guo, J. P., Li, Y., Cohen, J. B., Li, J., Chen, D. D., Xu, H., Liu, L., Yin, J. F., Hu, K. X., and Zhai, P. M.: Shift in the temporal trend of boundary layer height in China using long-term (1979–2016) radiosonde data, Geophys. Res. Lett., 46, 6080–

630   6089, https://doi.org/10.1029/2019GL082666, 2019.

Guo, J. P., Zhang, J., Yang, K., Liao, H., Zhang, S. D., Huang, K. M., Lv, Y. M., Shao, J., Yu, T., Tong, B., Li, J., Su, T. N., Yim, S. H. L., Stoffelen, A., Zhai, P. M., and Xu, X. F.: Investigation of near-global daytime boundary layer height using high-resolution radiosondes: first results and comparison with ERA5, MERRA-2, JRA-55, and NCEP-2 reanalyses, Atmos. Chem. Phys., 21, 17,079–17,097, https://doi.org/10.5194/acp-21-17079-2021, 2021.

635   Hahn, C., Warren, S., and London, J.: The effect of moonlight on observation of cloud cover at night, and application to cloud climatology, J. Clim., 8, 1429–1446, https://doi.org/10.1175/1520-0442(1995)008<1429:TEOMOO>2.0.CO;2, 1995.

Hahn, C. J., Rossow, B., and Warren, S. G.: ISCCP cloud properties associated with standard cloud types identified in individual surface observations, J. Clim., 14, 11–28, https://doi.org/10.1175/1520-0442(2001)014<0011:ICPAWS>2.0.CO;2, 2001.

640   Hersbach, H., Bell, B., and Berrisford, P., et al.: The ERA5 global reanlysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Huang, H. Y., and Margulis, S. A.: Impact of soil moisture heterogeneity length scale and gradients on daytime coupled land-cloudy boundary layer interactions, Hydrol. Progress., 27, 1988–2003, https://doi.org/10.1002/hyp.9351, 2013.

Hollars, S., Fu, Q., Comstock, J., Ackerman, T.: Comparison of cloud-top height retrievals from ground-based 35 GHz MMCR

645   and GMS-5 satellite observations at ARM TWP Manus site, Atmos. Res., 72, 169–186, https://doi.org/10.1016/j.atmosres.2004.03.015, 2004.

Houghton, J.T. Meira Filho, L. G., Callander, B. A., Harris, N., Kattenberg. A., and Maskell, K.: Climate change 1995: the science of climate change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change. United Kingdom, 1996.

650   IPCC: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]. Cambridge University Press, 2021.

655   Johnson, R. H., Rickenbach, T. M., Rutledge, S. A., Ciesielski, P. E., and Schubert, W. H.: Trimodal characteristics of tropical convection, J. Clim., 12, 2397–2418, https://doi.org/10.1175/1520-0442(1999)012<2397:TCOTC>2.0.CO;2, 1999.

Kato, S, Rose, F. G., Sun-Mack, S., et al.: Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties, J. Geophys. Res., 116: D19209, https://doi:10.1029/2011JD016050, 2011.

660 Kim, S. W., Chung, E. S., Yoon, S. C., Sohn, B. J., and Sugimoto, N.: Intercomparisons of cloud-top and cloud-base heights from ground-based Lidar, CloudSat and CALIPSO measurements, Int. J. Remote Sens., 32, 1179–1197, https://doi.org/10.1080/01431160903527439, 2011.

Lawson, R. P., and Cooper, W. A.: Performance of some airborne thermometers in clouds, J. Atmos. Ocean. Tech., 7, 480–494, https://doi.org/10.1175/1520-0426(1990)007<0480: POSATI>2.0.CO;2, 1990.

665 Lawrence, M. G.: The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications, Bull. Am. Meteorol. Soc., 8, 225–234, https://doi.org/10.1175/BAMS-86-2-225, 2005.

Lee, X.: Fundamentals of boundary-layer meteorology. Springer International Publishing. Retrieved from https://www.springer.com/gp/book/97833196, 2018.

Lee, Y. J., Titov, D. V., Ignatiev, N. I., Tellmann, S., Patzold, M., andPiccioni, G.: The radiative forcing variability caused by
670   the changes of the upper cloud vertical structure in the Venus mesosphere, Planet. Space Sci., 113, 298–308, https://doi.org/10.1016/j.pss.2014.12.006, 2015.

Loeb, N. G., Kato, S., Su, W. Y., Wong, T., Rose, F. G., Doelling, D. R., Norris, J. R., and Huang, X. L.: Advances in understanding top-of-atmosphere radiation variability from satellite observations. Surv. Geophys., 33, 359−385, https://doi.org/10.1007/s10712-012-9175-1, 2012.

675 Li, D., Liu, Y. Z., Shao, T. B., Luo, R., and Tan, Z. Y.: Assessment of cloud base height product from ERA5 reanalysis using ground-based observations, Chinese J. Atmospheric Sci., https://doi.org/10.3878/j.issn.1006-9895.2208.22109, 2022.

Li, Z., Guo, J., Ding, A., Liao, H., Liu, J., Sun, Y., Wang, T., Xue, H., Zhang, H., Zhu, B.: Aerosol and boundary-layer interactions and impact on air quality, Natl. Sci. Rev., 4, 810–833, https://doi. org/10.1093/nsr/nwx117, 2017.

Liu, C., Wang, T. J., Rosenfeld, D., Zhu, Y. N., Yue, Z. G., Yu, X., Xie, X. D., Li, S., Zhuang, B. L., Cheng, T. T., and Niu,
680   S. J.: Anthropogenic effects on cloud condensation nuclei distribution and rain initiation in East China, Geophys. Res. Lett., 47, e2019GL086184, https://doi.org/10.1029/2019GL086184, 2020.

Liou, K. N.: Influence of cirrus clouds on weather and climate process: A global perspective, Mon. Wea. Rev., 114, 1167–1199, https://doi.org/10.1175/1520-0493(1986)114<1167: IOCCOW>2.0.CO;2, 1986.

Luo, Y., Zhang, R., and Wang, H.: Comparing occurrences and vertical structures of hydrometeors between eastern China and
685   the Indian monsoon region using CloudSat/CALIPSO data, J. Clim., 22, 1052–1064, https://doi.org/10.1175/2008jcli2606.1, 2009.

Mace, G. G., Ackerman, T. P., Minnis, P., and Young, D. F.: Cirrus layer microphysical properties derived from surface-based millimeter radar and infrared interferometer data, J. Geophys. Res., 103, 23,207–23,216, https://doi.org/10.1029/98JD02117, 1998.

690 Matveev, L. T.: Chapter 7 Vertical structure of a cloud field in Clouds dynamics, Atmospheric and oceanographic sciences library, 198–230, https://doi.org/10.1007/978-94-009-6360-3, 1981.

Mauger, G. S., and Norris, J. R.: Assessing the impact of meteorological history on subtropical cloud fraction, J. Clim., 23, 2926–2940, https://doi.org/10.1175/2010JCLI3272.1, 2010.

Miao, H., Wang, X. C., Liu, Y. M., and Wu, G. X.: An evaluation of cloud vertical structure in three reanalyses against CloudSat/cloud-aerosol lidar and infrared pathfinder satellite observations, Atmos Sci Lett., 20, e906, https://doi.org/10.1002/asl.906, 2019.

Miller, M. A., Jensen, M. P., and Clothiaux, E. E.: Diurnal cloud and thermodynamic variations in the stratocumulus transition regime: A case study using in situ and remote sensors, J. Atmos. Sci., 55, 2294–2310, https://doi.org/10.1175/1520-0469(1998)055<2294:DCATVI>2.0.CO;2, 1998.

Minnis, P., Yi, Y. H., Huang, J. P., and Ayers, K.: Relationships between radiosonde and RUC-2 meteorological conditions and cloud occurrence determined from ARM data, J. Geophys. Res., 110, 1–19, https://doi.org/10.1029/2005JD006005, 2005.

Moran, K. P., Martner, B. E., Post, M. J., Kropfli, R. A., Welsh, D. C., Widener, K. B.: An unattended cloud-profiling radar for use in climate research, Bull. Am. Meteorol. Soc., 79, 443–455, https://doi.org/10.1175/1520-0477(1998)079<0443: AUCPRF>2.0.CO;2, 1998.

Monteith, J. L., and Unsworth, M. H.: Principles of Environmental Physics, Third Ed. AP, Amsterdam, 2008.

Murray, F. W.: On the computation of saturation vapour pressure, J. Appl. Meteorol. Clim., 6, 203-204, https://doi.org/10.1175/1520-0450(1967)006<0203: OTCOSV>2.0.CO;2, 1967.

Meehl, G. A., and Washington, W. M.: Cloud albedo feedback and the super greenhouse effect in a global coupled GCM, Clim. Dynam., 11, 399–411, https://doi.org/10.1007/BF00209514,1995.

Nam, C., Bony, S., Dufresne, J. L., and Chepfer, H.: The 'too few, too bright' tropical low-cloud problem in CMIP5 models, Geohpys. Res. Lett., 39, L21801, https://doi.org/10.1029/2012GL053421, 2012.

Naud, C. M., Muller, J.-P., and Clothiaux, E. E.: Comparison between active sensor and radiosonde cloud boundaries over the ARM southern great plains site, J. Geophys. Res., 108, D4, https://doi.org/10.1029/2002JD002887, 2003.

Naud, C. M., Muller, J.-P., Clothiaux, E. E., Baum, B. A., and Menzel, W. P.: Intercomparison of multiple years of MODIS, MISR and radar cloud-top heights, Ann. Geophys., 23, 2415–2424, https://doi.org/1432-0576/ag/2005-23-2415, 2005.

Noel, V., Chepfer, H., Chiriaco, M., Yorks, J.: The diurnal cycle of cloud profiles over land and ocean between 51° S and 51° N, seen by the CATS spaceborne lidar from the International Space Station, Atmos. Chem. Phys., 18, 9457-9473, https://doi.org/10.5194/acp-18-9457-2018, 2018.

Nowak, D., Ruffieux, D., Agnew, J. L., and Vuilleumier, L.: Detection of fog and low cloud boundaries with ground-based remote sensing systems, J. Atmos. Ocean. Technol., 25, 1357–1368, https://doi.org/10.1175/2007JTECHA950.1, 2008.

Oreopoulos, L., Cho, N., and Lee, D.: New insights about cloud vertical structure from CloudSat and CALIPSO observations, J. Geophys. Res.-Atmos., 122, 9280–9300, https://doi.org/10.1002/2017JD026629, 2017.

Pietrowicz, J. A., and Schiermeir, F. A.: Observational evidence of systematic radiosonde temperature sensing anomalies, J. Appl. Meteorol. Clim., 17, 1572–1575, https://doi.org/10.1175/1520-0450(1978)017<1572: OEOSRT>2.0.CO;2, 1978.

Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riedi, J. C., and Frey, R. A.: The MODIS cloud products: Algorithms and examples from Terra, IEEE T. Geosci. Remote, 41, 459–473, https://doi.org/10.1109/TGRS.2002.808301, 2003.

Poore, K., Wang, J., and Rossow, W.: Cloud layer thicknesses from a combination of surface and upper-air observations, J.
730   Clim., 8, 550–568, https://doi.org/10.1175/1520-0422(1995)008<0550: CLTFAC>2.0.CO;2, 1995.

Prein, A. F., Langhans, W., Fosser, G., et al.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, Rev. Geophys., 53(2): 323-361, https://doi:10.1002/2014RG000475, 2015

Randall, D., Khairoutdinov, M., Arakawa, A., and Grabowski, W.: Breaking the cloud parameterization deadlock, B. Am. Meteorol. Soc., 84, 1547–1564, https://doi.org/10.1175/BAMS-84-11-1547, 2003.

735 Ramanathan, V., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E., and Hartmann, D.: Cloud-radiative forcing and climate: Results from the earth radiation budget experiment, Science, 243, 57–63, 1989.

Reddy, N. N., Ratnam, M. V., Basha, G., and Ravikiran, V.: Cloud vertical structure over a tropical station obtained using long-term high-resolution radiosonde measurements, Atmos. Chem. Phys., 18, 11,709–11,727, https://doi.org/10.5194/acp-18-11709-2018, 2018.

740 Seidel, D. J., Zhang, Y. H., Beljaars, A., Golaz, J.-C., Jacobson, A. R., and Medeiros, B.: Climatology of the planetary boundary layer over the continental United States and Europe, J. Geophys. Res., 117, D17, https://doi.org/10.1029/2012JD018143, 2012.

Sharma, S., Dass, A., Mishra, A. K., Singh, S., and Kumar, K.: A decadal climatology cloud vertical structure over the Indo-Gangetic Plain using radiosonde and radar observations, 266, 105949, https://doi.org/10.1016/j.atmosres.2021.105949,
745   2022.

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A., O'Connor, E., Rossow, W. B., Durden, S. L., Miller, S. D., Austin, R. T., Benedetti, A., Mitrescu, C., The CloudSat Science Team: The CloudSat mission and the A-Train: a new dimension of space-based observation of clouds and precipitation. Bull. Amer. Meteor. Soc., 83, 1771–1790, https://doi.org/10.1175/BAMS-83-12-1771, 2002.

750 Stephens, G. L.: Cloud feedbacks in the climate system: A critical review, J. Clim., 18(2), 237–273, https://doi.org/10.1175/JCLI-3243.1, 2005.

Stephens, G. L., Li, J. L., Wild, M., Clayson, C. A., Leob, N., Kato, S., L'Ecuyer, T., Stackhouse, P. W., Lebsock, M., and Andrews, T.: An update on Earth's energy balance in light of the latest global observations, Nat. Geos., 5, 691–696, https://doi.org/10.1038/NGEO1580, 2012.

755 Stubenrauch, C. J., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., Di Girolamo, L., Getzewich, B., Guignard, A., Heidinger, A., Maddux, B. C., Menzel, W. P., Minnis, P., Pearl, C., Platnick, S., Poulsen, C., Riedi, J., Sun-Mack, S., Walther, A., Winker, D., Zeng, S., and Zhao, G.: Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX radiation panel, B. Am. Meteorol. Soc., 94, 1,031–1,049, https://doi.org/10.1175/BAMS-D-12-00117.1, 2013.

760 Solomon, S. D., Qin, D., Manning, M., Chen, Z., Marquis, M., Avery, K. B., Tignor, M., and Miller, H. L.: Climate Change 2007: The Physical Science Basis: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, 996, 2007.

Subrahmanyam, K.V., and Kumar, K. K.: CloudSat observations of multi layered clouds across the globe, Clim. Dyn., 49, 765   327–341, https://doi.org/10.1007/s00382-016-3345-7, 2017.

Sugimoto, N., Matsui, I., Shimizu, A., Pinandito, M., and Sugondo, S.: Climatological characteristics of cloud distribution and planetary boundary layer structure in Jakarta, Indonesia revealed by lidar observation, Geophys. Res. Lett., 27, 2909–2912, https://doi.org/10.1029/2000GL011544, 2000.

Sun, G. R., Li, Y. Y., and Liu, L. H.: Why is there a titled cloud vertical structure associated with the northward advance of 770   the East Asian summer monsoon, Atmos. Sci. Lett., 20, e903, https://doi.org/10.1002/asl.903, 2019.

Takle, E. S., Roads, J., Rockel, B., Gutowski, W. J., Arritt, R. W., Meinke, I., Jones, C. G., and Zadra, A.: Transferability intercomparison - An opportunity for new insight on the global water cycle and energy budget, Bull. Am. Meteorol. Soc., 88, 375–384, https://doi.org/10.1175/BAMS-88-3-375, 2007.

Trenberth, K. E., Fasullo, J. T., and Kiehl, J.: Earth's global energy budget, B. Am. meteorol. soc., 90, 311–323, 775   https://doi.org/10.1175/2008BAMS2634.1, 2009.

Viúdez-Mora, A., Costa-Surós, M., Calbó, J., et al.: Modeling atmospheric longwave radiation at the surface during overcast skies: The role of cloud base height, J. Geophys. Res.: Atmos., 120(1): 199-214, https://doi:10.1002/2014JD022310, 2015.

Vogelezang, D.H.P., and Holtslag, A. A. M.: Evaluation and model impacts of alternative boundary-layer height formulation, Bound. Lay. Meteorol., 81, 245–269, https://doi.org/10.1007/BF02430331, 1996.

780 Wall, C., Zipser, E., and Liu, C. T.: An investigation of the aerosol indirect effect on convective intensity using satellite observations, J. Atmos. Sci., 71, 430-447, https://doi.org/10.1175/JAS-D-13-0158.1, 2014.

Wang, J., and Rossow, W.: Determination of cloud vertical structure from upper-Air observations, J. Appl. Meteorol., 34, 2243–2258, https://doi.org/10.1175/1520-0450(1995)034<2243:DOCVSF>2.0.CO;2, 1995.

Wang, J., Rossow, W. B., and Zhang, Y. C.: Cloud vertical structure and its variations from a 20-yr global rawinsonde dataset, 785   J. Clim., 13, 3041–3056, https://doi.org/10.1175/1520-0442(2000)013<3041:CVSAIV>2.0.CO;2, 2000.

Wang, S., Naud, C. M., Kahn, B. H., Wu, L. T., and Fetzer, E. J.: Coupling of precipitation and cloud structures in oceanic extratropical cyclones to large-scale moisture flux convergence, J. Clim., 31, 9565–9584, https://doi.org/10.1175/JCLI-D-18-0115.1, 2018.

Waliser, D. E., Li, J-L. F., Woods, C. P., Austin, R. T., Bacmeister, J., Chern, J., Genio, A. D., Jiang, J. H., Kuang, Z., Meng, 790   H., Minnis, P., Platnick, S., Rossow, W. B., Stephens, G. L., Sun-Mack, S, Tao, W.-K. Tompkins, A. M., Vane, D. G., Walker, C., and Wu, D.: Cloud ice: A climate model challenge with signs and expectations of progress, J. Geophys. Res., 114, D00A21, https://doi.org/10.1029/2008JD010015, 2009.

Weare, B. C.: Insights into the importance of cloud vertical structure in climate, Geophys. Res. Lett., 27, 907–910, https://doi.org/10.1029/1999GL011214, 2000.

795   Wielicki, B. A., Cess, R. D., King, M. D., Randall, D. A., and Harrison, E. F.: Mission to planet Earth: Role of clouds and radiation in climate, Bull. Amer. Meteor. Soc., 76, 2125–2154, https://doi.org/10.1038/NGEO1580, 1995.

Willen, U., Crewell, S., Baltink, H.K., and Sievers, O.: Assessing model predicted vertical cloud structure and cloud overlap with radar and lidar ceilometer observations for the Baltex Brige Campagin of CLIWA-NET, Atmos. Res., 75, 227–255, https://doi.org/10.1016/j.atmosres.2004.12.008, 2005.

800   Wild, M.: New Directions: A facelift for the picture of the global energy balance, Atmos. Environ., 55, 366–367, https://doi.org/10.1016/j.atmosenv.2012.03.022, 2012.

Winker, D. M., Hunt, W. H., and McGill, M. J.: Initial performance assessment of CALIOP, Geophys. Res. Lett., 34, L19803, https://doi.org/10.1029/2007GL030135, 2007.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the
805   CALIPSO mission and CALIOP data processing algorithms, J. Atmos. Ocean. Tech., 26, 2310–2323, https://doi.org/10.1175/2009JTECHA1281.1, 2009.

WMO. Meteorology - A three-dimensional science: Second session of the commission for aerology, WMO Bull, 4, 134-138, 1957.

Xi, B. X., Dong, X. Q., Minnis, P., and Khaiyer, M. M.: A 10-year climatology of cloud fraction and vertical distribution
810   derived from both surface and GOES observations over the DOE ARM SGP site, J. Geophys. Res.-Atmos., 115, D12124, https://doi.org/10.1029/2009JD012800, 2010.

Xu, G. R., Zhang, W. G., Wan, X., and Wang, B.: Cloud occurrence frequency and cloud liquid water path for non-precipitating clouds using ground-based measurements over central China, J. Atmos. Solar-Terr. Phys., 215, 105575, https://doi.org/10.1016/j.jastp.2021.105575, 2021a.

815   Xu, Z. Q., Chen, H. S., Guo, J. P., and Zhang, W. C.: Contrasting effect of soil moisture on the daytime boundary layer under different thermodynamic conditions in summer over China, Geophys. Res. Lett., 48, e2020GL090989, https://doi.org/10.1029/2020GL090989, 2021b.

Xu, H., Guo, J. P., Li, J., Liu, L., Chen, T. M., Guo, X. R., Lyu, Y. M., Wang, D., Han, Y., Chen, Q., and Zhang, Y.: The Significant Role of Radiosonde-measured Cloud-base Height in the Estimation of Cloud Radiative Forcing, Adv. Atmos.
820   Sci., 38, 1552−1565, https://doi.org/10.1007/s00376-021-0431-5, 2021c.

Zhang, G. J.: Lagrangian study of cloud properties and their relationships to meteorological parameters over the US southern Great Plains, J. Clim., 16, 2700–2716, https://doi.org/10.1175/1520-0442(2003)016<2700:LSOCPA>2.0.CO;2, 2003.

Zhang, Y., Zhang, L. J., Guo, J. P., Feng, J. M., Cao, L. J., Wang, Y., Zhou, Q., Li, L. X., Li, B., Xu, H., Liu, L., An, N., and Liu. H.: Climatology of cloud-base height from long-term radiosonde measurements in China, Adv. Atmos. Sci., 35,
825   158−168, https://doi.org/10.1007/s00376-017-7096-0, 2018.

Zhang, Y., Zhou, Q., Lv, S., Jia, S., Tao, F., Chen, D., and Guo, J.: Elucidating cloud vertical structures based on three-year Ka-band cloud radar observations from Beijing, China, Atmos. Res., 222, 88–99, https://doi.org/10.1016/j.atmosres.2019.02.007, 2019.

Zhang, J. Q, Chen, H. B., Li, Z. Q., Fan, X. H., Peng, L. A., Yu, Y., and Cribb, M.: Analysis of cloud layer structure in Shouxian, China using RS92 radiosonde aided by 95 GHz cloud radar, J. Geophys. Res.-Atmos., 115, D00K30, https://doi.org/10.1029/2010JD014030, 2010.

Zhang, J. Q., Chen, H. B., Bian, J. C., Xuan, Y. J., Duan, Y. J., and Cribb, M.: Development of cloud detection methods using CFH, GTS1, and RS80 radiosondes, Adv. Atmos. Sci., 29, 236–248, https://doi.org/10.1007/s00376-011-0215-4, 2012.

Zhang, Y. Y., and Klein, S.: Factors controlling the vertical extent of fair-weather shallow cumulus clouds over land: investigation of diurnal-cycle observations collected at the ARM southern great plains site, J. Atmos. Sci., 70, 1297–1315, https://doi.org/10.1175/JAS-D-12-0131.1, 2013.

Zhang, J. Q., Li, Z. Q., Chen, H. B., Yoo, H., and Cribb, M.: Cloud vertical distribution from radiosonde, remote sensing, and model simulations, Clim. Dyn., 43, 1129–1140, https://doi.org/10.1007/s00382-014-2142-4, 2014a.

Zhang, Y. H., Seidel, D. J., and Zhang, S. D.: Trends in planetary boundary layer height over Europe, J. Clim., 26, 10,071–10,076, https://doi.org/10.1175/JCLI-D-13-00108.1, 2014b.

Zhou, Q., Zhang, Y., Jia, S. Z., Jin, J. L., Lv, S. S., and Li, Y. N.: Climatology of cloud vertical structures from long-term high-resolution radiosonde measurements in Beijing. Atmosphere, 11, 401, https://doi.org/10.3390/atmos11040401, 2020.

Zhou, R., Wang, G., and Zhaxi, S. L.: Cloud vertical structure measurements from a ground-based cloud radar over the southeastern Tibetan Plateau, Atmos. Res., 258, 105629, https://doi.org/10.1016/j.atmosres.2021.105629, 2021.

845 **Tables**

Table 1. Summary of altitude-dependent thresholds of RH

Altitude Range	Height-Resolving RH Thresholds		
	min-RH	max-RH	inter-RH
0-2km	84%	94%	82%
2-6km	80%	92%	78%
6-12km	78%	88%	70%
>12km	70%	80%	70%

Figures

850

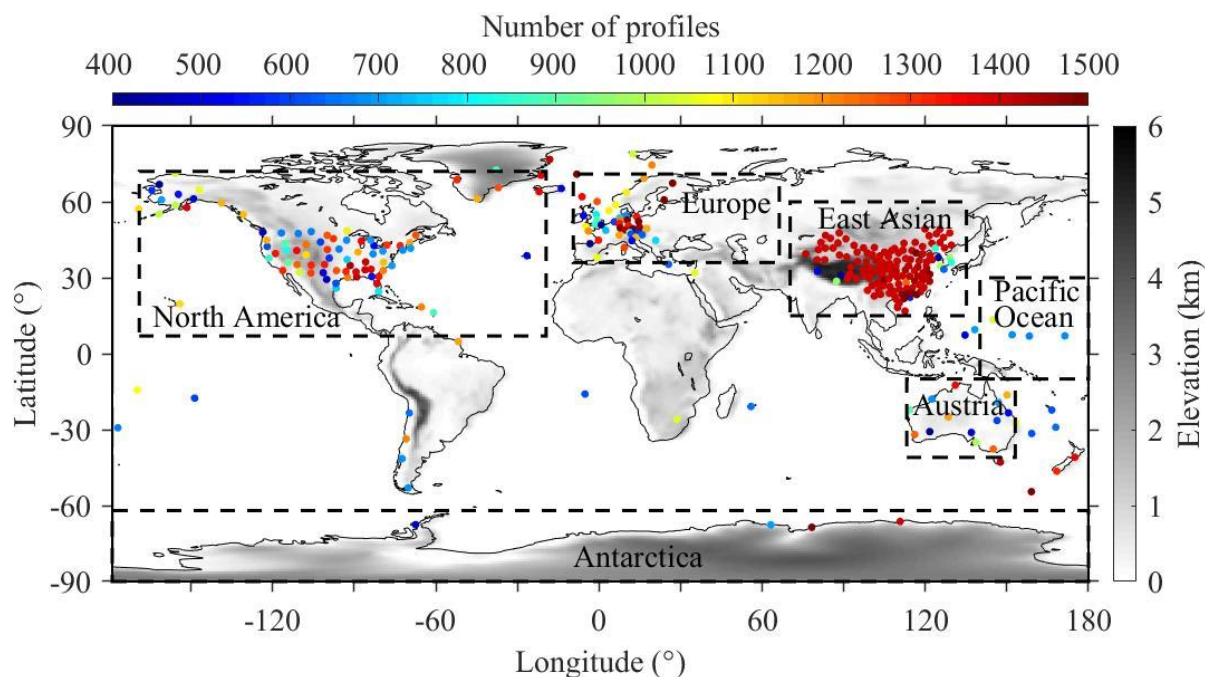



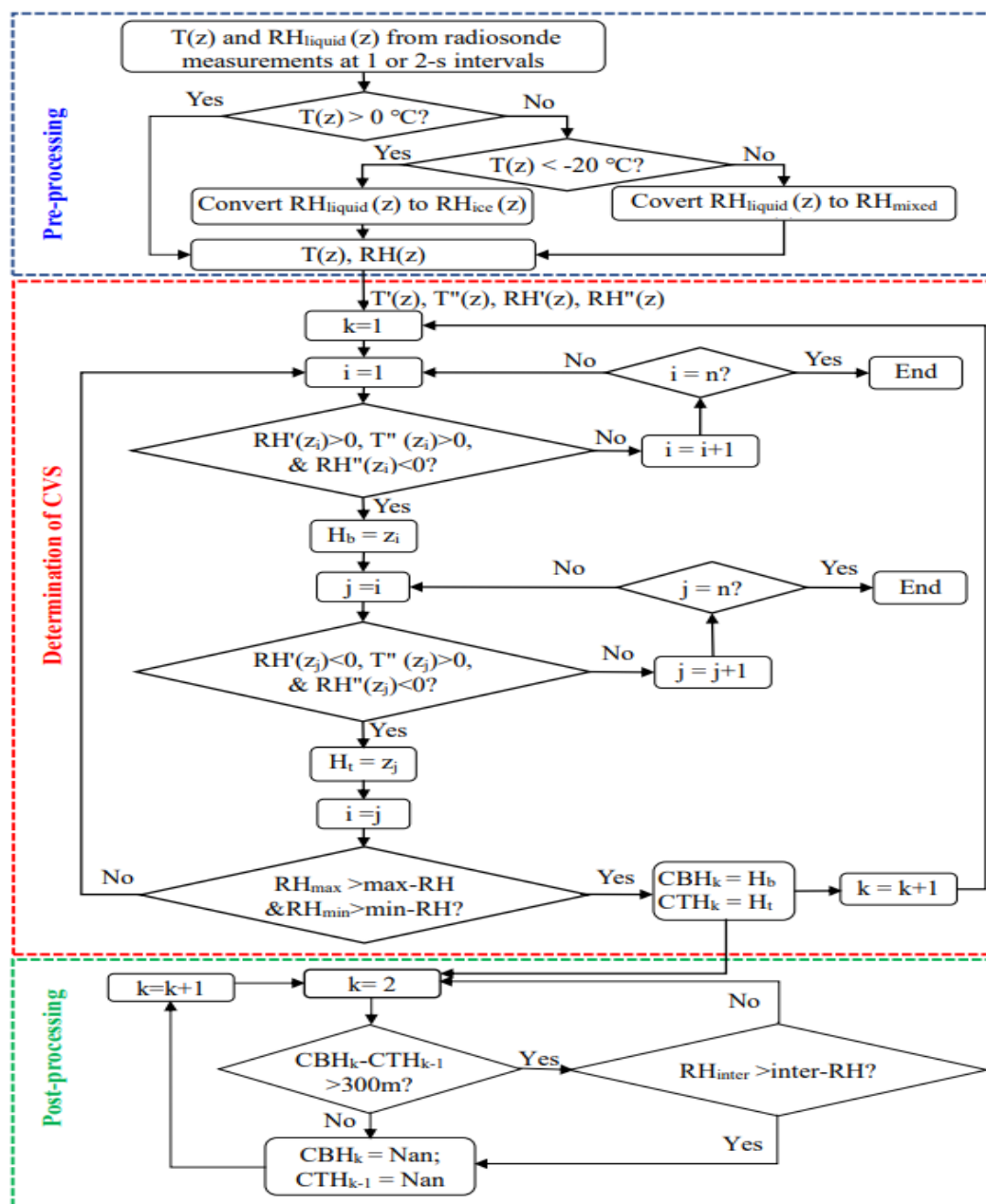
Figure 1. Geographic distribution of the number of profiles (colored full circle) for the near-global high-resolution radiosonde observation stations from 2018 to 2019. Also shown is the elevation of each radiosonde site in grey-scale shading. Black rectangles denote observation coverages at six regions of interest (i.e., Europe, North America, East Asia, Austria, Pacific Ocean, and
855 **Antarctica). The numbers of radiosonde stations within Europe, North America, East Asia, Austria, Pacific Ocean, and Antarctica are 44, 150, 120, 13, 5, and 4, respectively. The numbers of radiosonde stations in the Northern Hemisphere (NH) and Southern Hemisphere (SH) are 338 and 36, respectively.**

k: the number of cloud layer; H_{bi}: base height of ith moist layer; H_{ti}: top height of ith moist layer; RH_{min}: the maximum of RH in moist layer; RH_{max}: the minimum of RH in moist layer;

Figure 2. Flow chart showing the determination of the cloud vertical structure (CVS) using the vertical profiles of air temperature (T) and relatively humidity (RH) from high-resolution radiosonde data and the altitude-dependent thresholds of RH defined in Table 1.

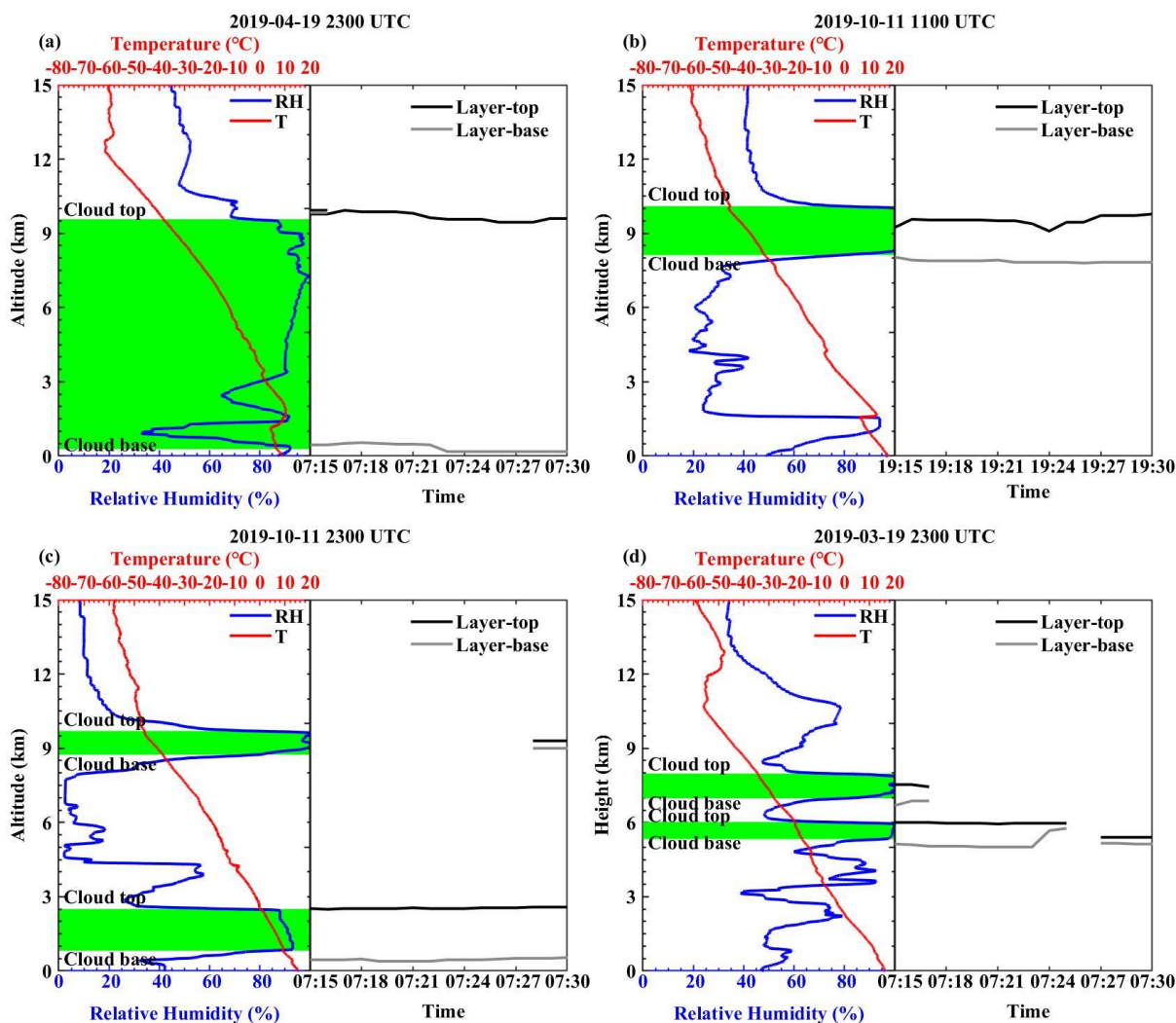



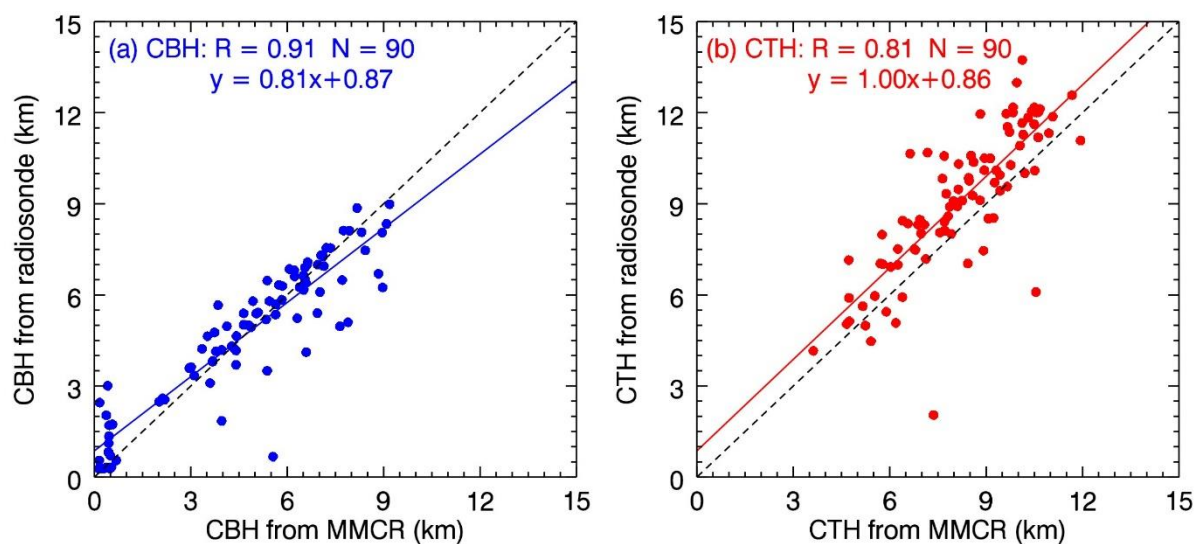
Figure 3. Examples of the detection of CVS by (left) high-resolution radiosonde and (right) Ka-band millimeter wavelength cloud radar (MMCR) at Beijing site for the four selected cases, (a, b) one-layer clouds, and (c, d) two-layer clouds. Green shading represents the cloud layers retrieved from radiosonde. In each subfigure (left), the blue and red solid line represent the RH and T profile, respectively.

875

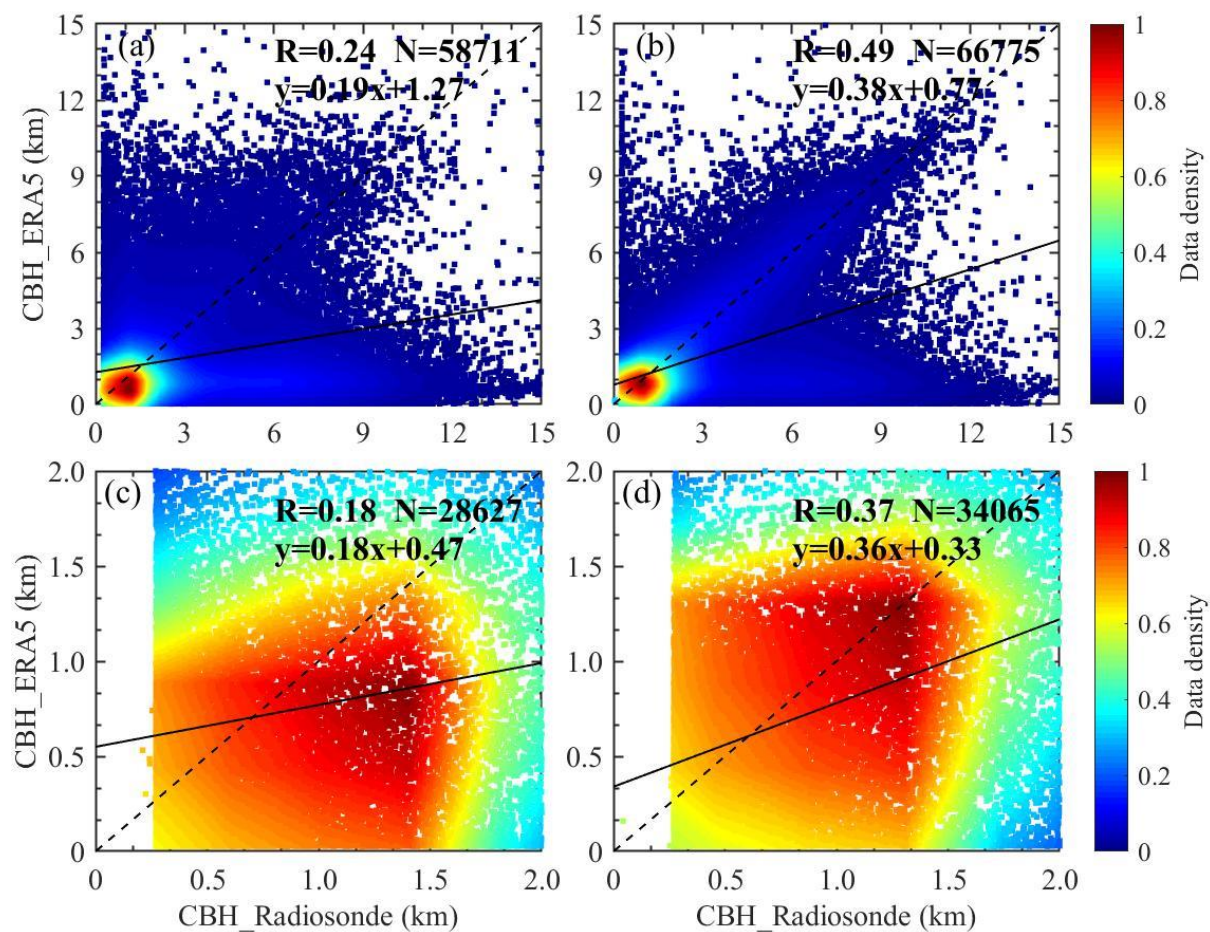
Figure 4. Scatter plots of the radiosonde derived (a) cloud base heights (CBH) and (b) cloud top heights (CTH) versus those from the MMCR at Beijing site during the year of 2019. R represents the correlate coefficient and N represents the sample number. The dashed line and solid line in each panel denote the 1:1 line and the linear regression line, respectively.

880

Figure 5. Scatter plots of the comparisons of the radiosonde derived CBHs with those from ERA5 for all the data at (a) 0000 and (b) 1200 UTC, and for the cases that both CBHs are less than 2.0 km at (c) 0000 and (d) 1200 UTC during the period of 2018–2019. Each point represents one measurement at 0000 or 1200 UTC. R represents the correlate coefficient and N represents the sample number. The linear regression equation is also given, and the regression line is marked in black solid line. The dash black line is 1:1 885 **line.**

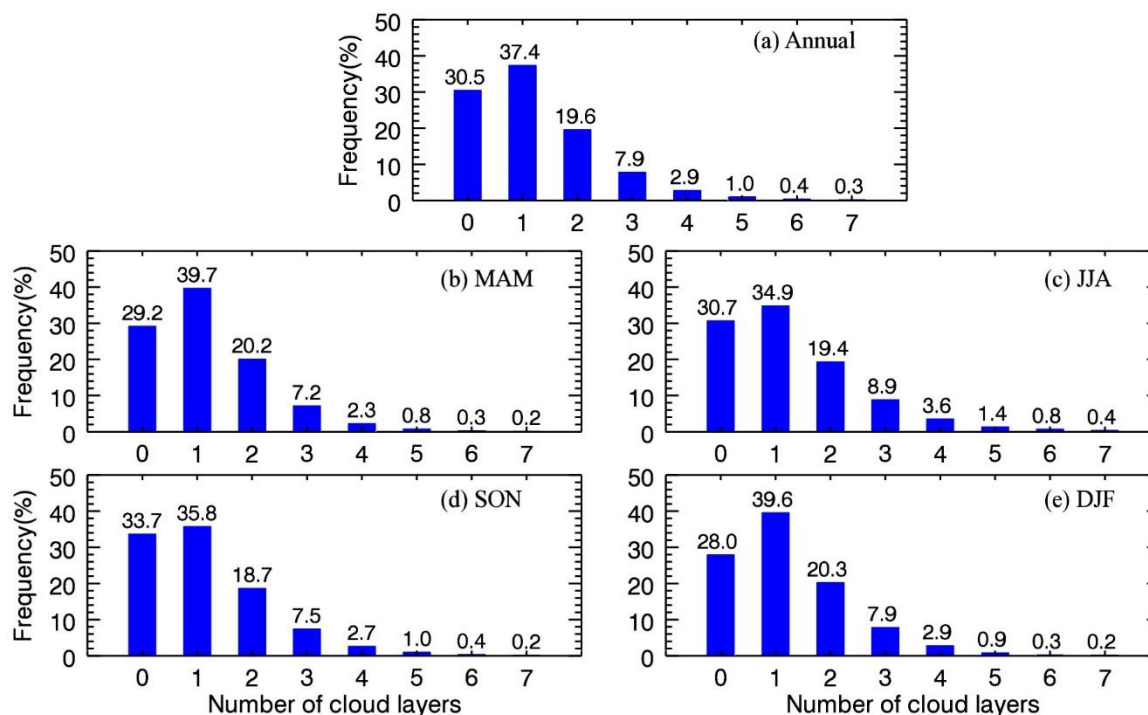



Figure 6. Near-global mean occurrence frequencies of clouds with a variety of number of layers ranging from 0 to 7 as detected by high-resolution radiosonde measurements at 1200 UTC during the period of 2018–2019: (a) annual, (b) March–April–May (MAM), (c) June–July–August (JJA), (d) September–October–November (SON), and (e) December–January–February (DJF). Also marked is the probability for the specified cloud type at the top of each bar.

890

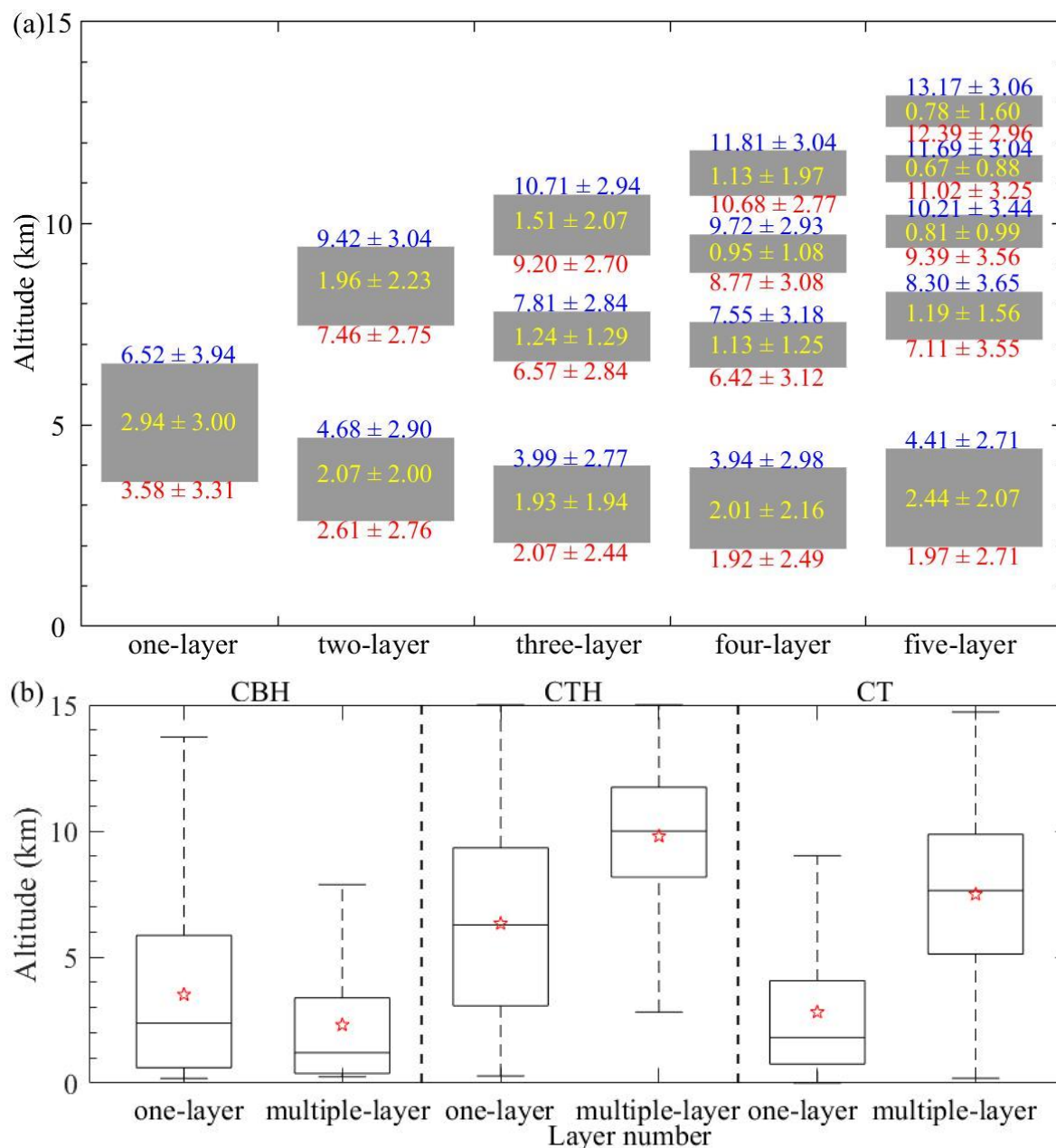



Figure 7. Near-global annual mean (a) vertical locations of one-, two-, three-, four-, and five-layer clouds and (b) boxplot of CVS (CBH, CTH, and CT) for one- and multi-layer clouds at 1200 UTC during the period of 2018–2019. The mean ± one standard deviation values of CBH, CTH, and total CT for each cloud type are also marked in (a).

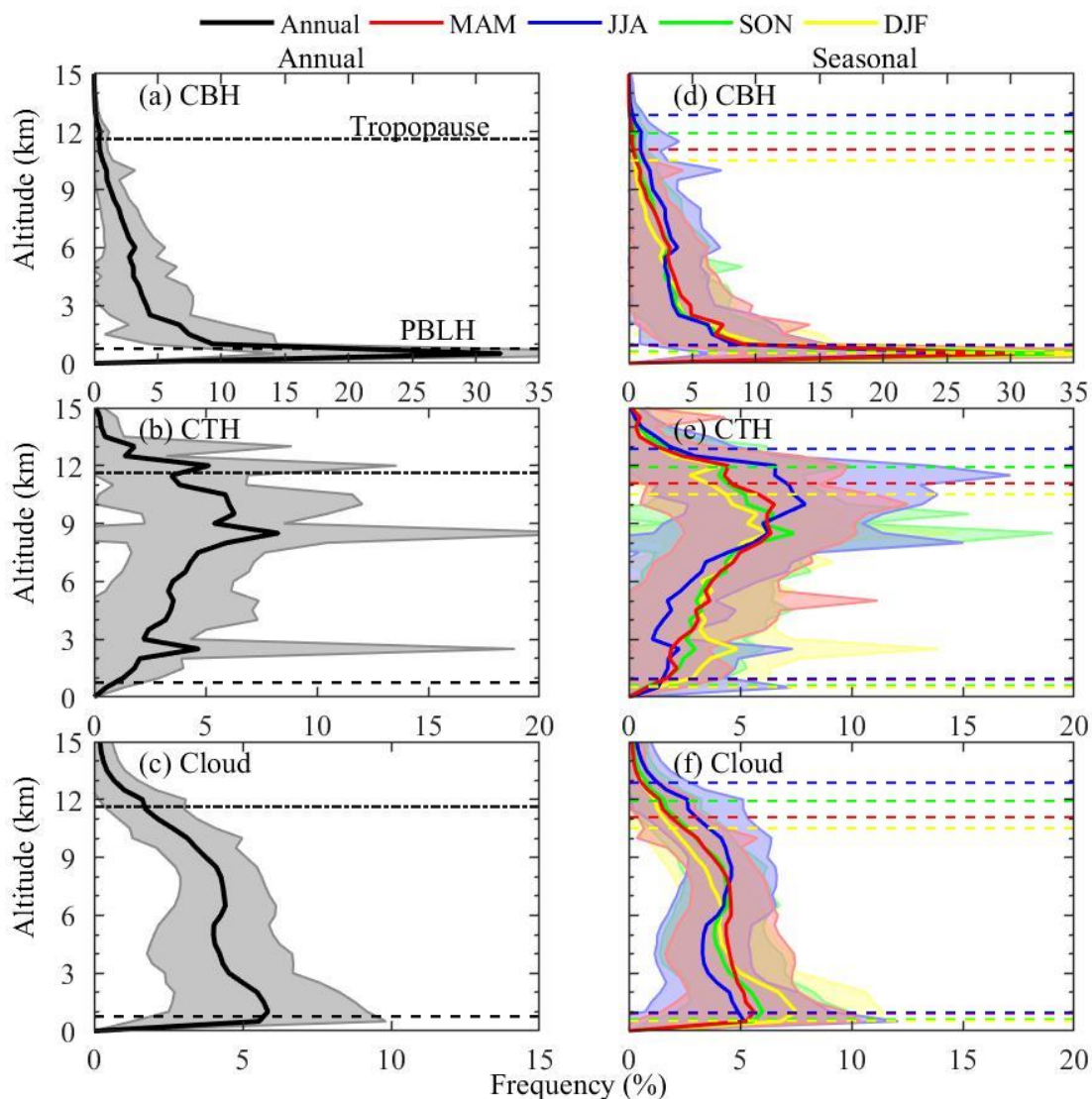



Figure 8. Near-global mean vertical distributions of (a-c) annual and (d-f) seasonal occurrence frequencies of CBHs, CTHs, and clouds as detected by radiosonde data at 1200 UTC during the period of 2018–2019, respectively. The annual, MAM, JJA, SON, and DJF are marked in black, red, blue, green, and yellow, respectively. Samples are vertically divided with a resolution of 500 m. The percentage for a given altitude is defined as the ratio of cloudy samples on that altitude to all cloudy samples. The solid lines are the mean values and shadows are the one standard deviation at annual or a given season. The planetary boundary layer height (PBLH) is determined with the method proposed by Vogelezang and Holtslag (1996), marked in dot-hyphen, and the tropopause is defined with the method from WMO (1957), marked in hyphen. The determination of PBLH and tropopause are detailed in the Supplementary Information.

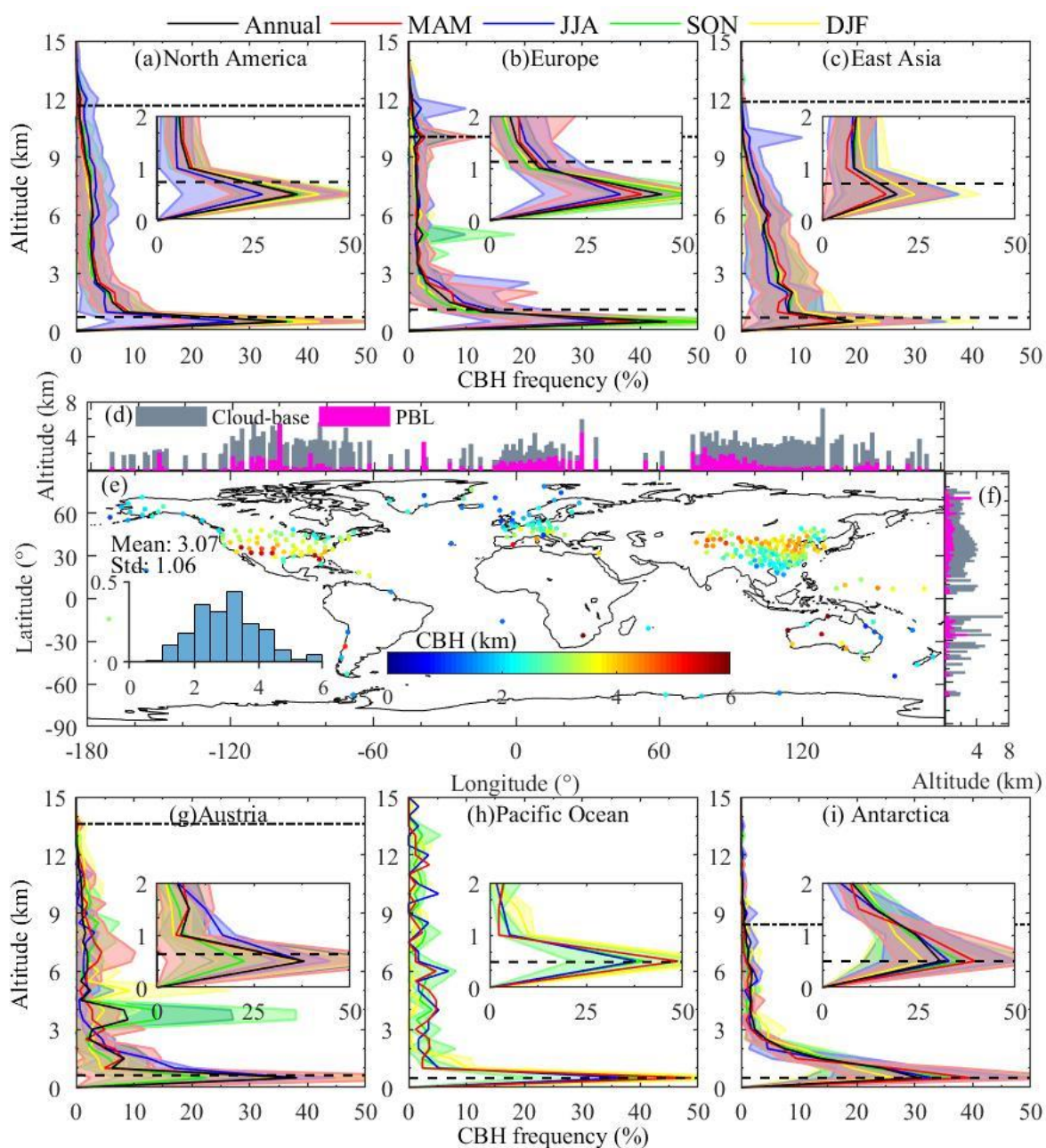



Figure 9. Regional mean vertical distributions of the occurrence frequencies of CBHs at 1200 UTC during the period of 2018–2019.
910 **The altitude resolved annual and seasonal averaged occurrence frequencies of CBHs are displayed in (a-c, g-i) over six regions of interest, including North America, Europe, East Asia, Austria, Pacific Ocean, Antarctica. Also shown are the near-global geographic distribution of the annual mean CBH (e), with the histogram of the probability distribution for CBH in the inset and the corresponding meridional (d) and zonal (f) means overlaid with the mean PBLH.**

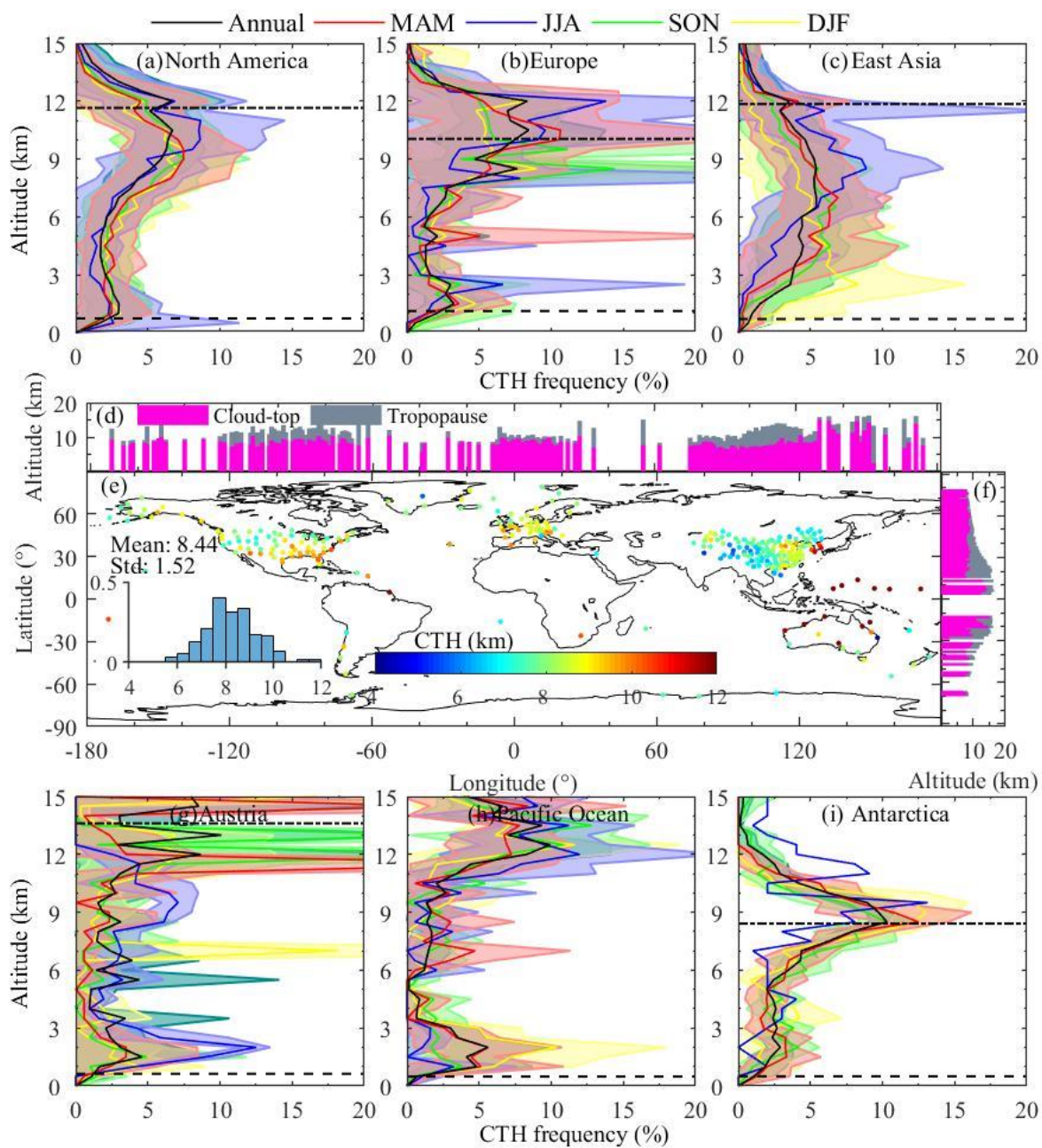



Figure 10. Similar as Figure 9, but for the occurrence frequencies of CTHs at 1200 UTC during the period of 2018–2019.

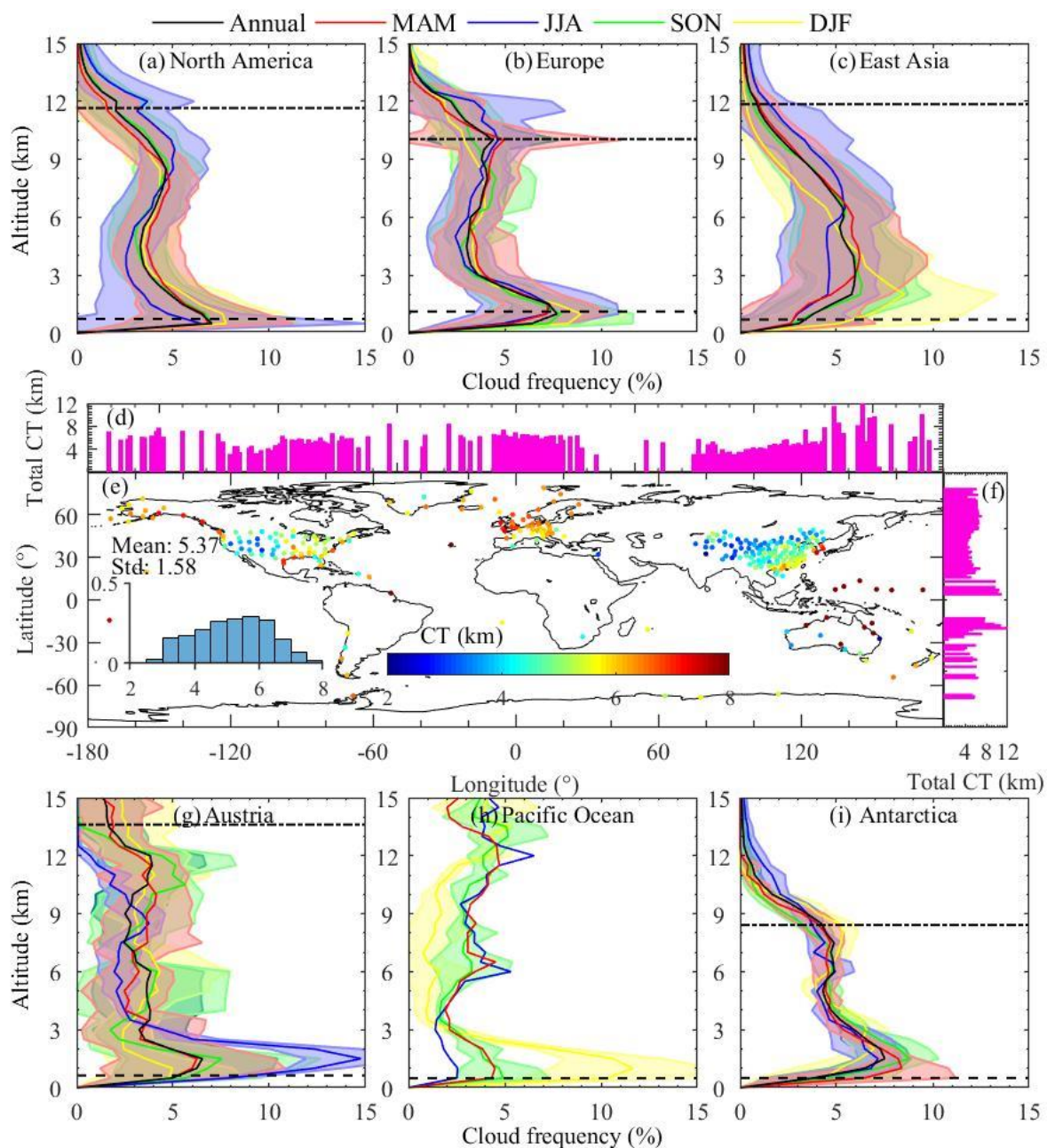
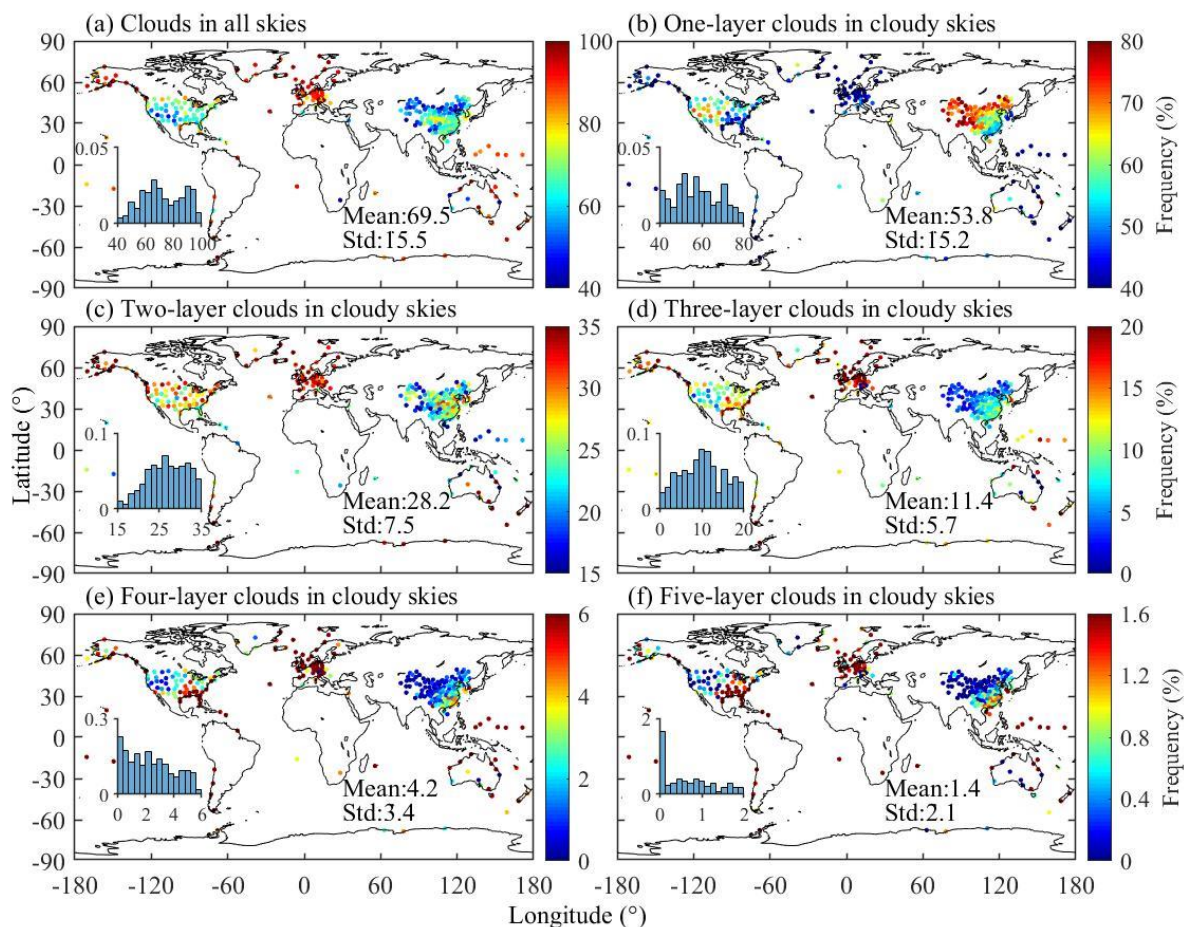



Figure 11. Similar as Figure 9, but for the occurrence frequencies of clouds at 1200 UTC during the period of 2018–2019.

920 **Figure 12. The geographic distributions of the occurrence frequencies of (a) clouds in all skies and (b–f) one-, two-, three, four-, and five-layer clouds in cloudy skies at 1200 UTC during the period of 2018–2019. It should be noted that the range of the color bar differ a lot in order to improve the visual interpretation. Also shown are the histograms of probability distributions for the cloud occurrence frequencies in each panel.**

925

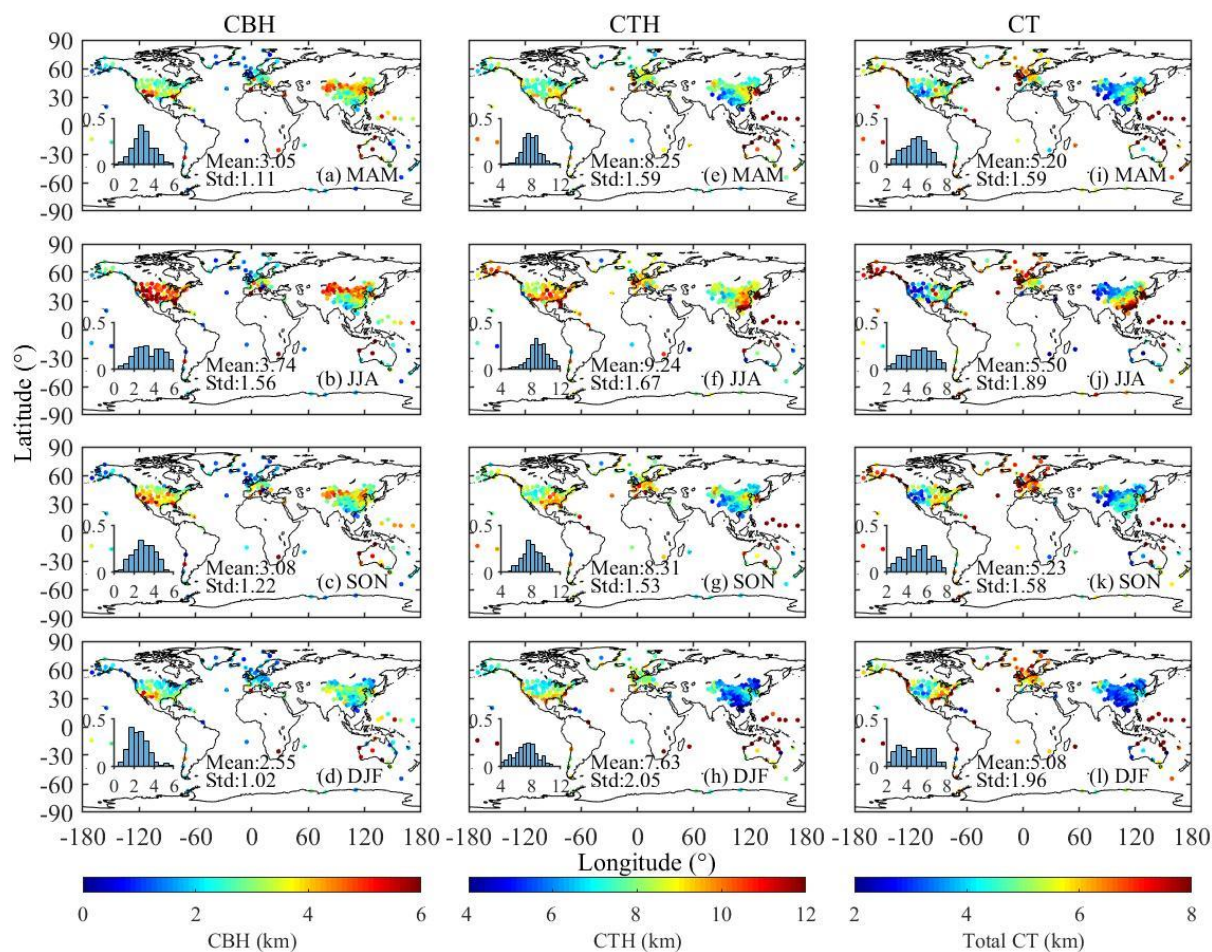



Figure 13. The geographic distributions of the seasonal mean CBH (a-d), CTH (e-h), and CT (i-l) at 1200 UTC during the period of 2018–2019. Also shown are the histograms of probability distributions for the CVS in each panel.

930

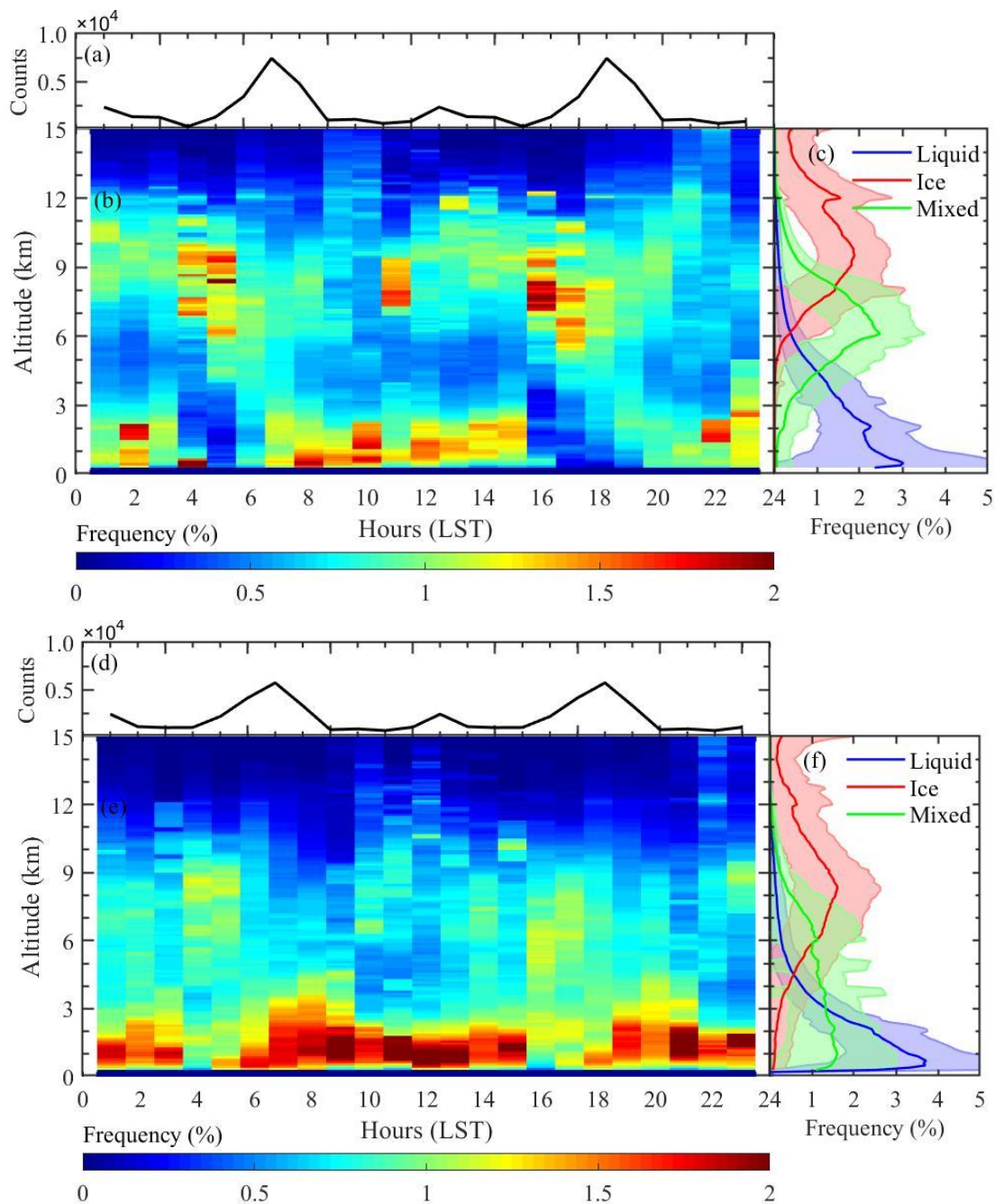



Figure 14. Diurnal variations of all clouds and altitude-resolved clouds as detected by all radiosondes at 0000 and 1200 UTC in boreal summer (JJA, a) and b)) and boreal winter (DJF, d) and e)) during the period of 2018–2019. Also shown are the vertical probability distributions of liquid (in blue), ice (in red), and liquid-ice mixed (in green) clouds in boreal summer and boreal winter (c, f).

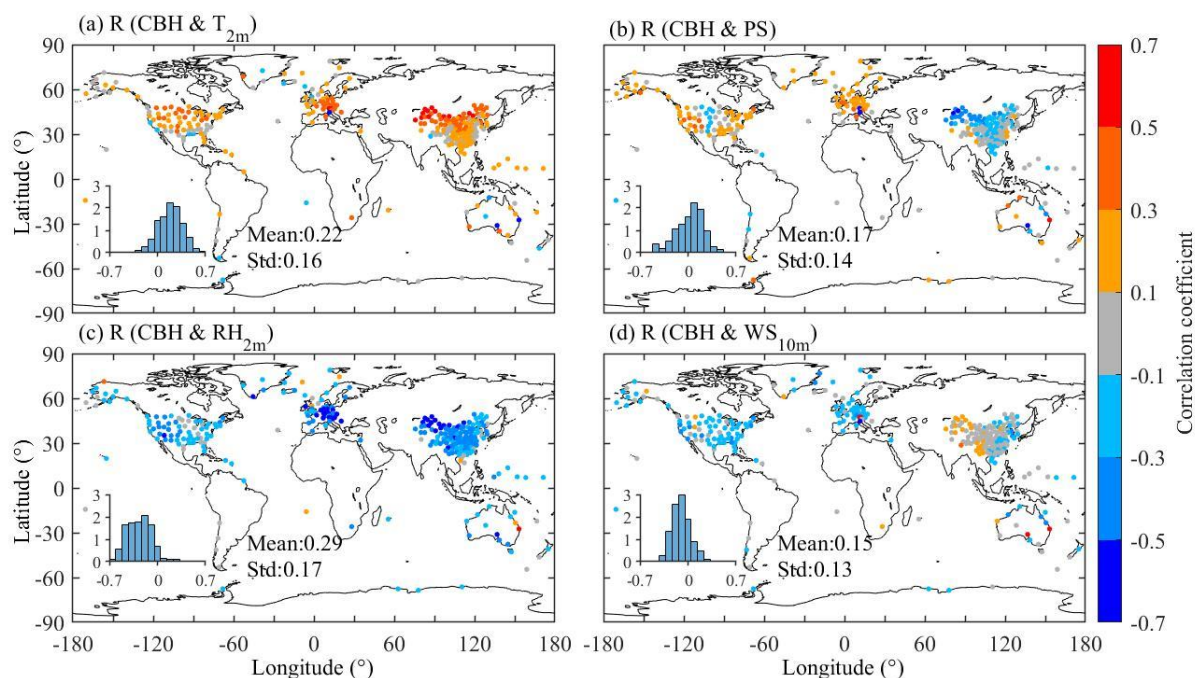



Figure 15. Geographic distributions of the correlation coefficients (Rs) between radiosonde derived CBH and surface meteorological variables: (a) 2m air temperature (T_{2m}), (b) surface pressure (PS), (c) 2m relatively humidity (RH_{2m}), and (d)10m wind speed (WS_{10m}) at 1200 UTC during the period of 2018–2019. Also shown are the histograms of probability distributions for their corresponding R values in each panel.

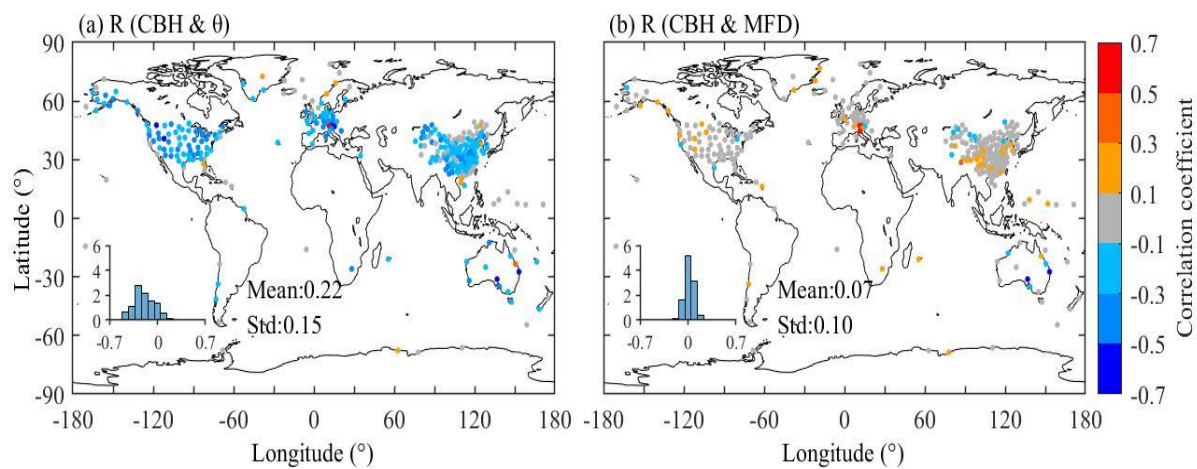



Figure 16. The same as Figure 15, but for the correlations between CBH and (a) soil water content (θ) and (b) moist flux divergence (MFD) at 1200 UTC during the period of 2018–2019.

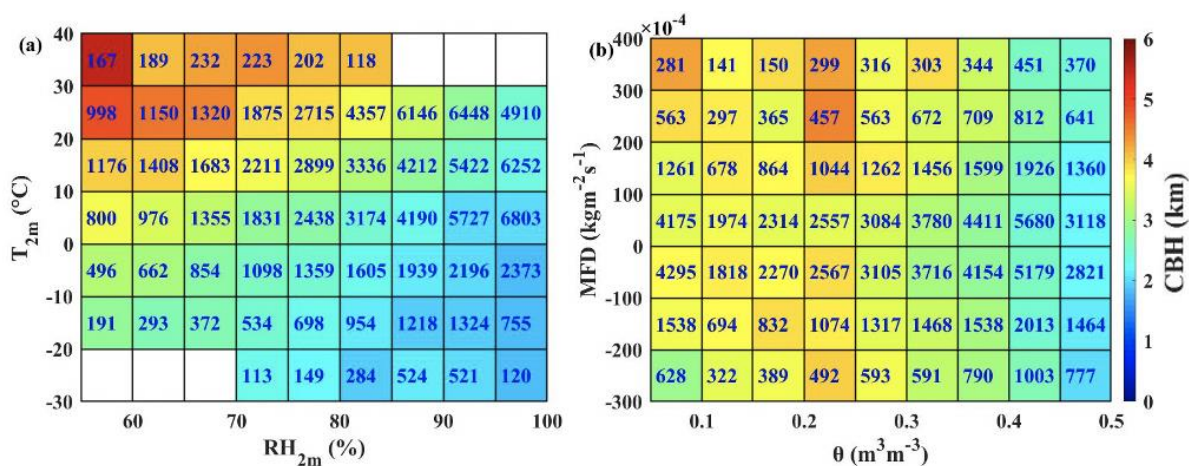
945

950

Figure 17. Joint dependences of CBH on (a) T_{2m} and RH_{2m}, (b) θ and MFD at 1200 UTC during the period of 2018–2019. The number labelled in each cell represents its corresponding sample size.