Investigation of the phase distribution and related parameters in mixed-phase clouds using satellite observations and model simulations

Zur Erlangung des akademischen Grades einer
DOKTORIN DER NATURWISSENSCHAFTEN (Dr. rer. nat.)
von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M. Sc. Olimpia Bruno
aus Messina

Tag der mündlichen Prüfung:	27.10.2023
Referentin:	Prof. Dr. Corinna Hoose
Korreferent:	Prof. Dr. Jan Cermak

To Demetrio, my husband and best supporter,
and to Hypatia, the first woman scientist.

Abstract

One of the largest uncertainties in numerical weather prediction and climate models is the representation of mixed-phase clouds. In mixed-phase clouds, hydrometeors consisting of ice and supercooled liquid water (i.e., water below 0°C) can coexist. In the mixed-phase temperature range (from −40°C to 0°C), ice-nucleating particles (INPs, e.g., mineral dusts, biological aerosol particles) are required for glaciation to be possible. The partitioning into liquid and ice depends not only on the INPs, but also, for example, on cloud dynamics and ice multiplication processes, which in turn influence the lifetime and precipitation type of these clouds and the Earth-atmosphere energy balance at local and global scales. In this thesis, the ice and liquid partitioning in mixed-phase clouds is analysed: first, using and comparing 4 years of collocated and non-collocated observations from both active and passive satellite sensors; then, using and comparing observations from passive satellite sensors, global climate models and global storm-resolving models; finally, to tune global climate models for a more realistic representation of mixed-phase clouds. The thermodynamic phase partitioning of clouds, here represented by the supercooled liquid fraction (SLF), is analysed for different cloud types between 60° North and South, and also stratified by latitude and continental and maritime regions. Further investigations based on the geographical distribution of mixed-phase clouds and the liquid effective radius (r_{liq}) have been carried out to better understand the macro- and microphysical processes in mixed-phase clouds. Despite a phase and temperature mismatch between passive and active satellite retrievals, the datasets qualitatively agree on hemispheric, land-ocean and cloud-type dependencies, e.g., showing larger SLF in the Southern Hemisphere (SH) than in the Northern Hemisphere (NH), probably due to more INPs in the NH than in the SH, except for low-level continental mixed-phase clouds, which show the opposite result. Furthermore, the agreements found in this comparison highlight the potential of passive satellite retrievals to contribute to cloud-phase research. Comparing passive satellite retrievals, global climate models and global storm-resolving models, despite disagreements on the amount of mixed-phase clouds, the datasets mostly agree on their geographical location, mainly in mid-latitudes. All datasets show that SLF decreases with cloud top temperature, but models often fail to reproduce the hemispheric, land-ocean and cloud type dependencies of the observations: only one (out of two) of the observational and two (out of four) of the model datasets reproduce the previous result of greater SLF in the SH than in the NH, with the exception of low-level continental mixed-phase clouds. The observational datasets show that low-level mixed-phase clouds have larger $\overline{r_{liq}}$ over the ocean than over land, which can be related to the Twomey effect, but also that large values of $\overline{r_{liq}}$ are found in mountainous regions, which are dominant between 30° and 60° South. Orographic updrafts may explain both the presence of larger $\overline{r_{liq}}$ in low-level continental mixed-phase

clouds between 30° and 60° South and the lower SLF found for these clouds compared to the other regions. In addition, observations show an increase in $\overline{r_{liq}}$ with decreasing SLF, possibly due to the collision of droplets and their subsequent aggregation into larger droplets as they ascend in the cloud. Some cloud types show a decrease in the SLF-$\overline{r_{liq}}$ bin frequency with decreasing SLF, which may be caused by droplet glaciation due to collision with other droplets or with ice crystals in the mixed phase temperature range. Using SLF calculated from observations at the top of and inside mixed-phase clouds as a reference, global climate models were tuned by modifying the efficiency of the Wegener-Bergeron-Findeisen process, the INPs concentration, and the glaciation temperature of cloud particles detrained from convective updrafts. After model tuning, the simulated SLF at the top and in the inside mixed-phase clouds were in good agreement with the observations.

Zusammenfassung

Eine der größten Unsicherheiten in numerischen Wettervorhersagen und Klimamodellen ist die Darstellung von Mischphasenwolken. In Mischphasenwolken können Hydrometeore, die aus Eis und unterkühltem flüssigen Wasser (d. h. Wasser unter 0°C) bestehen, koexistieren. Im Mischphasentemperaturbereich (von −40°C bis 0°C) sind eisbildende Partikel (engl. ice nucleating particles, INPs, z. B. Mineralstäube, biologische Aerosolpartikel) notwendig, damit eine Vereisung möglich ist. Die Aufteilung in Flüssigkeit und Eis hängt nicht nur von den INPs ab, sondern zum Beispiel auch von der Wolkendynamik und den Eisvermehrungsprozessen, die wiederum die Lebensdauer und Niederschlagsart dieser Wolken und die Energiebilanz zwischen Erde und Atmosphäre auf lokaler und globaler Ebene beeinflussen. In dieser Arbeit wird die Eis- und Flüssigkeitsverteilung in Mischphasenwolken analysiert: zunächst im Vergleich unter Verwendung von 4 Jahren kollokierter und nicht kollokierter Beobachtungen von aktiven und passiven Satellitensensoren; dann im Vergleich unter Verwendung von Beobachtungen passiver Satellitensensoren, globaler Klimamodelle und globaler sturmauflösender Modelle; schließlich zur Anpassung globaler Klimamodelle für eine realistischere Darstellung von Mischphasenwolken. Die thermodynamische Phasenaufteilung von Wolken, hier dargestellt durch den Anteil der unterkühlten Flüssigkeit (engl. supercooled liquid fraction, SLF), wird für verschiedene Wolkentypen zwischen 60° Nord und Süd analysiert und auch nach Breitengrad sowie kontinentalen und maritimen Regionen unterteilt. Weitere Untersuchungen auf der Grundlage der geografischen Verteilung von Mischphasenwolken und des effektiven Radius von Wolkentröpfchen (eng. liquid effective radius, r_{liq}) wurden durchgeführt, um die makro- und mikrophysikalischen Prozesse in Mischphasenwolken besser zu verstehen. Trotz eines Phasen- und Temperaturunterschiedes zwischen passiven und aktiven Satellitenretrievals stimmen die Datensätze qualitativ in Bezug auf hemisphärische, Land-Ozean- und Wolkentypen-Abhängigkeiten überein, z. B. zeigen sie einen größeren SLF in der SH als in der NH, was wahrscheinlich auf mehr INPs in der Nordhalbkugel (engl. Northern Hemisphere, NH) als in der Südhalbkugel (eng. Southern Hemisphere, SH) zurückzuführen ist, mit Ausnahme von kontinentalen Mischphasenwolken in niedriger Höhe, die das gegenteilige Ergebnis zeigen. Darüber hinaus unterstreichen die in diesem Vergleich gefundenen Übereinstimmungen das Potenzial passiver Satellitenmessungen, einen Beitrag zur Erforschung der Wolkenphasen zu leisten. Beim Vergleich passiver Satellitenmessungen, globaler Klimamodelle und globaler sturmauflösender Modelle stimmen die Datensätze trotz Unstimmigkeiten über den Anteil der Mischphasenwolken größtenteils über ihre geografische Lage überein, hauptsächlich in den mittleren Breiten. Alle Datensätze zeigen, dass die SLF mit der Temperatur der Wolkenobergrenze abnimmt, aber die Modelle können die hemisphärischen, Land-Ozean- und Wolkenarten-Abhängigkeiten

der Beobachtungen oft nicht reproduzieren: nur einer (von zwei) der Beobachtungsdatensätze und zwei (von vier) der Modelldatensätze reproduzieren das vorherige Ergebnis einer größeren SLF in der SH als in der NH, mit Ausnahme der kontinentalen Mischphasenwolken in niedriger Höhe. Die Beobachtungsdatensätze zeigen, dass tiefliegende Mischphasenwolken über dem Ozean größere $\overline{r_{liq}}$ aufweisen als über dem Land, was mit dem Twomey-Effekt zusammenhängen kann, aber auch, dass große Werte von $\overline{r_{liq}}$ in Gebirgsregionen zu finden sind, die zwischen 30° und 60° Süd dominieren. Orographische Aufwinde können sowohl das Vorhandensein größerer $\overline{r_{liq}}$ in kontinentalen Mischphasenwolken in niedriger Höhe zwischen 30° und 60° Süd als auch die für diese Wolken im Vergleich zu den anderen Regionen festgestellte niedrigere SLF erklären. Darüber hinaus zeigen die Beobachtungen einen Anstieg von $\overline{r_{liq}}$ mit abnehmender SLF, was möglicherweise auf die Kollision von Tröpfchen und ihre anschließende Aggregation zu größeren Tröpfchen beim Aufstieg in der Wolke zurückzuführen ist. Bei einigen Wolkentypen nimmt die Häufigkeit der SLF-$\overline{r_{liq}}$-Bins mit abnehmendem SLF ab, was durch Tröpfchenvereisung aufgrund von Kollisionen mit anderen Tröpfchen oder mit Eiskristallen im Temperaturbereich der Mischphase verursacht werden kann. Anhand der aus Beobachtungen berechneten SLF an der Oberseite und im Inneren von Mischphasenwolken wurden globale Klimamodelle angepasst, indem die Effizienz des Wegener-Bergeron-Findeisen-Prozesses, die INPs-Konzentration und die Vereisungstemperatur von Wolkenpartikeln, die von konvektiven Aufwinden abgezogen werden, verändert wurden. Nach der Modellabstimmung stimmten die simulierten SLF an der Oberseite und im Inneren der Mischphasenwolken gut mit den Beobachtungen überein.

Contents

1. Introduction

Since ancient times, human beings have shown interest in meteorology: in some of the oldest references, such as the Indian philosophical writings Upanishads (e.g., Sivananda, 1985, Krishnananda, 1984) dating back to 3000 B.C., discussions about the formation of clouds and rain and the existence of seasonal cycles as a result of the Earth's rotation and revolution are included. This is a clear indication that the weather influenced life, just as it does today: agriculture, disease, social life and, more generally, all natural ecosystems were and are strictly dependent on the weather.

Over time, meteorology has been subject to continuous improvement. In particular, over the last century, scientists have been able to represent the climate system through global models, progressing from simple and uncoupled models which simulated the atmosphere, the land surface, the ocean, and the sea ice, to more complex systems where aerosols, the carbon cycle, the dynamic of vegetation, the atmospheric chemistry, and the land ice have also been coupled to the previous systems (Cubasch et al., 2013). Despite these advances, the modelled radiative forcing (i.e., the difference between the incoming and outgoing radiation from the Earth) still contains many uncertainties related to aerosols and their influence on clouds (Myhre et al., 2013; Boucher, 2013). Aerosols can affect the Earth's radiation budget through several processes: for example, by acting as cloud condensation nuclei or ice nucleating particles, aerosols can affect the cloud phase and its distribution within clouds, they can influence the way radiation is scattered, and they can affect the number and size of droplets in clouds.

Some clouds are composed exclusively of liquid droplets: they are called *warm clouds* and their temperature is above 0°C; some clouds can contain ice crystals: they are called *cold clouds* and their temperature is below 0°C. There is a category of clouds belonging to the cold clouds in which both liquid and ice phases can coexist: the *mixed-phase clouds*, where the presence of both thermodynamic phases can occur for temperatures between about –38°C and 0°C (Pruppacher and Klett, 1997). The different scenarios in which the coexistence of liquid droplets and ice crystals can develop have been summarised by Korolev (2007): Depending on the updraft and downdraft velocity of the liquid and ice cloud particles, (1) both could grow as long as water vapour is available (Korolev and Mazin, 2003), (2) the ice crystals could grow at the expense of the liquid droplets through the Wegener-Bergeron-Findeisen (WBF; Wegener, 1926; Bergeron, 1928; Findeisen, 1931) process, or (3) they could both evaporate (Korolev and Mazin, 2003) (Chapter 2).

Datasets derived from active and passive satellite sensors and from global climate models (GCMs) and global storm-resolving models are used in this thesis to explore the different scenarios of mixed-phase clouds (Chapter 3).

Several studies have shown that mixed-phase clouds occur regardless of season, can be found in different locations, and can be associated with different cloud types (Korolev et al., 2017). Observations of mixed-phase clouds include active (e.g., Zhang et al., 2010; Tan et al., 2014; Cesana and Storelvmo, 2017) and passive satellite (e.g, Coopman et al., 2019; Noh et al., 2019; Tan and Storelvmo, 2019), airborne in situ (e.g., Korolev, 2008; Costa et al., 2017; Barrett et al., 2020), ground-based (e.g., Henneberger et al., 2013; Yu et al., 2014; Gierens et al., 2020) and airborne remote sensing measurements (e.g., Wang et al., 2012; Plummer et al., 2014).

In particular, in Tan et al. (2014), mixed-phase clouds were statistically investigated in terms of the supercooled cloud fraction (SLF), defined as the ratio of the frequency of supercooled liquid pixels to the total frequency of supercooled liquid and ice pixels in the cloud within 2° latitude by 5° longitude grid-boxes, at several isotherms between -10°C and -30°C, distinguishing cases in the Northern Hemisphere (NH) and Southern Hemisphere (SH), as well as cases over ocean and over land. The study included about five years of data from NASA's spaceborne lidar, CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) Level 2 Vertical Feature Mask (VFM) versions 3.01 and 3.02, and determined the relationship between the cloud phase and several aerosol types. They found that dust aerosols might strongly influence the supercooled cloud fraction by acting as ice nucleating particles (INPs), illustrating how important the atmospheric aerosol composition can be for the cloud phase. In addition, a larger supercooled cloud fraction was found in the SH than in the NH, which may be caused by the presence of more land in the NH, where efficient INPs originate. This result may also explain why a larger supercooled cloud fraction was found over the ocean than over land.

In the first part of the present work, a statistical approach as in Tan et al. (2014) is used to quantify the phase distribution of mixed-phase clouds using SLF at cloud top on isotherms. In addition, the SLF distribution is analysed on a near-global scale between 60° North and South and stratified by latitude and continental/maritime regions. The International Satellite Cloud Climatology Project (ISCCP) cloud classification (Rossow and Schiffer, 1999) is used to distinguish different cloud types according to their cloud optical thickness and cloud top pressure in nine categories: cirrus, cirrostratus, deep convective, altocumulus, altostratus, nimbostratus, cumulus, stratocumulus, and stratus. This study includes collocated and non-collocated data from passive (Advanced Very High Resolution Radiometer — AVHRR) and active (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation — CALIPSO) polar-orbiting satellite sensors, with the intention of identifying robust signals despite differences, facilitating the potential identification of common features based on different sources and algorithms. As passive sensors offer the advantage of long-period records with near-global daily coverage, three AVHRR-based datasets are compared with the CALIPSO-based dataset. This part of the work is presented as a validation study and is described in Chapter 4.

As mixed-phase clouds are not yet fully understood, and consequently weather and climate models have difficulty representing them (Forbes and Ahlgrimm, 2014; Cesana et al., 2015; McCoy et al., 2016), a comparison between observational and model datasets would provide an opportunity to identify which

aspects are well simulated by models and which need improvement. In Cesana et al. (2015), 16 general circulation models were used to characterise and evaluate the vertical partitioning of the cloud phase and compared with observations from the GCM-Oriented CALIPSO Cloud Product (CALIPSO-GOCCP). They show that the global frequency of occurrence of the phase ratio (PR), defined as the ratio of ice clouds to total clouds (ice + liquid), reaches 90% for 13 out of 16 models at temperatures too warm and altitudes too low compared to the observations, probably related to the scarcity of supercooled liquid water compared to ice at high altitudes. They also show how important it is for the models to include the WBF process to reproduce the observations, and that models with more complex microphysics fit the observations better. The dataset comparison includes 2D histograms plotting the frequency of occurrence of PR against temperature.

In the second part of the present work, observations and models are compared using a similar approach to Cesana et al. (2015). In particular, 2D histograms plotting the frequency of occurrence of SLF at the cloud top against the cloud top temperature (SLF-CTT joint histograms) are used to compare the liquid-to-ice phase transition of passive satellite sensors (AVHRR and Moderate Resolution Imaging Spectroradiometer — MODIS), GCMs and global storm-resolving models. With the aim of finding out how the physics behind cloud formation and the environment in which clouds develop are related to the thermodynamic phase, this part of the work includes the geographical distribution of all clouds as well as of mixed-phase clouds. As the second part of this thesis includes further investigations related to the first part (Chapter 4), the ISCCP cloud classification (Rossow and Schiffer, 1999) is also used here to distinguish nine cloud types and, in addition, the SLF-CTT histograms are normalised to isotherms (CFAD (Contoured frequency by altitude diagram)-like histograms) for better comparison with the SLF on isotherms used in the first part of this thesis. This part of the work is described in Chapter 5.

The thermodynamic phase in clouds can depend on several factors: e.g., temperature, pressure, cloud dynamics (e.g., strength of updrafts and downdrafts, entrainment/detrainment), type and number concentration of ice nucleating particles, cloud microphysics (e.g., supersaturation, type of nucleation, size, shape, growth, evaporation, and fall velocity of liquid droplets and ice crystals), electricity in cloud, and orography. Cloud top information provided by satellites is limited to only a few of these factors. However, the size of liquid droplets and ice crystals is often provided and can be linked to micro- and macrophysical cloud properties. Previous observational studies include variations of cloud effective radius with the cloud (top) height (Qin et al., 2018; Ewald et al., 2019; Fu et al., 2022; Pörtge et al., 2023), the cloud (top) temperature (Yuan and Li, 2010; Martins et al., 2011; Coopman et al., 2020, 2021), the cloud water path (Jia et al., 2019; Liu et al., 2022), the cloud (top) phase (Coopman et al., 2021), aerosol-related properties (Jia et al., 2019; Liu et al., 2022), and the rain rate (Qin et al., 2018).

The third part of this work extends the investigations of the previous parts to the liquid cloud effective radius ($\overline{r_{liq}}$) for different cloud types using, for the first time, joint SLF-$\overline{r_{liq}}$ histograms. These histograms allow the identification of possible microphysical cloud processes such as the WBF process and the simultaneous growth (or evaporation) of liquid droplets and ice crystals, while also providing the size-

frequency distribution of $\overline{r_{liq}}$. By combining the information provided by the cloud and $\overline{r_{liq}}$ geographical distributions with SLF-$\overline{r_{liq}}$, SLF-CTT and CFAD-like histograms, possible interpretations of the cloud dynamics can be hypothesized. This part of the work is described in Chapter 6.
Because measurements in the polar regions are unreliable due to the presence of sea ice (King et al., 2004) and the high solar zenith angle (Grosvenor and Wood, 2014), the studies described so far in this thesis are limited to latitudes up to 60°. However, an analysis based on an annual cycle of cloud and radiation measurements from the Surface Heat Budget of the Arctic (SHEBA) program in the Beaufort and Chukchi Seas during 1997-98 in Shupe and Intrieri (2004) shows that satellite measurements of the cloud phase can be used to study the Arctic surface energy budget and the importance for models to reproduce well the cloud phase in order to simulate it realistically.
The fourth part of this work focuses on both the Arctic region (from 66° to 82° N) and the whole globe (up to 82°, i.e., the highest latitude passed by CALIPSO) and is described in Chapter 7. In this chapter, observations derived from the active satellite sensor CALIPSO are not only compared with GCMs, but also used to improve the GCMs by tuning them. This part of the work has been done by colleagues at the University of Oslo. Model tuning is a fundamental component of model improvement in the way GCMs represent the climate. Improving models is directly related to improving the representation of the physical processes built into them (Jakob, 2010), and some parameters may be more appropriate than others for tuning models (Mauritsen et al., 2012). For the first time, GCMs are tuned not only with information about cloud phase distribution inside the clouds, but also at the top of the clouds. This chapter is the result of my collaboration with the University of Oslo.
The work presented in this thesis aims to provide answers to the following questions:

1. Do mixed-phase clouds exhibit regional dependencies? And if they do, are the regional dependencies cloud-type-dependent? Can the passive satellite sensor AVHRR contribute to cloud-phase research? (Chapter 4)

2. How well do global climate and global storm-resolving models represent the amount, the geographical distribution, and the phase partitioning in mixed-phase clouds globally and regionally? (Chapter 5)

3. How well do global climate and global storm-resolving models represent the microphysics involved in mixed-phase clouds? (Chapter 6)

4. What can be done to improve models with respect to the simulation of mixed-phase clouds? (Chapter 7)

Background information on cloud macro- and microphysics, cloud optical properties, and remote sensing of clouds is provided in Chapter 2, while a description of the datasets and the methods used throughout the thesis is presented in Chapter 3. As Chapter 7 contains results obtained in collaboration with other

researchers, specifics about model setups and the methods are reported in that chapter. Finally, chapter 8 contains a summary of the findings presented in this thesis, conclusions, and the outlook.

2. Clouds and remote sensing

Around 67% of the Earth is estimated to be covered by clouds (King et al., 2013). Most of them are found in the *troposphere*, i.e., the atmospheric layer closest to the Earth's surface, which extends to about 12 km altitude ($\sim$17 km in the tropics and $\sim$8 km in the polar regions (Garcia, 2003)). The top of the troposphere, called the *tropopause*, is defined as the height at which the temperature reaches its minimum before remaining constant in the vertical in the *stratosphere* and increasing again until a maximum in temperature defines the top of the stratosphere, i.e., the *stratopause*. Other atmospheric layers follow vertically upwards. However, only the lowest part of the troposphere, which extends to about 1-3 km and is called the *planetary boundary layer* (PBL), is affected by the transfer of heat and momentum from the Earth's surface.

Clouds are not only found at different heights, depending on the processes they undergo, but also with different horizontal extents. For this reason, three scales can be defined to distinguish phenomena that extend beyond $\sim$2000 km (*synoptic scale*), phenomena ranging between 20 and 2000 km (*mesoscale*), and phenomena between 0.2 and 20 km (*convective scale*).

Depending on their shape and their horizontal and vertical extent, clouds can be organised into categories, each of which has been assigned a name. This cloud classification was originally developed by Luke Howard (Howard, 1803a,b,c) and is today used in modified form by the World Meteorological Organization. The different cloud types are illustrated in Fig. 2.1.

The contents of this chapter are taken from Lamb and Verlinde (2011) and Houze (2014) unless otherwise stated.

2.1. Cloud types

2.1.1. Low-level clouds

All the low-level clouds (cumulus, stratocumulus, stratus, and fog - not shown in Fig. 2.1) either occur within the PBL or have a strong interaction with it. They are considered "low" because their base is usually below 2 km. They can be sorted into *cumuliform* clouds (cumulus and cumulonimbus), which have strong updrafts and therefore a towering aspect, and *stratiform* clouds (fog, stratus, and stratocumulus), which are mostly horizontally extended and have little or no vertical motion.

Cumuliform clouds can be divided into precipitating and non-precipitating. *Cumulus* clouds are non-precipitating clouds with a horizontal and a vertical extent of at most one kilometre when they are isolated systems. Their shape resembles a cauliflower with well-defined edges. The temperature at the cloud base

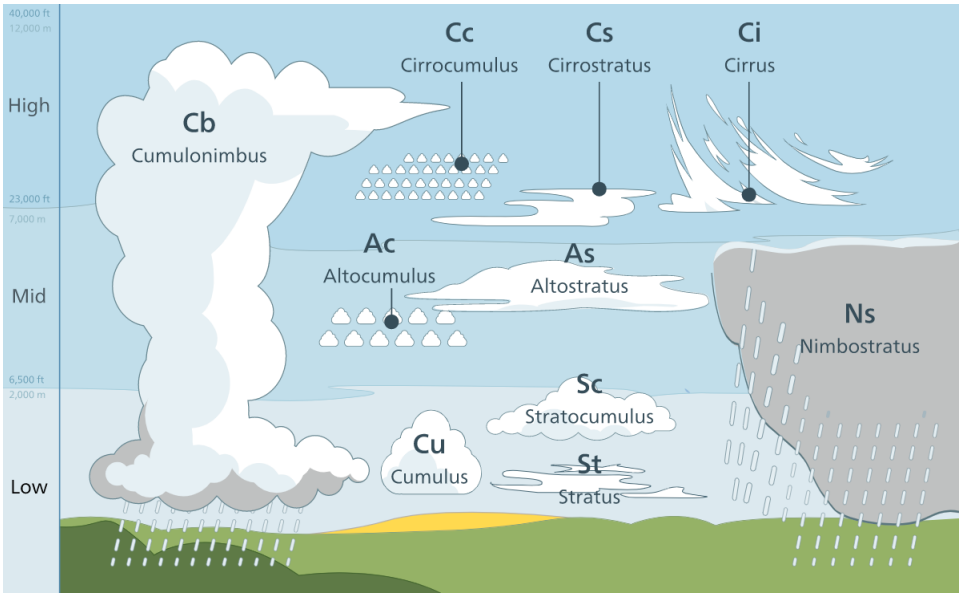

Figure 2.1.: Cloud-type classification in levels and genera. This illustration has been created for Coton, the cloud identification guide for mobile, by Valentin de Bruyn, and is licensed under CC BY-SA 3.0

can be warmer than 5-10°C, depending on region and season, while at the cloud top, it can reach –15°C in winter time (Rangno, 2003). However, if they continue to grow by convection, so-called *cumulus congestus* can form and their vertical extent can reach several kilometres. As cumulus congestus clouds grow, they can become *cumulonimbus* clouds, i.e. precipitating clouds with the shape of a high tower. In a mature phase, the upper part of a cumulonimbus is no longer characterised by well-defined edges due to the presence of ice, and is often shaped like an *anvil* or a wide plume. As these clouds can extend to the tropopause and the updrafts normally cannot penetrate the stable stratosphere, their tops appear flattened. The temperature of cumulonimbus clouds ranges from about 10°C at the base to about −65°C at the top, when the mature stage has been reached.

In stratiform clouds, the vertical motions are weaker than in cumuliform clouds. *Fog* occurs when the air at ground level meets certain conditions and forms clouds that touch the ground. Depending on the processes involved, this phenomenon is termed *steam fog*, *radiation fog* or *advection fog*. *Stratus* and *stratocumulus* clouds extend no more than 1 km vertically and can reach about 1000 km horizontally. Although they usually consist of liquid droplets, their temperature can drop to about –10°C or lower and ice can form (Rangno, 2015). Because of their horizontal extent, their complete observation from the ground is difficult, if not impossible, and satellites are the most reliable way to observe their large-scale features. Both of them rarely produce precipitation except drizzle. Stratocumulus clouds can be distinguished from stratus clouds by the presence of recognisable elements, such as a cell structure or long lines called *cloud streets*. It is common to find a unique cloud system made of both stratus and stratocumulus clouds.

2.1.2. Mid-level clouds

Mid-level clouds are characterised by a cloud base above 2 km and include only altostratus and altocumulus clouds. *Altostratus* clouds appear as a uniform layer of clouds with some inhomogeneities and transparent areas that allow solar radiation to pass through. At Aspendale, Australia, temperatures of these clouds have been observed to range from about $-20°C$ to about $-6°C$ in August (Platt, 1977). They are distinguished from stratus clouds by their higher base height and more often by the presence of liquid droplets and ice crystals. The precipitation produced by these clouds does not reach the ground. The deepening and thickening of altostratus clouds can form *nimbostratus* clouds, which have large extent not only in the horizontal but also in the vertical, potentially reaching the highest level of the troposphere as cumulonimbus clouds do. In terms of precipitation, unlike stratus, stratocumulus, and altostratus clouds, nimbostratus clouds produce significant and extensive precipitation, limiting horizontal visibility. In *altocumulus* clouds it is possible to distinguish elements. They have been observed between January and February in Capo Verde, Africa, at temperatures between $-36°C$ and $-5°C$ (Ansmann et al., 2009). Sometimes they can be similar to stratocumulus clouds and sometimes to small cumulus clouds, but with a higher base. Another form of altocumulus is represented by long rolls.

2.1.3. High-level clouds

High-level clouds include three genera of *cirriform* clouds and are found in the temperature range between $-20°C$ and $-85°C$. They consist almost completely of ice particles which evaporate slowly due to the sufficiently low saturation vapour pressure of ice. At these altitudes, the winds are strong and advect the particles far from the location where clouds formed, giving cirriform clouds their typical filamentous shape. *Cirrus* appear as isolated patches or narrow bands. *Cirrostratus* are veil-like layers that can cover part or all of the sky. *Cirrocumulus* is composed of small and regularly distributed elements.

2.2. Cloud microphysics

The formation of clouds involves a phase change from vapour to liquid or ice, and therefore the need to overcome an energy barrier before the system reaches a new equilibrium.

2.2.1. Processes in warm clouds

To form, clouds need a supersaturated environment. If the moist air reaches supersaturation (saturation ratio greater than 100%) with respect to liquid water, the water vapour will condense in small liquid droplets forming clouds. Theoretically, this process can take place in the absence of aerosols and ions (*homogeneous nucleation*) or with them (*heterogeneous nucleation*). However, homogeneous nucleation of liquid droplets is an unrealistic process because the supersaturation would have to reach several hun-

dred per cent. For this reason, cloud nucleation in nature is exclusively heterogeneous, exploiting the presence of so-called *cloud condensation nuclei* (CCNs).

2.2.1.1. Heterogeneous nucleation of liquid droplets

The presence of CCNs leads to a decrease of embryos radii, i.e., the equilibrium state between the droplet and the vapour can be reached under realistic conditions depending on the temperature and the cooling rate of the moist air, the size distribution, the concentration, and the nature of the particles (Mason 1957). To nucleate, the water embryos with radius r require energy. This energy is expressed by the Gibbs free energy ΔE:

$$\Delta E = 4\pi r^2 \sigma - \frac{4}{3}\pi r^3 n k_B T ln\left(\frac{e}{e_s}\right), \tag{2.1}$$

where σ is the surface tension of water, n is the number of water molecules in the embryo per unit volume, k_B is Boltzmann's constant, T is the temperature, e is the vapour pressure, and e_s is the saturation vapour pressure of air at the same temperature. The stability of the formed embryo is reached at a critical radius r_d, which can be obtained from the eq. 2.1 under the condition $\partial(\Delta E)/\partial r = 0$:

$$r_d = \frac{2\sigma}{n k_B T ln(S)}, \tag{2.2}$$

with saturation ratio $S = \frac{e}{e_s}$.

2.2.1.2. CCN activation

Increasing the ambient humidity, nucleated embryos grow and can be activated as CCN if a critical radius r_c is reached. The relationship between the ambient saturation ratio S and the embryo's radius r_d is described by the *Köhler equation*:

$$S - 1 \approx \frac{a}{r_d} - \frac{b}{r_d^3}, \tag{2.3}$$

where the term $\frac{a}{r_d}$ (*Kelvin term*) takes into account the curvature of the droplet surface and the term $\frac{b}{r_d^3}$ (*Raoult term*) depends on the chemical composition of the solute. The graphical dependence of S on r is shown in Fig. 2.2.

An increase in S due to the addition of water vapour in the environment or by cooling causes the particle to grow until S reaches a critical value S_c and the particle size reaches a critical radius r_c (eq. 2.4). From that point on, the particle can grow without the need to further increase S.

$$r_c = \sqrt{\frac{3b}{a}} \tag{2.4}$$

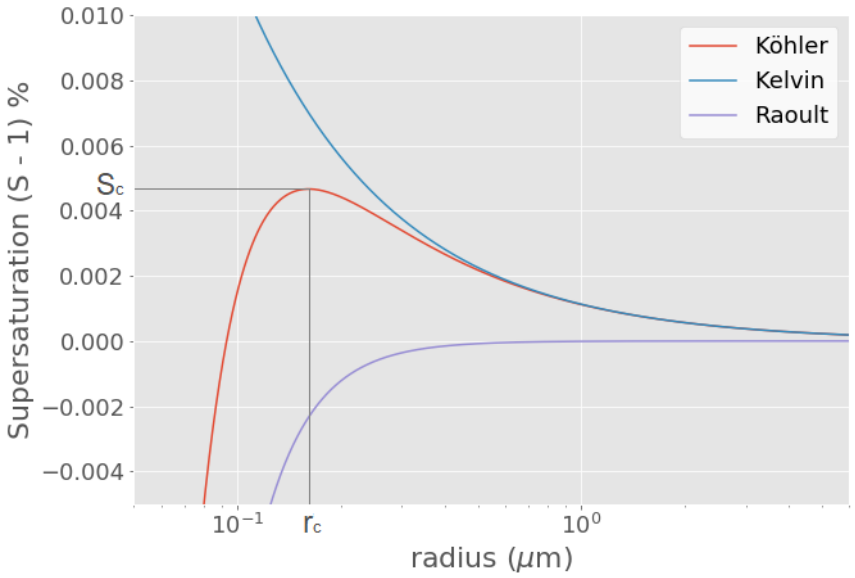

Figure 2.2.: Köhler curve for ammonium bisulphate ($(NH_4)HSO_4$) with dry radius of 0.02 µm, acting as CCN, with the contributions from Kelvin and Raoult terms.

2.2.1.3. Growth of liquid droplets

The droplets can grow by deposition or collision and coalescence.

Growth by deposition, called *condensational growth*, occurs when the supersaturation of the environment in which the droplet has formed exceeds S(r) (Fig. 2.2). If the supersaturation of the environment is less than S(r), the reverse process, *evaporation*, occurs. The condensational growth of a liquid droplet in terms of variation in time t of its mass m is expressed by the following equation:

$$\frac{\mathrm{d}m}{\mathrm{d}t} = 4\pi r D \left[\rho_v(\infty) - \rho_v(r)\right], \tag{2.5}$$

where D is the diffusion coefficient, ρ_v is the water vapour density, given here for a droplet at an infinite distance (representing the environment) and at the surface of the droplet with radius r. During its growth, the droplet releases latent heat. In the case of a saturated environment, the Clausius-Clapeiron equation is used:

$$\frac{1}{e_s}\frac{\mathrm{d}e_s}{\mathrm{d}T} = \frac{L}{R_v T}, \tag{2.6}$$

where L is the latent heat and R_v is the gas constant of water vapour for a unit mass. After some calculations, the growth rate of a droplet in relation to the radius r is obtained:

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{1}{r}\frac{S-1}{F_k + F_d}, \tag{2.7}$$

where F_k is associated with heat conduction and F_d is associated with water vapour diffusion. From this equation it can be seen that condensational growth is faster for smaller droplets and slower for larger droplets.

The term *collision* indicates contact between droplets. Causes of collision include aerodynamic processes, gravitational processes (droplet sedimentation), and electrical forces. Considering a water droplet in a flow, smaller and slower-moving droplets will follow the streamlines around it, while larger droplets with more inertia will be more likely to collide. The collision efficiency $E(r_L, r_s)$ depends on the dimension of both the collected (assumed to be the smaller, with radius r_s) and the colliding (assumed to be the larger, with radius r_L) droplets:

$$E(r_L, r_s) = \frac{x^2}{(r_L + r_s)^2}, \tag{2.8}$$

where x is the critical distance between the centres of the drops required for the collision to occur.

Once two or more droplets collide, three possible scenarios follow: the droplets rebound due to surface tension; the droplets break into several smaller droplets; *coalescence* takes place due to the merging of two or more droplets into a larger one. In this case, if coalescence occurs with several smaller droplets, continuous collection can be considered and the mass of the larger droplet will change over time according to the equation:

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \pi (r_L + r_s)^2 E_c(r_L, r_s) [u_L - u_s] \omega_L, \tag{2.9}$$

where u_L and u_s are the velocities of the larger and the smaller droplets, respectively, $E_c(r_L, r_s) = \varepsilon E(r_L, r_s)$ is the collection efficiency, ε is the coalescence efficiency, r_s becomes the mean radius of the small collected droplets, and ω_L is the collectable liquid water concentration, assuming the cloud droplet population as a continuum:

$$\omega_L = \frac{4}{3}\pi\rho_l \int_{r_{min}}^{r_L} r^3 n(r) \mathrm{d}r, \tag{2.10}$$

with ρ_l the density of water and $r_{min} \simeq 5$ µm as the minimum size for collection. After calculations, the increase in radius with time can be derived:

$$\frac{\mathrm{d}r_L}{\mathrm{d}t} = \frac{u_L E_c LWC}{4\rho_l}, \tag{2.11}$$

where LWC is the liquid water content of the considered volume. In contrast to condensational growth, which was faster for smaller droplets, the collection of droplets is faster for larger droplets as $u_L \propto r_L$.

2.2.2. Processes in cold clouds

If the moist air is saturated with respect to ice and below 0°C, ice particles can form in clouds. This has several consequences for that system (Mason, 1957):

- ice particles grow at the expense of liquid droplets, modifying the thermodynamic balance of the cloud and enhancing the precipitation generation efficiency;
- the cloud buoyancy and its turbulent motions are increased by the additional latent heat release;
- for equivalent particles size and concentration, shortwave radiation scattering and longwave radiation emission are less efficient in ice particles than in liquid drops;
- because of the shape heterogeneity of ice particles, different particles can show different physical characteristics;
- ice particles change the electrification of clouds;
- ice particles can reach much larger sizes than liquid drops because of their more stable structure, and consequently be more massive and faster falling than liquid drops, leading to damaging hydrometeors.

As with the nucleation of liquid droplets, the nucleation of ice crystals can be homogeneous or heterogeneous. The presence of so-called *ice nucleating particles* (INPs), gives the possibility to the cloud to glaciate at temperatures higher than –38°C (but still lower than 0°C) through vapour deposition, immersion freezing and contact freezing (Pruppacher and Klett, 1997). A schematic representation of nucleation processes is shown in Fig. 2.3.

2.2.2.1. Homogeneous ice nucleation

For temperatures below 0°C, it is more likely that water vapour will condense into water droplets rather than ice crystals, since Ostwald's rule of stages (Ostwald, 1897) states that the nucleating phase is not necessarily the most thermodynamically stable, but rather the one closest in free energy to the original phase (parent phase). As an unrealistically high supersaturation is required, homogeneous deposition nucleation is not observed in nature. However, if supercooled liquid droplets are already present, homogeneous ice nucleation in liquid droplets can take place. In this case, the reorientation of the molecules to form the hydrogen-bonded crystal structure of ice makes ice nucleation a possible process, which is further favoured by decreasing temperatures and increasing the efficiency of the crystallisation process as supercooling increases. The rate of ice nucleation increases exponentially with decreasing temperature, so that at about –38°C all droplets are frozen.
To nucleate, ice embryos with radius r_i (defined as the radius of the inscribed sphere) require energy, which is expressed by the Gibbs free energy ΔE_x:

$$\Delta E_x = A_{ix}\sigma_{ix} - n_i V_i \Delta\mu_{xi} - A_i n \sigma_{ix}\frac{\sigma_{nx} - \sigma_{in}}{\sigma_{ix}}, \tag{2.12}$$

where the subscripts n, i, and x refer to the nucleus, ice, and parent phase (vapour or liquid), respectively, A_{ix} is the area of the interface between the embryo and the parent phase, σ_{ix} is the interfacial free energy

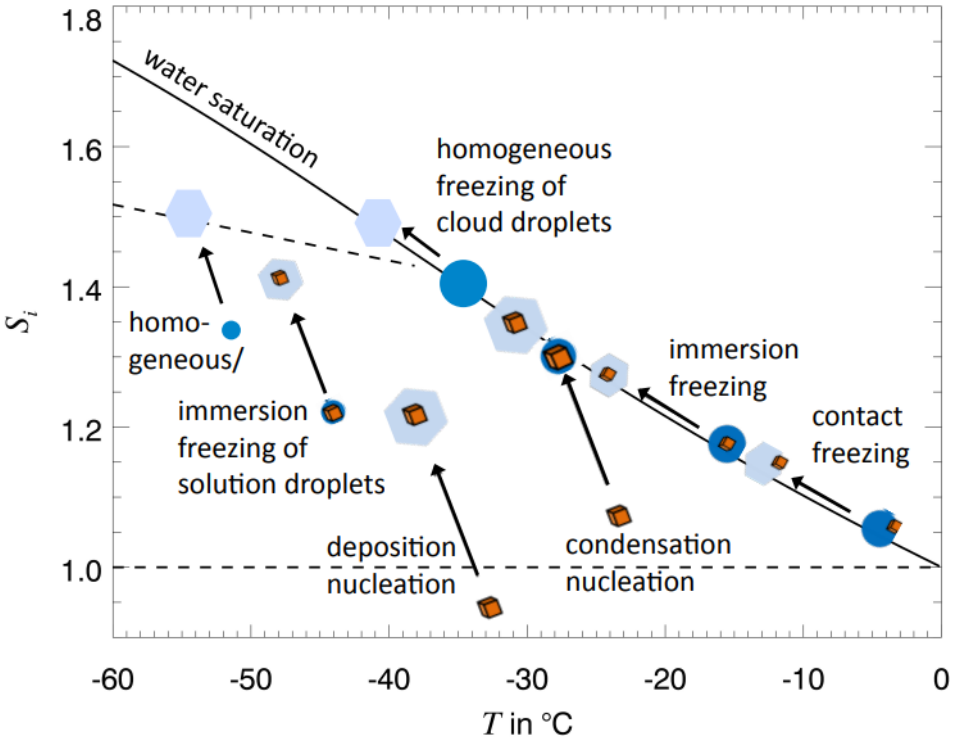

Figure 2.3.: Schematic representation of nucleation processes, from Hoose and Möhler (2012), distributed under CC BY 3.0.

between ice and parent phase, n_i is the molar density of ice, V_i is the volume of the ice embryo, $\Delta\mu_{xi}$ is the chemical potential of the parent phase relative to that of ice, A_{in} is the area of the interface between the embryo and the nucleus, σ_{nx} is the interfacial free energy between the nucleus and parent phase, and σ_{in} is the interfacial free energy between ice and the nucleus. The stability of the formed embryo is reached at a critical radius r_x, which can be obtained from the eq. 2.12 under the condition $\partial(\Delta E_x)/\partial r = 0$:

$$r_x = \frac{2\sigma_{ix}}{n_i \Delta\mu_{xi}}, \tag{2.13}$$

with n_i as the molar density of ice.

2.2.2.2. Heterogeneous ice nucleation

If INPs are present, they can lower the energy barrier that prevents ice formation. The different mechanisms leading to heterogeneous nucleation of ice (Vali et al., 2015) are:

- *deposition ice nucleation*: INP particles are dispersed in water vapour and the air is supersaturated with respect to ice but subsaturated with respect to water. In this case, the water vapour settles on the INPs and nucleates;

- *condensation freezing*: as deposition freezing, but the air is supersaturated with respect to both ice and water. For this reason, between the vapour and the ice phase, the liquid phase occurs through the formation of liquid droplets. This can be the case when updrafts transport liquid droplets until ice forms because of the reached low temperatures;

- *contact freezing*: INPs impact the surface of liquid droplets, which become ice. The molecular mechanism is still not understood (Hussain and Haji-Akbari, 2021);

- *immersion freezing*: INPs are either already inside the liquid droplets together with the CCNs that allowed the droplet condensation, or they collide with liquid droplets by Brownian diffusion, or they impact with liquid droplets falling at different velocities. In all these cases, water molecules cluster on active sites on INPs. The liquid droplets freeze if the INP size is large enough to generate an embryo exceeding the critical size to overcome the energy barrier (eq. 2.13).

2.2.2.3. Collisional processes and secondary ice formation

Turbulence and differences in fall speeds among ice particles can cause collisions. Possible consequences of collisions are *aggregation*, *rebound*, and *secondary ice production*.
The coalescence efficiency between two ice particles depends on temperature. Between –10°C and 0°C an ice particle can develop a thin quasi-liquid layer on its surface which is stickier approaching 0°C. This layer makes the bonding between two ice crystals possible and is reinforced by the deposition of water vapour around the contact point. However, between –17°C and –12°C another maximum of coalescence efficiency is observed, corresponding to the formation of ice structures with dendritic features. Since the resulting aggregated crystal will have higher coalescence efficiency, the probability of these dendritic crystals to aggregate is high.
In clouds with temperatures below 0°C, the presence of supercooled liquid droplets is also plausible as well as the collision between particles with different phases. This process is called *riming*.
The interaction between ice particles or ice and liquid particles can also produce secondary ice through several ice multiplication processes (Korolev et al., 2022). When two ice particles collide, they may break into several fragments, depending on the impact, the directional velocity, the size, and the shape of the particles. At temperatures between –7.5°C and –2.5°C, the *rime-splintering* (also known as Hallet-Mossop, Mossop and Hallett (1974)) process occurs: when graupel particles collide with droplets larger than 25µm, the thermal expansion due to the droplet glaciation can cause shattering of ice. Furthermore, a rimed droplet frozen by contact with the ice particle releases latent heat which is also transferred to the ice crystal: this thermal shock can cause fragmentation of the latter. Another ice multiplication process occurs in the presence of liquid water trapped inside a growing ice shell. The expansion of ice due to further freezing increases the pressure inside the ice shell, and the ice shell may shatter to relieve the internal pressure (Korolev and Leisner, 2020). Another source of secondary ice is the heterogeneous nucleation of liquid droplets around freezing droplets, followed by freezing. All of these ice multiplication

processes can affect cloud evolution, as each ice fragment may become an embryo for further glaciation processes.

2.2.3. Mixed-phase clouds and possible evolution scenarios

Clouds in which water in all three phases coexist are called mixed-phase clouds and can be found in the temperature range between –38°C and 0°C. In relation to the equilibrium vapour pressure over liquid (e_w), over ice (e_i), and the in-cloud vapour pressure (e), three possible scenarios describe the evolution of mixed-phase clouds (Korolev, 2007):

1. $e > e_w > e_i$: liquid droplets and ice particles compete for water vapour and both grow as long as sufficient water vapour is available. This condition may be true in mixed-phase clouds when ascending with strong updraft velocity $u > u_*$ (Korolev and Mazin, 2003):

$$u_* = \frac{e_w - e_i}{e_i} N_i \overline{r_i} \eta, \tag{2.14}$$

   where N_i and $\overline{r_i}$ are, respectively, the number concentration and the mean radius of ice particles, while $\eta = \eta(T,P)$ is a temperature- and pressure-dependent parameter.

2. $e_w > e > e_i$: ice particles grow at the expense of liquid droplets. This condition is known as *Wegener-Bergeron-Findeisen* (WBF) *process* (Wegener, 1926; Bergeron, 1928; Findeisen, 1931) and may occur when moderate updraft or downdraft velocity in the cloud is $u_0 < u < u_*$, with:

$$u_0 = \frac{e_i - e_w}{e_w} N_w \overline{r_w} \chi, \tag{2.15}$$

   where N_w and $\overline{r_w}$ are, respectively, the number concentration and the mean radius of liquid droplets, and $\chi = \chi(T,P)$;

3. $e_w > e_i > e$: both ice particles and liquid droplets evaporate. This is possible during entrainment or mixing of dry air at the cloud boundaries or in strong downdrafts when $u < u_0$ (Korolev and Mazin, 2003). Depending on N_w, N_i, r_w, and r_i, ice particles may sublimate before the total evaporation of liquid droplets, and the mixed-phase cloud would be a liquid cloud prior to its complete evaporation.

The process described in (1) is typical of convective clouds, such as cumulus, cumulonimbus, and nimbostratus clouds. In the latter two, the process described in (3) may also be involved. The WBF process described in (2) occurs in stratiform clouds, mainly in stratus and stratocumulus clouds (Korolev et al., 2017). While a mixed-phase cloud can become an ice cloud through the WBF process, ice clouds can become mixed-phase clouds if liquid droplets are activated in updrafts (Heymsfield, 1977).

2.2.4. Arctic clouds and challenges

The Arctic region consists mainly of ocean and sea ice. The following description of the processes that characterise clouds in this region can be found in Siebesma et al. (2020). Due to the high reflectivity of its surfaces (high albedo), the net energy loss of the Arctic for latitudes above 70° N is, on annual and spatial average, of ~ 100 W m^{-2}. To compensate for this energy loss, atmospheric and oceanic circulations transport energy from lower latitudes into the Arctic. This energy is advected by large-scale motions (such as synoptic-scale storms) that move warmer air in the meridional direction. Smaller contributors to this energy transport are ocean energy transport, subsidence, and latent heat release. The resultant Arctic atmospheric circulation combined with long-wave cooling produces the permanent Arctic inversion. The long-wave cooling to space takes place either at the cloud top or at the Arctic surface due to the presence of moisture; the atmospheric emissivity is lower than at lower latitudes because of the relatively dry air. For this reason, the advected warm air is first cooled at lower levels and this process generates the Arctic temperature inversion, which causes the persistence of low-level stratocumulus clouds in the Arctic. This persistent cloud layer forms as the moist, warm air transported from lower latitudes cools off and condensates because of the underlying Arctic ice and the Arctic Ocean.

Arctic stratocumulus clouds are typically mixed-phase or consist of supercooled liquid water. The cloud-top long-wave cooling causes the convective mixing that sustains the Arctic stratocumulus clouds. In particular, air parcels near the top of the cloud cool and descend because of the negative buoyancy. This mixing produces condensation which releases latent heat that partially counteracts the long-wave cooling of the mixed layer. However, a second source of moisture is needed to maintain these clouds, which is often found in the entrainment of air at the top of the cloud. In fact, the moist, warm air advected into the Arctic from lower latitudes causes an increase of the specific humidity with altitude, in particular in the lowest 1-2 km of the troposphere. The entrainment of air at the top of the cloud, therefore, acts as a source of moisture.

A second small temperature inversion is often present below the Arctic stratocumulus clouds, so that the surface is not always affected by the turbulence generated by the cloud layer. A weakly turbulent stably stratified layer below the cloud layer due to the wind shear near the surface inhibits the turbulent exchange between the cloud layer and the surface, i.e., the cloud layer is de-coupled from the surface.

Polar-orbiting satellites pass over the Arctic (and the Antarctic), but only provide measurements up to $\sim 80°$ (see Sec. 2.4.1). Another difficulty in detecting clouds is the low contrast in temperature and reflectivity when comparing clouds with sea ice or snow-covered surfaces. Arctic clouds contain less water than clouds in other regions due to the very dry air, making their detection challenging. Analysis of satellite observations is further complicated by variation between polar day and night as well as by large zenith angles during daytime. All these issues affect passive sensors such as AVHRR, despite calibrations, but it is possible to benefit from long data records. Arctic clouds are more likely to be detected by active sensors such as CALIOP but with lower temporal resolution.

2.3. Optical properties of clouds

The presence of clouds produces local and global warming and cooling effects which are directly dependent on the cloud thermodynamic phase, the concentration, the size distribution, the shape of ice crystals and liquid droplets, and the cloud thickness. Liquid and ice particles absorb and scatter the light differently: While the radiative transfer in liquid clouds mainly depends on the spectral properties of water and the droplet sizes and concentration, the geometry of ice particles, together with the ice crystal concentration uncertainties, and their size spectra, complicates the optical properties of ice clouds (Cotton and Anthes, 1992). The following sections describe the optical properties of clouds in more detail, and are taken from Siebesma et al. (2020), Petty (2004), and Jackson (1998).

2.3.1. From Maxwell's equations to optical properties of clouds

Light is an electromagnetic wave. This was postulated by James Clerk Maxwell, who published the final form of Maxwell's equations in his paper "A Dynamical Theory of the Electromagnetic Field" in 1865. Maxwell's equations describe the propagation of radiation in space and time as an oscillation of both electric and perpendicular magnetic fields propagating with the same periodicity. In matter, Maxwell's equations are:

$$\vec{\nabla} \cdot \vec{D} = \rho, \tag{2.16}$$

$$\vec{\nabla} \cdot \vec{B} = 0, \tag{2.17}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \tag{2.18}$$

$$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}, \tag{2.19}$$

where $\vec{E}$ and $\vec{H}$ are respectively the electric and the magnetic fields, $\vec{D}$ is the electric displacement, $\vec{B}$ is the magnetic induction, ρ and $\vec{J}$ are respectively the macroscopic (free) charge and the current density, and t is time. All the vector fields in this subsection are space and time-dependent (for example, $\vec{E} = \vec{E}(\vec{r},t)$). The vector fields in Maxwell's equations are related by:

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P}, \tag{2.20}$$

$$\vec{H} = \frac{1}{\mu_0}\vec{B} - \vec{M}, \tag{2.21}$$

where ε_0 and μ_0 are respectively the electric permittivity and the magnetic permeability of vacuum, $\vec{P}$ is the electric polarization, and $\vec{M}$ is the magnetization.

Maxwell's equations are only valid for small volumes: in this way, the physical properties of the medium vary continuously. For a monochromatic parallel beam of light in a homogeneous medium without sources, Maxwell's equations are solved by the electromagnetic field:

$$\vec{E}(\vec{r},t) = \vec{E}_0\, e^{i(\vec{k}\cdot\vec{r}-\omega t)}, \tag{2.22}$$

$$\vec{H}(\vec{r},t) = \vec{H}_0\, e^{i(\vec{k}\cdot\vec{r}-\omega t)}, \tag{2.23}$$

where $\vec{E}_0$ and $\vec{H}_0$ partially represent the amplitudes of the electric and magnetic fields respectively and are constant complex vectors, ω is the angular frequency defined as $\omega = 2\pi f$, with frequency f, and $\vec{k} = \vec{k}_R + i\vec{k}_I$ is the complex wave vector, which is linked to ω by the dispersion relation:

$$\vec{k}\cdot\vec{k} = \varepsilon\mu\omega^2, \tag{2.24}$$

where ε and μ are the electric permittivity and the magnetic permeability of a medium, respectively.
For a homogeneous wave, $\vec{k}_R$ and $\vec{k}_I$ propagate in the same direction and the equation above becomes:

$$(|\vec{k}_R| + i|\vec{k}_I|)^2 = \varepsilon\mu\omega^2 \tag{2.25}$$

or

$$|\vec{k}_R| + i|\vec{k}_I| = \omega\sqrt{\varepsilon\mu} = \frac{\omega}{c}n, \tag{2.26}$$

with the speed of light $c = \frac{1}{\sqrt{\varepsilon_0\mu_0}}$ and the refractive index $n = n_R + in_I$, whose complexity directly derives from k, can be written as:

$$n = \sqrt{\frac{\varepsilon\mu}{\varepsilon_0\mu_0}} = \frac{c}{c'}, \tag{2.27}$$

with $c' = \frac{1}{\sqrt{\varepsilon\mu}}$. For a non-absorbing medium, $\vec{k}_I = 0$ and, therefore, n is pure real; in this case, c' can be considered as the phase speed of the wave in the medium. If the medium is absorbing, the bulk absorption coefficient can be derived from equations 2.22 (or 2.23), which depends on the imaginary part only of the refraction index:

$$\beta_{abs}(\lambda) = \frac{2\omega}{c}n_I(\lambda) = \frac{4\pi}{\lambda}n_I(\lambda). \tag{2.28}$$

From Fig. 2.4 can be seen that n_I is very small at visible wavelengths (from 0.4 µm to 0.7 µm, part of the grey "Solar window"), therefore the absorption (eq. 2.28) is negligible and the transfer of solar radiation is dominated by scattering, while at thermal infrared wavelength n_I is $\sim 10^7$-times larger and absorption and emission are the dominant processes.
The absorption coefficient in eq. 2.28 is involved in the computation of the radiation intensity $I = I(x)$ once the incident radiation with intensity I_0 has passed through a medium with depth x:

$$I(x,\lambda) = I_0 e^{-\beta_{abs}(\lambda)x}, \tag{2.29}$$

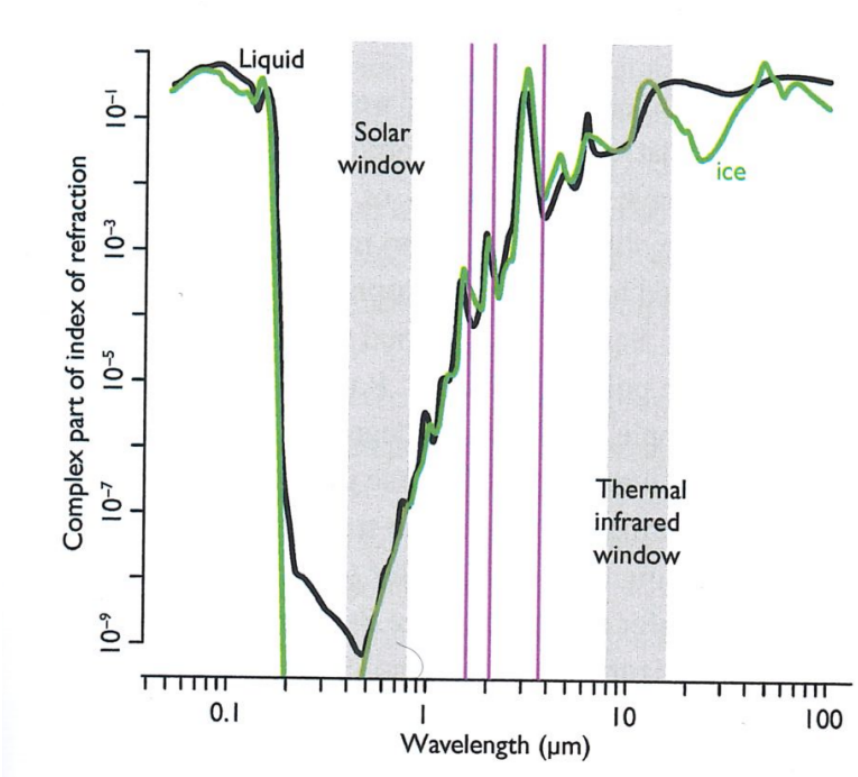

Figure 2.4.: Imaginary part of the refractive index n_I of pure water in liquid and ice phase. A black line is used for liquid water and a green line for ice. ©Cambridge University Press (Siebesma et al., 2020). Reproduced with permission of The Licensor through PLSclear.

that can be written in terms of the transmittance T as the fraction of the incident radiation intensity transmitted through the medium and the incident radiation:

$$T(x,\lambda) = \frac{I(x,\lambda)}{I_0} = e^{-\beta_{abs}(\lambda)x}, \tag{2.30}$$

also known as Lambert Beer's law.

Irrespective of the cloud phase, the incident solar, terrestrial, and atmospheric radiation will interact with clouds by being absorbed, reflected, and transmitted. The intensity of the interaction between radiation and cloud particles depends on the geometric cross-section of the particles C_{geo} and the wavelength λ of the radiation. The interaction between particles and radiation leads to the extinction of some of the incident radiation in conformity with the extinction cross-section $C_{ext}(r/\lambda) = Q_{ext}(r/\lambda)C_{geo}(r/\lambda)$, where Q_{ext} is the extinction efficiency and is a function of the ratio of the particle radius r and the wavelength λ. For typical sizes of cloud particles ($\mathcal{O}$(10μm) for liquid droplets and $\mathcal{O}$(100μm) for ice crystals), considering wavelength ranges of 0.4-0.8 μm for solar radiation and 8-12 μm for thermal radiation, interactions between liquid cloud particles and solar radiation as well as ice crystals and both

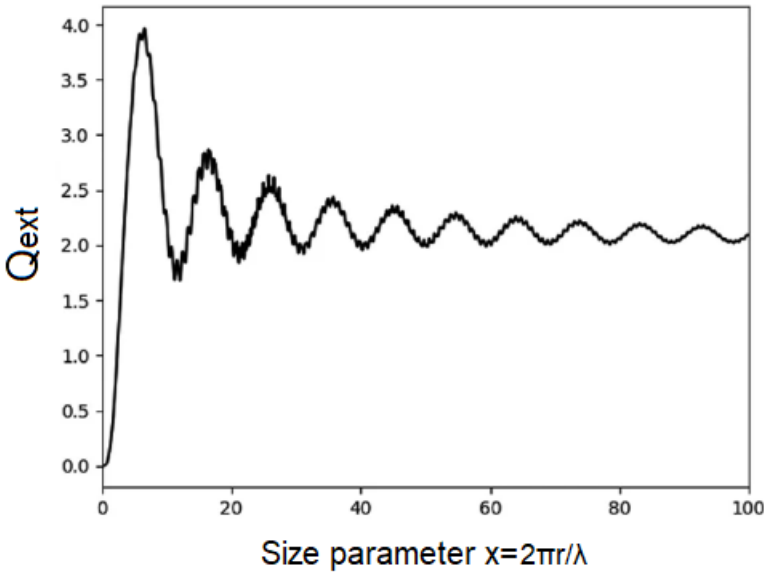

Figure 2.5.: The extinction coefficient Q_{ext} as a function of size parameter x for sphere with refractive index $n = 1.33 + 0.001i$. ©Springer Nature Switzerland AG (Adapted from Efremenko and Kokhanovsky (2021)). Reproduced with permission of The Licensor through PLSclear.

thermal and solar radiation are very efficient as $r \gg \lambda$, and $Q_{\text{ext}} \simeq 2$ (Fig. 2.5). Because the radiation interacts with a distribution N(r) of cloud particles, the bulk extinction coefficient is given by:

$$\beta_{\text{ext}}(\lambda) = \int_0^\infty C_{\text{ext}}(r/\lambda)N(r)\mathrm{d}r, \tag{2.31}$$

where C_{ext} has unit of squared length and β_{ext} of inverse length.
In analogy with the extinction coefficient in eq. 2.31, the scattering coefficient is defined as:

$$\beta_{\text{sca}}(\lambda) = \int_0^\infty C_{\text{sca}}(r/\lambda)N(r)\mathrm{d}r. \tag{2.32}$$

Extinction, scattering, and absorption coefficients are linked by the relation:

$$\beta_{ext} = \beta_{sca} + \beta_{abs}, \tag{2.33}$$

and define the single-scattering albedo ω_0, i.e., the relative importance of scattering:

$$\omega_0 = \frac{\beta_{sca}}{(\beta_{sca} + \beta_{abs})} = \frac{\beta_{sca}}{\beta_{ext}}. \tag{2.34}$$

Scattered radiation propagates in preferential directions defined by the size, shape, and imaginary part of the refractive index of the particles. The angular redistribution of the scattered radiation with scattering

angle Θ is given by the scattering phase function $P(\lambda,\Theta)$, which is a unitless probability distribution normalized so that $\frac{1}{4\pi}\int_1^{-1} P(\mu)d\mu = 1$, where $\mu = \cos\theta$, θ is the scattering angle between the incident and the scattered radiation, and $\theta = 0$ refers to forward scattering.
The fraction of light scattered into the forward hemisphere ($\mu > 0$), relative to the direction of the incident radiation, defines the asymmetry parameter g:

$$g = \frac{\int_1^{-1} \mu P(\mu)\mathrm{d}\mu}{\int_1^{-1} P(\mu)\mathrm{d}\mu}. \tag{2.35}$$

If $g = 1$, the scattering is fully forward; in case $g = 0$, there is isotropic scattering. Compared to the wavelength of radiation, small particles have asymmetry parameter close to zero, while larger particles have larger asymmetry parameters. Typical values of g for cloud droplets in the visible range are ~ 0.85 (AMS, 2012).

2.3.2. From cloud optical properties to cloud radiative properties

Clouds are complex systems with spatial and temporal variability. An approximation usually applied to be able to deal with this problem is to consider cloud properties changing only along the vertical dimension z (at a given time) and, more in general, a plane parallel atmosphere. It is convenient to define the unitless vertical coordinate τ, known as optical depth:

$$\tau(\lambda) = \int_0^z \beta_{ext}(\lambda, z)\mathrm{d}z, \tag{2.36}$$

where $z = 0$ and $\tau = 0$ are considered at the top of the atmosphere, while $\tau = \tau^*$ is the total atmospheric optical depth at the bottom of the atmosphere. Different cloud types can have different optical thicknesses. The International Satellite Cloud Climatology Project (ISCCP) uses global measurements from geostationary and polar-orbiting satellites and defines nine cloud types with a classification based on cloud top pressure and τ. In the ISCCP cloud classification, typical τ values of cumulus, altocumulus, and cirrus clouds are below 3.6; larger values of τ, between 3.6 and 23, are associated with stratocumulus, altostratus, and cirrostratus; stratus, nimbostratus, and deep convective clouds typically have $\tau > 23$. In section 3.3.0.1, the ISCCP cloud type classification is explained in more detail.
Considering the atmosphere as adjacent layers, it is possible to consider the radiation traveling through and interacting with the single layers. For each layer, the total rate of energy transported by radiation per unit area, i.e., the irradiance F, can be split into upward and downward components describing the energy flux normal to a horizontal (constant τ) plane:

$$\begin{aligned} F^{\uparrow}(\tau) &= \int_0^{2\pi}\int_0^{1} I(\tau,\mu,\phi)\mu \ \mathrm{d}\mu \ \mathrm{d}\phi \\ F^{\downarrow}(\tau) &= \int_0^{2\pi}\int_0^{-1} I(\tau,\mu,\phi)\mu \ \mathrm{d}\mu \ \mathrm{d}\phi \end{aligned} \tag{2.37}$$

where the intensity of the radiation $I(\tau,u,\phi)$ flows in a particular direction given by the cosine of the polar angle u (for upward transport, $u>0$), the cosine of the azimuthal angle ϕ, at a specific location indicated by τ. The dependence of τ on the wavelength λ is omitted for brevity.
The radiation intensity can be reduced by scattering and absorption and increased by scattering and emission. These changes are described in the differential form of the radiative transfer equation:

$$\mu\frac{\mathrm{d}I(\tau,\mu,\phi)}{\mathrm{d}\tau} = -I(\tau,\mu,\phi)+S(\tau,\mu,\phi), \tag{2.38}$$

where $S(\tau)$ is the source function, defined by:

$$S(\tau) = \frac{\omega_0(\tau)}{4\pi}F_{Sun}P_{Sun}(\tau,\Theta_{Sun})e^{-\tau/\mu_0}+(1-\omega_0(\tau))B(T)+\frac{\omega_0(\tau)}{4\pi}\int_0^{2\pi}\int_{-1}^{1}I(\tau,\mu',\phi')P(\tau,\Theta)\,\mathrm{d}\mu'\,\mathrm{d}\phi', \tag{2.39}$$

with:

- the first term on the r.h.s. indicating the incoming solar radiation flux F_{Sun} at the top of the atmosphere, with transmittance $e^{-\tau/\mu_0}$ and scattering angle Θ_{Sun} depending on the direction of the incoming radiation (μ_0,ϕ_0) and the outgoing radiation (μ,ϕ);

- the second term on the r.h.s. being the isotropic emission term with the Planck function $B(T)$, i.e., the spectral distribution of a black body emitted at temperature T:

$$B(\lambda,T) = \frac{2hc^2}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda k_B T}}-1}, \tag{2.40}$$

  where h is Planck's constant and k_B is the Boltzmann constant;

- the third term on the r.h.s. being the incident radiation intensity interacting with the layer at optical depth τ and with scattering angle Θ depending on the direction of the incoming radiation (μ',ϕ') and the outgoing radiation (μ,ϕ).

The difficulty of solving the eq. 2.38 is represented by the presence of the radiation intensity $I(\tau,\mu,\phi)$ on both sides of the equation. Different methods have been developed for solving this equation, which include the "two-stream" approximation (Schuster, 1905), the discrete ordinate method (Chandrasekhar, 1960), the adding method (van de Hulst and Irvine, 1963), and the iteration of formal solution (Herman and Browning, 1965).

2.4. Remote sensing of clouds

Remote sensing is the science that aims to detect and interpret measurements of radiance reflected and emitted by a body or by the atmosphere at a certain distance. The difference between the Earth's outgoing radiation (i.e., thermal radiation emitted by the surface-atmosphere system or solar radiation reflected at

the Earth's surface or scattered in the atmosphere) through the clear sky and clouds gives information about the cloud radiative effect and is used to evaluate clouds, which usually appear colder than the Earth in terms of brightness temperature. The sensors used for remote sensing can be passive, if the measured radiance is from an external source of radiation (terrestrial, solar, or atmospheric), or active, if the source of radiation is self-provided. These sensors can be installed, for example, on the ground, on aircraft or on board satellites travelling in different orbits.

2.4.1. Satellite orbits

Polar orbits, also known as "low Earth orbits" (LEO), are at low altitudes ($\sim$ 700 km), and their inclination of $\sim 1^\circ$ d^{-1} (per day) fixes the equatorial crossing local time. A more appropriate way to call them is "sun-synchronous" orbits as they do not pass over the poles but up to about 80° (Stephens, 1994).
To stay in an orbit around the Earth, satellites must balance both the centripetal acceleration and the gravitational force (Stephens, 1994):

$$mv^2/R = Gm_E m/R^2, \tag{2.41}$$

where R is the orbit radius, v is the tangential velocity, m is the mass of the satellite, m_E is the mass of the Earth, and G is the gravitational constant. From eq. 2.41, it is possible to determine the speed at which the satellite travels in its orbit:

$$v = \sqrt{Gm_E/R} \tag{2.42}$$

and the period of the orbit:

$$T_{orb} = 2\pi\sqrt{R^3/Gm_E}. \tag{2.43}$$

From the last equation can be deduced that the higher the orbit, the longer the period of the orbit. If the satellite is far enough away from the Earth that its period matches the Earth's period, it will remain in a fixed position above the Earth in a geostationary orbit.

2.4.2. Passive measurements

Passive sensors take advantage of the radiation emitted naturally by the Earth, the Sun, and the atmosphere and how it interacts with a medium, such as clouds, along its propagation direction.
In the thermal infrared (between approximately 8 µm and 14 µm), the imaginary part of the refractive index of water is high (water mostly absorbs, Fig. 2.4) while the single-scattering albedo of water is consequently low (eq. 2.34). Under these conditions, eq. 2.38 can be simplified and its integral is (Siebesma et al., 2020):

$$I_{IR}(\tau_{abs}, \mu, \phi) = I_{IR}(\tau_0)e^{-(\tau_{abs}-\tau_0)/\mu} + \int_{\tau_0}^{\tau} e^{-(\tau_{abs}-\tau'_{abs})/\mu} B(T(\tau'_{abs}))\mathrm{d}\tau'_{abs}, \tag{2.44}$$

where τ_{abs} is the absorption optical depth, τ_0 is the optical depth at the boundary, and it is assumed that the downwelling infrared radiation at the top of the atmosphere is almost absent, while the upwelling radiation at the surface is determined by the product of the Planck function at the surface temperature and the bulk surface emissivity. With passive sensors, the radiation intensity I_{IR} is measured and eq. 2.44 can be inverted to retrieve the brightness temperature $B(T)$, i.e., the temperature that a perfect black body would have to emit the observed radiation intensity. If the cloud emissivity can be calculated, $B(T)$ can be converted to a physical temperature and the cloud top pressure can be calculated once the temperature profile as a function of pressure is known.

Another important cloud parameter is the thermodynamic phase of clouds, which can be determined using, among others, thresholds based on the brightness temperature in the infrared window (Pavolonis and Heidinger, 2004; Pavolonis et al., 2005) or the spectrally dependent difference in the imaginary refractive index of liquid water and ice. The more accurate the cloud phase, the more certain the estimate of the cloud optical thickness (eq. 2.31). To simplify this process, a wavelength at which liquid and ice do not absorb (solar) radiation can be chosen, so that $\omega_0 \simeq 1$. In this case, eq. 2.38 turns into (Siebesma et al., 2020):

$$\begin{aligned} I_{solar}(\tau,\mu,\phi) = I_{solar}(\tau_0)e^{-(\tau-\tau_0)/\mu} + \frac{1}{4\pi}\int_{\tau_0}^{\tau} e^{-(\tau-\tau')/\mu} \\ \times\left(\int_0^{2\pi}\int_0^1 P(\tau',\Theta)I_{solar}(\tau',\mu',\phi')\,\mathrm{d}\mu'\,\mathrm{d}\phi'\right)\mathrm{d}\tau'. \end{aligned} \tag{2.45}$$

In this way, the radiative transfer equation depends solely on $P(\Theta)$ and τ, with the assumption that solar radiation at the top of the atmosphere travels in a single direction ($\mu_0 > 0, \phi_0$). Equation 2.45 can be inverted to retrieve τ. Because of the variability in ice particles and, therefore, the higher sensitivity of $P(\Theta)$ to ice particle size distribution, the optical thickness of ice clouds has more uncertainty than that of liquid clouds.

Another important cloud parameter is the cloud effective radius. To determine it, a wavelength λ can be selected for which absorption by a clear atmosphere is negligible but absorption by condensed water is considerable. Considering that reflection must come from the interfaces and absorption from the interior of the medium, measurements at non-absorbing wavelength are sensitive to $\int_0^\infty r^2 n(r)dr$, which is proportional to τ, while measurements at absorbing wavelength are sensitive to $\int_0^\infty r^3 n(r)dr$. For spherical cloud droplets, the cloud effective radius r_{eff} is defined by the World Meteorological Society as the "size distribution of liquid water drops, assimilated to spheres of the same volume". This variable can be calculated using the following equation:

$$r_{eff} = \frac{\int_0^\infty r^3 n(r)\mathrm{d}r}{\int_0^\infty r^2 n(r)\mathrm{d}r}. \tag{2.46}$$

In addition, Nakajima and King (1990) introduced lookup tables for ice and liquid clouds to retrieve r_{eff} and τ simultaneously using simultaneous measurements at non-absorbing and absorbing spectral channels (e.g., 0.75 µm and 2.16 µm respectively).
Passive sensors include the Advanced Very High-Resolution Radiometer (AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS).

2.4.2.1. AVHRR - Advanced Very High-Resolution Radiometer

The Advanced Very High-Resolution Radiometer (AVHRR) was one of the first instruments onboard satellites to take measurements of atmospheric parameters. Mainly used for clouds and sea-surface temperature studies, the AVHRR has channels in the visible (0.6 µm) and infrared (0.9, 1.6, 3.7, 10.8, and 12 µm) spectrum, with a horizontal resolution of about 1 km in the visible and about 4 km in the infrared (Stephens, 1994).
AVHRRs can be found onboard the polar-orbiting satellites from NOAA-NASA (National Oceanic and Atmospheric Administration - National Aeronautics and Space Administration) and MetOp-EUMETSAT (Meteorological Operational Satellites - European Organisation for the Exploitation of Meteorological Satellites).
Depending on the satellite's equatorial crossing time, the AVHRRs onboard the NOAA satellites are named the AVHRR-AM (crossing the equator in the early morning on the ascent) and the AVHRR-PM sensors (crossing the equator in the afternoon). The swath width of the AVHRRs is about 2700km, so global coverage is provided twice daily from both ascending and descending tracks. The satellites orbit at about 850km above the Earth's surface. The measurements have a spatial resolution of about 1km at the nadir and 6km at the edge of the scan (Klaes, 2018).
AVHRR radiance measurements are used as inter-satellite calibrators: data from each geostationary satellite and the afternoon NOAA polar orbiter are compared; then, all radiance measurements are normalized to a single satellite; finally, this normalization is adjusted to that of NOAA-7 in July 1983.
The cloud detection scheme exploits both visible reflection (VIS) and emitted radiation (IR). It is based on the assumption that clear scenes are less variable, darker in the VIS, and warmer in the IR intensities than cloudy scenes. First, the spatial and temporal variation of IR and VIS intensities are statistically collected. Then the global clear sky composite is derived from the IR and VIS measurements, assuming that the intensities vary less in clear scenes than in cloudy scenes. The uncertainties in the estimate of clear-sky intensities are then compared with the differences between the estimated clear-sky intensities and the measured intensities. If these differences are greater than the uncertainties and result in colder IR and brighter VIS, then the pixel is considered cloudy. After the clear-sky/cloudy pixel classification, the measured intensities are compared to radiative transfer model calculations and then converted into cloud optical thickness and cloud top pressure. Measurements during nighttime use only IR intensities and no cloud optical thickness can be retrieved. Nevertheless, IR variations can provide information on cloud top pressure as they are associated with the cloud top brightness temperature (Stephens, 1994).

2.4.2.2. MODIS - Moderate Resolution Imaging Spectroradiometer

The Moderate Resolution Imaging Spectroradiometer (MODIS) was first flown onboard NASA's Terra satellite, launched in 1999. In 2002, the MODIS instrument was also launched on board NASA's Aqua satellite. With both Terra and Aqua MODIS, it is possible to acquire images of the entire Earth in 1 or 2 days using 36 spectral bands with wavelengths ranging from the visible (0.4 μm) to the infrared (14.4 μm) spectrum (Kelkar, 2007). The Terra-Aqua coupled system, with equatorial crossing times of, respectively, 10:30 a.m. and 1:30 p.m., optimises cloud-free imaging while minimising the optical effects of glare and shadow associated with morning and afternoon sunlight. The orbital altitude of about 705 km and the scanning pattern of $\pm 55^{\circ}$ result in a swath width of 2330 km. MODIS data products are provided in three spatial resolutions, i.e., 250 m, 500 m, and 1000 m.

A spectroradiometer is based on a monochromator with a diffraction grating. The incoming radiation is collimated to the monochromator by a concave mirror and diffracted into its spectral components by the diffraction grating before being focused onto a detector by a second concave mirror. Grating rotation allows the spectrum to be scanned while the electrical signal is recorded by the detector.

Calibration of MODIS makes use of a solar diffuser and reflective bands: if the reflectance of the diffuser's surface and the angle of the incoming solar radiation are known, the radiance of the diffuser can be estimated and the reflective bands can be calibrated. Calibration takes place when Terra is close to the North Pole or when Aqua is close to the South Pole, i.e., when solar radiation is available and hits the diffuser at a certain angle. As the polar regions are reached after a nighttime view of the Earth from MODIS, an attenuation screen is used to calibrate the 9 bands which would otherwise saturate if sampling were done in full sunlight. Once the satellite is travelling on the sunlit side of the Earth, the solar diffuser is protected from solar radiation. Because of the constant solar intensity, any change measured by MODIS must be related to degradation of either MODIS or the solar diffuser (Nishihama et al., 1997).

2.4.3. Active measurements

Active sensors, unlike passive sensors, send their own radiation in pulses and measure the intensity and the delay of the signal reflected back to the sensor. This means that the measurement is based on scattering, which is the dominating process in the visible and microwave energy spectrum. The sensors used for active measurements are lidar (Light Detection and Ranging) and radar (RAdio Detection And Ranging) instruments. Their fields of view are very narrow, so most of the backscattered radiation is mainly scattered only once before returning in exactly the same direction as it was sent. For this specific case with scattering in a single direction $P(\Theta = \pi)$, eq. 2.38 depends only on the distance d travelled by the backscatter and changes in (Siebesma et al., 2020):

$$\begin{aligned} I_{back}(d) &= \frac{K}{d^2}\gamma_{back}(d) exp\left(-2\alpha \int_0^d \beta_{ext}(d')\mathrm{d}d'\right) \\ &= \frac{K}{d^2}\gamma_{back}(d) exp\left(-2\alpha\tau\right), \end{aligned} \tag{2.47}$$

where K contains information about the instrument as well as the pulse strength and the receiver gain, $\gamma_{back} = \beta_{sca}(d) \times P(\Theta = \pi, d)/4\pi$ is the backscatter coefficient, and α is the multiple-scattering factor ($= 1$ for single scattering, < 1 for multiple scattering). For large optical depth ($\alpha\tau \gtrsim 3$) the signal is fully attenuated and the measurement is not reliable, if present. This is a significant limitation of lidar in exploring optically thick clouds.

The number and size of particles can be deduced from the scattered and extinguished signals, while the shape of the particle is given by polarisation. When radiation is reflected by spherical particles with single scattering, its polarisation remains unchanged; if the polarisation ratio (ratio between vertical polarisation and horizontal polarisation) is reduced, the particles are not spherical and therefore they can be ice or aerosols (for lidar only, see Sec. 2.4.3.1).

While, on the one hand, this makes phase identification by active sensors more reliable than phase identification by passive sensors, on the other hand, the current active sensors have a limited single near-nadir viewing direction and, therefore, they provide much smaller samples than those provided by passive sensors.

2.4.3.1. Backscatter Lidar CALIOP - Cloud-Aerosol Lidar with Orthogonal Polarization

Lidar systems have been in use since the early 1900s: around 1938, pioneering work was performed in France using pulsed light to determine the cloud heights (Bureau, 1946). There are several categories of lidar that differ in the properties they measure. The atmospheric backscatter lidar is currently the most widely used lidar and it differs from the other lidar types in that it uses the absorption of the gas that the backscattering laser travels through to obtain information about possible aerosol and cloud hydrometeors present in the volume of atmosphere travelled. The lidar transmitter and its receiver telescope are aligned so that the telescope can collect the backscattered radiation. The transmitted narrow laser beam undergoes some divergence as it propagates. However, the field of view of the receiver is large enough to prevent signal loss. In this way, the lidar volumes of a few cubic meters, defined by short pulse lengths ($\sim 10^{-8}$s) and small beam divergence ($\sim 10^{-3} - 10^{-4}$ radians), become tens of kilometers (Stephens, 1994). The calibration of the lidar signal consists of the calibration of the coefficients α_{back} and β_{ext} in eq. 2.47.

For the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), which uses two wavelengths specifically at 532 nm and 1064 nm, this calibration consists of three steps (Hostetler et al., 2006):

- First, the backscatter coefficient for the 532 nm parallel channel is calculated, assuming that the atmosphere at mid- and high latitudes is free of aerosols at an altitude of about 30-34 km. This backscatter coefficient is compared with the molecular backscatter coefficient calculated at low latitudes, where an important aerosol contribution is expected. Knowing the molecular number density and the estimates of the molecular backscatter cross-section, the molecular backscatter coefficients can be derived (Reagan et al., 2002).

- Then, the backscatter coefficient for the 532 nm perpendicular channel is calculated. As the free air signal is not strong enough to calibrate the perpendicular channel, the 532 nm parallel channel is used for calibration.

- Similarly to the previous step, the calibration of the 532 nm parallel and perpendicular channels is used to calibrate the 1064 nm channel, as its signal would be too low in the mid-stratosphere to make the calibration measurement reliable. To achieve this, specific cirrus clouds are selected: in fact, cirrus cloud particles are large and, therefore, the ratio of the 532 nm and 1064 nm backscatter coefficients and the ratio of the 532 nm and 1064 nm extinction coefficients are both close to 1. The 1064 nm channel is calibrated by comparing the 1064 nm backscatter signal with the calibrated 532 nm cirrus backscatter measurements (Reagan et al., 2002).

To avoid a lower signal-to-noise ratio in daylight conditions, the 532 nm parallel channel calibration coefficient is estimated on the night side of the orbit. These estimates are stored and used for interpolations/extrapolations at any point in the orbit. If detector gains or boresight operations are changed, only calibration coefficients generated from data collected after the change are used.
CALIOP is onboard CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) satellite, with an orbital altitude of about 705 km. The footprint diameter of CALIOP is 70 m, while the horizontal and vertical resolutions are 333 m and 30 m, respectively, below 8 km altitude and 1 km and 60 m, respectively, between 8 km and 20 km. The primary product of CALIOP is the cloud top height, which is converted to temperatures using model data from the Goddard Earth Observing System, Version 5 (GEOS-5) vertical profiles. The cloud phase is then obtained from the particle depolarisation ratio δ of the backscattered light $\left(\delta = \frac{\beta_{\perp}}{\beta_{\parallel}}\right)$.

2.4.4. Definition of observation data processing levels

Satellite-based Earth observations can be provided at different "levels" by the responsible institutions. The levels from 0 to 4 range from full-resolution raw data to geographical composites of processed measurements to more processed variables derived from multiple measurements of previous levels. Although the levels are not fully standardised, the commonly used levels are summarised in table 2.6.

Level of product	Institution(s) typically responsible	Indicative content	How produced	Science application
L0	Space agency or meteorological agency	Raw telemetry: timings, counts, instrument data, etc	Downlinked to receiving stations and consolidated	Usually none
L1	Space agency or meteorological agency	Calibrated radiances (and/or counts and gain parameters), with location, time, and viewing geometry	Calculated from L0 using in-flight calibration results, and platform navigation	Basis for retrieval of geophysical variables
L2	Space agency, meteorological agency, research organisation, commercial	Estimates of geophysical variables on the spatio-temporal sampling pattern of the L1 radiances ('swath data')	Retrieved, combining radiances (may also exploit auxiliary datasets)	Full spatial resolution analyses of retrieved variable, e.g. for process study
L3	Meteorological agency, research organisation, commercial	L2 data transformed to a fixed spatio-temporal sampling ('grid') often at reduced spatio-temporal resolution.	Aggregation in space and/or time, e.g. by averaging available L2 data within grid	Model testing, analysis of change and variability
L4	Meteorological agency, research organisation	Spatio-temporally complete fields on a regular grid	Gap filling of L2 and/or L3 data by interpolation in space and/or time, perhaps combining data from more than one sensor	Model testing, prescribed field for simulations, convenient analysis of change and variability

Figure 2.6.: Summary of the characteristic of level 0 (L0) to level 4 (L4). From Mittaz et al. (2019), distributed under CC BY 3.0.

3. Datasets and methods

For this work, several datasets from both observations, global climate models (GCMs), and global storm-resolving models were analysed. The observations include both passive (**Cloud_cci** (Cloud Climate Change Initiative)) **AVHRR-PMv2**, **CLARA-A2** (CM SAF (Climate Monitoring Satellite Application Facility) cLoud, Albedo and surface Radiation dataset from AVHRR data - Edition 2), **Cloud_cci AVHRR-PMv3**, **MCD06COSP_D3_MODIS** (MODIS (Aqua/Terra) Cloud Properties Level 3 daily) and active (**CALIOP**) satellite sensor-based datasets. The GCM datasets analysed are **CAM6-Oslo** (Community Atmosphere Model Version 6) of the Norwegian Earth System Model version 2 (NorESM2) including the MODIS-COSP2 (CFMIP (Cloud Feedback Model Intercomparison Project) Observation Simulator Package version 2.1.0) simulator, and three global storm-resolving models from the DYAMOND (DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains) Winter project: **ICON** (ICOsahedral Nonhydrostatic), **SCREAM** (Simple Cloud-Resolving E3SM (Energy Exascale Earth System Model) Atmosphere Model), and **GEOS** (Goddard Earth Observing System). The information of the analysed datasets is summerised in Table 3.1. As the study includes an analysis of mixed-phase clouds and the passive satellite sensor-based datasets provide only cloud-top information, the following cloud parameters are considered for this study: cloud-top phase (CPH) considered either as a liquid or ice pixel (for the AVHRR-based datasets and CALIOP), or as liquid (CLW) and ice (CLI) cloud fraction (for MCD06COSP_D3_MODIS and CAM6-Oslo), or as the mass mixing ratio of ice (q_i) and liquid water (q_l, for the DYAMOND datasets); cloud-top pressure (CTP); cloud-top temperature (CTT); cloud optical thickness (COT); liquid cloud effective radius (r_{liq}). While CPH, CLW and CLI, and q_i and q_w are used for cloud phase distribution, CTP and COT are used for cloud type classification. Furthermore, COT is related to the water content of clouds and thus to the size of cloud droplets, while r_{liq} can provide information on aerosols and cloud microphysics.

3.1. Observations

3.1.1. AVHRR-based datasets

Cloud_cci AVHRR-PMv2, CLARA-A2, and Cloud_cci AVHRR-PMv3 are based on measurements from the passive satellite sensor AVHRR onboard polar-orbiting NOAA satellites. The AVHRR observations provide daily global cloud top information (Level 3, see Sec. 2.4.4), organised into global latitude-longitude composites with a spatial resolution of 0.05°x0.05°.

Dataset	Sensor/scheme	Source	Spatial resolution	2D/3D	Cloud parameters	Cloud phase	Reference
Cloud_cci v2	AVHRR (passive)	NOAA-19	$0.05° \times 0.05°$	2D	CPH, CTT, CTP, COT, r_{liq}	liquid or ice	Stengel et al. (2017)
CLARA-A2	AVHRR (passive)	NOAA-19	$0.05° \times 0.05°$	2D	CPH, CTT, CTP, COT, r_{liq}	liquid or ice	Karlsson et al. (2017)
Cloud_cci v3	AVHRR (passive)	NOAA-19	$0.05° \times 0.05°$	2D	CPH, CTT, CTP, COT, r_{liq}	liquid or ice	Stengel et al. (2020)
CALIOP v4.20	CALIOP (active)	CALIPSO	$0.05° \times 0.05°$	3D	CPH, CTT, CTP, COT	liquid or ice	Hu et al. (2009)
MCD06COSP	MODIS (passive)	Aqua/Terra	$1° \times 1°$	2D	CLW, CLI, CTP, COT, r_{liq}	liquid and ice cloud fraction	Pincus et al. (2022)
CAM6-Oslo	two moment	model + simulator	$1.875° \times 2.5°$	2D	CLW, CLI, CTT, CTP, COT, r_{liq}	liquid and ice cloud fraction	Seland et al. (2020)
ICON	single moment	model	5 km	3D	q_i, q_w, CTT, CTP	cloud mass fraction of ice and liquid water	Reinert et al. (2023)
SCREAM	two-moment	model	3.25 km	3D	q_i, q_w, CTT, CTP, COT	cloud mass fraction of ice and liquid water	Caldwell et al. (2021)
GEOS	single moment	model	3 km	3D	q_i, q_w, CTT, CTP, r_{liq}	cloud mass fraction of ice and liquid water	Putman and Suarez (2011)

Table 3.1.: Overview of the datasets involved in this work.

3.1.1.1. Cloud_cci AVHRR-PMv2

In Cloud_cci AVHRR-PMv2 (hereafter Cloud_cci v2, Stengel et al., 2017, ATBD-CC4CLv5.1, 2019), the binary liquid or ice cloud top phase is generated using a threshold decision tree (Fig. 3.1, Pavolonis and Heidinger, 2004, Pavolonis et al., 2005) in which the infrared brightness temperature undergoes a series of spectral tests. The assigned phases are: "switch to liquid", "fog", "liquid", "supercooled", "switch to ice", "opaque ice", "cirrus", "overlap", and "probably opaque". The first four contribute to the final "liquid" phase, while the remaining five contribute to the final "ice" phase. The determination of CTT is obtained by interpolation of CTP and atmospheric profiles from the ECMWF (European Centre for Medium-Range Weather Forecasts) ERA-Interim reanalysis. If some ice pixels are found for temperatures above 0°C or some liquid pixels are found for temperatures below −40°C, the cloud top temperature is involved in the phase reassignment.

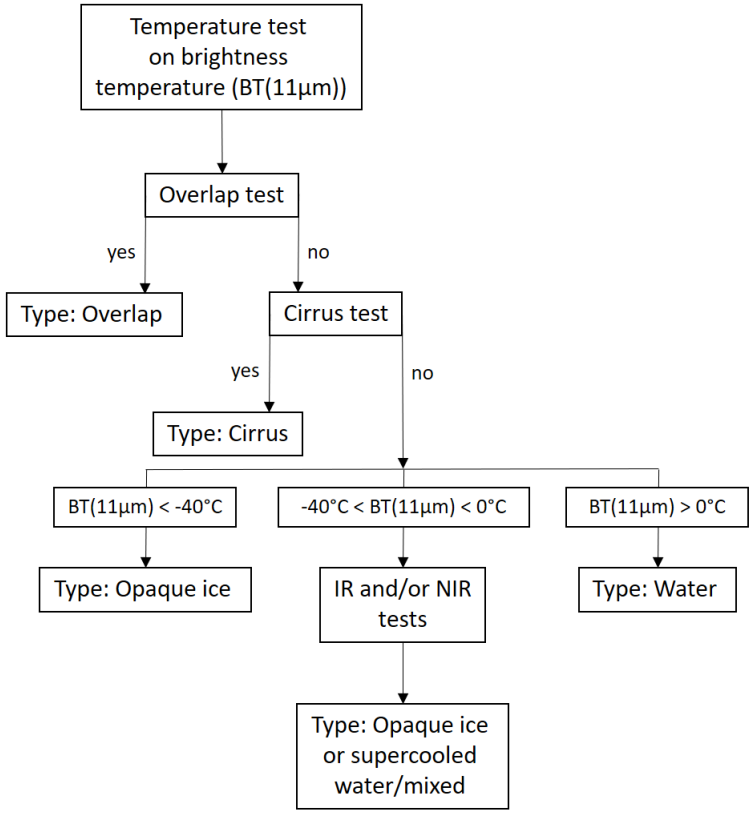

Figure 3.1.: Representation of the cloud phase algorithm based on Pavolonis et al. (2005).

3.1.1.2. CLARA-A2

In CLARA-A2 (Karlsson et al., 2017; ATBD-CPP_AVHRR, 2016), as in Cloud_cci v2, the cloud top phase is either liquid or ice and is generated by a threshold decision tree (Pavolonis and Heidinger, 2004; Pavolonis et al., 2005), that differs for some thresholds from those used in Cloud_cci v2. CTP and CTT

are retrieved using two different algorithms, depending on the cloud type: For opaque clouds, CTP is derived from the best fit between the simulated (from NWC SAF PPS (Nowcasting Satellite Application Facility Polar Platform System) cloud software) and the measured brightness temperature at 11µm, so that CTT is derived from the simulated vertical profiles; for semi-transparent cirrus and fractional water cloud pixels, a scatter plot technique is applied from which a CTT valid for these clouds is derived, so that CTP is derived from the simulated vertical profiles. As in Cloud_cci v2, a reclassification based on cloud top temperature is performed if ice pixels are found for temperatures warmer than 0°C or if liquid pixels are found for temperatures colder than −40°C.

3.1.1.3. Cloud_cci AVHRR-PMv3

For the Cloud_cci AVHRR-PMv3 retrieval scheme (hereafter Cloud_cci v3, Stengel et al., 2020, ATBD-CC4CLv6.2, 2019), AVHRR measurements were collocated with CALIOP cloud phase. An artificial neural network trained by the CALIOP cloud phase is used to classify CPH into a binary (liquid or ice) flag. Optimal estimation is used to produce CTP, COT, and the liquid effective radius r_{liq}, by fitting a physically consistent surface/cloud/atmosphere model to the satellite observations. As for Cloud_cci v2, the CTT is obtained by interpolation of CTP and atmospheric profiles from the ECMWF ERA-Interim reanalysis, and possible ice pixels with a temperature above 0°C and liquid pixels with a temperature below −40°C are again classified using the cloud top temperature.

3.1.2. CALIOP

For this study, CALIOP Level 2 (see Sec. 2.4.4) Cloud Layer Data in version 4.20 (Hu et al., 2009) was used, which provides data sampled in granules (orbital segments) with a spatial resolution of 5 km ($\sim 0.05°$ at the equator, as AVHRR). Because of its very narrow footprint ($\sim$ 70 m), it takes almost a month to collect data for global coverage. The dataset provides vertical distributions of clouds and their properties in layers, with each layer containing only one thermodynamic phase (i.e., either liquid or ice). The cloud phase retrieval (Fig. 3.2) consists of two steps: first, an initial cloud phase is determined by a depolarization-backscatter threshold approach, that distinguishes water clouds from ice clouds with “randomly-oriented” ice (ROI) and "horizontally-oriented” ice (HOI); next, a spatial coherence method is applied to separate light backscattered by spherical liquid droplets or by randomly oriented ice particles from light backscattered by horizontally oriented ice particles. Misclassifications of the cloud phase for lower layers of multi-layer clouds are corrected by estimating the probability that they are water or ice. This takes into account the temperature, the depolarisation ratio, and its standard deviation statistics. As for the AVHRR-based datasets, a phase reassignment is applied for ice pixels at temperatures warmer than 0°C and for liquid pixels or pixels with an unknown phase at temperatures colder than −40°C.

CALIOP can retrieve up to an optical thickness of ~ 5 in the cloud (Karlsson and Håkansson, 2018). For this study, only “medium” and “high” cloud-aerosol discrimination scores and “medium” and “high” cloud phase confidence scores are considered.

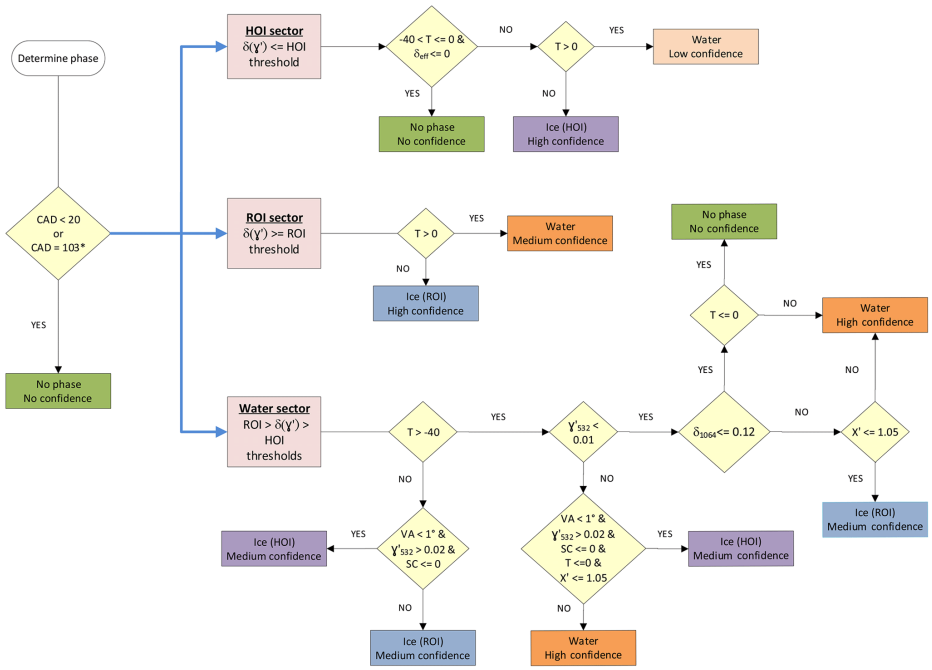

Figure 3.2.: Cloud phase algorithm of CALIOP v4 including the previous step based on the cloud-aerosol discrimination (CAD) algorithm (Liu et al., 2019). *CAD* refers to CAD scores, *ROI* refers to the randomly oriented ice, *HOI* refers to the horizontally oriented ice, T (expressed in °C) refers to the 532 nm centroid temperature of a cloud layer, γ_{532} is the integrated attenuated backscatter at 532 nm, δ is the depolarization ratio using the channels at 532 nm or both 1064 nm and 532 nm channels, *VA* is the near-nadir viewing angle, *SC* is the spatial coherence test result, and X' is the layer-integrated attenuated 1064/532 nm backscatter coefficient ratio. From Avery et al. (2020), distributed under CC BY 4.0.

3.1.3. MCD06COSP_D3_MODIS

MCD06COSP_D3_MODIS (hereafter MCD06COSP, Pincus et al., 2022; NASA, 2022) is a product specifically designed to facilitate the comparison of MODIS observations and climate model output obtained by applying the MODIS-COSP simulator (more about the COSP simulator in section 3.2.1.1). MCD06COSP provides daily global latitude-longitude composites with a spatial resolution of 1°x1°, derived by averaging Level 2 MODIS cloud top properties (5x5 km) and cloud optical properties (1x1 km) in each Level 3 gridcell. This dataset combines daily timeframes of MODIS Terra and MODIS Aqua cloud top, cloud mask, and cloud optical retrieval data, providing several cloud top-related parameters. For better agreement with the COSP simulator data, the optical properties of the clouds in MCD06COSP are retrieved using the 3.7 µm channel.

The "partly cloudy" cloud optical property parameters have not been included in this study as they are less reliable than the "cloudy" parameters. The ice or liquid amount in clouds is expressed by the ice

or liquid cloud fraction while the sum of these does not always correspond to the total cloud fraction as undetermined phases may also be included.
The cloud phase algorithm (Fig. 3.3, Marchant et al., 2016) uses a majority voting logic applied after four primary tests involving the CTT at 1 km, the cirrus detection test at 1.38 µm from the MOD35 cloud mask, IR cloud phase at 1 km test and cloud effective radius tests derived from the 3.7 µm channel.
Some issues have been identified for radii smaller than 0.4 µm for undetermined phase and liquid water clouds. These Level 2 data are masked or removed from the Level 3 data.

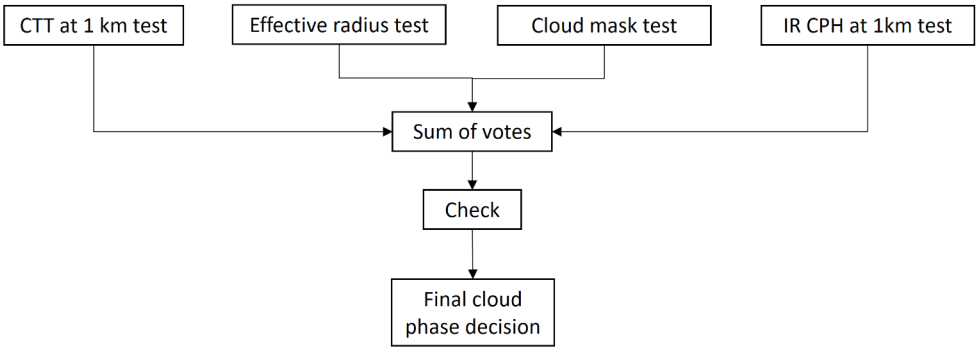

Figure 3.3.: Representation of the cloud phase algorithm of MCD06COSP based on Marchant et al. (2016) with the appropriate modifications reported in NASA (2022).

3.2. Models

3.2.1. NorESM2 — CAM6-Oslo

The NorESM2 model (Seland et al., 2020) is the second generation of the coupled Earth System Model developed by the Norwegian Climate Center. The atmospheric model component of NorESM2 is based on the CAM6 version of CESM2.1 (Community ESM 2.1), and uses the hydrostatic finite-volume dynamical core on a regular latitude-longitude grid with $1.9° \times 2.5°$ horizontal resolution and 32 hybrid-pressure vertical levels extending to a rigid top at $\sim$ 40 km, i.e., 3.6 hPa. NorESM2 uses the OsloAero5.3 (Kirkevåg et al., 2018) aerosol scheme and the aerosol-radiation-cloud interactions from Kirkevåg et al. (2013). A two-moment bulk cloud microphysics scheme is used. Output parameters were calculated for 4 years (from June 2019 to May 2013) after a 3-month model spin-up to allow for atmospheric adjustments (Shaw et al., 2022). To facilitate the comparison between observations and simulations, horizontal winds and surface pressure were nudged to ERA-Interim reanalysis data (Dee et al., 2011). We will refer to this model as CAM6-Oslo.

3.2.1.1. COSP2 simulator

The COSP simulator (Bodas-Salcedo et al., 2011) is a software tool developed by the Cloud Feedback Model Intercomparison Project (CFMIP) community. It simulates multiple active and passive satellite observations using gridbox mean vertical profiles of emissivity, temperature, cloud optical thickness, hydrometeor mixing ratios of clouds and precipitation, and humidity from each model in which it is implemented. A schematic representation of its main features is shown in Fig. 3.4. The gridbox means are downscaled to make them comparable to the higher-resolution satellite products: the gridbox mean vertical profiles are broken down into sub-columns that are comparable in size to a satellite pixel; they are then passed to the satellite instrument simulators that simulate the satellite retrieval that could produce that profile; finally, the simulated satellite retrievals are collected by statistical modules (averages or histograms) present in COSP or in the instrument simulator that generate diagnostics ready for comparison with similar statistics from satellite observations. For this study, COSP outputs are daily averages and atmospheric conditions are nudged to reanalysis data. The COSP simulator version 2.1.0 (COSP2, Swales et al., 2018) for MODIS has been implemented in CAM6-Oslo. The MODIS simulator (Pincus et al., 2012) in the COSP package determines cloud detection and calculates cloud optical depth by using simple integration at sub-column scale. For each sub-column, extinction-weighted integration is used to estimate cloud top pressure and cloud phase. Particle size is then computed using a pseudo-inversion: near-infrared fluxes at the top of the atmosphere are computed using polynomial fits of optical properties as a function of (height-dependent) cloud particle size; differences between the real fluxes and simulated fluxes are then minimised by varying the particle size. The cloud optical properties in the MODIS-COSP simulator are simulated at 3.7 µm.

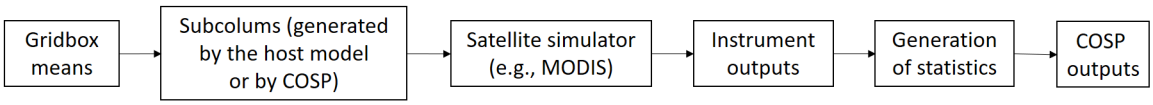

Figure 3.4.: Schematic representation of COSP.

3.2.2. DYAMOND Winter project

The DYAMOND initiative aims to compare global models at storm-resolving scales (spatial resolution of 5 km or less). It consists of three experiments: the global atmosphere-only models run during the boreal summer from the 1st of August 2016 for 40 days, and both the global atmosphere-only and the global coupled atmosphere-ocean models run during the boreal winter from the 20th of January 2020 for 40 days (Duras et al., 2021). All models included in these projects are initialised with a global meteorological analysis taken from ECMWF. To make the comparison between observations and models as consistent as possible, the Winter experiment of ICON, SCREAM, and GEOS was selected for this study. In fact, only the Winter experiment includes a model dataset providing COT (SCREAM), needed

to classify cloud types, and a model dataset providing r_{liq} (GEOS), needed to compare the cloud microphysics with the other datasets included in this work. Thus, the following variables are available for analysis: cloud pressure, cloud temperature, and mass mixing ratio of cloud ice or liquid water; in addition, SCREAM provides COT, while GEOS provides r_{liq}. For consistency, the atmosphere-only experiments of all models are included in this study, as only ICON and GEOS were also run as coupled atmosphere-ocean models. The data are not provided on a latitude-longitude grid.

3.2.2.1. ICON

The ICON model (Reinert et al., 2023; Zängl et al., 2015) is a global and regional numerical weather prediction model developed in cooperation between the German Weather Service (DWD) and the Max-Planck Institute for Meteorology in Hamburg (MPI-M). The ICON horizontal grid consists of a spherical icosahedral Arakawa-C grid in which an initial root is divided into n sections (Rn) and then k times into bisections (Bk), for a final customisable RnBk grid. The ICON vertical grid is height-based: the vertical layers near the surface follow the terrain, while they gradually become constant-height layers as the distance from the terrain increases. The vertical grid consists of n full levels (with level = 1 at the top of the atmosphere), where the vertical wind component is computed, and n+1 half levels (with level = n+1 at the terrain level, such as the Earth's surface), where the density of air, the virtual potential temperature, the vertical wind component, and the horizontal velocity component normal to the triangle edges are computed. The integration in time uses a two-time-level predictor-corrector scheme, fully explicit in the horizontal and implicit for the terms characterising the sound wave propagation. In the DYAMOND experiment, the setup of the atmosphere-only ICON run consists of an R02B09 grid, corresponding to a spatial resolution of 5 km and 20971520 grid columns, with 90 vertical full levels and a time resolution of three hours. A single-moment microphysics scheme ("graupel scheme"; D. Klocke, personal communication, 2022) has been used to predict the specific mass content of liquid and ice water in form of rain, graupel, and snow.

3.2.2.2. SCREAM

The SCREAM model (Caldwell et al., 2021) originates from the E3SM project, developed by the US Department of Energy to estimate the future climate of the energy sector using high-performance computing. The SCREAM horizontal grid consists of a spherical cubic grid with 1024x1024 spectral elements on each face (ne1024); each spectral element has a uniformly spaced 2x2 grid for a total of 25165824 grid columns with a horizontal resolution of 3.25 km. The vertical grid consists of 128 layers, with a sponge layer (above 19km) in the top 14 layers, and a model top at 2.25 hPa (40 km). The model uses explicit Runge-Kutta schemes in the horizontal direction and a diagonally-implicit Runge-Kutta scheme in the vertical direction. For each timestep, a nonlinear equation involving vertical velocity and geopotential is solved several times. The time resolution of the output is three hours. The double-moment cloud mi-

crophysics scheme is based on the Predicted Particle Properties (P3) scheme of Morrison and Milbrandt (2015), using only one population of ice particles growing from small embryos into large and possibly rimed snowflakes, while liquid droplets are sorted into cloud and rain droplets. Vapour deposition, sublimation, and the WBF process follow Gettelman and Morrison (2015) with an efficient transition from liquid to ice.

3.2.2.3. GEOS

The GEOS model (Putman and Suarez, 2011) is an Earth system model composed of modules that can be linked together to address different aspects of Earth science. With a non-hydrostatic finite-volume dynamical core, the horizontal grid of the GEOS model consists of a cubed-sphere staggered-D grid developed in collaboration with the Geophysical Fluid Dynamics Laboratory at NOAA, while the vertical grid is expressed in height-based, terrain-following Lagrangian control volume coordinates (similar to an eta coordinate system), where level = 1 corresponds to the top of the atmosphere while the lowest level corresponds to the Earth's surface. Pressure is expressed in hybrid-sigma coordinates. The integration in time uses an incremental analysis update (Bloom et al., 1996). Each edge of a cubed-sphere face has 2880 grid cells for a total of 49766400 grid columns with 181 vertical levels and a horizontal resolution of 3 km. The outputs are generated with a time resolution of 1 hour. A single-moment (6-class) cloud microphysics scheme with global carbon emissions for chemistry transport is used.

3.3. Methods

This work provides a statistic of the supercooled liquid fraction (SLF) in mixed-phase clouds and consists of three parts: a validation study comparing datasets from active and passive satellite sensors, a comparison of satellite and GCM products, and the improvement of GCMs by tuning them to observations. The following sections contain a description of the first two parts, while the third part is described in Chapter 7.

3.3.0.1. Cloud-type classification

If the datasets provide the COT, clouds are sorted as in Fig. 3.5 using thresholds for cloud top pressure (CTP= $[0, 440, 680, 1000]$ hPa) and cloud optical thickness (COT= $[0, 3.6, 23, 379]$), according to the International Satellite Cloud Climatology Project (ISCCP) classification (Rossow and Schiffer, 1999). Each COT-CTP combination is identified by the cloud name (e.g, cumulus, altocumulus, etc.) assigned in Rossow and Schiffer (1999), bearing in mind that a classification of cloud types based solely on COT and CTP has limitations, as shown in Hahn et al. (2001): low-level clouds cannot be distinguished by COT because they overlap in the COT-CTP joint histograms; nimbostratus, cumulonimbus, and thick altostratus cannot be distinguished because of their COT-CTP distributions; cirriform clouds are also

difficult to separate in COT-CTP histograms because of overlap in their distributions; altocumulus and thin altostratus should also be considered as a single cloud group in COT-CTP histograms.
Not all datasets provide the COT needed to sort the clouds into different types; ICON and GEOS do not provide the COT but their clouds can be sorted by using CTP to classify clouds into high-, mid-, and low-level clouds defined by CTP thresholds of $[0, 440, 680, 1000]$ hPa. Since SCREAM provides the COT for each vertical level, the column integral of the COT for each pixel is considered.

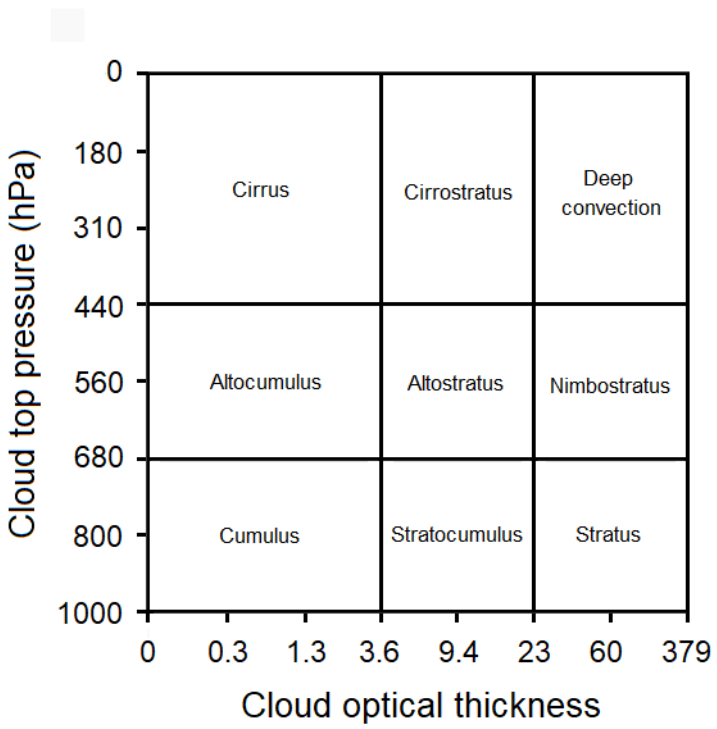

Figure 3.5.: ISCCP cloud-type classification following Rossow and Schiffer (1999).

3.3.1. Validation study: AVHRR vs CALIOP

The first part of this study is dedicated to the comparison of datasets based on passive (AVHRR: Cloud_cci v2, CLARA-A2, Cloud_cci v3) and active (CALIOP) satellite sensors, with the aim of identifying possible phase and temperature mismatches and thus understanding the potential and the limitations of the datasets. For a more reliable comparison, both collocated and non-collocated data were compared. The collocated data include cloudy pixels retrieved by all datasets within 3 minutes and 5 km and represent 0.02% of the AVHRR non-collocated data. For the AVHRR-derived datasets, all cloudy pixels with COT < 0.3 are filtered out from the analysis (Stengel et al., 2015). The highest cloud layer below an accumulated COT = 0.3 is considered the cloud top for CALIOP to mimic the AVHRR view for better comparability (Fig. 3.6). For CALIOP, pixels retrieved as cloudy with "no confidence" or "low confidence" are filtered out, as well as cloudy pixels with phase "no confidence" or "low confidence".

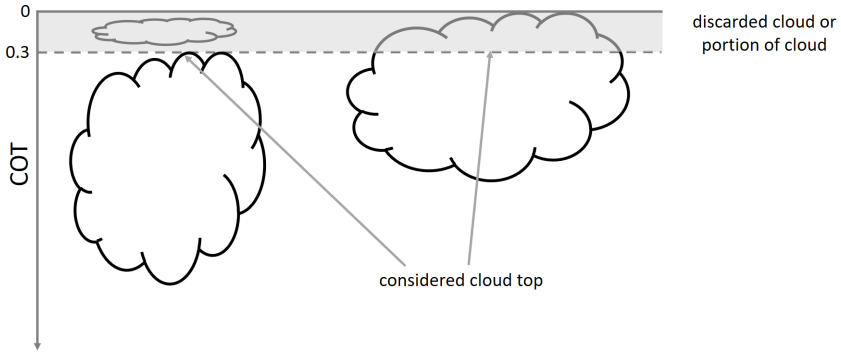

Figure 3.6.: Representation of the cloud layer considered as cloud top for CALIOP.

The study is limited to latitudes frame 60° N to 60° S, termed near-global, thus excluding satellite data with low confidence due to the presence of sea ice (King et al., 2004) and the high solar zenith angle (Grosvenor and Wood, 2014). In addition to the near-global analysis, continental and marine regions are considered separately, as well as different latitudes with the following band separation: Northern Hemisphere (NH) from 60° N to 30° N; Southern Hemisphere (SH) from 30° S to 60° S; Tropics from 30° N to 30° S. The analysis covers 4 years (from 1 June 2009 to 31 May 2013), chosen such that it includes the measurements provided by the newest AVHRR/3 instrument on board the latest NOAA satellite (NOAA-19). This avoids possible consistency problems due to sensor calibration differences with AVHRR on board previous versions of NOAA satellites.
The comparison is based on the SLF calculated at pixel scale on isotherms from −50°C to 5°C, with a step of 1°C, using the following equation:

$$SLF_{n^\circ\,pixel} = \frac{\text{n}^\circ\text{liquid pixels}}{\text{n}^\circ\text{liquid pixels} + \text{n}^\circ\text{ice pixels}} \tag{3.1}$$

For this study, only daytime data were used, as these provide the optical properties of clouds.

3.3.2. Comparison between observations and models

The second part of this work compares satellite (Cloud_cci v3, MCD06COSP), GCM (CAM6-Oslo) and global storm-resolving model (ICON, SCREAM, GEOS) datasets. The validation study described in section 3.3.1 with results in Chapter 4 shows that the passive satellite sensor AVHRR can contribute to the cloud phase research. As Cloud_cci v3 is the most recent AVHRR-based dataset analysed in Chapter 4, only this dataset is included for the comparison with models. MCD06COSP is also included to facilitate the comparison between observations and CAM6-Oslo in which the MODIS-COSP simulator has been implemented.

The analysis of the satellite datasets and CAM6-Oslo covers 4 years (from 1 June 2009 to 31 May 2013), as the validation study, while the DYAMOND models have only been analysed for February 2020 — the Winter project of DYAMOND consists of 40 days, 10 days as spin-up period and 30 days of analysis.
Furthermore, the analysis of the datasets providing COT is limited to daytime measurements, as the COT involved in the cloud-type classification can only be detected by satellite sensors in the visible range. To remove very thin clouds for which retrievals are more uncertain, only satellite measurements with COT > 0.3 and effective radius > 4 μm are used. For consistency with MCD06COSP, the cloud top in the MODIS-COSP simulator in CAM6-Oslo is defined as the highest cloud layer with an overlying COT = 0.3. While Cloud_cci v3, MCD06COSP and CAM6-Oslo provide the cloud top information in 2-dimensional composites extended in latitude and longitude (as the satellite view), the DYAMOND datasets provide the 3-dimensional view of the atmosphere, so a top-down approach is used to locate the cloud top. In ICON and GEOS, the cloud top is defined as the highest atmospheric layer for which $q_w + q_i > 10^{-8}$ kg kg^{-1} (Costa-Surós et al., 2020); in SCREAM, the cloud top is defined as the highest atmospheric layer for which COT > 0. Since the COT in SCREAM is only provided for daytime data, mimicking satellite retrievals based on solar wavelengths, a further analysis was performed by replacing the COT condition with that of Costa-Surós et al. (2020) (hereafter SCREAM-CS), thus using daytime and nighttime data to classify clouds into height levels, for a better comparison of the cloud amount with ICON and GEOS.
ICON, SCREAM, and GEOS were also analysed at the coarser resolutions 0.5°, 1°, and 2° to better investigate the liquid-to-ice transition at the cloud top. The resolution of 0.5°, in particular, allows a more consistent comparison with Cloud_cci v3. Due to the absence of a regular horizontal grid in the DYAMOND datasets, it is not possible to create gridboxes by aggregating constant numbers of gridpoints, so the coarsening of the spatial resolution is done by rounding the latitude and longitude values to the closest 0.5°, 1°, and 2° interval, respectively, in which the cloud variables are averaged. These gridboxes must contain pixels whose warmest and coldest temperature difference does not exceed 20°C.
This study includes both all cloudy pixels and cloudy pixels (or gridboxes) which are considered mixed-phase, i.e., with temperature between -40°C and 0°C, and 0 < SLF < 1 (or $q_w > 0$ and $q_i > 0$ in ICON, GEOS, and SCREAM-CS). The SLF derivation is described in section 3.3.2.1.
This study focuses on the geographical distribution of cloud types and both SLF-CTT and SLF-$\overline{r_{liq}}$ joint histograms of cloud types in the mixed-phase. A detailed description of these products can be found in section 3.3.2.3.

3.3.2.1. SLF derivation

In Cloud_cci v3, as reported in Sec. 3.1.1.3, the thermodynamic phase of a single cloudy pixel can be either liquid or ice, so a reduction in spatial resolution is required to compute the SLF within gridboxes: the original latitude-longitude daily composites made by 3000x7200 pixels are reorganised into 500x1200 gridboxes of 6x6 pixels, each of which is considered cloudy only if it contains at least two cloudy pixels

(i.e., more than $\sim 5.56\%$ of the gridbox contains clouds) with COT > 0.3 and effective radius > 4 µm. The sensitivity to the number of cloudy pixels in a gridbox required to include it in the analysis was also tested using gridboxes with at least five cloudy pixels. Furthermore, the gridboxes are only included in the analysis if the largest difference between the coldest and the warmest cloud top temperature of all valid pixels is less than 20°C; using the cloud top temperature as a proxy for the cloud top height, gridboxes with clouds of very different heights (such as low- and high-level clouds (Fig. 4.5)) are filtered out. Finally, gridboxes are termed "mixed-phase" if they include coexisting liquid and ice pixels, which must have a cloud top temperature between -40°C and 0°C. For each gridbox the SLF is then calculated using equation 3.1, which will satisfy the condition 0 < SLF < 1 due to the coexistence of liquid and ice phases.

The model datasets and MCD06COSP provide information on both liquid and ice content or cloud fraction in each pixel. For this reason, no gridboxes are needed to compute the SLF, as the individual pixels can already contain mixed phase. This will also allow the ability of the models to represent mixed-phase clouds and the phase transition from liquid to ice to be tested. Therefore, the SLF is calculated as the ratio of the liquid cloud fraction and the sum of the liquid and ice cloud fractions (for MCD06COSP and CAM6-Oslo, eq. 3.2) or as the ratio of the mass mixing ratio of liquid water and the sum of the mass mixing ratios of ice and liquid water (for ICON, SCREAM, and GEOS, eq. 3.3).

$$SLF_{cloud\,fraction} = \frac{\text{liquid cloud fraction}}{\text{liquid cloud fraction} + \text{ice cloud fraction}} \tag{3.2}$$

$$SLF_{mass\,fraction} = \frac{q_w}{q_w + q_i} \tag{3.3}$$

In Hoose et al. (2018), a similar approach was used to generate cloud top liquid fraction information from model data, deriving both $SLF_{mass\,fraction}$ and $SLF_{n^\circ\,pixel}$, the second to make the case study comparable to passive satellite sensor products; from this study, a warm bias at SLF ~ 0.5 of 1°C to 2°C was found for $SLF_{n^\circ\,pixel}$ compared to $SLF_{mass\,fraction}$.

When coarsening the spatial resolution in ICON, SCREAM, and GEOS, the gridboxes for mixed-phase clouds are selected as for Cloud_cci v3: the cloud top temperature must be between -40°C and 0°C; the $SLF_{mass\,fraction}$ (eq. 3.3) calculated at each latitude-longitude coordinate is averaged in each gridbox ($\overline{SLF_{mass\,fraction}}$), and included in the analysis only if the condition $0 < \overline{SLF_{mass\,fraction}} < 1$ is satisfied. To compare the DYAMOND models at different resolutions, $\overline{SLF_{mass\,fraction}}$ is collected on isotherms using for each gridbox its averaged CTT.

3.3.2.2. K-means clustering for the cloud type classification

For some datasets, k-means clustering is used. The clustering of clouds into groups called *regimes* has been a recurring research theme over the last few decades. Some work is based on vertical velocity (Norris and Weaver, 2001; Bony et al., 2004), on sea level pressure (Tselioudis et al., 2000), or on sea

surface temperature and large-scale vertical motion (Ringer and Allan, 2004; Williams et al., 2003, 2006), some others are based on CTP, COT, and total cloud cover (Jakob and Tselioudis, 2003; Rossow et al., 2005; Williams et al., 2005; Williams and Tselioudis, 2007; Oreopoulos et al., 2017; Schuddeboom et al., 2018; Zhang et al., 2018). Recent studies have used clustering algorithms as simple and useful techniques to obtain cloud regimes from observational and model datasets; these algorithms include k-means clustering (Jakob and Tselioudis, 2003; Rossow et al., 2005; Williams et al., 2005; Williams and Tselioudis, 2007; Oreopoulos et al., 2017; Tan et al., 2018) and self-organising map (Schuddeboom et al., 2018; Zhang et al., 2018).

While the clouds from almost all datasets analysed in this work can be identified on a pixel-by-pixel basis using COT-CTP joint histograms, a different approach is required for Cloud_cci v3 and for the DYAMOND datasets at coarser spatial resolution. For each gridbox in Cloud_cci v3 and in SCREAM, a COT-CTP joint histogram is obtained by collecting the COT and CTP values of non-discarded pixels, and thus each gridbox-related COT-CTP joint histogram potentially contains the information of two or more pixels. The same approach is applied to the gridboxes of ICON and GEOS, considering histograms of only CTP distributions. A clustering technique is used to classify the gridboxes into cloud types or heights. Here, k-means clustering is used for reasons explained in the next paragraph.

The k-means algorithm (Anderberg, 1973) is a method used to sort data (points or vectors) into clusters resulting from the minimum sum of squared Euclidean distances between each cluster and its nearest cluster mean called centroid. Although this algorithm has limitations, these are well known and can be overcome to improve clustering results. These limitations occur, for example, when the distance between clusters is too large and therefore the centroids cannot move between clusters, or when stable clusters prevent the movement of centroids (e.g., high-density regions may accumulate more centroids than necessary, with the risk of losing less dense clusters). Furthermore, another difficulty in cluster analysis is the "correct" number of clusters into which the data should be sorted; in fact, this information is not known a priori. A better initialisation method and the iteration of the k-means with different initial solutions seem to solve the main limitations of this algorithm (Fränti and Sieranoja, 2019). Initialising 42 centroids for Cloud_cci v3 covers all possible COT-CTP combinations (excluding those for COT < 0.3 as one of the applied filters, the possible COT-CTP combinations are 6 times 7, from COT= $[0.3, 1.3, 3.6, 9.4, 23, 60, 379]$ and CTP= $[0, 180, 310, 440, 560, 680, 800, 1000]$ hPa bins in Fig. 3.5, respectively) and makes the k-means algorithm deterministic. In this way, the risk of neglecting some cloud types is prevented, clusters with low density are not lost unless they are empty, and there is no need to repeat the k-means clustering since the initial centroids would be the same, and hence the final clusters too. Furthermore, as gridboxes potentially contain more than one COT-CTP combination and include neighbouring clouds, the distance between these gridboxes is reduced compared to the distance between gridboxes with only one COT-CTP combination, and therefore centroids can move easily between clusters. Similarly to Cloud_cci v3, 7 prescribed initial centroids are used to cluster CTP histograms for ICON, SCREAM-CS, and GEOS (7 as the number of CTP bins), while 49 initial clusters (7

COT times 7 CTP combinations) are used for SCREAM as there is no reason to filter out bins with COT < 0.3 in this model.

The final clusters need to be merged to obtain the final 9 cloud types of Fig. 3.5 or to organise them into 3 cloud heights. For this, the most frequent COT-CTP combination in each of the resulting clusters is used to determine to which cloud type the clusters belong to. After this step, each of the 9 final COT-CTP joint histograms and 3 CTP histograms is individually normalised and represents a cloud type (see Sec. 3.3.2.3, eq. 3.5).

3.3.2.3. Geographical distributions and SLF-CTT(CTP)-r_{liq} dependency

Once the cloud classification is complete and the SLF has been computed, the geographical distribution, and both the SLF-CTT and SLF-r_{liq} joint histograms can be derived for each cloud type.

The geographical distribution of a cloud type H_{geo} is normalised pixel by pixel with the total number of available observational or model output time steps (including cloud-free pixels) for the specific latitude-longitude position:

$$H_{geo} = \frac{\sum_{i,j} \text{time step}_{ij} \text{ (cloud type)}}{\sum_{i,j} \text{time step}_{ij}}, \tag{3.4}$$

where i and j represent the position in latitude and longitude of the geographical distribution. If a particular location represented by a pixel were always covered by clouds, the sum of the frequency H_{geo} of all cloud types for that pixel would be 1. This normalisation allows to display the differences in cloudiness between regions and between cloud types.

In the SLF-CTT joint histograms, the SLF changing with CTT is shown for all cloud types, with an SLF bin width of 0.05 and a CTT bin width of 1°C. As the CTT is not provided in MCD06COSP, these histograms are replaced by SLF-CTP joint histograms, with CTP ranging from 0 hPa to 1000 hPa with a bin width of 25 hPa and keeping the SLF bin width of 0.05. These histograms have been normalised, for each cloud type, with their own total frequency of occurrence ($H_{k,l}$, eq. 3.5, with $N_{k,l}$ the absolute frequency of occurrence at bin combination (k,l), k referring to CTT — or CTP — and l to SLF), which shows the relative frequency of occurrence of SLF-CTT combinations for each cloud type, and using an approach similar to CFADs (Contoured Frequency by Altitude Diagrams, $H_{CFADk,l}$, eq. 3.6). CFAD histograms were introduced by Yuter and Houze (1995) and later used in other forms for other studies (e.g., Fu et al., 2003; Chen et al., 2016; Ewald et al., 2019; Shen et al., 2022). In CFAD histograms, the normalisation is applied at each altitude level. Using the CTT as a proxy for the cloud top height, a CFAD-like normalisation was applied for each CTT interval. The resulting histograms show the SLF distribution at each CTT and correspond therefore closely to the CFAD histograms.

$$H_{k,l} = \frac{N_{k,l}}{\sum_{k,l} N_{k,l}} \tag{3.5}$$

$$H_{CFADk,l} = \frac{N_{k,l}}{\sum_k N_{k,l}} \tag{3.6}$$

The SLF-$\overline{r_{liq}}$ joint histograms show how $\overline{r_{liq}}$ varies with SLF in all cloud types, which can be used to interpret cloud dynamics and microphysics. These histograms have been normalised individually for each cloud type (similarly to eq. 3.5, with k referring to $\overline{r_{liq}}$) and show the frequency of occurrence of SLF-$\overline{r_{liq}}$ combinations.

To highlight differences in CTT (or CTP), SLF, and $\overline{r_{liq}}$ between NH and SH as well as land and ocean for the different cloud types, the difference of the joint histograms is calculated for specific cases. A significance test has been applied to the resulting joint histograms, showing dots where the bin combinations show a land/ocean or NH/SH difference that can be considered statistically significant. The significance test used is based on the p-value (Wilks, 1995), which indicates the probability that observed differences between groups are due to chance. The p-value has been computed in the highest resolution possible, i.e., for each bin combination and the eight surrounding bins. For this reason, no significance test can be applied to the bin combinations at the edges of the histograms, which therefore do not show any dots. The p-value ranges between 0 and 1: values close to 0 indicate that the observed difference is unlikely to be due to chance, whereas values close to 1 indicate no difference between the groups or differences due to chance. The closer the p-value is to 0, the more significant the difference between the two groups. The threshold value for significance can be arbitrary; in this work, a threshold value of 0.05 is considered (Fisher, 1925). Significance tests have often been misused in science, confusing statistically significant results with scientifically significant results (Dahiru, 2008; Amrhein et al., 2019). Being aware of this, and considering that the significance test cannot be applied to the edges of the histograms, the differences that pass the significance test are treated with special attention, without discarding a priori the differences that do not pass the significance test.

4. Comparison between AVHRR and CALIOP observations of the cloud phase[1]

Observations are essential to improve our knowledge of cloud systems. Polar-orbiting satellite sensors provide global information on clouds and can be divided into *active* and *passive*. As these sensors operate in different ways (see section 2.4), the cloud properties they provide may differ. In mixed-phase cloud research, special attention is given to the thermodynamical phase of clouds from $-40°$C to $0°$C, as in this temperature range clouds can consist of both liquid droplets and ice particles (see section 2.2.3). Furthermore, mixed-phase cloud research investigates how the cloud phase is related to microphysical parameters and cloud types, how it is distributed in clouds, and whether there are latitudinal or regional variations (e.g., Tan et al. 2014; Tan and Storelvmo 2019; Coopman et al. 2020, 2021; Yang et al. 2022). This chapter presents a comparison between datasets derived from passive (AVHRR, see section 3.1.1) and active (CALIOP, see section 3.1.2) polar-orbiting satellite sensors between $60°$ North and $60°$ South (hereafter "near-global"), based on measurements of cloud top phase and cloud top temperature (CTT), sorted in addition to different cloud types (see section 3.3) and stratified by latitude and marine/continental regions. This comparison will highlight possible biases between AVHRR and CALIOP datasets, as well as possible regional and cloud-type dependencies.

4.1. Temperature and phase bias

As a first step, possible cloud top temperature and phase biases between CALIOP and AVHRR-based datasets (i.e., Cloud_cci v3, Cloud_cci v2, and CLARA-A2) are investigated using collocated data (see section 3.3.1) for a more robust interpretation of the results presented in this chapter.

Figure 4.1 compares point-by-point the cloud top temperature retrieved by AVHRR-based datasets and CALIOP over the entire tropospheric temperature range and for three cloud top phase combinations: same phase, CALIOP retrieving ice while Cloud_cci v3, v2, and CLARA-A2 retrieving liquid, and vice versa. The contour lines indicate the areas (A, B, and C) where the frequency of occurrence per 1K x 1K bin is greater than 240 (i.e., greater than 0.01% of 2085991 cloudy pixels within the histogram in Fig. 4.1(a); the percentages change with the AVHRR-based dataset considered, as some bin combinations may be outside the histogram ranges). This threshold highlights areas of agreement and disagreement between the datasets, and separates the area where the sensors retrieve the same phase into two regions

[1]This section includes text and figures from Bruno, O., C. Hoose, T. Storelvmo, Q. Coopman, and M. Stengel, 2021: Exploring the cloud top phase partitioning in different cloud types using active and passive satellite sensors. *Geophysical Research Letters*, **48 (2)**, e2020GL089 863, https://doi.org/10.1029/2020GL089863.

(A and B). Because of the small contribution given by the collocated pixels with CALIOP retrieving ice and AVHRR-based datasets liquid (turquoise shaded pixels), there is no contour line for this phase combination. This figure shows a CTT bias of the AVHRR-based datasets compared to CALIOP mainly at lower temperatures (region B), and a frequent disagreement on phase (with AVHRR-based datasets retrieving ice where CALIOP retrieves liquid) in the mixed-phase temperature range (between -40°C and 0°C, region C).

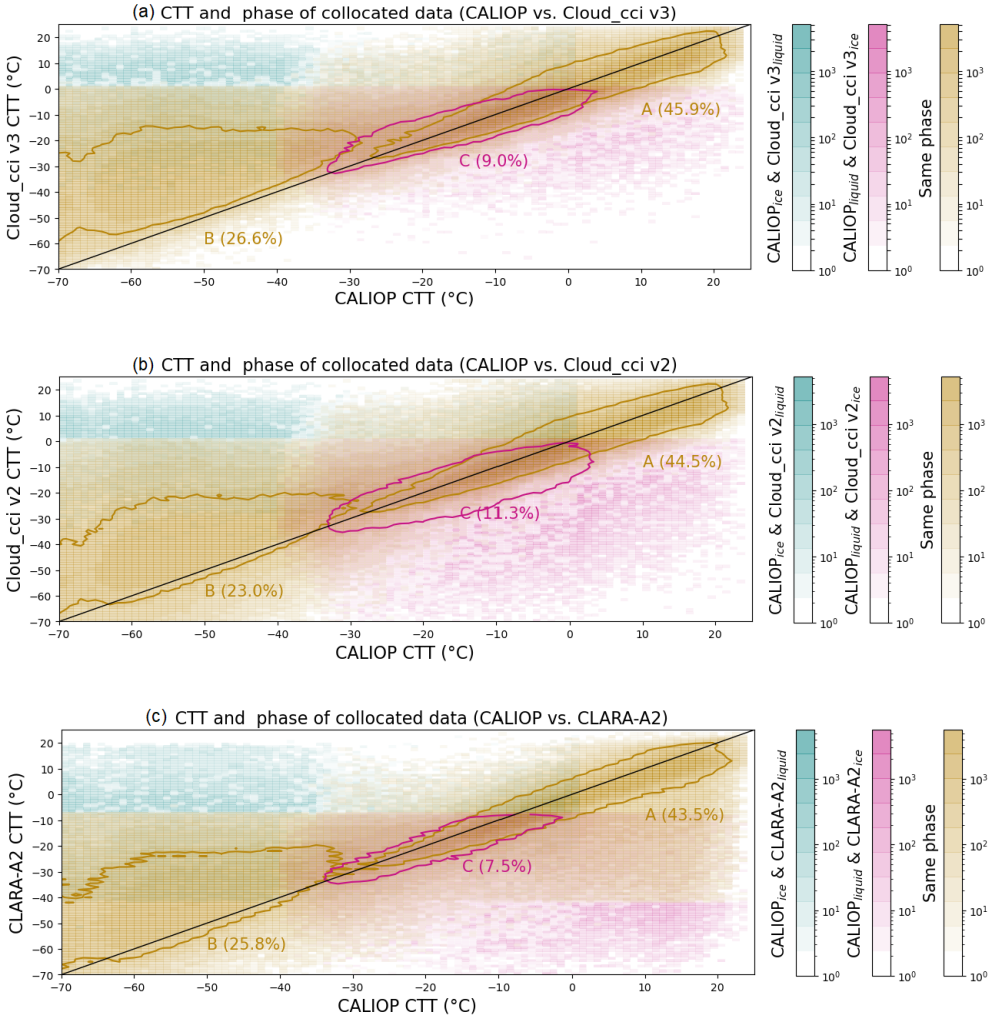

Figure 4.1.: Point-by-point comparison of CTT and cloud top phase for (a) CALIOP and Cloud_cci v3, (b) CALIOP and Cloud_cci v2, and (c) CALIOP and CLARA-A2 using the collocated data, where the brightness of the bins (1K x 1K) represents the absolute frequency of occurrence and the different colors represent the different combinations of retrieved phase. The contour lines encompass bins with a frequency greater than 240, while the percentages refer to the contoured areas and represent the relative amount of cases within the contour lines with respect to the total cloudy pixels in the histogram.

Region A, including part of the cases with the same phase, incorporates cases with good agreement in temperature and cases where the largest temperature difference is about 10°C. The percentage of cloudy pixels in region A is higher than those in B and C for all AVHRR-based datasets, representing

45.9% of total cloudy pixels within the histogram for Cloud_cci v3, 44.5% for Cloud_cci v2, and 43.5% for CLARA-A2. A clear warm bias in CTT from AVHRR-based datasets with respect to CALIOP is indicated by a portion of region B. This region, including pixels with the same phase, represents 26.6% of total cloudy pixels within the histogram for Cloud_cci v3, 23% for Cloud_cci v2, and 25.8% for CLARA-A2. Finally, region C, with 9% of total cloudy pixels within the histogram for Cloud_cci v3, 11.3% for Cloud_cci v2, and 7.5% for CLARA-A2, refers to pixels retrieved as liquid in CALIOP and as ice in AVHRR-based datasets. It includes, similar to region A, cases with good agreement in temperature and cases where the largest temperature difference is about 10°C. Nevertheless, region C contributes principally to a phase mismatch, discussed in more detail below.

The phase mismatch can also be seen in Fig. 4.2, where the geographical distribution of SLF is shown for non-collocated data of all datasets in the temperature range between −40°C and 0°C and on different isotherms: While CALIOP shows mainly liquid cloud tops, both Cloud_cci v2 and v3 show mainly ice cloud tops, while CLARA-A2 agrees better with CALIOP at warmer temperatures (see also Fig. 4.3).

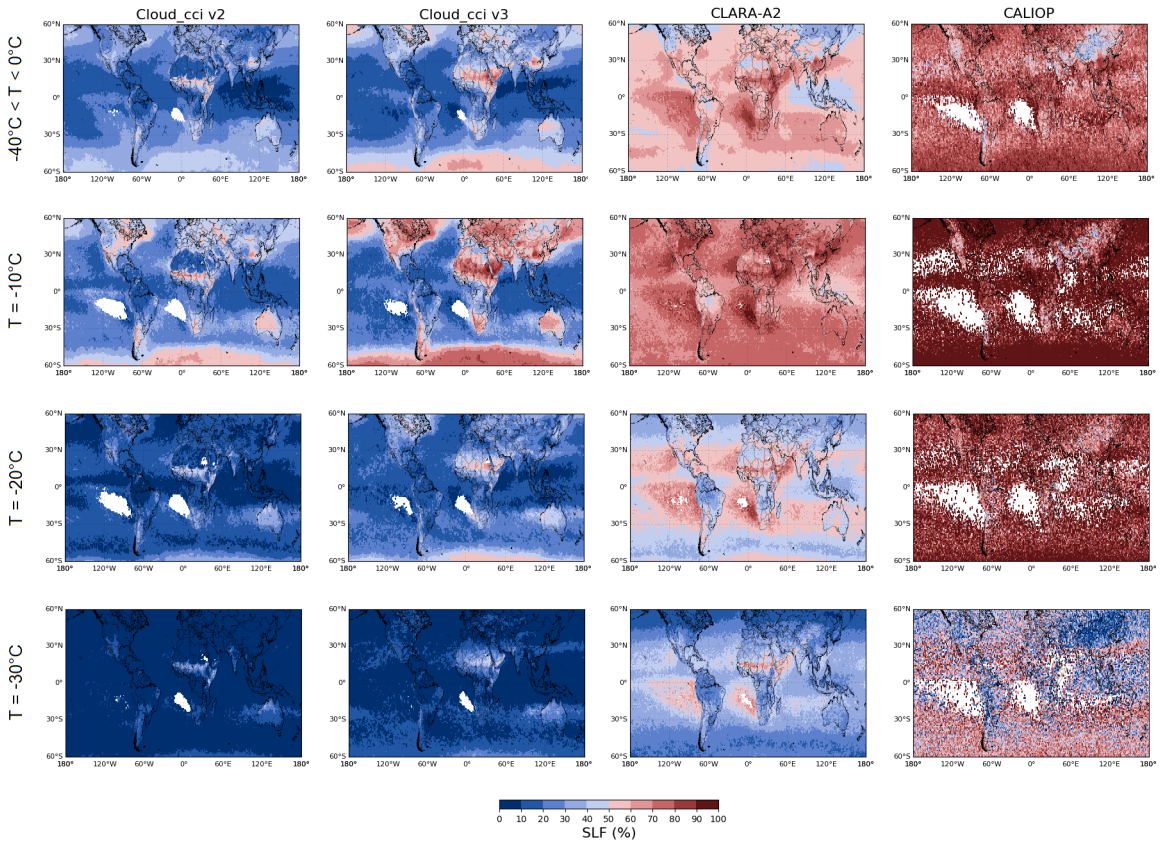

Figure 4.2.: Near-global geographical distribution of mean SLF within the temperature range from −40°C to 0°C (first row) and on the isotherms at −10°C, −20°C, and −30°C for the analysed datasets.

Discussion

Many factors can contribute to the discrepancies between CALIOP and AVHRR. One of the most important is the difference in the sensors, the former being passive and the second active: While the AVHRR has problems detecting multi-layered clouds with small COT, leading to misclassification of the cloud top phase, CALIOP can detect multilayer clouds with optical thicknesses up to 5 (Karlsson and Håkansson, 2018; Young et al., 2018), which could also lead to misclassifications. In the processing of the data, cloud edges were not excluded, as this was not possible for the AVHRR-based datasets, representing a possible source of misclassification for low- and mid-level clouds (Pavolonis et al., 2005). Moreover, a possible phase change of a detected cloud top would cause a modification of COT, and therefore a possible misclassification into an optically thicker or thinner cloud category, modifying the SLF of another cloud type. Some of these issues have also been presented in Cesana et al. (2019) for shallow cumulus and stratocumulus clouds, emphasising that errors in retrieving cloud phase, cloud optical thickness, and cloud top height can result in cloud type misclassifications. Furthermore, the filter we apply to the optical thickness may not be sufficient to ensure that the cloud data are treated in the same way. In Stengel et al. (2015), CALIOP's liquid cloud fraction resulted closer to the AVHRR-based dataset CLAVR-x (Cloud from AVHRR Extended) than to other AVHRR-based datasets. One reason for this was that the CLAVR-x algorithms used a priori information based on CALIOP climatologies for ice clouds. This in turn prevented phase and CTT from being retrieved independently.

4.2. Comparison of collocated and non-collocated datasets

After the comparison of collocated data, a comparison of near-global collocated and non-collocated data is presented in this section to address how the phase and temperature biases spread in 4-year analysis (from 1 June 2009 to 31 May 2013) in terms of SLF on isotherms.
Figures 4.3(a-b) show a comparison of collocated and non-collocated data, respectively, between Cloud_cci v2, CLARA-A2, Cloud_cci v3, and CALIOP.

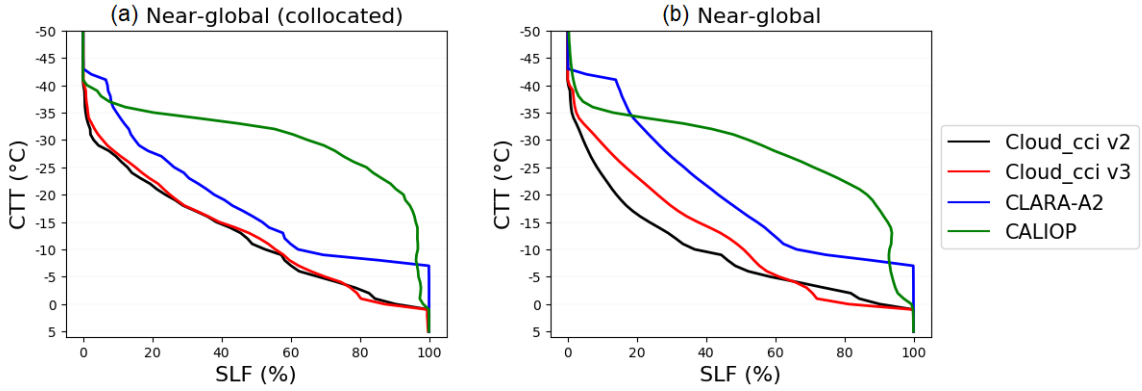

Figure 4.3.: Comparison of supercooled liquid fractions (SLFs) vs. cloud top temperature (CTT) for collocated (a) and non-collocated (b) near-global (60° N to 60° S) data. Different colors represent different datasets.

The difference in SLF and the associated CTT among the datasets stands out in these figures, and in particular the gap between the three AVHRR-based datasets and CALIOP, up to about 25°C or SLF of about 80% at a fixed temperature.

Discussion

The disagreement between datasets in Fig. 4.3 can be understood with figure 4.1, which shows both the CTT bias at lower temperatures and the frequent disagreement on phase in the mixed-phase temperature range (see section 4.1). Only a part of region A, where AVHRR-based datasets and CALIOP retrieves the same phase within 10° temperature difference, is within the mixed-phase temperature range (around 13% of total cloudy pixels within the histogram for Cloud_cci v3). Region B, albeit including only around 0.7% of total cloudy pixels within the histogram between -40°C and 0°C when comparing CALIOP and Cloud_cci v3, includes many cases contributing only to the SLF computation for Cloud_cci v3 (counting about 11% of total cloudy pixels within the histogram considering the mixed-phase temperature range for only Cloud_cci v3), which are excluded from the SLF computation in CALIOP because they are outside its mixed-phase temperature range. The percentage of cases in C changes only a little if considering solely temperatures between -40°C and 0°C (for example, from 9% to 8.7% for Cloud_cci v3). For AVHRR datasets retrieving liquid while CALIOP retrieving ice (turquoise region in Fig. 4.1), CLARA-A2 revealed more cases than Cloud_cci v2 and v3 for temperatures below -30°C and down to -41°C; For this reason, CLARA-A2 shows SLF around 7% between -40°C and -41°C using the collocated data in Fig. 4.3(a) and around 17% using the non-collocated data in Fig. 4.3(b). A quantitative analysis of the differences between CALIOP and Cloud_cci v2 and v3 can be found in Stengel et al. (2020): While any phase bias of Cloud_cci v2 and v3 with respect to CALIOP has nearly vanished for COTs of approximately 0.15 into the clouds, there is still a significant bias at COT $= 1$ for the cloud top height (CTH) of ice clouds, to which CTT is linked. As a consequence, CTH is usually retrieved from levels below the levels used to retrieve the phase, so that the retrieved CTT can be warmer than the effective temperature of the assigned cloud top phase, agreeing with our results. As the collocated data represent only about 9.5% of the CALIOP and 0.02% of the AVHRR non-collocated data, the temperature and phase biases found in the collocated data, which clarify the differences between the datasets in Fig. 4.3(a), contribute more to the differences in the non-collocated data (Fig. 4.3(b)).

4.3. Regional dependencies in AVHRR and CALIOP datasets

After the comparison of near-gobal collocated and non-collocated data, the SLF using non-collocated data is now presented in figure 4.4 to highlight possible differences between 30° and 60° North and South, and between marine and continental regions.

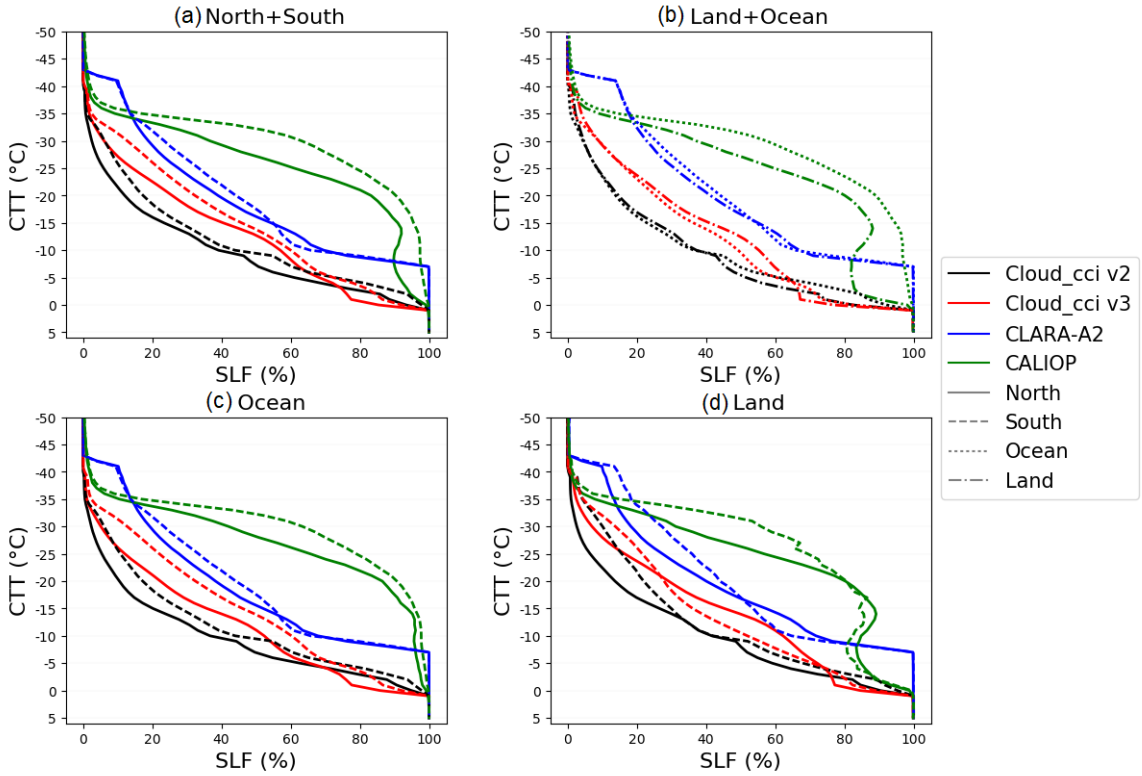

Figure 4.4.: Comparison of supercooled liquid fractions (SLFs) vs. cloud top temperature (CTT) for non-collocated data constraining the mid-latitude in the Northern Hemisphere (NH) and in the Southern Hemispheres (SH) (a), land and ocean (b), and mid-latitudes in NH and SH for only marine (c) and only continental (d) regions. Different colors represent different datasets; different line types represent different regions. Data analysed between 60° N and 60° S; mid-latitudes between 30° and 60°.

Figure 4.4(a) shows, for all datasets, larger SLF in SH than in NH (with an average difference for single datasets between 1.7% and 9.8%), in agreement with e.g. Tan et al. (2014) and Villanueva et al. (2021). This result might be explained by the larger size of the continental area and therefore the prevalence of continental aerosol with the ability to act as INPs in the NH. In Fig. 4.4(b), CALIOP shows clearly larger SLF over ocean than over land (with an average difference of 11.2%), but the AVHRR-based datasets do not agree with CALIOP for the entire temperature range. While larger SLF in SH than in NH is confirmed when constraining the analysis to marine pixels (Fig. 4.3(c), with an average difference for single datasets between 2.1% and 7.3%), it is confirmed only for specific temperature ranges over land (Fig. 4.3(d)), generally for T < −23°C. Further analyses using different cloud types were necessary to understand the origin of this last feature, principally due to the low-level clouds, which occur at warmer temperatures than the clouds at higher levels (see Section 4.4, Fig. 4.5).

4.4. Regional dependencies of cloud types in AVHRR and CALIOP datasets

For a better understanding of what may characterise the results presented in the previous section (4.3), this section provides a regional comparison of SLF for different cloud types.

Figure 4.5 shows the near-global SLF distribution for different cloud types and different datasets. The cloud types have been grouped into high-, mid-, and low-level clouds taking into account the temperature ranges that the datasets have in common at the three heights individually. This figure is derived from SLF-CTT distributions, for which the least frequent cases in CTT (frequency of occurrence lower than 2% with respect to the maximum of each distribution in temperature) have been filtered out. Similarly to Figures 4.3 and 4.4, in Fig. 4.5 the systematically lower SLF in AVHRR compared to CALIOP is found. A further outcome can be identified in this figure for every height-level: same-height clouds showed SLF increasing with COT. This is consistent in all datasets with a few exceptions.

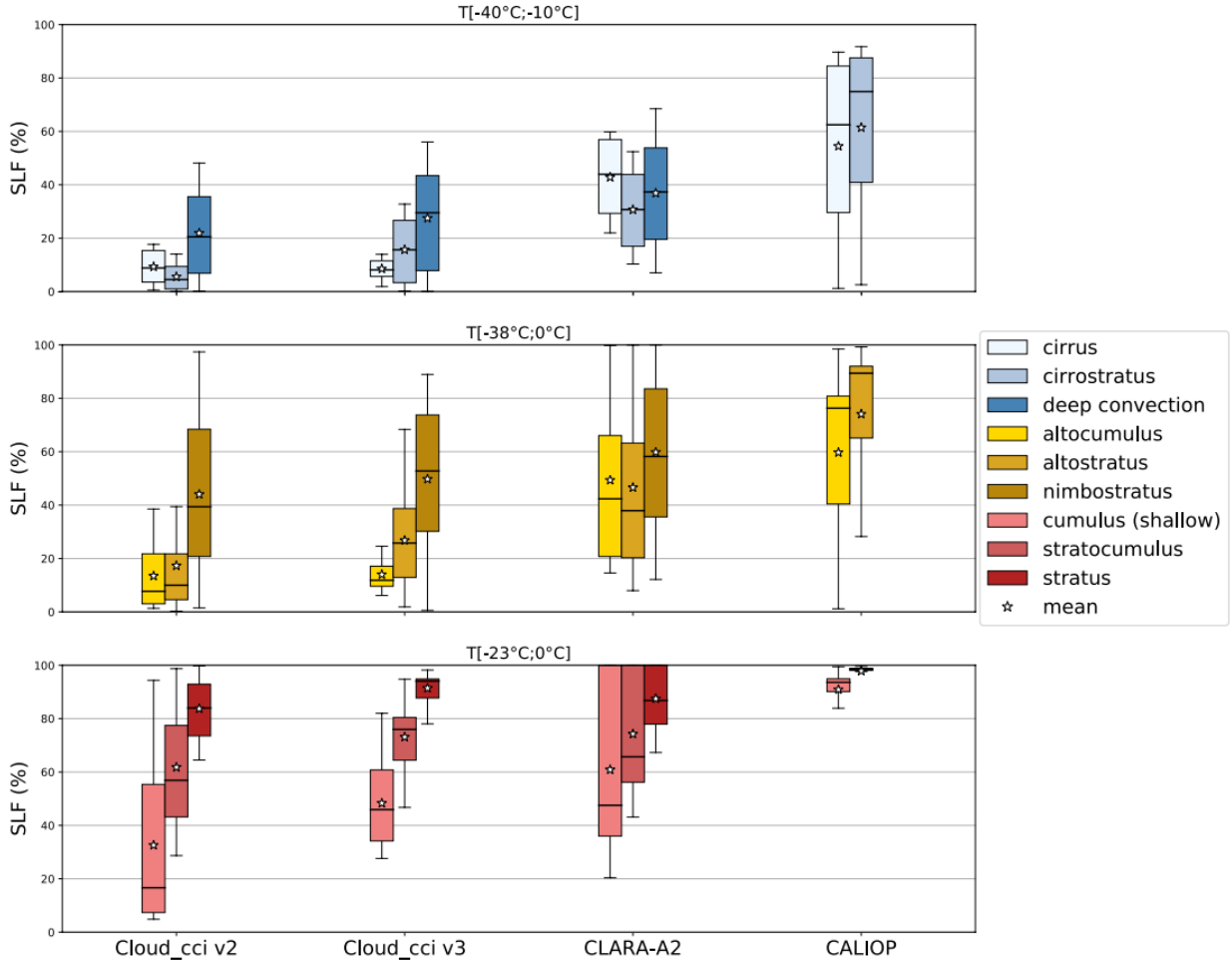

Figure 4.5.: Boxplot of the supercooled liquid fraction (SLF) for different cloud types sorted in three height levels. Clouds at the same height-level share the same cloud top temperature range, specified at the top of each panel. The different datasets are separated by columns and every color corresponds to one cloud type. The boxes extend from the lower to upper quartile values of the data, whereas the whiskers show the entire range of the data. The horizontal lines within the boxes represent the median of the distributions, while the stars represent their mean values.

Figure 4.6 condenses the SLF into average values for different cloud types, land and ocean, NH and SH. All datasets confirm generally larger SLF for optically thicker clouds at the same height-level.

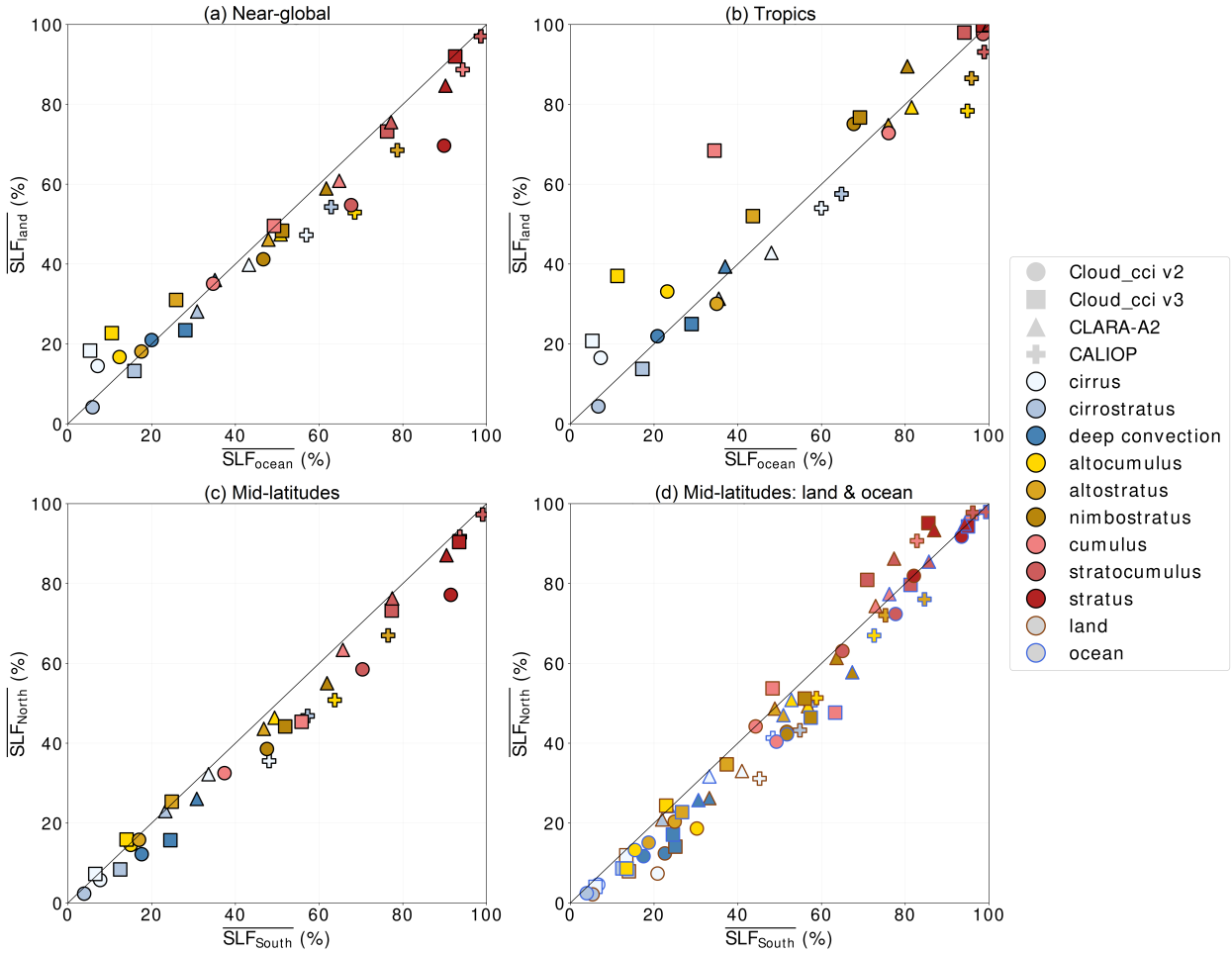

Figure 4.6.: Comparison of mean SLF for different cloud types, considered in the temperature ranges they have in common at the same height-level and for each subplot individually, for near-global marine and continental pixels (a), tropical marine and continental pixels (b), and mid-latitudes in NH and SH (c), with the further separation of marine and continental regions (d). Different markers identify different datasets, filling colors distinguish the cloud types, while edge colors refer to continental or marine pixels in (d).

Figures 4.6(a)-(c) reveal larger SLF over ocean than over land and in SH than in NH (in agreement with Coopman et al. 2020), respectively, for most of the cloud types with the exception of some high- and mid-level clouds in Cloud_cci v2 and v3, especially in Fig. 4.6(a). In the Tropics (Fig. 4.6(b)), most of the clouds show larger SLF over ocean than over land, except for again some high- and mid-level clouds in Cloud_cci v2 and v3, and low-level clouds in Cloud_cci v2. Finally, figure 4.6(d) shows that, in the mid-latitudes, separating the marine and continental pixels in NH and SH, SLF is larger in SH than in NH, with the exception of most of the low-level clouds over land (cumulus, stratocumulus, and stratus clouds), for which the opposite occurs. As continental clouds are separated from marine clouds for this plot, and the SLF is averaged over the isotherms common to both continental and marine clouds, the SLF values may be shifted compared to Fig. 4.6(c).

This study has been conducted using the collocated data too (figures A.3 and A.4), confirming the main results of the shown findings, although with more noise.

Discussion

Although clouds containing more droplets than ice particles result in higher optical thickness, we cannot exclude an influence of the cloud dynamics on both COT and SLF to explain the results in Fig. 4.5. For example, optically thicker clouds tend to have stronger updrafts and consequently higher supersaturation values, which may inhibit the glaciation process (Korolev, 2007), potentially lowering the glaciation temperature in clouds and causing the presence of more supercooled liquid water than ice. However, it is not possible to determine which process can explain the result obtained from this study.
As the continental area between 30° and 60° is less extensive in the SH than in the NH, a larger SLF in the SH than in the NH may be due to fewer INPs in the SH. However, continental low-level clouds show smaller SLF in the SH than in the NH. Considering that the common temperature range of the analysed continental low-level clouds goes from -15°C to 0°C (Fig. A.1), the result presented in Fig. 4.6 shows agreements with Villanueva et al. (2020), that found lower ice content in clouds in NH than in SH for T = $(-15° \pm 6°)$C, possibly because of the larger amount of feldspar in the SH. Another explanation for it could also be the higher density of particles acting as CCN in the NH, resulting in smaller droplet sizes, which might limit secondary ice formation (Mossop, 1980). These speculations are partially supported by further previous studies: Some anthropogenic aerosols such as black carbon, sulfate, and organic aerosols, do not act as efficient INPs but are efficient CCNs (Hoose and Möhler, 2012); model outputs have shown that sulfate aerosol and black carbon have the highest mass concentration in the lower troposphere of the NH (Liu et al., 2009), where they act as CCN (Boucher and Lohmann, 1995), whereas they act as INPs only at very high altitudes over the Tropics and the polar regions (Liu et al., 2009). Indeed, Tan et al. (2014) found that dust (as mineral desert dust), polluted dust (as dust mixed with urban pollution and biomass-burning smoke), and smoke (as biomass-burning aerosols, principally made of soot and organic carbon) are mainly distributed in the Tropics and in the NH. Synoptic conditions experienced by low-level clouds in these regions may also impact their thermodynamic phase.

4.5. Conclusions

This chapter presented a validation study comparing datasets derived from passive (AVHRR-based) and active (CALIOP-based) satellite sensors. The results presented in this chapter are summarised here:

- Collocated data show a warm bias of the AVHRR CTT compared to CALIOP and a phase mismatch (some liquid cloud tops in CALIOP are retrieved as ice in AVHRR-based datasets). This may be due to the different sensors, one passive and one active.

- For clouds with temperatures from -40°C to 0°C at the same height-level, the supercooled liquid fraction (SLF) increases with cloud optical thickness. This result may depend on different processes: clouds having more droplets than ice particles (i.e., higher SLF) result in higher opti-

cal thickness; optically thicker clouds may have stronger updrafts, which may inhibit glaciation processes, thus causing the presence of more supercooled liquid water than ice particles.

- The supercooled liquid fraction is higher in the Southern Hemisphere than in the Northern Hemisphere between 30° and 60°, probably due to fewer INPs in the SH, except for low-level continental clouds where the opposite is true. Several reasons may explain this result: the presence of specific aerosols acting as INPs in the SH over continents; the higher density of particles acting as CCN in the NH, resulting in smaller droplet sizes, which may inhibit secondary ice processes; synoptic conditions.

- Despite the phase and temperature mismatch, the passive and active satellite retrievals are in qualitative agreement regarding SH-NH, land-ocean, and cloud type dependencies of the cloud top phase, showing that the passive satellite sensor AVHRR can contribute to the cloud-phase research once its limitations have been taken into account.

5. Comparison between observational and model datasets

Weather and climate models are used to reproduce and predict atmospheric processes such as cloud formation and evolution. However, mixed-phase clouds are still not well represented because they are not fully understood (Forbes and Ahlgrimm, 2014; Cesana et al., 2015; McCoy et al., 2016). Observations provide a great opportunity to explore clouds and identify features that can be simulated by models. In this chapter, a comparison between observational and model datasets highlights how well the cloud phase distribution is represented in models, and shows the current limitations of models in simulating cloud microphysics. In Chapter 4 the potential of passive satellite sensors in cloud physics research was presented. Because of this so far underused potential, Cloud_cci v3 and MCD06COSP, derived from the passive sensors AVHRR and MODIS respectively, were chosen for this study: while Cloud_cci v3 is the most recent AVHRR-based dataset used for the study presented in Chapter 4, MCD06COSP allows a direct comparison between observations and models, as the COSP simulator for MODIS has been implemented on CAM6-Oslo, one of the model datasets included in this study (see sections 3.1.3 and 3.2.1.1 for MCD06COSP and COSP, respectively). The model datasets used for comparison are CAM6-Oslo, ICON, SCREAM, and GEOS. The comparison between the datasets is based on the fraction of all and mixed-phase clouds, their COT-CTP joint histograms (see section 3.3.0.1) in which the clouds are sorted according to Rossow and Schiffer (1999) or CTP-based histograms when COT is not available, their geographical distribution, and their SLF-CTT and CFAD-like joint histograms (see section 3.3.2.3). In this study, clouds are considered mixed-phase if the temperature is between -40°C and 0°, and contain both liquid and ice phases (see section 3.3.2.1). Except for Cloud_cci v3, all datasets provide pixel-wise information on both liquid and ice cloud fraction or mass mixing ratio, allowing SLF to be computed pixel-wise (see section 3.3.2.1). Therefore, the results for Cloud_cci v3 consist of gridboxes of 6x6 pixels in which SLF can be computed. As there are no significant differences in the results when comparing gridboxes with at least two or five cloudy pixels (unless mentioned), the results associated with gridboxes with at least two cloudy pixels are shown in this work. For a more consistent comparison between the datasets, this chapter also includes the analysis at 0.5° resolution of ICON, SCREAM, and GEOS. A sensitivity study analysing ICON, SCREAM, and GEOS at coarse resolutions of 0.5°, 1°, and 2° is also included in this chapter. In the datasets with data collected in gridboxes, k-means clustering is used for cloud type classification (3.3.2.2). For a better interpretation of the results shown in Chapter 4, the comparison between datasets presented here focuses on latitudes between 60°N and 60°S, and more specifically on the land-ocean and NH-SH (between 30° and 60°) differences.

To make this chapter more readable, the figures show only a subset of all cloud types and are sometimes limited to one observational dataset and one model dataset, while the full set of plots related to this chapter can be found in the Appendix.

5.1. Differences in the near-global amount of all and mixed-phase clouds

The use of different sensors and models can lead to differences in the amount of clouds. The total amount of all and mixed-phase clouds retrieved from the datasets between 60° North and 60° South is summarised in Table 5.1. The total amount of all and mixed-phase clouds is defined by the Relative Frequency of Occurrence (RFO) for all clouds (RFO_{all}) and mixed-phase clouds (RFO_{MP}) as follows:

$$RFO_{all(MP)} = \frac{\text{n°(mixed-phase) cloudy pixels}}{\text{n°[cloudy + invalid cloudy + clear sky] pixels}} \tag{5.1}$$

where the pixels are replaced by gridboxes in Cloud_cci v3.

	Cloud_cci v3	MCD06COSP	CAM6-Oslo	SCREAM	SCREAM-CS	ICON	GEOS
all	(59.2%) 52.9%	87.3%	71.7%	81.3%	70.7%	65.5%	72.1%
MP	6.2%	55.0%	16.7%	3.1%	3.4%	6.9%	14.7%
MP/all	11.7%	63.0%	23.3%	3.5%	4.8%	10.5%	20.4%

Table 5.1.: Total amount of all ("all") and mixed-phase ("MP") clouds retrieved by all datasets included in this study between 60° North and 60° South, from 1 June, 2009 to 31 May 2013 for Cloud_cci v3, MCD06COSP, and CAM6-Oslo and for February 2020 for ICON, SCREAM(-CS), and GEOS. In the third row, "MP/all" indicates the amount of mixed-phase clouds relative to the amount of all clouds. In Cloud_cci v3 the amount of cloudy boxes and mixed-phase boxes is provided; a pixel-wise amount of "all" clouds is also provided in parenthesis.

The main results from Table 5.1 are:

- The total cloud amount is more consistent between the datasets than the amount of mixed-phase clouds, which is partly due to resolution and the definition of "mixed phase";
- The amount of all clouds retrieved by Cloud_cci v3 decreases from 59.2% to 52.9% when gridboxes are used;
- Cloud_cci v3 retrieves the least amount of all clouds (52.9%);
- MCD06COSP retrieves the highest amount of both all (87.3%) and mixed-phase (55.0%) clouds;
- SCREAM shows the lowest amount of mixed-phase clouds (3.1%);
- CAM6-Oslo, ICON, SCREAM-CS, and GEOS capture a near-global amount of "all" clouds ranging from 65.5% to 72.1% close to previous satellite studies (from 61.9% to 67.2% for afternoon satellites; PVIR, 2020), while SCREAM and MCD06COSP overestimate it;

- The amount of all and mixed-phase clouds closest to the observations of Cloud_cci v3 is provided by ICON.

Discussion

Cloud_cci v3 provides the least amount of clouds, probably partly due to the limited number of channels used by the AVHRR sensor, and partly due to the regridding process, which involves the application of several filters (e.g., gridboxes are considered cloudy only if the difference between the warmest and the coldest cloudy pixels is within 20°C). In fact, the total number of cloudy pixels in Cloud_cci v3 is 59.2%, while the percentage drops from 52.9% to 48.2% for cloudy boxes and from 6.2% to 6.1% for mixed-phase boxes when gridboxes with at least 5 cloudy pixels are used for analysis instead of 2 cloudy pixels. Cumulus clouds (Table 5.2) are the largest contributor to the decrease for all cloudy boxes (Table 5.1), with 18.0% using gridboxes with at least 2 cloudy pixels and 14.9% using gridboxes with at least 5 cloudy pixels.

MCD06COSP has the highest fraction of both all (87.3%) and mixed-phase (55.0%) clouds. Although this may be explained by a higher sensitivity of the retrieved data due to a larger number of channels than those used by AVHRR, the temporal and spatial averaging of MCD06COSP may also contribute to the cloud amounts in Table 5.1. In fact, Aqua and Terra measurements are combined to produce MCD06COSP, giving as cloudy also locations where only one of the satellites detects clouds, and averaged in space to a resulting spatial resolution of 1°.

SCREAM shows the lowest amount of mixed-phase clouds (3.1%). This result may be due to the higher spatial resolution compared to other datasets (the spatial resolution dependence of the DYAMOND datasets is further discussed in section 5.4). Since SCREAM does not show the lowest amount of clouds also for all clouds, the lowest amount of mixed-phase clouds may be due to a very efficient Wegener-Bergeron-Findeisen (WBF) process (see section 2.2.3).

Between 2003 and 2010, a near-global (between 60° North and South) average cloud cover ranging from 61.9% to 67.2% has been retrieved by different afternoon satellite sensors (PVIR, 2020). Only CAM6-Oslo, ICON, SCREAM_CS, and GEOS come close to this percentage considering all clouds between 60° South and 60° North, while SCREAM and MCD06COSP overestimate it. The possible reasons for the overestimation of MCD06COSP have already been discussed. In SCREAM, this may be due to the method used to detect the top of the clouds as the highest level with COT > 0, which limited the analysis to the simulated daytime clouds (as in Cloud_cci v3, MCD06COSP, and CAM6-Oslo), excluding the nighttime clouds.

	2-pixels	5-pixels
cirrus	1.1%	1.0%
cirrostratus	4.7%	4.1%
deep convective	1.4%	1.9%
altocumulus	1.4%	1.2%
altostratus	6.3%	6.0%
nimbostratus	0.2%	0.4%
cumulus	18.04%	14.9%
stratocumulus	18.5%	17.5%
stratus	1.3%	1.2%

Table 5.2.: Total amount of all ("all") cloudy boxes retrieved by Cloud_cci v3 between 60° North and 60° South, from 1 June, 2009 to 31 May 2013 for different cloud types using gridboxes with 2 and 5 cloudy pixels. K-means clustering has been used for the cloud-type classification (see Sec. 3.3.2.2).

Furthermore, only CAM6-Oslo, ICON, and GEOS show close values for mixed-phase clouds with an average of 17.9%, while MCD06COSP and Cloud_cci retrieved mixed-phase clouds of 55.0% and 6.2%, respectively. For selected regions and cloud types, previous studies based on observations have estimated the amount of mixed-phase clouds to be between $\sim$ 30% and 60% of the total cloud cover (metric comparable to "MP/all" in Table 5.1). However, the measurements of these observations were made in the Arctic (Shupe et al., 2006), in the former Soviet Union (Mazin, 2006), over the Southern Ocean (from 30° to 75° S and from 50° to 92° W, D'Alessandro et al. 2019), and globally (up to 82°, but only for midlevel stratiform clouds with liquid top, Zhang et al. 2010).
Future studies can benefit from this work, as it adds new references to the existing literature on the amount of mixed-phase clouds, which includes all cloud types up to 60° latitude.

5.2. Amount and geographical distribution of all cloud types

To better understand the differences found in the amount of all clouds shown in section 5.1, this section focuses on the geographical distribution of the different cloud types.
Figure 5.1 shows the cloud classification of all clouds using COT-CTP joint histograms and their geographical distributions for Cloud_cci v3. The other datasets are shown in figures: A.5 for MCD06COSP, A.6 for CAM6-Oslo, A.7 for SCREAM. The COT-CTP joint histograms are replaced by a cloud top height classification for GEOS, ICON, and SCREAM-CS (Figure A.8).
The datasets providing COT mostly agree on the cloud types with the highest RFO for the different heights: stratocumulus for low-level clouds, altostratus for mid-level clouds, and cirrostratus (cirrus for SCREAM) for high-level clouds. This is shown in Fig. 5.1 for Cloud_cci v3. The datasets show no specific bias in COT or CTP.

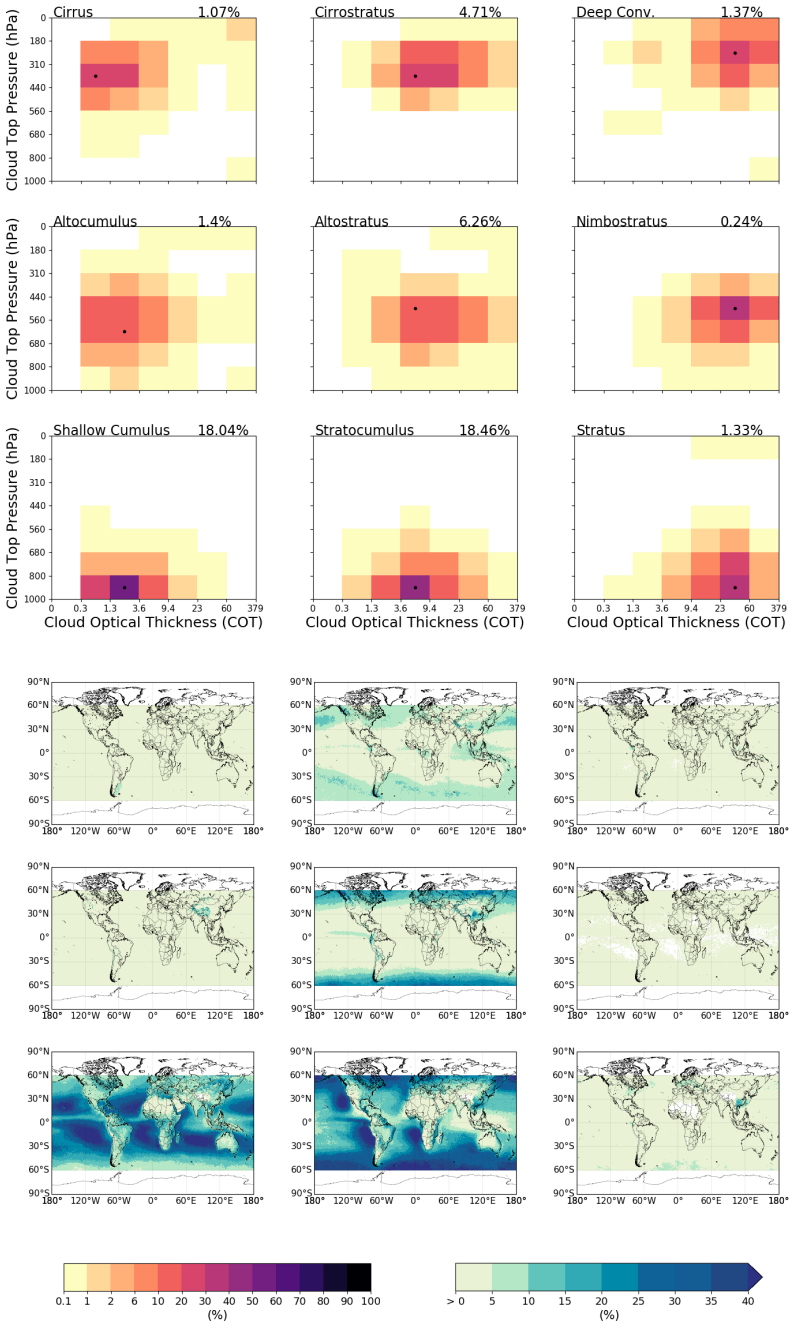

Figure 5.1.: This figure is composed of two parts. The upper panel shows COT-CTP joint-histograms for different cloud types in a 3x3 matrix for Cloud_cci v3. Each cloud category is identified by a name in the upper left corner of the histogram, while its Relative Frequency of Occurrence (RFO) is shown in the upper right corner. Each histogram is normalized to 1 and the single dot indicates the most frequent COT-CTP combination. The lower panel shows the geographical distribution of the different cloud types in the upper panel and is also shown in a 3x3 matrix for one-to-one correspondence with the upper panel. Each pixel in the geographical distributions is normalized to the total number of measurements taken in that pixel.

In Cloud_cci v3, the need to introduce gridboxes to investigate SLF for all cloud types leads to possible contamination of gridboxes with cloud types from neighbouring COT-CTP joint histograms. As

gridboxes are not required for the analysis of the other datasets, this contamination is only present in Cloud_cci v3. Examples of COT-CTP joint histograms for some cloud types, comparing Cloud_cci v3 and MCD06COSP, are shown in Fig. 5.2. This figure also shows that the geographical distribution of clouds at the same height differs depending on the dataset: for example, most of the marine cumulus clouds in Cloud_cci v3 are not retrieved as cumulus clouds by MCD06COSP but may be included in the stratocumulus clouds, while the mid-latitude stratus clouds retrieved by MCD06COSP are not retrieved as stratus clouds by Cloud_cci v3 but may be included in the stratocumulus clouds.

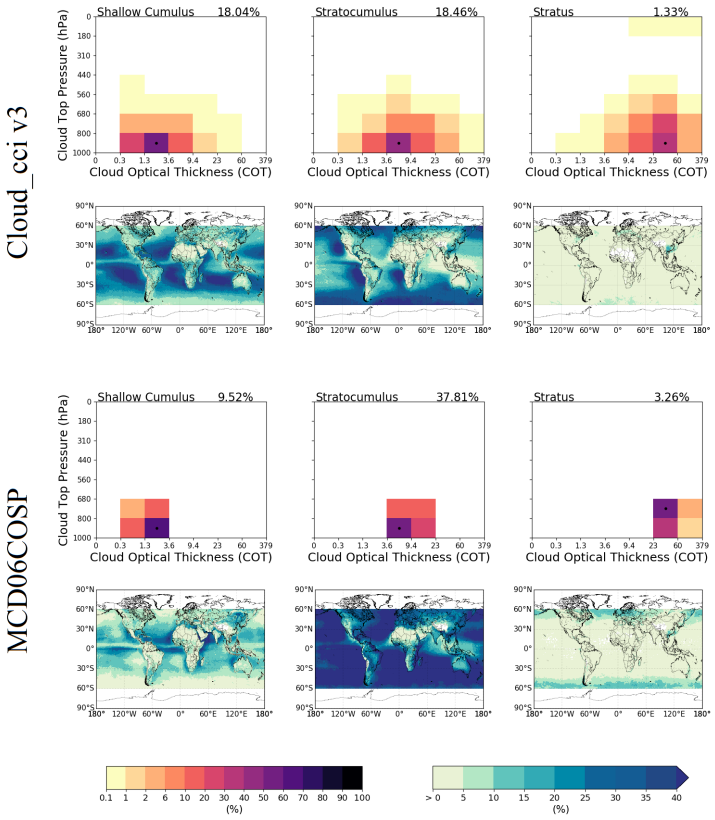

Figure 5.2.: COT-CTP joint-histograms and geographical distributions of low-level clouds (cumulus, stratocumulus, and stratus) for Cloud_cci v3 (upper panel) and MCD06COSP (lower panel). At the top of each histogram, the Relative Frequency of Occurrence (RFO) is shown.

To better compare the geographical distribution of clouds and their RFO of the analysed datasets, the cloud types at the same height are combined in Fig. 5.3 to obtain high-, mid-, and low-level clouds. The RFO (eq. 5.1) of the different clouds varies significantly between the datasets. While the satellites retrieve more low-level clouds (RFO = 37.83% and 50.59%) than mid- (RFO = 7.9% and 19.3%) and high-level clouds (RFO = 7.15% and 17.41%), the model results show a majority of high-level clouds (RFO between 27.17% and 52.06%). The datasets mostly agree on the regions where clouds at certain heights occur.

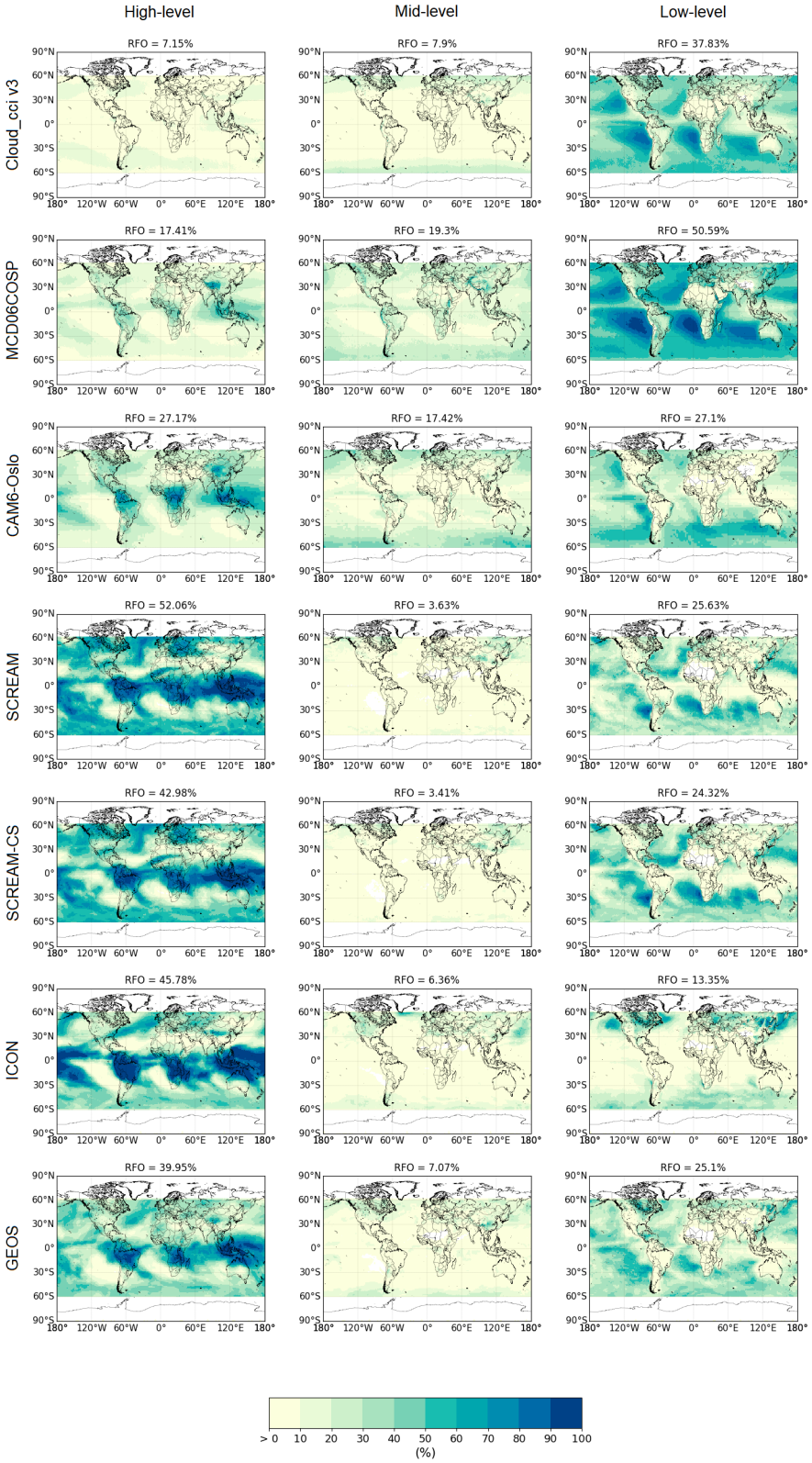

Figure 5.3.: Geographical distribution of all clouds for the different datasets combined at three height-levels: high-, mid-, and low-level. The Relative Frequency of Occurrence (RFO) is shown above each distribution. Each pixel in the geographical distributions is normalised to the total number of measurements taken in that pixel. Data from the DYAMOND datasets are represented at 1° spatial resolution for better visualization.

However, the geographical distributions also show that the same cloud type can be found in different locations depending on the dataset: for example, low-level clouds are more widely distributed over the western sides of the oceans in Cloud_cci v3 and MCD06COSP than in the models (e.g., in the Pacific Ocean near South America), ICON shows mid-level clouds in the Bering Sea, not present in the other datasets, and high-level clouds are present on the eastern side of the Pacific Ocean near North America in SCREAM, SCREAM-CS, and GEOS, with lower frequency in Cloud_cci v3 and MCD06COSP, but these are absent in CAM6-Oslo and ICON.

Discussion

The geographical distributions in Fig. 5.3 show that low-level clouds are mostly marine clouds while mid- and high-level clouds are found over both continents and oceans, consistent with other studies (Rossow et al., 2005; Williams and Tselioudis, 2007; Oreopoulos et al., 2017; Tan et al., 2018). In particular:

- High-level clouds are mainly found in the extratropics and over the equator, reaching a wider area that can cover the tropics, depending on the dataset. These areas are often subject to synoptic- and mesoscale events, i.e., tropical and extratropical cyclones, consisting mostly of cirrus and cirrostratus in the upper levels.

- Mid-level clouds, such as altostratus and altocumulus, can have different origins. Although they can form independently (e.g., altocumulus as high-based convective clouds), they can also form as remnants of cumulonimbus and nimbostratus clouds, part of the clouds that constitute cyclones. For this reason, they are mainly found in the same regions as high-level clouds.

- Low-level clouds are found mainly over the oceans, and especially over their eastern edges, where their formation is made possible by the atmospheric subsidence above a low sea surface temperature.

The differences in cloud locations between datasets, together with the other differences found in the RFO, suggest a possible misclassification of clouds using the observations, further showing the limitations of sorting clouds with simple COT-CTP joint histograms (see section 3.3.0.1, Hahn et al., 2001), and a possible wrong representation of the geometric structure of clouds in models.
Figure 5.2 shows that the analysed datasets can differ in the COT assigned to clouds of the same height, as well as in the cloud amount. The MCD06COSP retrieval for cloud optical properties (including COT), uses only one channel and this may contribute to misclassification of cloud types when considering clouds at the same height.

5.3. Amount and geographical distribution of mixed-phase cloud types

Figure 5.4 shows the cloud classification in COT-CTP joint histograms of mixed-phase boxes (see section 3.3.2.1) and their geographical distributions for Cloud_cci v3.

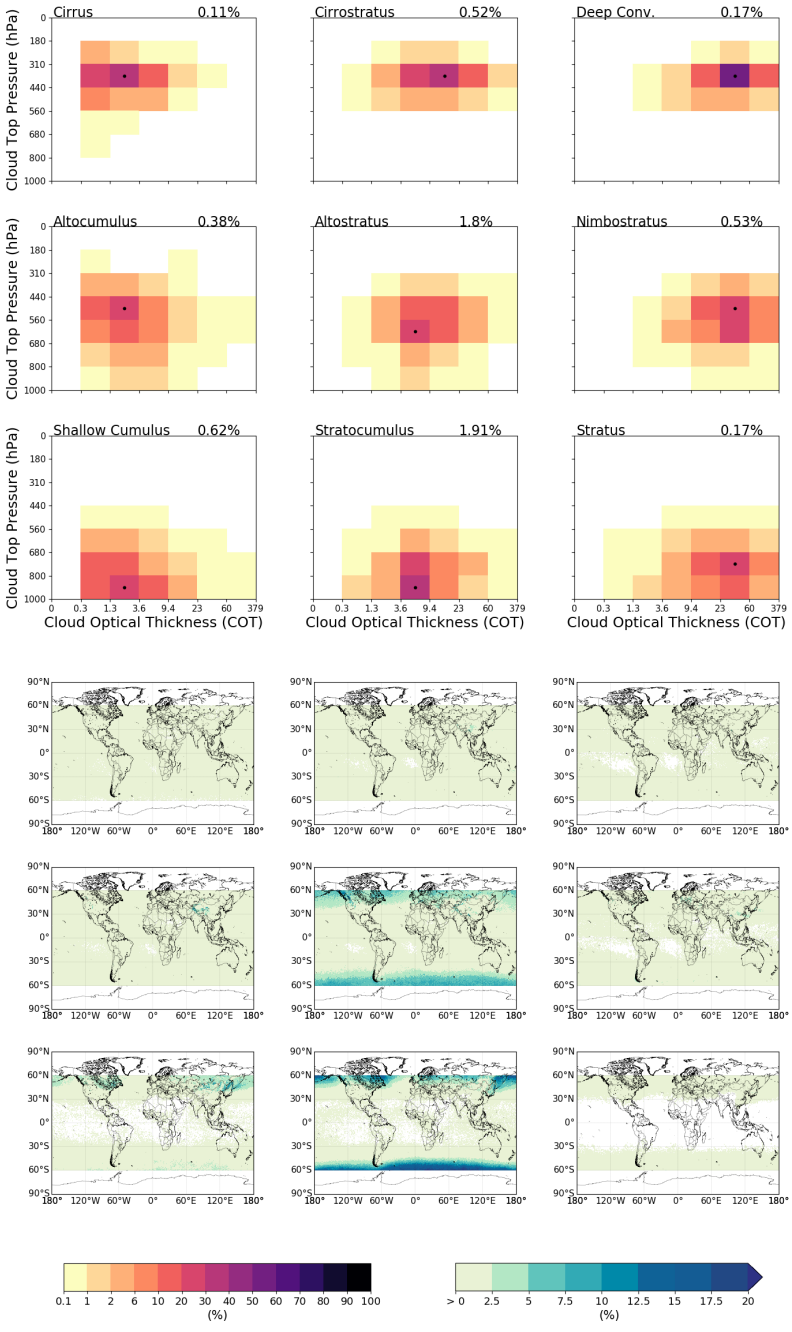

Figure 5.4.: This figure is composed of two parts. The upper panel shows COT-CTP joint-histograms for different cloud types in the mixed phase in a 3x3 matrix for Cloud_cci v3. Each cloud category is identified by a name in the upper left corner of the histogram, while its Relative Frequency of Occurrence (RFO) is shown in the upper right corner. Each histogram is normalized to 1 and the single dot indicates the most frequent COT-CTP combination. The lower panel shows the geographical distribution of the different cloud types in the upper panel and is also shown in a 3x3 matrix for one-to-one correspondence with the upper panel. Each pixel in the geographical distributions is normalized to the total number of measurements taken in that pixel. The colorbar for the geographical distribution has a smaller range than in figures including all clouds.

The other datasets are shown in figures: A.9 for MCD06COSP, A.10 for CAM6-Oslo, A.11 for SCREAM. The COT-CTP joint histograms are replaced by a cloud top height classification for GEOS, ICON and SCREAM-CS (Figure A.12).

As in section 5.2, COT-CTP joint histograms of Cloud_cci v3 may be contaminated by gridboxes with cloud types from neighbouring COT-CTP joint histograms. Stratocumulus clouds are the most abundant low-level mixed-phase clouds, altostratus clouds are the most abundant mid-level mixed-phase clouds, and cirrostratus clouds are the most abundant high-level mixed-phase clouds. Except for ICON and GEOS, which do not provide COT information, the other datasets agree with Cloud_cci v3 on this result (except for SCREAM, for which cirrus clouds are the most abundant high-level mixed-phase clouds).

All models systematically show an increase in CTP (and consequently a decrease in cloud top height) for the mixed-phase high-level clouds compared to all clouds. This is shown for cirrostratus clouds in CAM6-Oslo by the black dots in Fig. 5.5, as these represent the most abundant high-level cloud type found in the mixed-phase, in Figures A.7 and A.11 for SCREAM, and in Figures A.8 and A.12 for the other DYAMOND models.

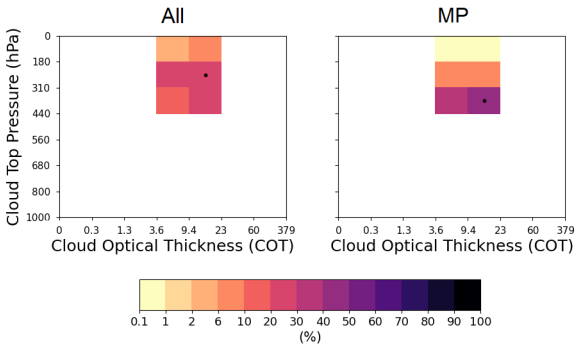

Figure 5.5.: CTT-CTP histograms for all (left) and mixed-phase (MP, right) cirrostratus clouds in CAM6-Oslo. The histograms are normalised individually. The colorbar indicates the relative frequency of occurrence for a given CTT-CTP bin.

Figure 5.6 shows the geographical distributions of the mixed-phase clouds at three heights for the different datasets. The main contributors to mixed-phase clouds are low-level clouds. This result is consistent across all datasets. While MCD06COSP shows a negligible contribution in cloud fraction at high and middle levels when comparing the clouds in the mixed-phase regime to all clouds (Fig. 5.3), the other datasets show a reduction in cloud fraction at all heights for mixed-phase clouds. The datasets agree on the geographical distribution of mixed-phase clouds, which are mainly located in the mid-latitudes, with the exception of MCD06COSP, for which a high frequency of mixed-phase clouds is also found in the tropics, regardless of the cloud type. However, the geographical distributions also show differences between the datasets: for example, ICON, SCREAM, SCREAM-CS, and GEOS show low-level clouds over the maritime area around Japan which are not present in Cloud_cci v3, MCD06COSP, and CAM6-Oslo.

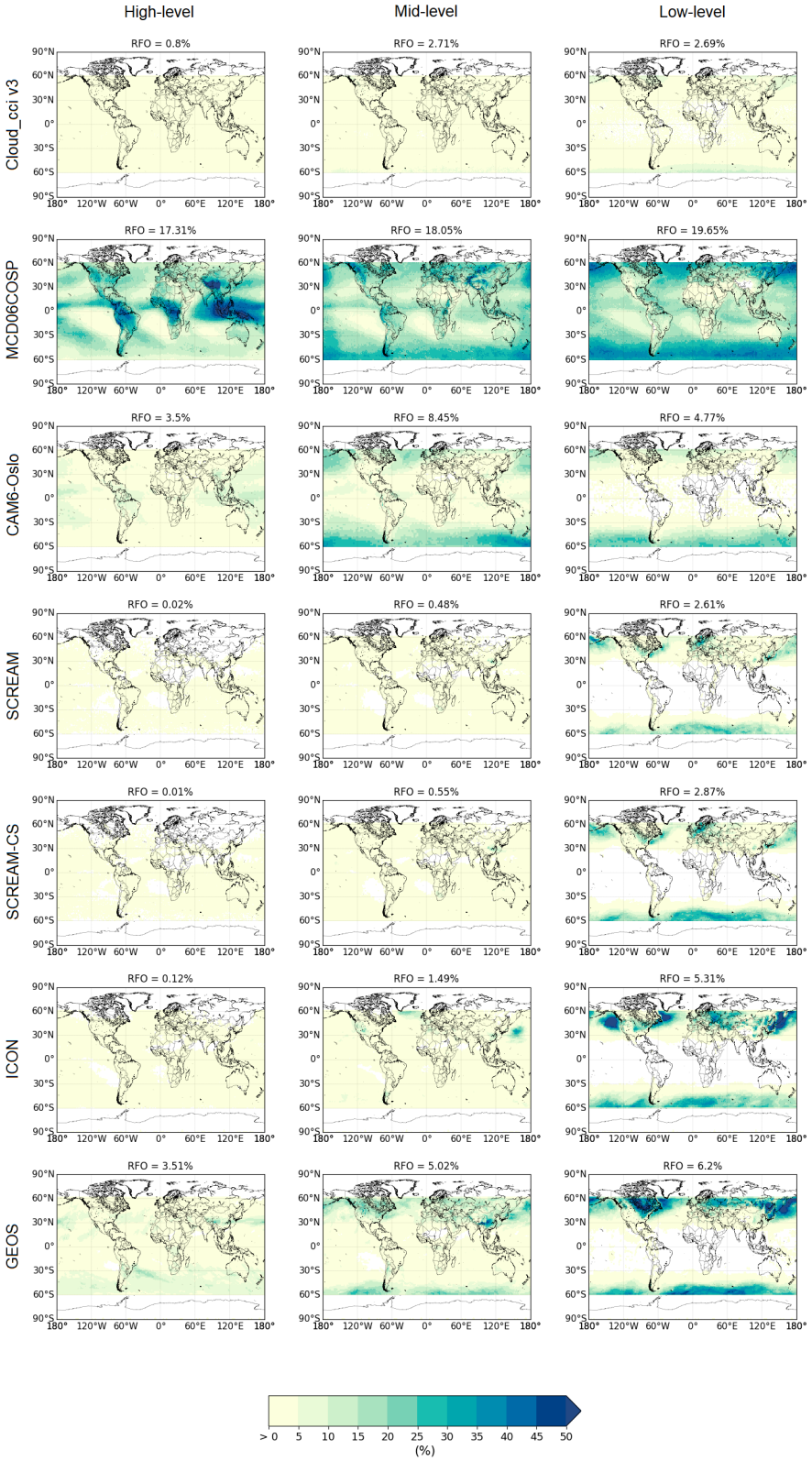

Figure 5.6.: Geographical distribution of mixed-phase clouds for the different datasets combined at three height-levels: high-, mid-, and low-level. The Relative Frequency of Occurrence (RFO) is shown above each distribution. Each pixel in the geographical distributions is normalised to the total number of measurements taken in that pixel. Data from the DYAMOND datasets are represented at 1° spatial resolution for better visualization. The colorbar has a smaller range than in Fig. 5.3.

Discussion

The reduction in cloud top height found for high-level mixed-phase clouds compared to all clouds, and hence the increase in cloud top temperature, may contribute to the presence of the liquid phase in high-level mixed-phase clouds in the temperature range between −40°C and 0°C. The presence of cirrus clouds in the mixed-phase regime may not be surprising, as previous studies have shown the presence of supercooled liquid droplets in this cloud type (Sassen, 1992).
The presence of mixed-phase clouds mainly at mid-latitudes is consistent with Hu et al. (2010), as supercooled liquid water has been observed mainly in both convective and stratiform clouds in these regions. However, MCD06COSP retrieves more mixed-phase clouds than the other datasets, which are also found in the tropics in areas where strong updrafts lead to the formation of high-level clouds with ice tops (e.g., the Intertropical Convergence Zone). This may indicate the presence of problems in the MCD06COSP cloud retrieval.
The differences between the datasets in the geographical distributions in Fig. 5.6 may imply different dynamics and therefore the possibility of regionally dependent cloud micro- and macrophysics within the same cloud type. As in section 5.2 for all clouds, the position of a cloud type in the mixed-phase varies between the datasets (e.g., only the DYAMOND models retrieve low-level mixed-phase clouds in the areas surrounding Japan, while MCD06COSP retrieves high-level mixed-phase clouds in the same area), again highlighting the limitations of the COT-CTP joint-histogram-based method (see section 3.3.0.1, Hahn et al. 2001) used to sort the clouds and a possible misclassification of clouds.

5.4. Comparison between coarsened spatial resolutions of DYAMOND datasets

Since the results from SCREAM and SCREAM-CS shown so far do not contain substantial differences, only SCREAM will be considered in the following sections, as SCREAM provides information on the COT, which allows the cloud type classification used for the observations.
In this paper, all analyses performed with ICON, SCREAM and GEOS are also shown at the coarser spatial resolution of 0.5° (ICON 0.5°, SCREAM 0.5° and GEOS 0.5°) to make comparisons with Cloud_cci v3 more consistent. However, due to the coarser resolution of MCD06COSP (1°) and CAM6-Oslo (∼2°), the DYAMOND models were also evaluated at the coarser spatial resolutions of 1° and 2°, and the results are presented in this section.
The cloud amount provided by SCREAM, ICON and GEOS changes with the coarser spatial resolution as shown in the table 5.3.
From the original to the coarser spatial resolution of 0.5°, the amount of all clouds decreases. For mixed-phase clouds, this only occurs for GEOS and SCREAM. Coarsening the spatial resolution from 0.5° to 1° and 2° decreases the amount of all and mixed-phase clouds for all DYAMOND datasets. While the "MP/all" ratio increases for ICON and GEOS with increasing spatial resolution, it decreases for SCREAM. Since the cloud amount decreases up to about 50% when the spatial resolution is coarsened

	SCREAM				ICON				GEOS			
	3 km	0.5°	1°	2°	5 km	0.5°	1°	2°	3 km	0.5°	1°	2°
all	81.3%	35.4%	29.2%	20.7%	65.5%	58.6%	44.1%	28.3%	67.2%	49.5%	36.5%	24.3%
MP	3.1%	0.7%	0.4%	0.1%	6.9%	11.5%	9.9%	7.2%	14.7%	8.9%	7.0%	5.0%
MP/all	3.5%	2.0%	1.4%	0.5%	10.5%	19.6%	22.5%	25.4%	20.4%	18.0%	19.2%	20.6%

Table 5.3.: Total amount of all ("all") and mixed-phase ("MP") clouds retrieved from the DYAMOND datasets between 60° North and 60° South from 1 June 2009 to 31 May 2013 for February 2020 at the original and at the coarser spatial resolutions of 0.5°, 1° and 2°. In the third row, "MP/all" indicates the amount of mixed-phase clouds relative to the amount of all clouds.

from 0.5° to 2°, only the results at 0.5° are considered in this paper (closer to the gridbox resolution of Cloud_cci v3 — 0. 3° — than the resolution of MCD06COSP — 1° — and CAM6-Oslo — $\sim 2°$) together with the results at the original resolution.

Figure 5.7 shows SLF averaged on isotherms with intervals of 1°C in ICON for clouds at different heights and at different spatial resolutions. Coarsening the spatial resolution in ICON mostly leads to a decrease of the SLF considering all cloud heights at a common temperature of -11°C. In GEOS, only the high-level clouds mostly show an increase in SLF when the spatial resolution is coarsened, while the opposite is true for the low clouds (Fig. A.14). In SCREAM (Fig. A.13), it is not possible to clearly observe whether the SLF increases or decreases with the resolution: although the majority of the isotherms show a decrease of the SLF with the resolution, the difference between the cases supporting this result and the cases contradicting it is about 3% comparing the resolutions 0.5° and 1° and about 10% comparing the resolutions 1° and 2°.

Discussion

Coarsening the spatial resolution to 0.5° reduces the amount of all clouds. At coarser resolution, the gridboxes are considered cloudy only if the difference between the warmest and the coldest cloudy pixels is within 20°C (see section 3.3.2.1); applying this filter excludes a portion of the clouds. While this may explain why there is a decrease in the amount of mixed-phase clouds in GEOS and SCREAM, the reason why the percentage of mixed-phase clouds and ICON increases from the original resolution to 0.5° may be due to a very efficient WBF process: the averaging of the supercooled liquid fraction in the gridboxes may contribute to increasing the mixed-phase boxes more than the temperature filter reduces them. The temperature filter excludes more and more cloudy gridboxes, coarsening the spatial resolution from 0.5° to 2° and decreasing the amount of all and mixed-phase clouds in all datasets.

ICON for temperatures below -11° and GEOS for high-level clouds show a decrease in SLF coarsening the spatial resolution. Coarsening the resolution may increase the probability of finding ice in a mixed-phase box, as the probability of finding ice below 0°C increases exponentially with decreasing temperature (see section 2.2.2.1).

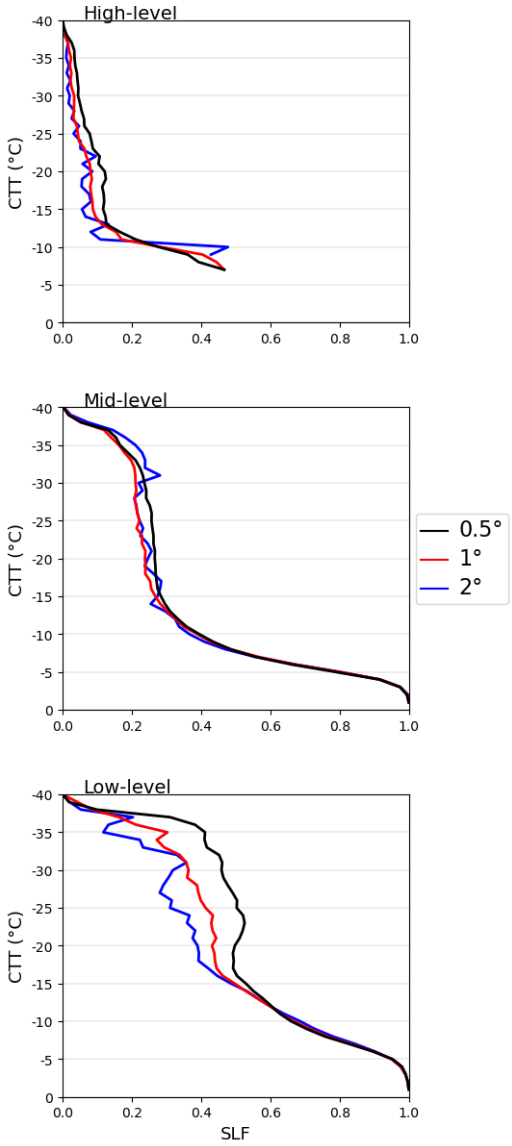

Figure 5.7.: SLF profiles on isotherms of high-, mid-, and low-level clouds for ICON at the spatial resolutions of 0.5°, 1°, and 2°.

5.5. CTT-height dependency of cloud types in the mixed phase and WBF process

In order to better understand the results in Chapter 4 about the SLF differences found for different regions, the following investigations are limited to mixed-phase clouds in the continental and marine regions at latitudes between 30° and 60°. In Chapter 4, the continental low-level mixed-phase clouds show a different hemispheric dependence from the other clouds when comparing their SLF. As shown in sections 5.2 and 5.3, stratocumulus clouds are the most abundant low-level clouds in the mixed-phase and of all clouds. As the continental area in the Northern Hemisphere is more extensive than in the Southern Hemisphere, and therefore provides more data, continental (stratocumulus) clouds in the NH are shown from this section onwards, unless the focus of the discussion is on other cloud types.

Figure 5.8 shows the SLF-CTT joint histograms for the different cloud types over continental regions in the NH for Cloud_cci v3. From this figure, the diversity of cloud types can be observed as each cloud has a unique SLF-CTT joint histogram. Furthermore, CTT decreases with increasing height level, as expected from the direct dependence of CTT on the cloud top height, and SLF decreases with decreasing CTT.

The cloud-type-dependent diversity of the SLF-CTT joint histograms is also found for the other regions and datasets: figures A.15 to A.17 for Cloud_cci v3, figures A.18 to A.21 for MCD06COSP using SLF-CTP joint histograms, figures A.22 to A.25 for CAM6-Oslo (A.26 to A.29 using SLF-CTP joint histograms for comparison with MCD06COSP), figures A.30 to A.33 for SCREAM, Fig. A.34 for ICON, Fig. A.35 for GEOS, figures A.36 to A.39 for SCREAM 0.5°, Fig. A.40 for ICON 0.5°, and Fig. A.41 for GEOS 0.5°.

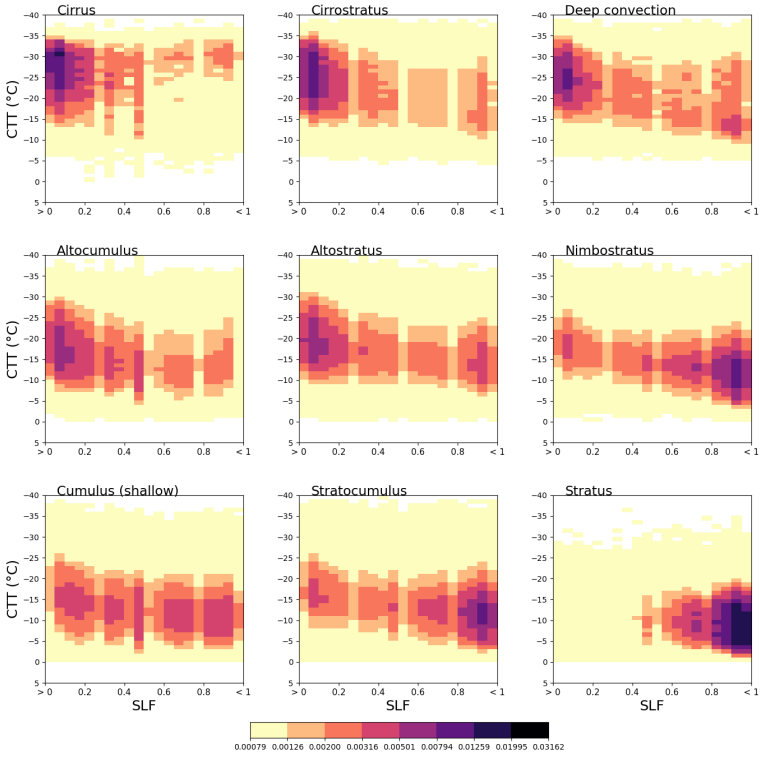

Figure 5.8.: SLF-CTT joint histograms for continental mixed-phase clouds in the NH for Cloud_cci v3. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

Figure 5.9 shows the SLF-CTT joint histograms for stratocumulus clouds over continental regions in the NH for all datasets. MCD06COSP does not provide the CTT, so the SLF-CTP joint histogram is used as a substitute, and the same type of histogram is included for CAM6-Oslo for comparison; low-level clouds are shown for ICON and GEOS.

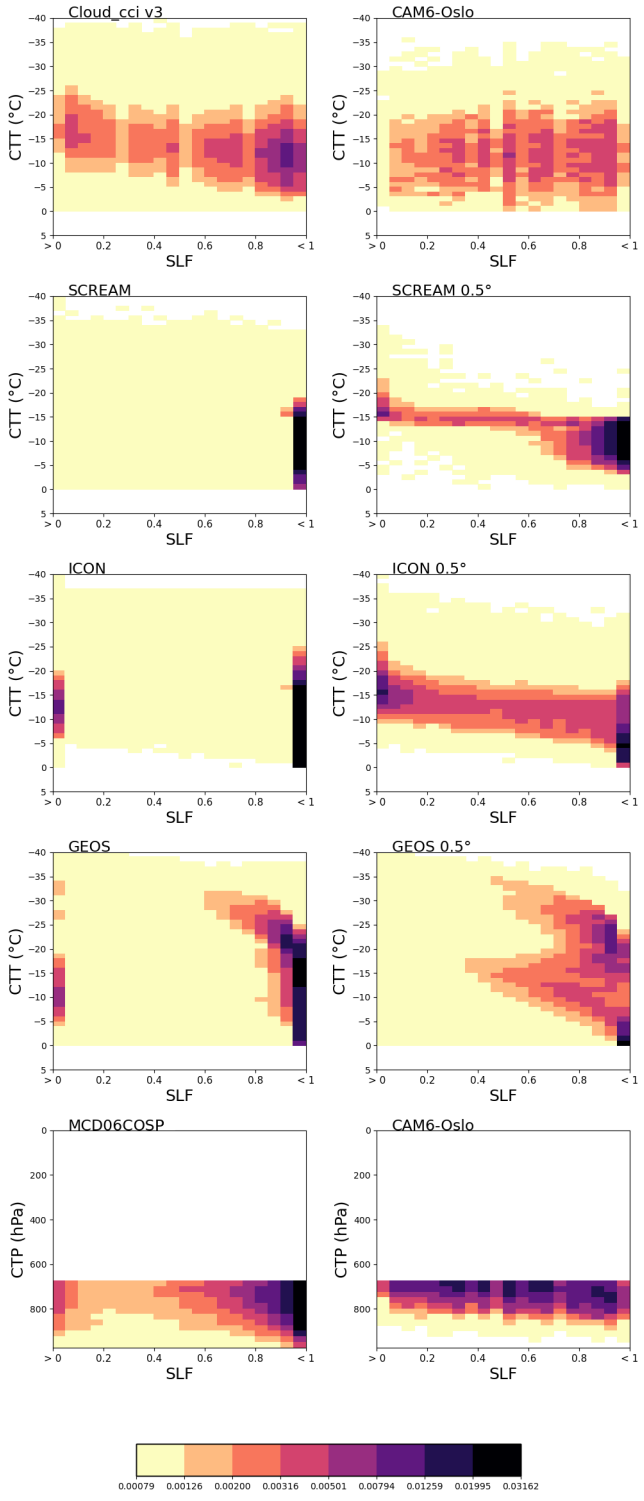

Figure 5.9.: SLF-CTT joint histograms for the continental mixed-phase stratocumulus clouds in the NH for all datasets. A SLF-CTP histogram replaces the SLF-CTT histogram for MCD06COSP, while both SLF-CTP and SLF-CTT histograms are shown for CAM6-Oslo. Stratocumulus is replaced by low-level clouds for ICON, ICON 0.5°, GEOS, and GEOS 0.5°. The colorbar indicates the cloud fraction for a given SLF-CTT or SLF-CTP bin relative to the individual histogram.

Compared to the observations and CAM6-Oslo, the DYAMOND datasets show fewer intermediate stages of SLF at the original spatial resolution, with the majority of the SLF-CTT frequency for SLF < 0.05 and SLF > 0.95 (0.6 for GEOS). While SCREAM and ICON do not show any intermediate SLF values for low- and mid-level clouds (Figures A.30 to A.33 for SCREAM and Fig. A.34 for ICON), but cover the whole SLF range for high-level clouds (Fig. 5.10), GEOS shows intermediate SLF values for almost all cloud types (Fig. A.35).

Intermediate SLF values are also found more frequently in the DYAMOND datasets at a resolution of 0.5° than at the original resolution, with the exception of the high-level clouds, whose SLF decreases when the resolution is coarsened to 0.5° (Fig. 5.10). Both Cloud_cci v3 and CAM6-Oslo show SLF > 0.4 for stratus clouds (Fig. 5.11), with SLF reaching lower values over land (Cloud_cci v3 only in the SH). From the DYAMOND datasets, SCREAM 0.5° shows SLF > 0.4 for marine stratus clouds only in the SH, while GEOS 0.5° shows SLF > 0.3 for marine low-level clouds in the SH (Fig. 5.12) and similar SLF values for the other geographical areas (Fig. A.41).

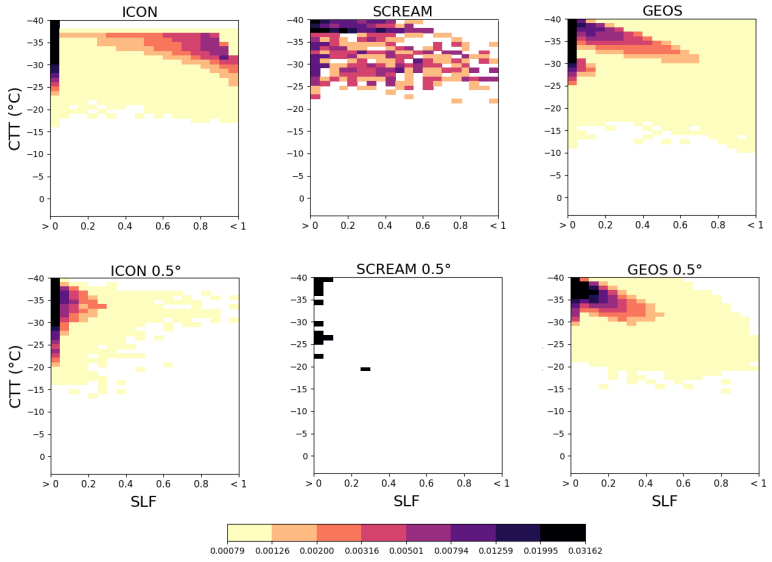

Figure 5.10.: SLF-CTT joint histograms for continental mixed-phase high-level (ICON and GEOS) and cirrostratus (SCREAM) clouds in the NH at the original resolution (top) and at 0.5° (bottom). Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

Discussion

A decrease of SLF with CTT is shown in Fig. 5.8 for Cloud_cci v3. In the temperature range between −40°C and 0°C, this could be driven by heterogeneous ice nucleation (section 2.2.2.3, Lamb and Verlinde 2011; Pruppacher and Klett 1997) as well as secondary ice formation and ice crystal growth at the expense of liquid droplets (WBF process). Even at the same height, different cloud types show different SLF-CTT dependencies, indicating individual cloud processes that could be driven by cloud dynamics or regional effects, e.g., different aerosol composition. One reason why the DYAMOND datasets, especially

ICON and SCREAM, display fewer intermediate SLF values (Fig. 5.9) may be the lower resolution of Cloud_cci v3 (due to the regridding process in gridboxes), MCD06MODIS and CAM6-Oslo (both providing averages), which makes intermediate SLF values more likely in the histograms of the latter datasets. This is confirmed by coarsening the spatial resolution of the DYAMOND datasets to 0.5°. In addition, comparing Cloud_cci v3 and the DYAMOND models at coarser resolution, more ICON 0.5° and SCREAM 0.5° than GEOS 0.5° may have an overly efficient WBF process leading to rapid cloud freezing once ice nucleation starts.

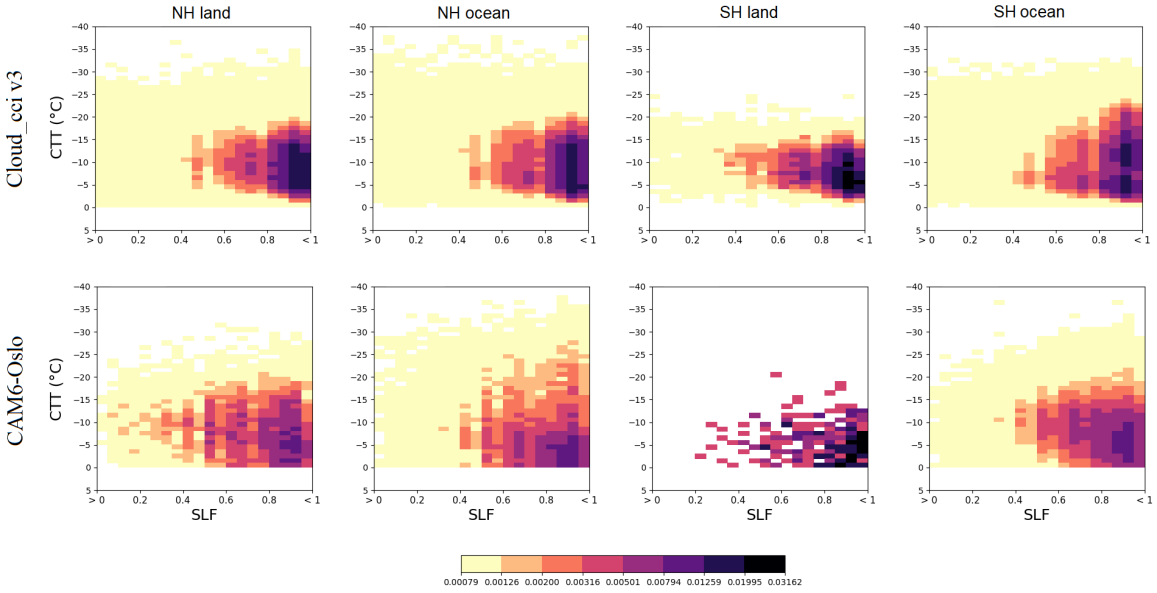

Figure 5.11.: SLF-CTT joint histograms for continental and marine mixed-phase clouds in the NH and SH for Cloud_cci v3 (top) and CAM6-Oslo (bottom). Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

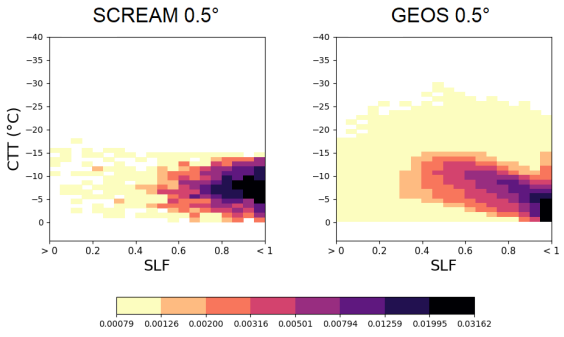

Figure 5.12.: SLF-CTT joint histograms for marine mixed-phase stratus clouds for SCREAM 0.5° (left) and low-level clouds GEOS 0.5° (right) in the SH. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

The importance of the efficiency of the WBF process in models to represent the mixed-phase clouds has been addressed, e.g., by Forbes and Ahlgrimm (2014) and Cesana et al. (2015), who respectively show that the ice deposition growth rate largely determines the liquid-ice partition in mixed-phase clouds, and that while the effects of the WBF process can be observed with satellite data, models may fail to determine the liquid-ice partition in mixed-phase clouds despite the WBF parameterisation. While GEOS shows intermediate SLF values for all cloud types (for some still fewer than Cloud_cci v3 (Fig. 5.9) and for some others extended over the whole SLF range (Fig. A.35)), the SLF is constrained to values below 0.05 and above 0.95 for low and mid-level clouds in ICON and SCREAM. The presence of intermediate SLF values for high-level clouds in ICON and SCREAM may be due to contamination by other cloud types and therefore to cloud-type misclassification, or it may be indicative of erroneous processes occurring exclusively in high-level clouds.

Coarsening the spatial resolution of the DYAMOND datasets leads to lower SLF values for high-level clouds (Fig. 5.10). Indeed, the temperature filter applied to the gridboxes limits the possibility of lower clouds contaminating the results for high-level clouds (see section. 3.3.2).

The absence of small values of SLF for mixed-phase stratus clouds in Cloud_cci v3 and CAM6-Oslo shown in Fig. 5.11 may be caused by very efficient ice formation processes bringing SLF to zero: considering that a consequence of riming can be the production of secondary ice by the rime-splintering process (Mossop and Hallett, 1974, which occurs at temperatures between $-3°C$ and $-8°C$) and that "a large amount of secondary ice would intensify the WBF process" (Korolev et al., 2022), the results that Cloud_cci v3 and CAM6-Oslo show for stratus clouds may be explained by these processes. However, only values for SLF > 0 and SLF < 1 are shown in Fig. 5.11. Therefore, inefficient ice formation processes in mixed-phase stratus clouds are more probable to explain the absence of low SLF values for this cloud type. The difference found between continental and marine regions may be due to the greater abundance of INPs over land than over the ocean, or to differences in the cloud dynamics, or both. If Cloud_cci v3 was correct in retrieving lower values of SLF for continental stratus clouds only in the SH, then the presence of specific INPs in the continental regions in the SH and/or differences in dynamics with respect to marine area could be related to the speculations made in section 4.4, and this would also imply that CAM6-Oslo fails to capture differences between NH and SH over continents. Although SLF-CTT histograms are not available for MCD06COSP, its SLF-CTP histograms for stratus clouds are similar to the SLF-CTP histograms of CAM6-Oslo.

Among the DYAMOND datasets, only SCREAM 0.5° and GEOS 0.5° show similar results (Fig. 5.12). While in GEOS 0.5° it is not possible to consider only stratus clouds but low-level clouds, in SCREAM only stratus clouds are considered, and two conclusions can be deduced: coarsening the spatial resolution to 0.5° may be necessary to fairly compare the DYAMOND datasets with the observations of Cloud_cci v3; SCREAM 0.5° may fail to properly represent macro- and microphysical processes in mixed-phase stratus clouds, as only marine stratus clouds in the SH are in agreement with observations.

5.6. Land and ocean contrast in the SLF-CTT joint histograms

The previous section showed that each cloud type is characterised by a unique SLF-CTT histogram. This section focuses on possible land-ocean contrasts. Cirrostratus and stratocumulus clouds are again considered for the comparisons, as they are the most common cloud types in the high and low clouds they represent. Figure 5.13 shows the histograms of these cloud types distinguished by location and their differences for Cloud_cci v3. While cirrostratus clouds are warmer over land than over the ocean in both hemispheres, stratocumulus clouds show this characteristic only in the SH, while the opposite is true in the NH.

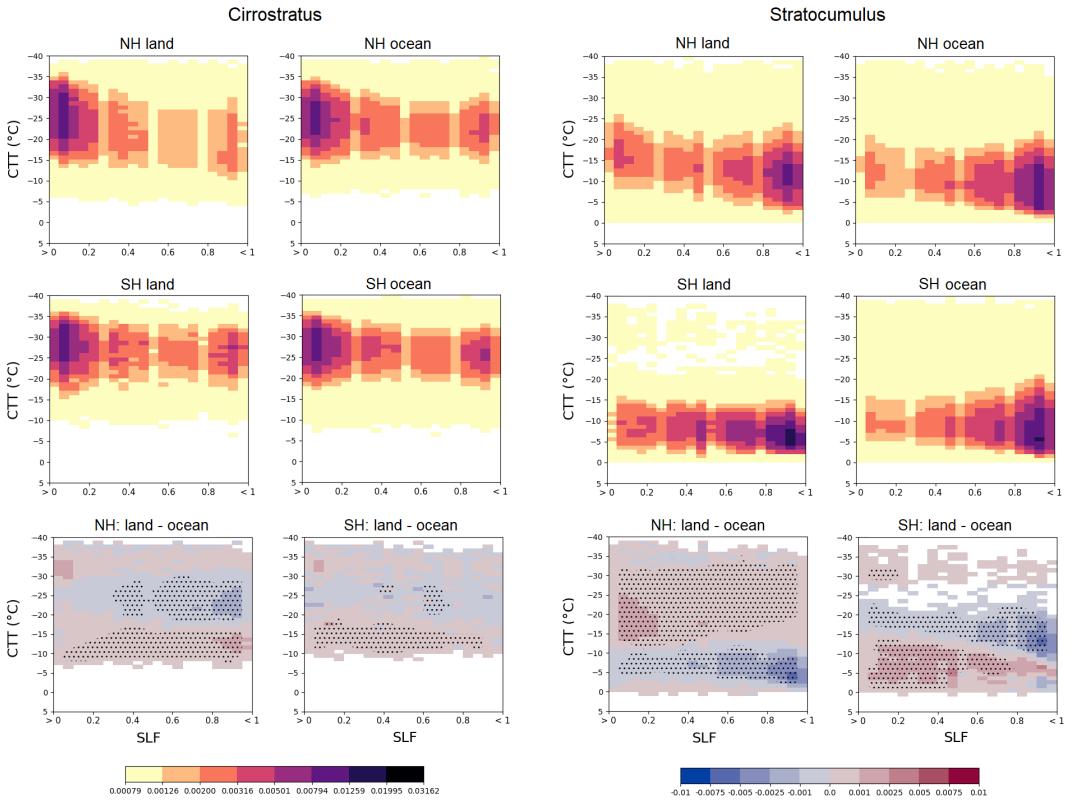

Figure 5.13.: In the top two rows: SLF-CTT joint histograms of cirrostratus (left) and stratocumulus (right) in the mixed-phase in NH and SH over land and ocean for Cloud_cci v3. Bottom: The difference of land minus ocean for NH (left) and SH (right). The dots represent the SLF-CTT combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar on the left indicates the cloud fraction for a given SLF-CTT bin relative to the individual histogram. The colorbar on the right indicates the difference between two SLF-CTT histograms for a given SLF-CTT bin.

Figure 5.14 shows the land/ocean contrast of cirrostratus (or high-level) clouds in both the NH and the SH for all datasets. CAM6-Oslo, GEOS, GEOS 0.5°, and ICON 0.5° (only in the SH) agree with Cloud_cci v3. MCD06COSP shows a similar contrast to CAM6-Oslo only in the SH.

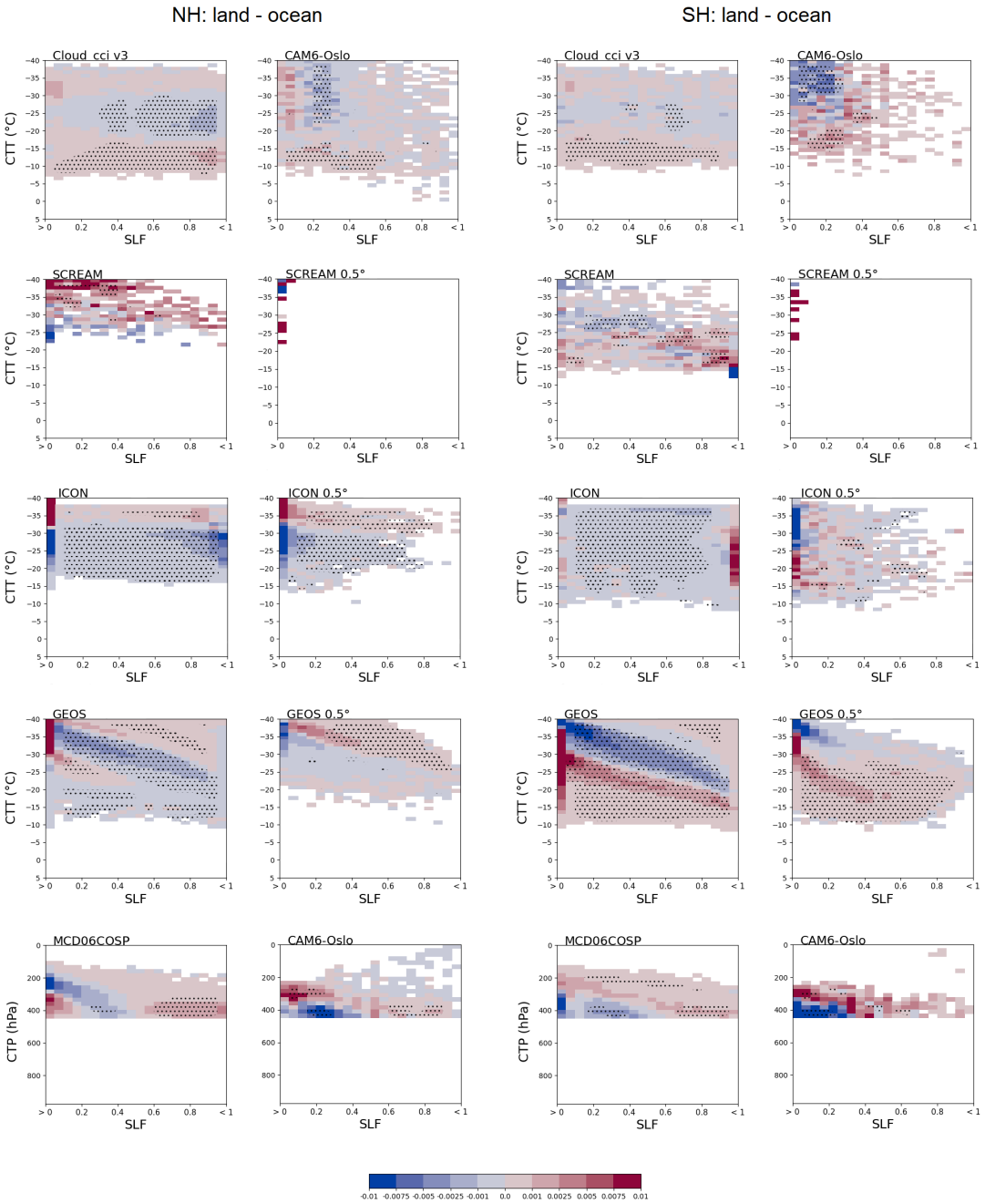

Figure 5.14.: Differences between continental and marine cirrostratus (or high-level) clouds in the NH (left) and SH (right) represented by SLF-CTT joint histograms for all datasets. A SLF-CTP histogram replaces the SLF-CTT histogram for MCD06COSP, while both SLF-CTP and SLF-CTT histograms are shown for CAM6-Oslo. Cirrostratus clouds are replaced by high-level clouds for ICON, ICON 0.5°, GEOS, and GEOS 0.5°. The dots represent the SLF-CTT combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar indicates the land minus ocean difference for a given SLF-CTT or SLF-CTP bin relative to the individual histogram.

Figure 5.15 shows the land/ocean contrast of stratocumulus (or low-level) clouds in both the NH and the SH for all datasets. Again, CAM6-Oslo, GEOS, GEOS 0.5°, and ICON 0.5° (only in the SH) agree

with Cloud_cci v3. For stratocumulus clouds, SCREAM 0.5° is also in agreement with Cloud_cci v3. MCD06COSP shows a land/sea contrast similar to CAM6-Oslo in both hemispheres.

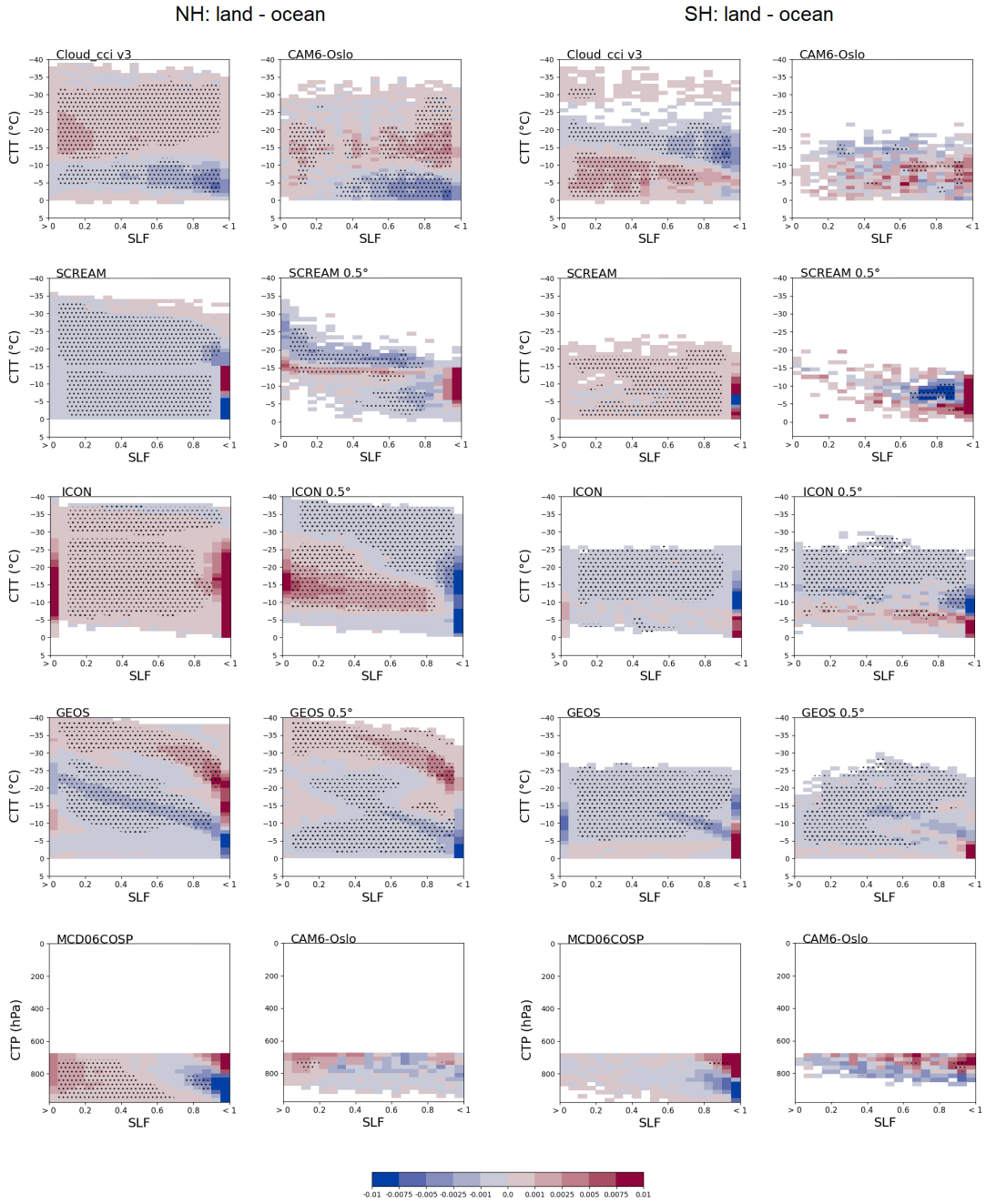

Figure 5.15.: Differences between continental and marine stratocumulus (or low-level) clouds in the NH (left) and SH (right) represented by SLF-CTT joint histograms for all datasets. A SLF-CTP histogram replaces the SLF-CTT histogram for MCD06COSP, while both SLF-CTP and SLF-CTT histograms are shown for CAM6-Oslo. Stratocumulus clouds are replaced by low-level clouds for ICON, ICON 0.5°, GEOS, and GEOS 0.5°. The dots represent the SLF-CTT combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar indicates the land minus ocean difference for a given SLF-CTT or SLF-CTP bin relative to the individual histogram.

Discussion

Low- and mid-level clouds are more influenced by the Earth's surface, i.e., by orography, aerosol transport, and the ground heated by the Sun (land has a lower heat capacity than water and heats up faster, producing more convection and clouds during the day). Therefore, turbulence and updrafts may characterise the dynamics of mid- and low-level clouds, leading to higher tops. The reason why it is only found in NH may be related to the continental geographic extent between 30° and 60°, which is much wider in the NH than in the SH and may be less influenced by the ocean.
While CAM6-Oslo and GEOS reproduce well the results found in Cloud_cci v3, ICON and SCREAM partially agree with the observations only at the coarser spatial resolution of 0.5°, supporting the importance of comparing datasets at similar spatial resolutions.

5.7. Hemispheric dependencies of cloud types on CTT-SLF joint histograms

The main result in section 4.4 was that all four datasets (from passive and active satellite sensors) showed higher SLF in the SH than in the NH for all cloud types, except for continental low-level clouds (and altocumulus clouds, for Cloud_cci v3 only), for which the opposite was found. In an attempt to compare the results in section 4.4 with models, CFAD-like histograms (see section 3.3.2.3) are used for consistency. Despite using a different method for this second part of the analysis, Cloud_cci v3 confirms the results found previously and shown in section 4.4.
This is shown in Fig. 5.16, where both CFAD-like histograms for different areas as well as their difference between NH and SH for continental and marine areas are shown for only cirrostratus and stratocumulus, which are representative for high- and mid-level clouds (excluding altocumulus) and low-level clouds (and altocumulus), respectively. For a given CTT < -15°C, cirrostratus clouds in the SH have higher SLF than in the NH over both land and the ocean, as shown by the turquoise colors at high SLF in the difference plot. The same holds only over marine regions for stratocumulus clouds, while the opposite is true over land.
Figure 5.17 shows the hemispheric contrast of cirrostratus (or high-level) clouds over land and ocean for all datasets. Only ICON and CAM6-Oslo show agreement with Cloud_cci v3 over both land and the ocean, and ICON especially at SLF extremes. SCREAM agrees with Cloud_cci v3 only over the ocean, while the histograms of SCREAM 0.5° cannot be evaluated due to lack of data. GEOS 0.5° shows agreement with Cloud_cci v3 only over the ocean and for CTT < -25°C. MCD06COSP shows similar results to CAM6-Oslo only over land.
Figure 5.18 shows the hemispheric contrast of stratocumulus (or low-level) clouds over land and the ocean for all datasets. CAM6-Oslo shows larger SLF in the SH than in the NH for marine clouds, in agreement with Cloud_cci v3, but also for continental clouds, in contrast to Cloud_cci v3. SCREAM agrees with Cloud_cci v3 over land and for CTT > -15° over the ocean, while SCREAM 0.5° agrees with Cloud_cci v3 only over land. The results of ICON are in contrast to Cloud_cci v3, while ICON

0.5° agrees with Cloud_cci v3 only over land for temperatures from −10°C to −5°C. GEOS agrees with Cloud_cci v3 over land only for temperatures from −10°C to −20°C and the ocean only for temperatures from −15°C to −25°C. GEOS 0.5° agrees with Cloud_cci v3 over land only for temperatures from 0°C to −10°C and over the ocean only for temperatures from −20°C to −25°C. MCD06COSP shows similar results to CAM6-Oslo over both land and the ocean, thus in agreement with Cloud_cci v3 only over the ocean.

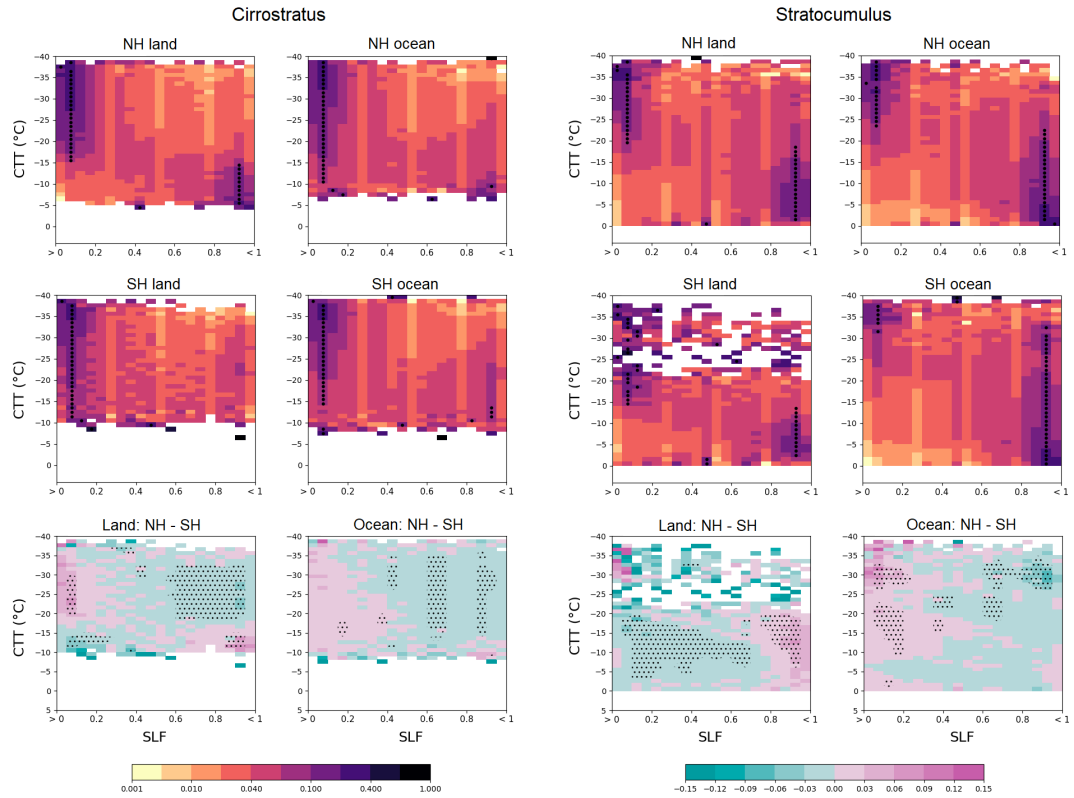

Figure 5.16.: In the top two rows: CFAD-like histograms of mixed-phase cirrostratus (left) and stratocumulus (right) in NH and SH over land and the ocean for Cloud_cci v3. Each line of the CFAD-like histograms shows a dot where the maximum SLF is located. Bottom: The difference between NH and SH for land (left) and the ocean (right). The dots represent the SLF-CTT combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar on the left indicates the cloud fraction for a given SLF-CTT bin relative to the individual histogram. The colorbar on the right indicates the difference between two SLF-CTT histograms for a given SLF-CTT bin.

The complete set of CFAD-like histograms can be found in the Appendix: figures A.42 to A.45 for Cloud_cci v3, figures A.46 to A.49 for MCD06COSP using SLF-CTP joint histograms, figures A.50 to A.53 for CAM6-Oslo (A.54 to A.57 using SLF-CTP joint histograms for comparison with MCD06COSP), figures A.58 to A.61 for SCREAM, Fig. A.62 for ICON, Fig. A.63 for GEOS, figures A.64 to A.67 for SCREAM 0.5°, Fig. A.68 for ICON 0.5°, and Fig. A.69 for GEOS 0.5°.

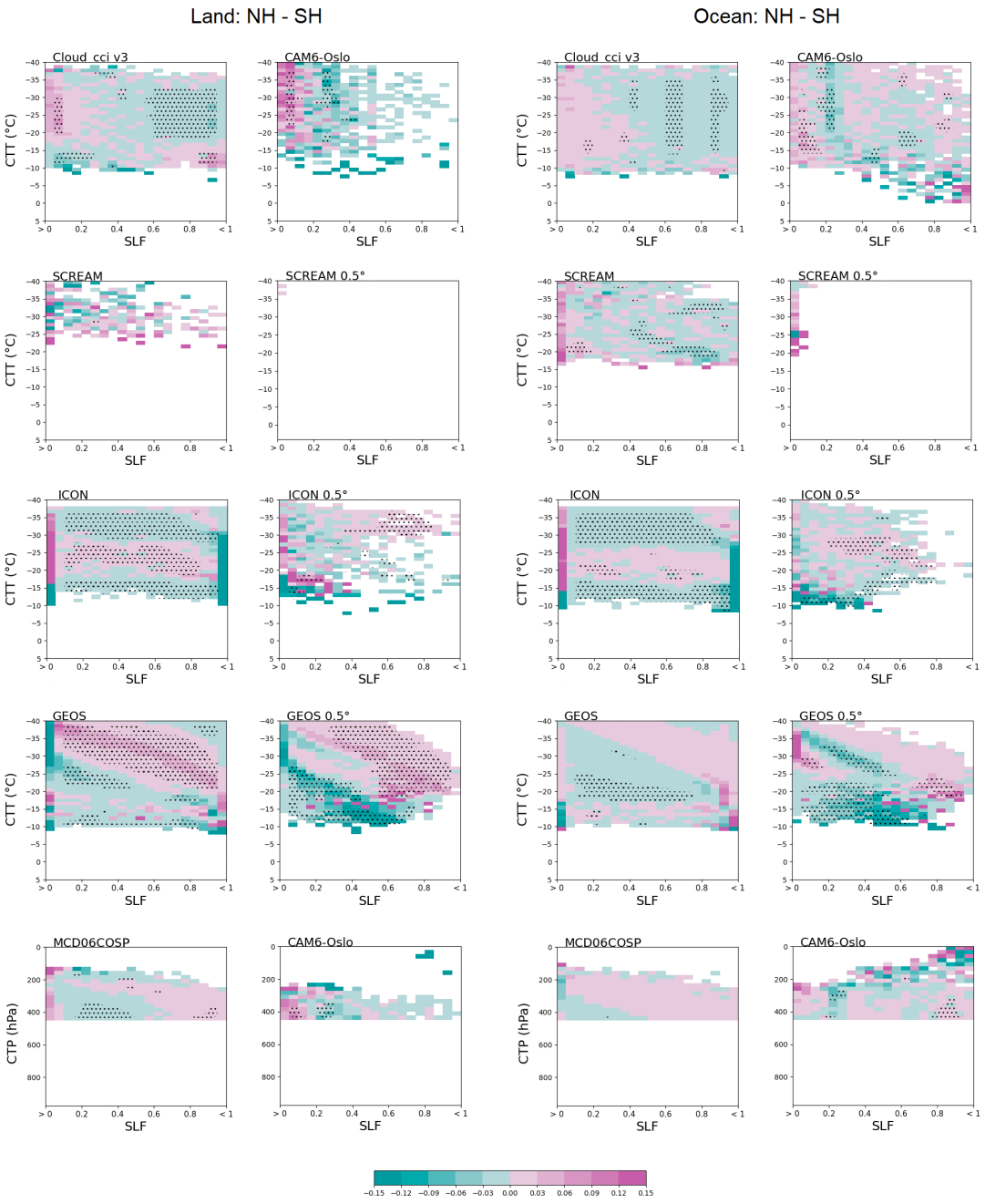

Figure 5.17.: Differencens between CFAD-like histograms of cirrostratus (or high-level) clouds in the NH and SH over land and the ocean for all datasets. CFAD-like histograms of MCD06COSP are derived from SLF-CTP histograms, while both SLF-CTP and SLF-CTT histograms are used to derive CFAD-like histograms of CAM6-Oslo. Cirrostratus clouds are replaced by high-level clouds for ICON, ICON 0.5°, GEOS, and GEOS 0.5°. The dots represent the SLF-CTT combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar indicates the land minus ocean difference for a given SLF-CTT or SLF-CTP bin relative to the individual histogram.

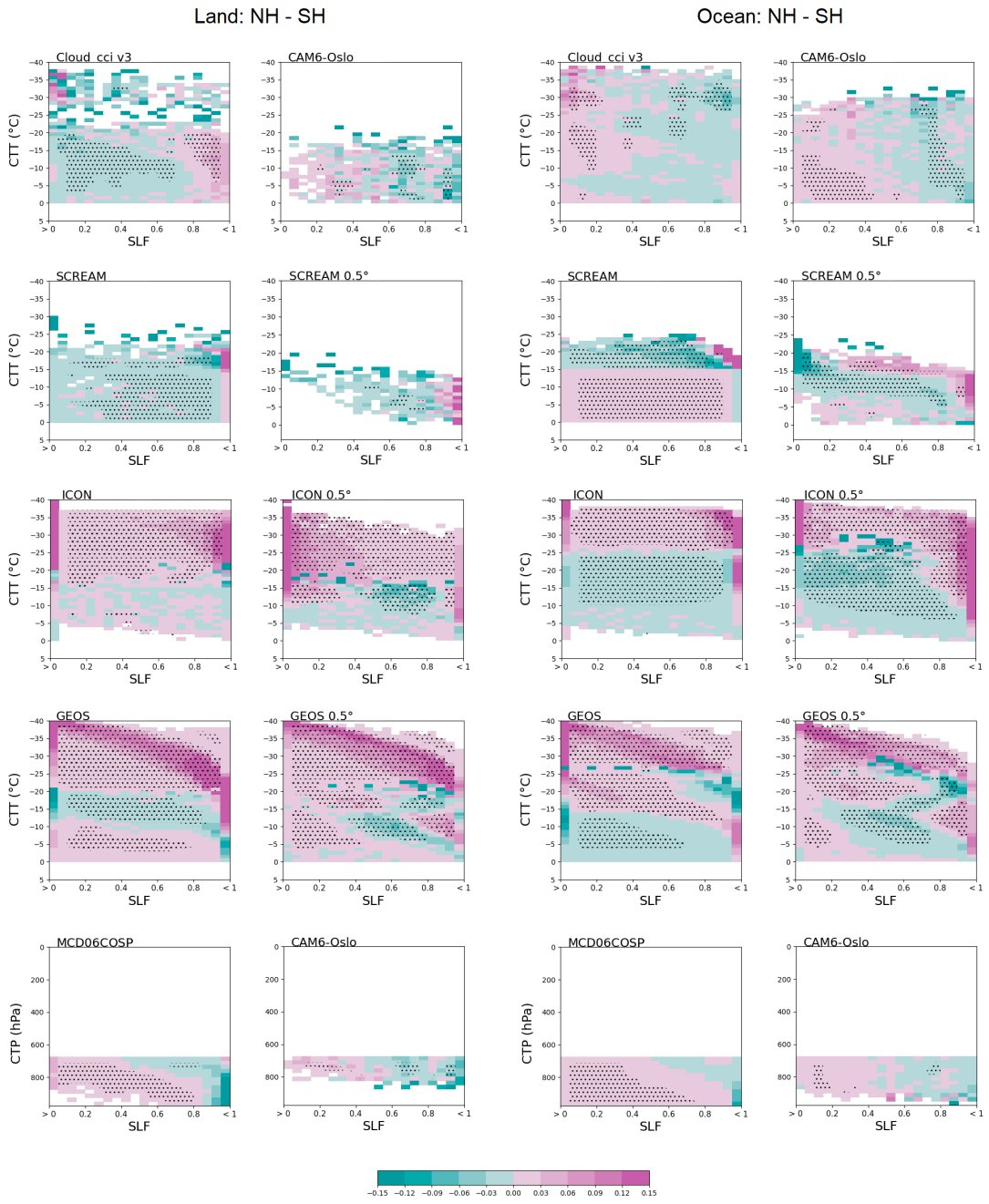

Figure 5.18.: Differencens between CFAD-like histograms of stratocumulus (or low-level) clouds in the NH and SH over land and ocean for all datasets. CFAD-like histograms of MCD06COSP are derived from SLF-CTP histograms, while both SLF-CTP and SLF-CTT histograms are used to derive CFAD-like histograms of CAM6-Oslo. Stratocumulus clouds are replaced by low-level clouds for ICON, ICON 0.5°, GEOS, and GEOS 0.5°. The dots represent the SLF-CTT combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar indicates the land minus ocean difference for a given SLF-CTT or SLF-CTP bin relative to the individual histogram.

Discussion

Differences between NH and SH in cloud dynamics and cloud microphysics of the simulated cloud types, as well as possible seasonal effects influencing the DYAMOND datasets analysed in this thesis, which are only available for February 2020, may explain the differences between the datasets shown in the results as well as the contrast with the observations. Furthermore, the number of datasets in agreement with observations is higher for high-level clouds than for low-level clouds, suggesting that cloud micro- and macrophysics in low-level clouds are more influenced by the Earth's surface than in high-level clouds, making the simulation of low-level clouds more complex. When comparing low-level clouds, the only agreement with observations was found for marine clouds, suggesting that continental low-level clouds may be difficult to represent due to possible heterogeneity in orography, aerosol composition, urban areas, etc.

5.8. Ice formation in marine low-level clouds at about -15°C in Southern Hemisphere

Some CFAD-like histograms show a trimodal SLF tendency, with a high frequency of occurrence at low SLF around −15°C and a high frequency of occurrence at high SLF for higher and lower temperatures. This is shown in Fig. 5.19 for cumulus (or low-level) marine clouds in the SH, the only region where Cloud_cci v3, CAM6-Oslo, SCREAM 0.5°, and GEOS 0.5° show this phenomenon. While in Cloud_cci v3 and CAM6-Oslo SLF reaches very low values close to zero at temperatures between −15°C and −28°C in the first and −11°C and −15°C in the second dataset, in SCREAM 0.5° SLF reaches 0.8 at about −7°C, and in GEOS SLF reaches 0.5 at about −13°C. While in Cloud_cci v3 this result occurs only for marine cumulus clouds in the SH, in CAM6-Oslo it occurs for marine cumulus and stratocumulus clouds in both hemispheres, in SCREAM 0.5° it occurs for all marine low-level and nimbostratus clouds in the SH, and in GEOS it occurs for both continental and marine low-level clouds in both hemispheres and for both continental and marine mid-level clouds only in the SH.

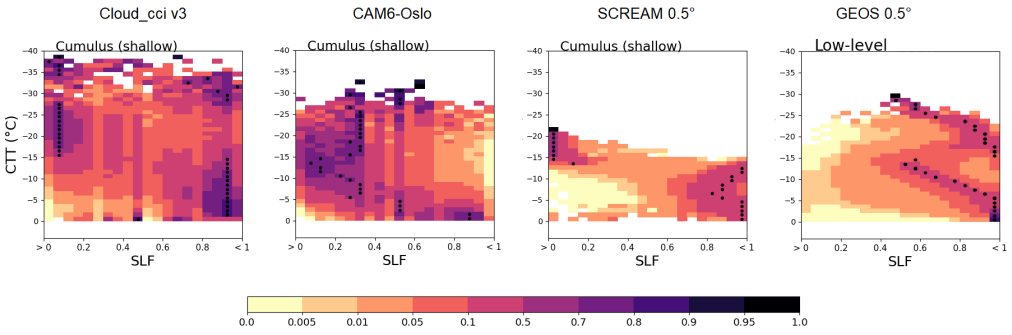

Figure 5.19.: CFAD-like histograms of marine mixed-phase cumulus clouds for Cloud_cci v3 and CAM6-Oslo and of marine mixed-phase low-level clouds for ICON in SH. The dots represent the highest frequency of occurrence for each isotherm. The colorbar indicates the cloud fraction for a given SLF-CTT bin relative to a single isotherm.

Discussion

In section 2.2.2.3, the formation of dendritic ice structures at temperatures between $-12°C$ and $-17°C$ was mentioned to support the high efficiency of coalescence processes occurring in this temperature range. In Libbrecht (2005) and Lamb and Verlinde (2011) the temperature range in which plate-like crystals (and hence dendrites) are most commonly found is between $-8°C$ and $-22°C$. The presence of dendritic crystals at these temperatures, which may also act as a source of secondary ice, may explain the high frequency of occurrence at low SLF in the temperature range between $-10°C$ and $-28°C$ in Fig. 5.19.
Evidence for a minimum in SLF at about $-15°$ is also present in previous studies analysing stratiform and mid-level clouds from observations in specific regions (Mignani et al., 2019; Zhang et al., 2019a; Silber et al., 2021; Danker et al., 2022), or globally over the ocean (Nagao and Suzuki, 2022), or distinguishing day/night, land/ocean, and seasonal contrasts in the tropics and extratropics (Riley and Mapes, 2009).
While it is not clear why high frequency of low SLF values in this temperature range is only found for some cloud types and in some regions in the datasets analysed in this work, the presence of this robust result also in previous literature leaves hope for the possibility of using passive satellite sensors to study cloud microphysics.

5.9. Conclusions

This chapter presented a comparison between observations (Cloud_cci v3 and MCD06COSP) and model datasets (CAM6-Oslo, ICON, SCREAM, and GEOS). The results presented in this chapter are summarised here:

- There is a large variability in the amount of clouds retrieved by the datasets (from 52.9% in Cloud_cci v3 to 87.3% in MCD06COSP for all clouds and from 55.0% for MCD06COSP to 3.1% in SCREAM for mixed-phase clouds), which can be explained by the different sensors, number of channels, and algorithms when comparing the observations, and the different algorithms, micro-, and macrophysics when comparing the models. CAM6-Oslo, ICON, SCREAM-CS, and GEOS retrieve a near-global amount of "all" clouds ranging from 65.5% to 72.1% close to previous studies based on satellite measurements (ranging from 61.9% to 67.2% for afternoon satellites; PVIR, 2020). This thesis provides new references on the amount of mixed-phase clouds, which includes all cloud types up to $60°$ latitude.

- When comparing all or mixed-phase clouds at high, middle, and low heights, the datasets mostly agree on the geographical position where the clouds occur. Possible misclassification of cloud types can occur when considering clouds at the same height but with different cloud optical thickness (COT). For all clouds, the observations mostly retrieve low-level clouds while the model datasets mostly contain high-level clouds. The model datasets show a systematic decrease in the

cloud top height (and hence, an increase in temperature) for the high-level mixed-phase clouds compared to all clouds; this may contribute to the presence of liquid phase in high-level mixed-phase clouds in the temperature range between $-40°$C and $0°$. Most of the datasets agree in retrieving mixed-phase clouds at low heights and at mid-latitudes, where supercooled liquid water has also been observed in previous studies (e.g., Hu et al., 2010).

- A validation study comparing ICON, SCREAM and GEOS at coarsened spatial resolutions ($0.5°$, $1°$ and $2°$) shows that the amount of all clouds decreases when the original spatial resolution is coarsened to $0.5°$. This occurs only for GEOS and SCREAM, when only the mixed-phase clouds are considered. The reduction in the number of clouds is due to a cloud top temperature (CTT) filter applied in the gridboxes, which are only included in the analysis if the difference between the warmest and coldest pixels is within 20°C. In ICON, the increase in the amount of mixed-phase clouds may be due to an over-efficient Wegener-Bergeron-Findeisen (WBF) process. The amount of clouds systematically decreases in all datasets coarsening the spatial resolution to $1°$ and $2°$, as the filter in CTT excludes more and more cloudy gridboxes. Comparing the datasets at coarser spatial resolutions using SLF on isotherms, ICON for all cloud heights at CTT $< -11°$C and GEOS for high clouds only show a decrease in SLF at coarser and coarser resolutions, possibly due to the higher probability of finding ice at coarser resolutions and decreasing temperatures below 0°C.

- The analysed datasets show different SLF-CTT joint histograms for each cloud type (or cloud height), indicating individual cloud processes, and a decreasing SLF with CTT: In the temperature range between $-40°$C and 0°C this could be driven by heterogeneous ice nucleation as well as secondary ice formation and ice crystal growth at the expense of liquid droplets (WBF process). In ICON and SCREAM, low- and mid-level clouds do not show intermediate SLF values. This may be due to an overly efficient WBF process. Intermediate SLF values are obtained by coarsening the spatial resolution to $0.5°$, but still less than in Cloud_cci v3. Cloud_cci v3 and CAM6-Oslo show only SLF > 0.4 for stratus clouds, while similar results are shown by GEOS $0.5°$ for low-level clouds and SCREAM $0.5°$ for marine stratus clouds only in the SH. Very efficient ice formation processes may cause this result: riming and secondary ice production by the rime-splintering process (Mossop and Hallett, 1974) may intensify the WBF process (Korolev et al., 2022), bringing SLF to zero.

- In Cloud_cci v3, CAM6-Oslo, GEOS, and GEOS $0.5°$, high-level clouds are warmer over land than over the ocean in both hemispheres, while low-level clouds are warmer over land than over the ocean only in the SH and colder over land than over the ocean in the NH. This result is reproduced only by ICON $0.5°$ and MCD06COSP and only in the SH. Low-level clouds are more influenced by the Earth's surface than high-level clouds, therefore, turbulence and updrafts may characterise the dynamics of low-level clouds, leading to higher tops. The reason why this result is limited to

the NH may be related to the continental geographic extent between 30° and 60°, which is much wider in NH than in SH and may be less influenced by the ocean.

- In Cloud_cci v3, high-level clouds have higher SLF in the SH than in the NH over both land and the ocean; CAM6-Oslo and MCD06COSP agree with this result only over land, while SCREAM agrees only over the ocean. Low-level clouds in Cloud_cci v3 show higher SLF in the SH than in the NH over the ocean, but the opposite over land; CAM6-Oslo, MCD06COSP, and SCREAM agree with Cloud_cci v3 only over the ocean. Other datasets agree only over land or ocean and for limited temperature ranges. The diversity of the results may depend on differences in cloud dynamics and cloud microphysics of the simulated cloud types, as well as possible seasonal effects affecting ICON, SCREAM, and GEOS (available only for February 2020). Furthermore, the number of datasets in agreement with observations is higher for high-level clouds than for low-level clouds, suggesting that the cloud micro- and macrophysics of low-level clouds are more influenced by the Earth's surface than for high-level clouds, making the simulation of low-level clouds more complex. When comparing low-level clouds, the only agreement with observations was found for marine clouds, suggesting that continental low-level clouds may be difficult to represent due to possible heterogeneity in orography, aerosol composition, urban areas, etc.

- Cloud_cci v3 and CAM6-Oslo show evidence of enhanced ice production at modest supercooling in marine cumulus clouds in the SH, SCREAM 0.5° in marine low-level and nimbostratus clouds in the SH, and GEOS 0.5° in both continental and marine low-level clouds in both hemispheres and in both continental and marine mid-level clouds only in the SH. In Cloud_cci v3, CAM6-Oslo, and GEOS 0.5°, this result occurs at temperatures around -15°C, in agreement with previous studies based on observations of stratiform and mid-level clouds, leaving hope for the possibility of using passive satellite sensors to study cloud microphysics.

6. Droplet size effects in mixed-phase clouds

In order to better understand how the results presented in the previous chapters are influenced by the cloud microphysics, a major focus is placed on the liquid effective radius of cloud droplets using, for different cloud types or heights, SLF-$\overline{r_{liq}}$ joint histograms (3.3.2.3) separated into continental and marine areas in the NH and the SH and $\overline{r_{liq}}$ geographical distributions. As ICON and SCREAM do not provide radius information, this section only includes a comparison between Cloud_cci v3, MCD06COSP, CAM6-Oslo, and GEOS.

6.1. Thermodynamic phase dependence of the liquid effective radius size

In Chapter 5 it was shown that the datasets mostly agree in having the highest amount of mixed-phase clouds in stratocumulus clouds. The continental area in the NH is more extensive than in the SH, resulting in more data and better statistics. For these reasons, Fig. 6.1 shows SLF-$\overline{r_{liq}}$ joint histograms of continental stratocumulus clouds (low-level clouds for GEOS and GEOS 0.5°) in the NH. With the exception of GEOS and GEOS 0.5°, all the other datasets show an increase in $\overline{r_{liq}}$ with a decrease in SLF, and this is consistent for all cloud types. This result is shown in Fig. 6.2 for Cloud_cci v3 in the NH over land, while in figures A.70 to A.72 for the other regions, figures A.73 to A.76 for MCD06COSP, figures A.77 to A.80 for CAM6-Oslo, Fig. A.81 for GEOS, and Fig. A.82 for GEOS 0.5°. Cloud_cci v3 shows a high frequency of $\overline{r_{liq}}$ between 12 and 13 µm. This is most likely due to the *a priori* value assigned to $\overline{r_{liq}}$, which is 12 µm. GEOS and GEOS 0.5° do not show the wide variability in radius size shown by the other datasets.

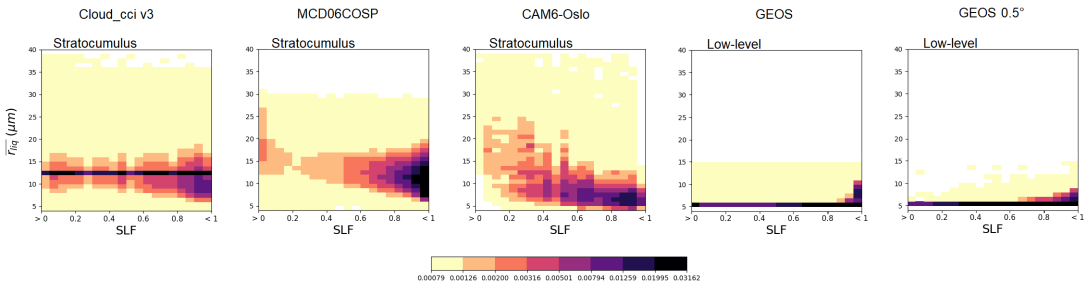

Figure 6.1.: SLF-$\overline{r_{liq}}$ joint histograms of continental mixed-phase stratocumulus clouds in the NH for the datasets Cloud_cci v3, MCD06COSP, CAM6-Oslo, and SLF-$\overline{r_{liq}}$ joint histograms of continental low-level clouds in the NH for GEOS and GEOS 0.5°. The colorbar indicates the cloud fraction for a given SLF-$\overline{r_{liq}}$ bin relative to the individual histograms.

In addition, the following results are shown in Fig. 6.2 for continental clouds in NH for Cloud_cci v3 (and in Figures A.70 to A.72 for the other areas):

- Cumulus, stratocumulus, stratus, and nimbostratus clouds show a decrease in the SLF-$\overline{r_{liq}}$ bin frequency with decreasing SLF;
- For stratus clouds (and less for stratocumulus clouds) the frequency of smaller droplets decreases with decreasing SLF;
- For high clouds, both the increase and decrease in $\overline{r_{liq}}$ are found with decreasing SLF at SLF < 0.2.

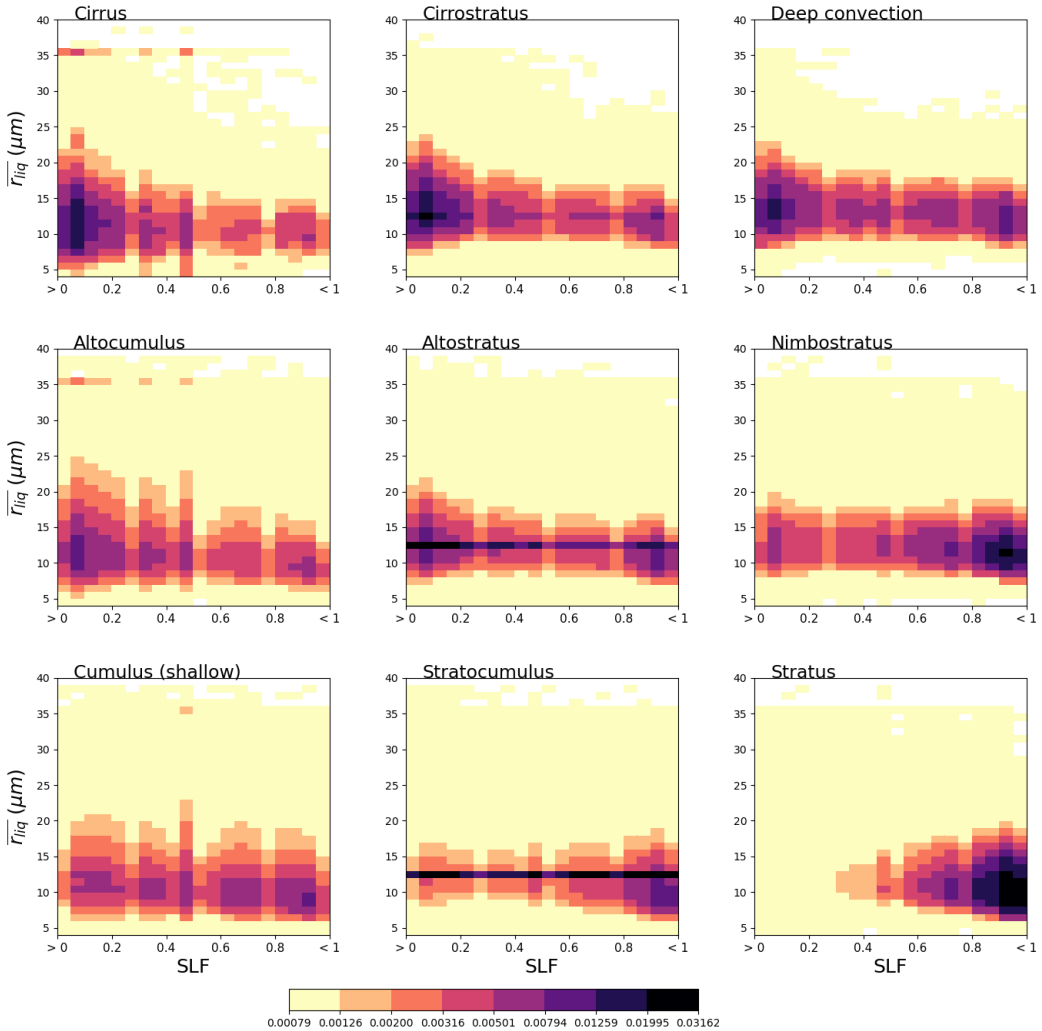

Figure 6.2.: SLF-$\overline{r_{liq}}$ joint histograms of different continental cloud types in the mixed phase in the NH for Cloud_cci v3. The colorbar indicates the cloud fraction for a given SLF-$\overline{r_{liq}}$ bin relative to the individual histograms.

Discussion

Coopman et al. (2020) has shown that the glaciation of clouds with larger droplets is more likely to occur at warmer temperatures. Chapters 4 and 5 have shown that SLF decreases from 1 to 0 with decreasing CTT; therefore, warmer temperatures can be associated with SLF approaching 1, and this may explain why smaller values of $\overline{r_{liq}}$ are found as SLF approaches 1 in Fig. 6.1. This result can be explained by several microphysical processes: in clouds with strong updrafts, liquid droplets and ice crystals can grow simultaneously at the expense of available water vapour (Korolev and Mazin 2003, see section 2.2.3), allowing the presence of large liquid droplets at low SLF; during ascent in the cloud (where a decrease in SLF is expected due to the decreasing CTT), droplets can collide and aggregate, resulting in larger droplets (Lamb and Verlinde, 2011); collisions in the mixed phase temperature range can cause heterogeneous glaciation and the formation of secondary ice (Korolev and Leisner, 2020), which can explain the decrease in SLF with the increase in $\overline{r_{liq}}$.
The results shown in Fig. 6.2 may be due to cloud processes related to specific cloud types:

- Cumulus, stratocumulus, and nimbostratus clouds can have relatively strong updrafts (Houze, 2014), thus the simultaneous growth of liquid droplets and ice crystals, as long as water vapour is available (Korolev and Mazin, 2003), is likely. As mentioned above, the larger droplets have a higher probability of colliding and eventually freezing or producing secondary ice, thus decreasing the SLF-$\overline{r_{liq}}$ bin frequency with decreasing SLF. Cloud type misclassification (Hahn et al., 2001) may explain why this also occurs for altostratus clouds in MCD06COSP and CAM6-Oslo, and why this result is present in nimbostratus clouds and not in deep convective clouds.

- For stratus clouds, some speculation was made in section 5.5 about the possibility that both the WBF process (Korolev, 2007) and very efficient glaciation processes (e.g., riming and secondary ice formation) as well as inefficient ice formation processes may drive the microphysics in these clouds. These speculations are supported by the SLF-$\overline{r_{liq}}$ joint histograms, as the frequency of smaller droplets in the stratus decreases with decreasing SLF, which could be related to the growth of ice at the expense of liquid droplets through the WBF process, causing the smaller droplets to evaporate faster than the larger droplets.

- For high clouds, while the increase in $\overline{r_{liq}}$ with low SLF may be due to strong updrafts (Korolev and Mazin, 2003) for deep convective clouds, the increase found in cirrus and cirrostratus clouds may be due to cloud misclassification (Hahn et al., 2001). The decrease in $\overline{r_{liq}}$ may be due to the WBF process (Korolev, 2007), which allows ice crystals to grow, reducing the water content of the liquid droplets and thus their size. It should be stressed again that the cloud type classification is based on the COT-CTP joint histograms, so the same cloud type may undergo different processes depending on the location where it forms or the atmospheric or dynamical conditions. For this reason, while both increasing and decreasing $\overline{r_{liq}}$ processes at lower SLF may occur in the same

cloud, it is also likely that different clouds classified as the same cloud type will be subject to different processes that are collected in only one histogram.

- In Chapter 4, a cloud-phase mismatch between AVHRR-based datasets and CALIOP was shown, in which the former retrieved some liquid clouds which CALIOP retrieved as ice clouds. It represented between 7.5% and 11.3% of the total collocated cloudy pixels, depending on the datasets. This implies that the liquid effective radius may be affected by errors due to incorrect phase assignment of cloudy pixels. In Stengel et al. (2017) a validation study between several Cloud_cci datasets is presented. It shows that several datasets derived from passive satellite sensors overestimate the occurrence of liquid clouds at the expense of ice clouds. The sensitivity of passive retrievals to optically very thin ice cloud layers above liquid cloud layers is very low and the filter on COT applied in this work, which removes clouds for COT < 3, may not be sufficient to overcome this problem. In addition, there could always be some ice in the primarily liquid pixels, and these large ice crystals may bias $\overline{r_{liq}}$.

- The decrease of $\overline{r_{liq}}$ with SLF in GEOS may indicate that some processes dominate other processes, (e.g., the higher probability that larger droplets have to collide and glaciate may cause this result).

6.2. Land-ocean contrast of the liquid effective radius

In section 5.6, the analysed datasets were compared for different regions using SLF-CTT joint histograms for different cloud types. A difference in CTTs in mixed-phase clouds was observed, with warmer CTTs over land than over the ocean for all clouds except the continental mid- and low-level clouds in the SH. A similar contrast between land and sea is also found for $\overline{r_{liq}}$ and is shown from Cloud_cci v3 in Fig. 6.3 for cirrostratus and stratocumulus clouds, the first representative of high-level clouds and the second representative of mid- and low-level clouds. While stratocumulus clouds are the most frequent low-level clouds retrieved in the mixed-phase temperature range, cirrostratus clouds are the most frequent high-level clouds. In particular, mid- and low-level clouds have, for the entire SLF range, larger $\overline{r_{liq}}$ over the ocean than over land in the NH, while the opposite occurs in the SH; high-level clouds show larger $\overline{r_{liq}}$ over the ocean than over land for all SLF values in both the NH and SH.
In MCD06COSP, continental clouds have smaller $\overline{r_{liq}}$ than marine clouds, regardless of the cloud type, with the exception of cumulus clouds in the NH where the opposite is true. Low-level clouds in CAM6-Oslo show results in agreement with Cloud_cci v3. These results are visible in Fig. 6.4. The small variability in $\overline{r_{liq}}$ of GEOS and GEOS 0.5° complicates the interpretation of the result, shown in Fig. 6.5. While GEOS shows mostly $\overline{r_{liq}}$ smaller over the ocean than over land in both hemispheres, GEOS 0.5° shows low-level clouds with smaller $\overline{r_{liq}}$ over continents than over the ocean in the NH and the opposite in the SH, in a better agreement with Cloud_cci v3.

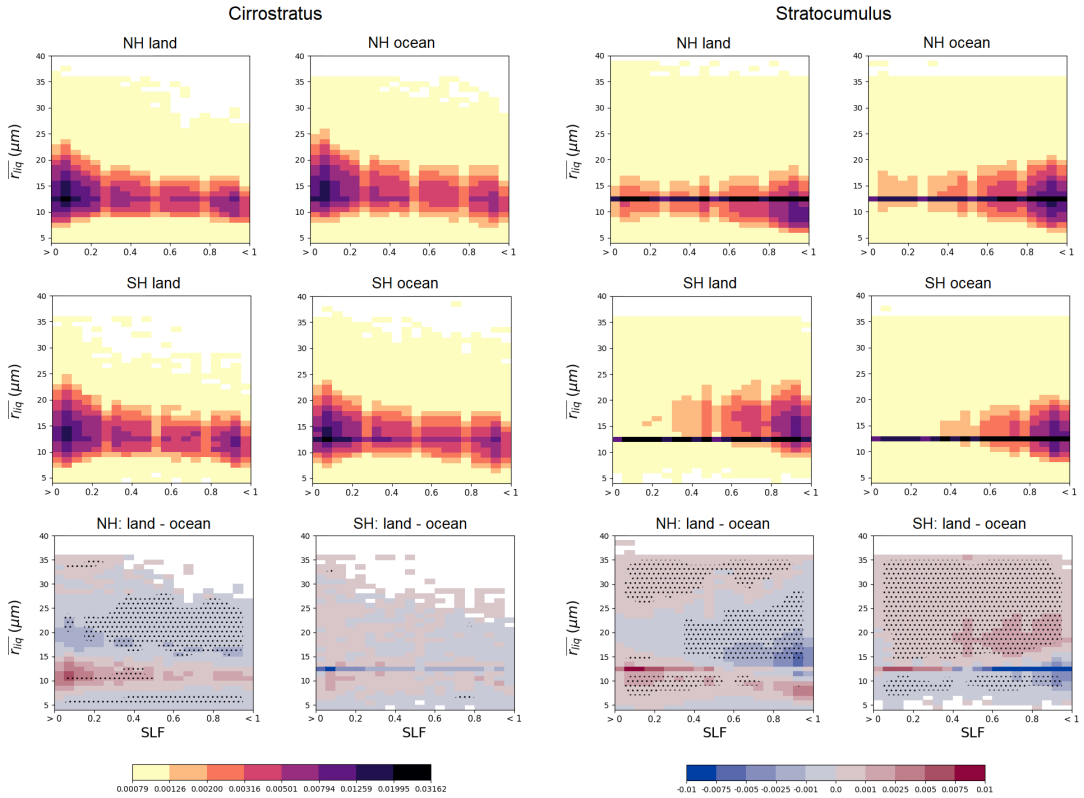

Figure 6.3.: In the top two rows: SLF-$\overline{r_{liq}}$ joint histograms of mixed-phase cirrostratus (left) and stratocumulus (right) in NH and SH over land and sea for Cloud_cci v3. Bottom: The difference of land minus ocean for NH (left) and SH (right). The dots represent the SLF-$\overline{r_{liq}}$ combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar on the left indicates the cloud fraction for a given SLF-$\overline{r_{liq}}$ bin relative to the individual histogram. The colorbar on the right indicates the difference between two SLF-$\overline{r_{liq}}$ histograms for a given SLF-$\overline{r_{liq}}$ bin.

Discussion

These results may indicate that low- and mid-level clouds are strongly influenced by the Earth's surface, which affects them differently depending on which hemisphere is considered, whereas high-level clouds are less influenced by the Earth's surface.

Low- and mid-level clouds in the NH may be mainly influenced by the aerosol concentration, which is greater over land than over the ocean (Lamb and Verlinde, 2011), leading to smaller $\overline{r_{liq}}$ over land than over the ocean (Twomey effect, Lamb and Verlinde 2011). In the SH, the opposite is observed; in this region, the orography may affect this result more: in particular, the continental regions between 30° and 60° are mainly mountainous, so orographic updrafts may cause the growth of supercooled liquid droplets (Korolev and Mazin, 2003), making the dynamics possibly a more important contributor to larger $\overline{r_{liq}}$ over land in the SH.

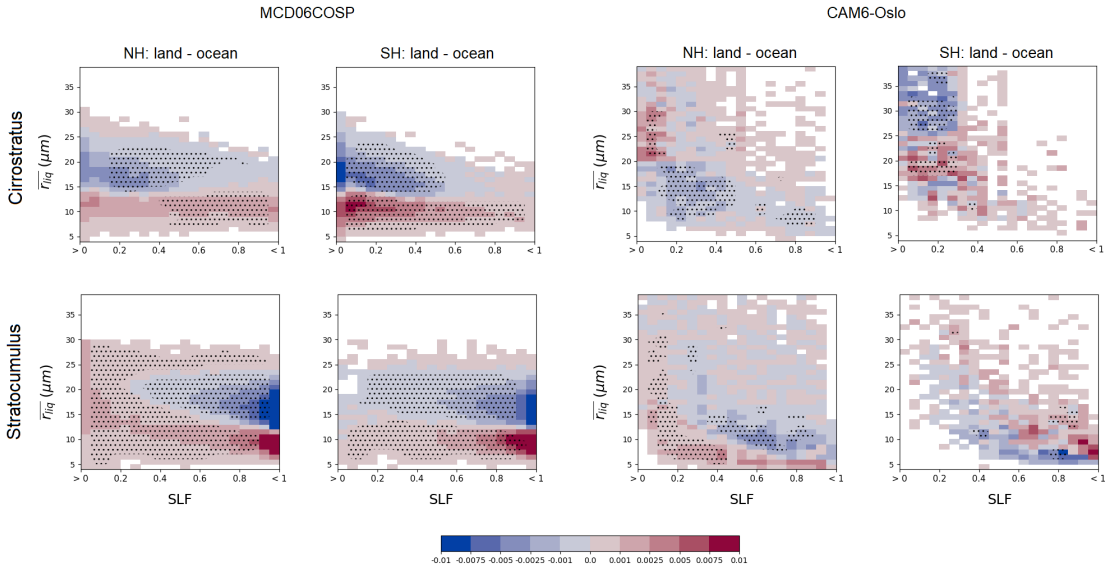

Figure 6.4.: Differences of SLF-$\overline{r_{liq}}$ histograms (land minus ocean) in the NH and SH for cirrostratus and stratocumulus clouds in MCD06COSP (left) and CAM6-Oslo (right). The dots represent the SLF-$\overline{r_{liq}}$ combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar indicates the difference between two SLF-$\overline{r_{liq}}$ histograms for a given SLF-$\overline{r_{liq}}$ bin.

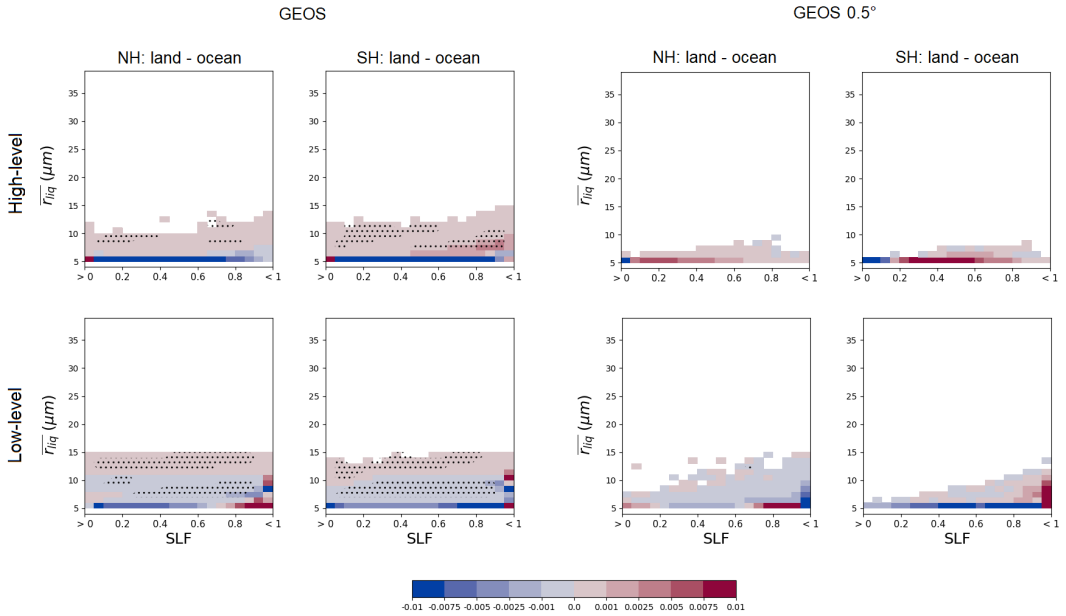

Figure 6.5.: Differences of SLF-$\overline{r_{liq}}$ histograms (land minus ocean) in the NH and SH for high- and low-level clouds in GEOS (left) and GEOS 0.5° (right). The dots represent the SLF-$\overline{r_{liq}}$ combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar indicates the difference between two SLF-$\overline{r_{liq}}$ histograms for a given SLF-$\overline{r_{liq}}$ bin.

Although high-level clouds do not qualitatively show differences in NH and SH, but rather a land-ocean dependence, the supercooled droplets over the ocean, which are larger and/or smaller than over conti-

nents, leave open questions about possible smaller-scale regional dependencies. Also, while smaller $\overline{r_{liq}}$ could be explained by evaporation of liquid droplets to allow ice crystal growth through the WBF process (Korolev, 2007) as in deep convective clouds, larger $\overline{r_{liq}}$ could be due to strong updrafts (Korolev and Mazin, 2003) expected in deep convective cloud systems but absent in cirrus and cirrostratus clouds. The regional coexistence of strong convective cloud systems and cirriform clouds should be further investigated to clarify whether the larger $\overline{r_{liq}}$ over the ocean may depend on specific micro- and macroscale processes in the cloud.

These results are only partially reproduced by the other datasets. MCD06COSP agrees with Cloud_cci v3 only for altostratus, nimbostratus, stratocumulus and stratus clouds. In general, MCD06COSP shows larger $\overline{r_{liq}}$ over the ocean than over land in both NH and SH, regardless of the cloud type. This could be due to the lower aerosol concentration over the ocean than over land, which together with the higher amount of water vapour leads to larger droplets. CAM6-Oslo land-ocean differences in radius are not always clear for high-level clouds, which complicates the comparison with Cloud_cci v3. However, mid- and low-level clouds in CAM6-Oslo are mostly consistent with Cloud_cci v3, with $\overline{r_{liq}}$ larger over land than over the ocean in the SH, and the opposite in the NH. However, it should be stressed again that MCD06COSP and the COSP simulator implemented in CAM6-Oslo use only one spectral channel to retrieve the optical properties of the clouds. Between GEOS and GEOS 0.5°, the land-ocean contrast in GEOS 0.5° is better in agreement with Cloud_cci v3, supporting the importance of comparing datasets at similar spatial resolution.

6.3. Hemispheric dependencies of cloud types on liquid effective radius

In section 4.4, the comparison of cloud types for different regions using passive and active observations showed higher SLF in the SH than in the NH, except for the continental low-level clouds, for which lower SLF values were found in the SH than in the NH. This result was confirmed by Cloud_cci v3 using a different method and partially by CAM6-Oslo and SCREAM in section 5.7.

In the current section, $\overline{r_{liq}}$-related hemispheric dependencies are discussed in order to improve the understanding of the results presented in sections 4.4 and 5.7.

Figure 6.6 shows SLF-$\overline{r_{liq}}$ histograms for different regions of cirrostratus — representative of mid- and high-level clouds — and stratocumulus clouds — representative low-level clouds — with their differences between NH and SH over continental and marine regions for Cloud_cci v3. Cirrostratus clouds show smaller $\overline{r_{liq}}$ for all SLF values in the NH than in the SH over land (excluding the smallest values of $\overline{r_{liq}}$ for which the histograms show very low occurrence), but the opposite over ocean, while stratocumulus clouds show smaller $\overline{r_{liq}}$ in the NH than in the SH for all SLF values over both land and ocean. From these results, only the NH-SH difference for continental stratocumulus clouds pass the significance test.

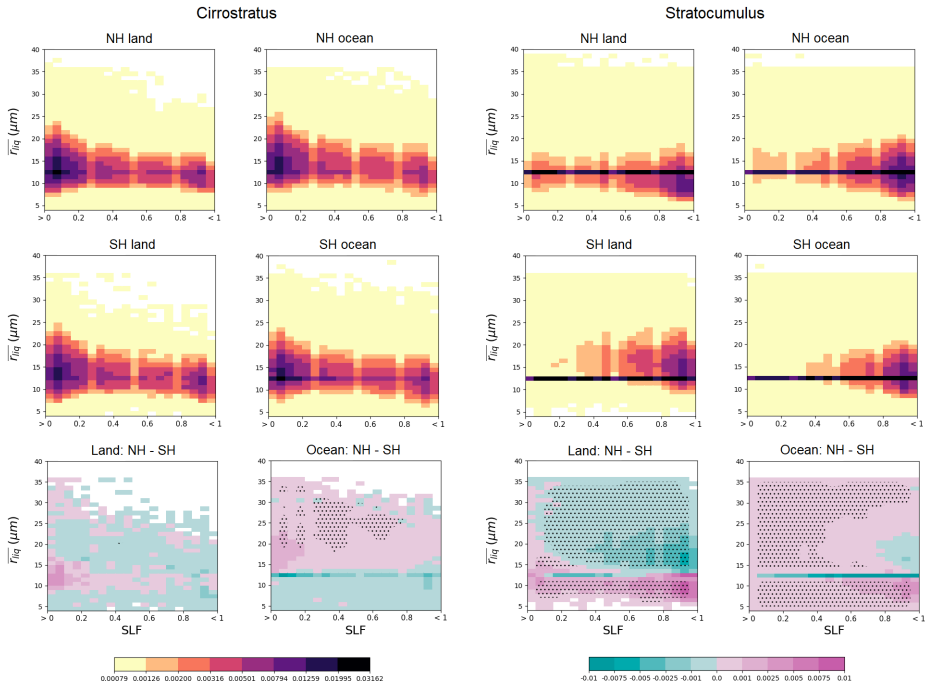

Figure 6.6.: In the top two rows: SLF-$\overline{r_{liq}}$ joint histograms of mixed-phase altostratus (left) and stratocumulus (right) in NH and SH over land and sea for Cloud_cci v3. Bottom: The difference of land minus ocean for NH (left) and SH (right). The dots represent the SLF-$\overline{r_{liq}}$ combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar on the left indicates the cloud fraction for a given SLF-$\overline{r_{liq}}$ bin relative to the individual histogram. The colorbar on the right indicates the difference between two SLF-$\overline{r_{liq}}$ histograms for a given SLF-$\overline{r_{liq}}$ bin.

MCD06COSP shows good agreement with Cloud_cci v3 for marine low- and high-level clouds and only partially over the continents. CAM6-Oslo shows good agreement with Cloud_cci v3 over land, but opposite results over the ocean. These results are shown in Fig. 6.7. While GEOS shows agreement with Cloud_cci v3 for low-level clouds, GEOS 0.5° shows agreement for both high- and low-level clouds (Fig. 6.8), with smaller $\overline{r_{liq}}$ in the NH than in the SH for both continental and marine low-level clouds but larger $\overline{r_{liq}}$ in the NH than in the SH for both continental and marine high-level clouds.

Discussion

Previous studies based on aircraft campaigns (Minikin et al., 2003; Hamilton et al., 2014; Williamson et al., 2021) have shown that the aerosol concentration in the NH is way much higher than in the SH at different altitudes and regardless of the surface type (land or sea). Smaller $\overline{r_{liq}}$ in the NH may be a direct consequence of this (Twomey effect, Lamb and Verlinde (2011)), occurring over both land and ocean for low-level clouds, but only over land for mid- and high-level clouds. The reason why mid- and high-level clouds show smaller $\overline{r_{liq}}$ in the NH than in the SH over land but not over the ocean may be related to other micro- and macrophysical processes contributing more than the Twomey effect at those heights: a possibility is that, over land, $\overline{r_{liq}}$ is smaller in the NH than in the SH also because of the mainly

mountainous orography in the SH (see Discussion in section 6.2), which may cause larger supercooled liquid droplets (Korolev and Mazin, 2003).

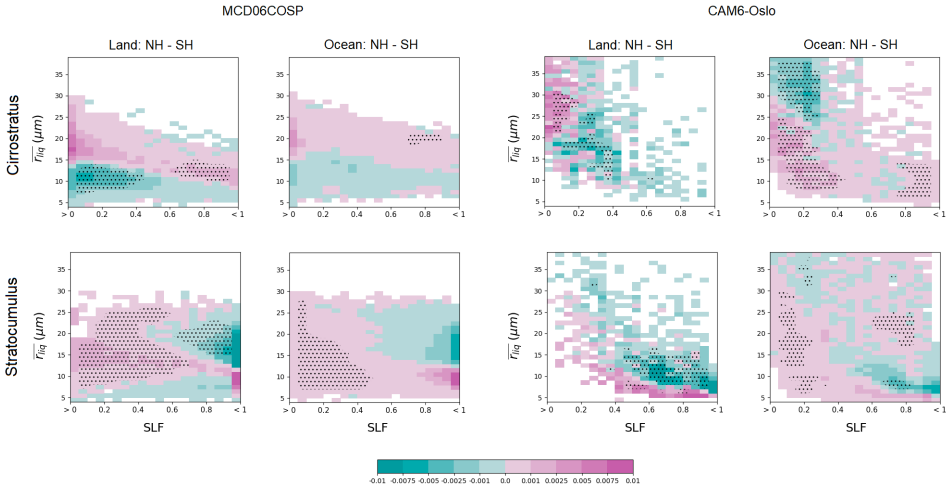

Figure 6.7.: Differences of SLF-$\overline{r_{liq}}$ histograms (NH minus SH) over continental and marine areas for cirrostratus and stratocumulus clouds in MCD06COSP (left) and CAM6-Oslo (right). The dots represent the SLF-$\overline{r_{liq}}$ combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar indicates the difference between two SLF-$\overline{r_{liq}}$ histograms for a given SLF-$\overline{r_{liq}}$ bin.

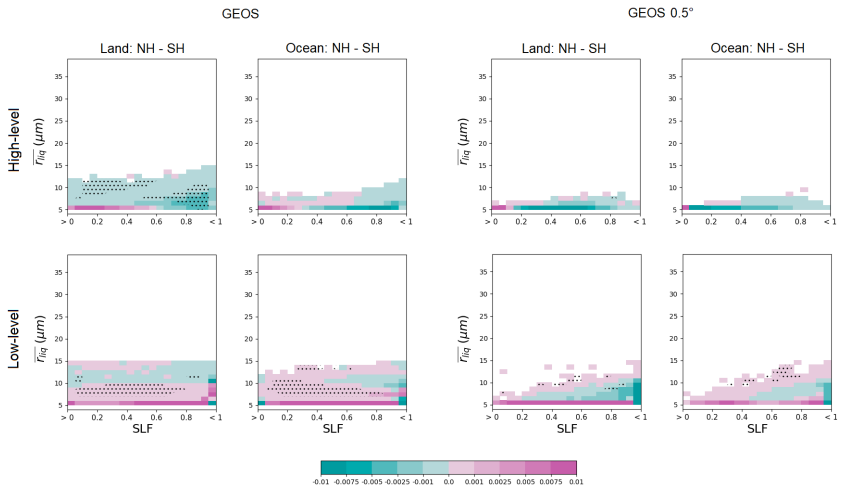

Figure 6.8.: Differences of SLF-$\overline{r_{liq}}$ histograms (NH minus SH) over continental and marine areas for high- and low-level clouds in GEOS (left) and GEOS 0.5° (right). The dots represent the SLF-$\overline{r_{liq}}$ combination where the p-value is less than 0.05, indicating a statistically significant test result. The colorbar indicates the difference between two SLF-$\overline{r_{liq}}$ histograms for a given SLF-$\overline{r_{liq}}$ bin.

6.4. Geographical distribution of liquid droplets in mixed-phase low-level clouds

Figure 6.9 shows the geographical distribution of $\overline{r_{liq}}$ for mixed-phase shallow cumulus, stratocumulus and stratus clouds in Cloud_cci v3, MCD06COSP and CAM6-Oslo, while Fig. 6.10 shows the geo-

graphical distribution of $\overline{r_{liq}}$ for mixed-phase low-level clouds in GEOS and GEOS 0.5° (all cloud types or heights are in figures: A.83 for Cloud_cci v3, A.84 for MCD06COSP, A.85 for CAM6-Oslo, A.86 for GEOS, and A.87 for GEOS 0.5°). In GEOS and GEOS 0.5°, the geographical distribution of $\overline{r_{liq}}$ is averaged on a geographical resolution of 1° for a better visualisation of the results.

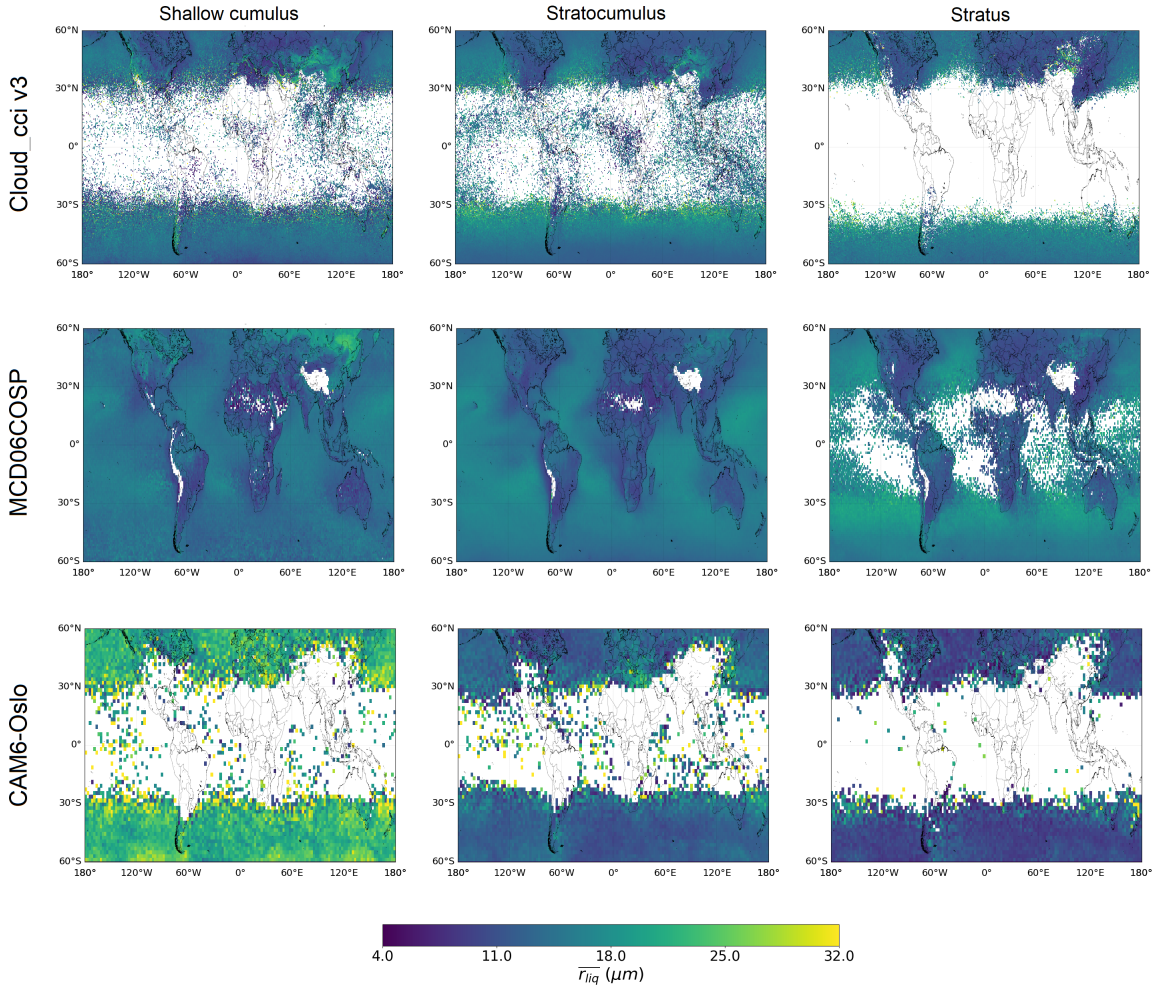

Figure 6.9.: Geographical distribution of $\overline{r_{liq}}$ in shallow cumulus (left), stratocumulus (centre), and stratus (right) clouds in the mixed phase for Cloud_cci v3 (top), MCD06COSP (middle), and CAM6-Oslo (bottom). The colorbar indicates the size (μm) of $\overline{r_{liq}}$.

Continental clouds in Cloud_cci v3 and MCD06COSP generally have smaller $\overline{r_{liq}}$ than marine clouds. However, larger values of $\overline{r_{liq}}$ are also found over continents, in particular in mountainous regions (e.g., the Andes Mountains in South America, and the North American Cordillera). CAM6-Oslo also shows larger values of $\overline{r_{liq}}$ over the Andes Mountains in South America but a less land-ocean contrast compared to the observations. In GEOS and GEOS 0.5° the land-ocean contrast is difficult to detect.
MCD06COSP is the only dataset showing mixed-phase clouds in the tropics (result already shown and discussed in section 5.3).

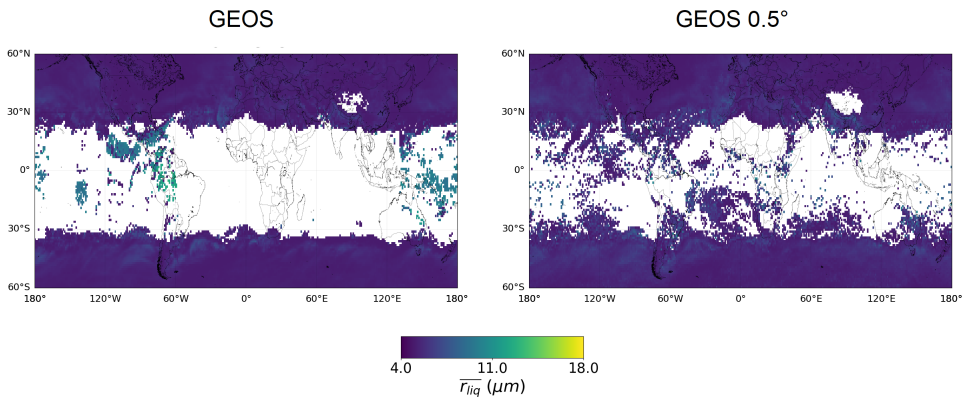

Figure 6.10.: Geographical distribution of $\overline{r_{liq}}$ in low-level clouds in the mixed phase for GEOS (left) and GEOS 0.5° (right). The colorbar indicates the size (μm) of $\overline{r_{liq}}$. The color scale is reduced compared to Fig. 6.9

In CAM6-Oslo and GEOS, with different orders of magnitude, large values of $\overline{r_{liq}}$ are found mainly in the tropics for low-level clouds. Furthermore, the model datasets do not show a strong land-ocean contrast as the observations do. GEOS and GEOS 0.5° show larger $\overline{r_{liq}}$ in mountainous regions in the NH (e.g., the Himalayas and the Rockies Mountains of western Canada) but not in the SH, where observations show larger $\overline{r_{liq}}$.

Discussion

Smaller $\overline{r_{liq}}$ in continental than marine regions may be an indirect effect of the aerosol number concentration, acting as CCN or INP, being greater over land than over the ocean (Lamb and Verlinde, 2011), so that more droplets compete for water vapour over land, growing less than over the ocean.
Continental larger values of $\overline{r_{liq}}$ found in mountainous regions support the hypothesis that orographic updrafts may cause droplets to grow to larger sizes. The continental area in the latitude range from 30° to 60° N is larger than the continental area in the latitude range from 30° to 60° S, which is mainly mountainous. This could explain why continental low-level clouds have larger $\overline{r_{liq}}$ in the SH than in the NH.
The regions where larger $\overline{r_{liq}}$ is found can vary with the dataset and the cloud type. For example, the continental region in the SH with larger droplets in Cloud_cci v3 is found in shallow cumulus clouds, while in MCD06COSP the larger droplets are found in stratus clouds; the continental region in the NH with larger droplets in Cloud_cci v3 includes northeastern China, while in MCD06COSP this region includes southern Russia and part of the east coast of China. These are significant differences that go beyond a possible misclassification of cloud types. A validation study with collocated data would help to further investigate these results and is left for future studies. However, it is possible to refer to the Product Validation and Intercomparison Report PVIR (2020), which includes a comparison of the Cloud_cci AVHRR-AM v3 and AVHRR-PM v3 datasets with the Cloud_cci MODIS-Terra and MODIS-Aqua datasets, respectively. In particular, the comparison of Cloud_cci AVHRR-PM v3 (referred to as

Cloud_cci v3 in this thesis) and MODIS-Aqua (afternoon satellites) shows: larger CTP in MODIS, also increasing with latitude; larger COT in MODIS for liquid clouds, also increasing with latitude and with a difference from AVHRR-PM v3 of more than 10 (but less than 20) at 60°; larger $\overline{r_{liq}}$ in MODIS; liquid cloud fraction in MODIS about 10% larger than in AVHRR-PM v3, globally and for all subregions considered in their study. Mismatches in COT, CTP and phase may be responsible for the differences found in Fig. 6.9, as the mismatch mentioned for $\overline{r_{liq}}$ does not seem to affect the results presented for low-level clouds, for which the opposite occurs (only in high- and mid-level clouds $\overline{r_{liq}}$ is larger in MCD06COSP than in Cloud_cci v3). However, while the validation study PVIR (2020) compares morning and afternoon satellites separately, our analysis includes MCD06COSP in the comparison, which combines observations from MODIS-Aqua and MODIS-Terra and uses only one spectral channel to retrieve the cloud optical properties. This may also contribute to the retrieved mixed-phase clouds in the tropics.
If the presence of larger $\overline{r_{liq}}$ over land in the SH than in the NH for low-level clouds found in Cloud_cci v3 and partially in MCD06COSP is taken as true, the results in sections 4.4 and 5.7 with Cloud_cci v3 showing a lower SLF over land in the SH than in the NH for the same cloud types could be explained by the higher probability of larger droplets grown by orographic updrafts, colliding with ice particles and the possibility to rime and produce secondary ice (Lamb and Verlinde, 2011). Although mountainous regions can add uncertainty to cloud retrievals in the presence of snow (PVIR, 2020; Karlsson et al., 2015; King et al., 2004), the improvements in the Cloud_cci v3 cloud detection algorithm, compared to the that in Cloud_cci v2 (used in Chapter 4) make cloud detection more reliable (PVIR, 2020).
In CAM6-Oslo and GEOS, relatively large values of $\overline{r_{liq}}$ are found mainly in the tropics, where the presence of mixed-phase low-level clouds is not expected. This result may indicate an unrealistic representation of the geometric structure of clouds in models, leading to misclassification of cloud types. The weaker land-ocean contrast found in the models may be due to balanced contributions from aerosol concentrations over the ocean, which are lower than over land (Lamb and Verlinde, 2011) and cause the growth of $\overline{r_{liq}}$, and different cloud dynamics over continents, including orography-related updrafts, which may cause the growth of $\overline{r_{liq}}$ over continents. GEOS and GEOS 0.5° show larger $\overline{r_{liq}}$ in mountainous regions in the NH (e.g., the Himalayas and the Rockies Mountains) but not in the SH, where observations show larger $\overline{r_{liq}}$,
In section 4.4, speculations about a possible diversity of INPs in the continental regions of the SH and NH were made to justify the lower values of SLF found in continental low-level clouds in the SH. This cannot be confirmed by the results presented in this Chapter, as neither the physical nor the chemical nature of the aerosols was investigated, and is left for future studies.

6.5. Conclusions

In this chapter, the comparison of SLF-$\overline{r_{liq}}$ joint histograms for satellite (Cloud_cci v3 and MCD06COSP) and model (CAM6-Oslo, GEOS, and GEOS 0.5°) datasets is presented to better understand cloud micro-

physical processes affecting clouds in continental and marine areas in the latitude ranges 30°-60° North and South, and better interpret the results shown in Chapters 4 and 5. The results presented in this chapter are summarised here:

- All datasets — except GEOS and GEOS 0.5° — show an increase in $\overline{r_{liq}}$ with a decrease in SLF. This may be due to the higher probability of clouds with larger droplets to glaciate: droplets can collide and aggregate as they ascend in the cloud, resulting in larger droplets; in the mixed-phase temperature range, colliding droplets can glaciate and cause secondary ice formation, leading to an increase in $\overline{r_{liq}}$ with a decrease in SLF. The disagreement of GEOS and GEOS 0.5° with this result for all cloud types regardless of the hemisphere and the surface type (land or ocean) may be due to the higher probability that larger droplets have to collide and glaciate.

- In Cloud_cci v3, low-level clouds have smaller $\overline{r_{liq}}$ over land than over the ocean in the NH, while the opposite occurs in the SH; high-level clouds show larger $\overline{r_{liq}}$ over the ocean than over land in both the NH and SH. In the NH, low-level clouds may have smaller $\overline{r_{liq}}$ over land than over the ocean because of the higher aerosol concentration over land (Twomey effect); in the SH, the predominantly mountainous orography may affect droplet growth more than aerosol concentration, causing $\overline{r_{liq}}$ to grow through strong updrafts. The results for high-level clouds require further investigation, particularly into possible regional dependencies and the micro- and macroscale processes occurring in these clouds. Except for GEOS, these results are partially reproduced by all datasets: in MCD06COSP, $\overline{r_{liq}}$ is larger over the ocean than over land, regardless of the hemisphere and the cloud type; CAM6-Oslo does not agree with Cloud_cci v3 for high-level clouds in the NH; GEOS 0.5° shows agreements with Cloud_cci v3 but the small variability of $\overline{r_{liq}}$ makes the interpretations of the results difficult.

- In Cloud_cci v3, low-level clouds show smaller $\overline{r_{liq}}$ in the NH than in the SH over both land and the ocean, while high- (and mid-level) clouds reproduce this result only over land and show the opposite over the ocean. The aerosol concentration, which is higher in the NH than in the SH, may cause smaller $\overline{r_{liq}}$ in the NH than in the SH (Twomey effect); mid- and high-level clouds may be affected by other processes; over land, in addition to the Twomey effect, the predominantly mountainous orography in the SH may contribute to larger $\overline{r_{liq}}$ than in the NH. These differences between land and ocean are not uniformly reproduced by all datasets: MCD06COSP shows agreement with Cloud_cci v3 for low- and high-level clouds over the ocean and only partially over land; CAM6-Oslo shows good agreement with Cloud_cci v3 over land, but not over the ocean; GEOS shows agreement with Cloud_cci v3 for low-level clouds, while GEOS 0.5° shows agreement for both high- and low-level cloud.

- Geographical distributions of $\overline{r_{liq}}$ from observations show: smaller $\overline{r_{liq}}$ over land than over the ocean in low-level clouds, probably due to the higher aerosol number concentration over land than

over the ocean (Twomey effect); over land, larger values of $\overline{r_{liq}}$ are found in mountainous regions, probably due to orographic updrafts which can make droplets grow larger; low-level mixed-phase clouds in the tropics in MCD06COSP, probably due to misclassification of cloud types and the use of only one channel to retrieve cloud optical properties. Models do not show a clear land-ocean contrast like observations: this may be due to different processes contributing to the growth of $\overline{r_{liq}}$ over both land (e.g., the presence of orography) and the ocean (e.g., the lower aerosol number concentration compared to continental regions).

- Mountainous regions can add uncertainty to cloud retrievals if snow is present (PVIR, 2020; Karlsson et al., 2015; King et al., 2004). Nonetheless, Cloud_cci v3 provides a more realistic representation of clouds compared to Cloud_cci v2.

7. Comparison of the global and Arctic cloud phase distribution from GCMs and satellite observations[1]

While the previous chapters (4, 5, and 6) focused on latitudes within the tropics and midlatitudes, this chapter focuses on the Arctic, a vast region where satellite sensors mostly have problems with measurements (see section 2.2.4) and on the whole globe.

7.1. Motivation

Tan and Storelvmo (2019) shows that a correction for the supercooled liquid fraction in mixed-phase clouds in global simulations affects the Arctic amplification and that the strength of the Arctic amplification depends on the cloud's microphysical properties.
While the fully-coupled simulations of Tan and Storelvmo (2019) examine how a global SLF correction affects the Arctic, the first part of this study uses atmosphere-only simulations to investigate how and why the Arctic is affected by local changes in cloud microphysics.
The cloud phase in mixed-phase clouds is not homogeneously distributed (Zhang et al., 2019b; Sokol and Storelvmo, 2022). In GCMs, an overly efficient WBF process can lead to an underestimation of liquid water in mixed-phase clouds (Komurcu and Storelvmo, 2014; Cesana et al., 2015; McIlhattan et al., 2017), as models represent clouds as a homogeneous mixture of liquid droplets and ice crystals (Korolev, 2007; Storelvmo et al., 2015). Improving model parameterisation would lead directly to model improvement (Jakob, 2010). For this reason, model tuning to observations becomes relevant for model improvement and is applied on a global scale in the second part of this study.

7.2. Datasets

For this study, satellite observations and GCMs are analysed from 1 June 2009 to 31 May 2013.
The observations used are CALIOP Level 2 Cloud Layer Data in version 4.20 (Hu et al., 2009, see section 3.1.2).
NorESM2-MM (NorESM2 Medium atmosphere-Medium ocean resolution, i.e., as NorESM2 — see section 3.2.1 — but with a higher horizontal resolution of 1.25° x 0.9375°) is used for the global study.

[1] This chapter is the result of my collaboration with the University of Oslo. More about these studies can be found in Shaw, J., Z. McGraw, O. Bruno, T. Storelvmo, and S. Hofer (2022) and Hofer, S., F. Hellmuth, M. Pietschnig, L. Hahn, I. Mostue, J. Shaw, O. Bruno, R. David, Z. McGraw, T. Carlsen, and T. Storelvmo (2023). My contributions to these publications are related to observational data curation and analysis, methodology, and writing (writing only in Hofer et al. (2023)).

Three models are used as GCMs for the Arctic analysis (more about Arctic clouds in section 2.2.4): two modified versions of NorESM2 and a modified version of the Community Earth System Model Version 2 (CESM2).

7.2.1. NorESM2 — CAM6-Oslo

A description of the NorESM2 model (Seland et al., 2020) is given in section 3.2.1, where it was named CAM6-Oslo. The CAM6-Oslo model differs from the version used for the analysis shown in the previous chapters in the two-moment bulk microphysics scheme, which is modified here to limit the in-cloud ice number concentration to the calculated concentration of available ice nuclei, i.e., although the Hallet-Mossop secondary ice scheme is not disabled, the ice number concentration cannot exceed 10^6 kg^{-1}.

7.2.2. CESM2 — CAM6

The Community Earth System Model Version 2 (CESM2, Danabasoglu et al. 2020) is the latest generation of the coupled climate/Earth system models developed by the National Center for Atmospheric Research (NCAR) and partners. It differs from CAM6-Oslo in the use of the MAM4 (Liu et al., 2016) aerosol scheme instead of OsloAero5.3 and in a different parameterisation of mid- and high-level ice clouds. As CESM2 also uses CAM6, this model will be named as CAM6.

7.2.3. NorESM2 — CAM6-OsloIce

A third model is used for this study, obtained by modifying CAM6-Oslo: CAM6-OsloIce. Specifically, the ice number limit at temperatures between -37°C and 0°C is disabled to restore heterogeneous and secondary ice production through the Hallet-Mossop process. To avoid strong secondary ice production due to the absence of a limitation on ice number, a limit on the ice number tendency variable from secondary ice production has been introduced. In CAM6-OsloIce, scavenged INPs are replaced by new INPs in the next model time step by transport and entrainment.

7.3. Methods

In this study, two SLF metrics are used: the cloud top SLF, which corresponds to the SLF at the highest cloud layer below COT = 0.3 (as in section 3.3), and the cloud bulk SLF, which is computed as the SLF in the upper layers with COT between 0 and 3 (Fig. 7.1). The SLF has been calculated on isotherms from -40°C to 0°C, with an increment of 5°C. The CALIOP observations included in this study cover the period from 1 June 2009 to 31 May 2013.

7.3.1. Comparison over the Arctic region

The analysed Arctic region consists of the area between 66° and 82° N. In order to minimise the root mean square error of both the cloud top and cloud bulk SLF, the time scale of the WBF process and the number of aerosols active as INPs were iteratively modified in the base models (CAM6-Oslo, CAM6, and CAM6-OsloIce). The efficiency of the WBF process and the INP number concentration in the fitted models are scaled by multipliers to obtain higher or lower SLF in order to better fit with the satellite observations.

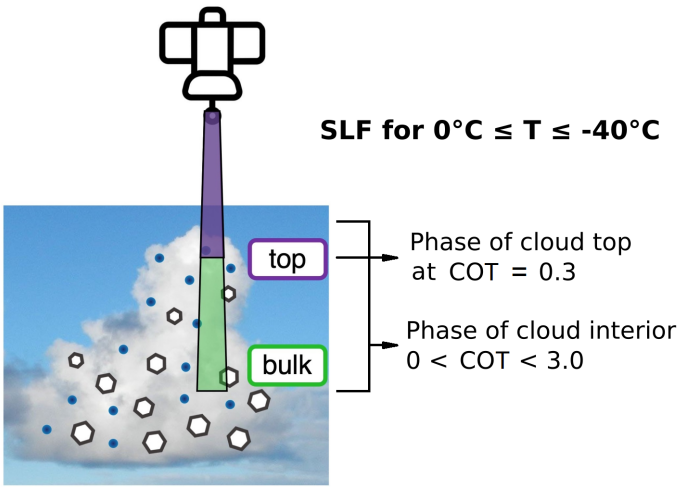

Figure 7.1.: Representation of the cloud top and the cloud interior (bulk) used to retrieve the cloud phase from CALIOP observations. This figure is adapted from Hofer et al. (2023)

Table 7.1 summarises the base and fitted models. Ice crystal size and number concentration variables in the constrained runs (termed "Fit" in Table 7.1) are consistent with the observations from the M-PACE experiment in northern Alaska in October 2004 (Prenni et al., 2009), indicating that ice crystals are reasonably reproduced by the models.

Run name	Model	Ice number limit (secondary ice limit)	WBF multiplier	INP multiplier
CAM6-Oslo	NorESM2	Yes (No)	1.0	1.0
CAM6	CESM2	Yes (No)	1.0	1.0
CAM6-OsloIce	NorESM2	No (Yes)	1.0	1.0
CAM6-Oslo Fit 1	NorESM3	Yes (No)	1.25	10.0
CAM6-OsloIce Fit 2	NorESM2	No (Yes)	0.5	0.05
CAM6-OsloIce Fit 3	NorESM2	No (Yes)	0.2	0.1
CAM6 Fit 4	CESM2	Yes (No)	1.0	100

Table 7.1.: List of the models with the original and with modified (named "Fit") prescribed parameterisations.

7.3.2. Global and regional comparison

NorESM2-MM was tuned by limiting the WBF process efficiency and the INP concentration, and by modifying the temperature-based linear relationship that controls the phase of cloud particles being detrained from convective updrafts in NorESM2: in the "Control" run, all particles detrained below $T_{Ice\_Detrainment} = -35°C$ from convection consist of ice, all particles detrained above $T_{Liq\_Detrainment} = -5°C$ consist of liquid, and these temperatures must satisfy the linear temperature ramp defining the fraction of ice to liquid particles as in eq. 7.1:

$$r = \frac{T_{Liq\_Detrainment} - T}{T_{Liq\_Detrainment} - T_{Ice\_Detrainment}} \tag{7.1}$$

By changing the appropriate parameters (Table 7.2), the differences between the observed and simulated SLF metrics (top and bulk) were reduced: globally (NorESM2 Global), only in the northern extratropics (from 30° to 90° N, NorESM2 N-ET), and only in the southern extratropics (from 30° to 90° S, NorESM2 S-ET).

	WBF multiplier	INP multiplier	$T_{Liq\_Detrainment}$	$T_{Ice\_Detrainment}$
Control	1	1	-5°	-35°
Global	0.001	0.5	-30°	-35°
N-ET	0.001	0.5	-20°	-35°
S-ET	0.001	0.5	-25°	-35°

Table 7.2.: List of the parameterisations in the "Control" and the modified runs.

7.4. Results

7.4.1. Model tuning over the Arctic region

Figure 7.2(a) shows the cloud-top and cloud-bulk SLF on isotherms from CALIOP and the base models. Between $-35°C$ and $-10°C$, the cloud top SLF is larger than the cloud bulk SLF for CALIOP, with the largest difference at $-20°C$: at this isotherm, the cloud top SLF is more than twice the cloud bulk SLF. At $0°C$ and $-5°C$, the cloud top SLF is still larger than the cloud bulk SLF for CALIOP, but within one standard deviation for these isotherms. Although all models also show that the cloud top SLF is larger than the cloud bulk SLF, they differ quantitatively from CALIOP: while both the cloud top and cloud bulk SLF of CAM6-Oslo show strong agreement with the observations, CAM6 overestimates the cloud top SLF and CAM6-OsloIce underestimates both the cloud top and cloud bulk SLF. The underestimation found in OsloIce is due to a high ice number concentration causing a rapid depletion of liquid water. CAM6-Oslo and CAM6-OsloIce differ mostly in SLF at temperatures below $-20°C$, i.e., where secondary ice processes are not active, showing that more freedom in heterogeneous nucleation processes

in CAM6-OsloIce leads to cloud changes and that the limitation of INPs in CAM6-Oslo is necessary to maintain liquid cloud tops at these temperatures.

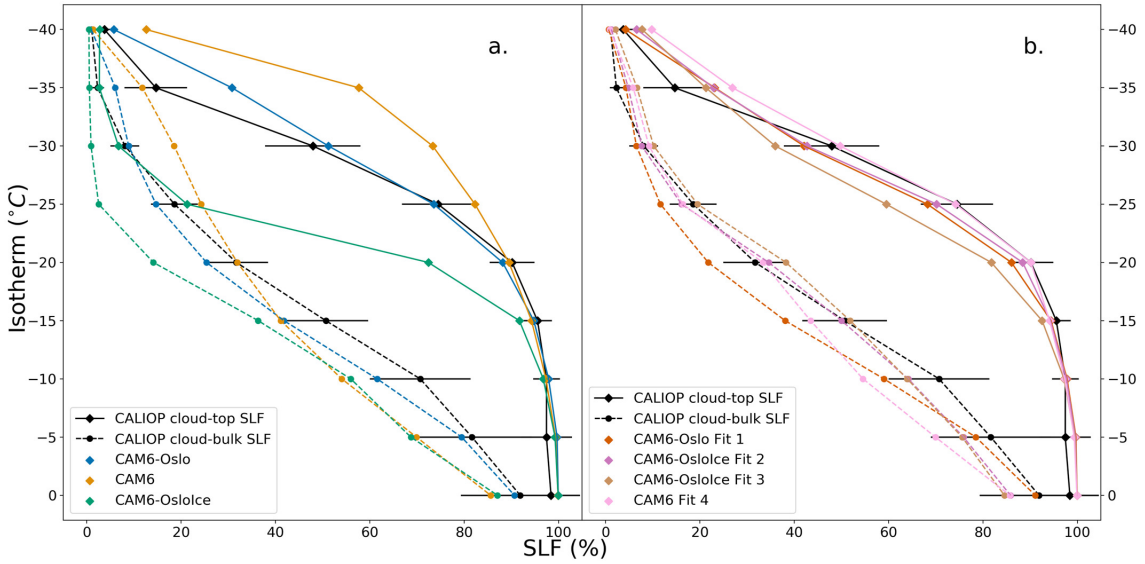

Figure 7.2.: Cloud-top and cloud-bulk supercooled liquid fraction (SLF) on isotherms for (a) base models and (b) fitted models. Error bars for CALIOP indicate one standard deviation. All values are area-weighted averages from 66° to 82°N. From Shaw et al. (2022), distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Licence.

Figure 7.2-b shows the cloud top and cloud bulk SLF on isotherms from CALIOP and the fitted models. The presence of more liquid water at the cloud top than in the cloud (cloud bulk) found in Arctic mixed-phase clouds using CALIOP is consistent with previous observational studies (Hobbs and Rangno, 1998; Pinto, 1998; Curry et al., 2000; Verlinde et al., 2007). This vertical phase structure is better represented by CAM6-Oslo than by CAM6, suggesting that high cloud parameterisations and model aerosol schemes play an important role in cloud phase determination.

7.4.2. Model tuning globally and regionally

The geographical distribution of the difference between SLF at the cloud top and at the cloud bulk on the -20°C isotherm (Fig. 7.3(a)) shows that this difference is positive worldwide and that the higher amount of liquid phase at the cloud top compared to the lower cloud layers is not only a feature of certain cloud types or regions (Carey et al., 2008; Zhang et al., 2010; Shaw et al., 2022; Yang et al., 2022), but extends worldwide and is independent of cloud type. The SLF difference is higher in the extratropics and lower in the tropics, where convection is stronger and can lead to well-mixed layers, hindering the formation of separate cloud phase layers.

Figure 7.3(b) shows that the CALIOP lidar can penetrate opaque clouds, i.e., clouds in which the lidar signal is quickly completely attenuated compared to other clouds, for 1.67 ± 0.49 km on average and at least 1 km at all latitudes. This information is important to affirm that the cloud bulk includes at least 1

km of cloud, and therefore the SLF associated with the cloud interior refers to different thermodynamic properties than the SLF at the cloud top.

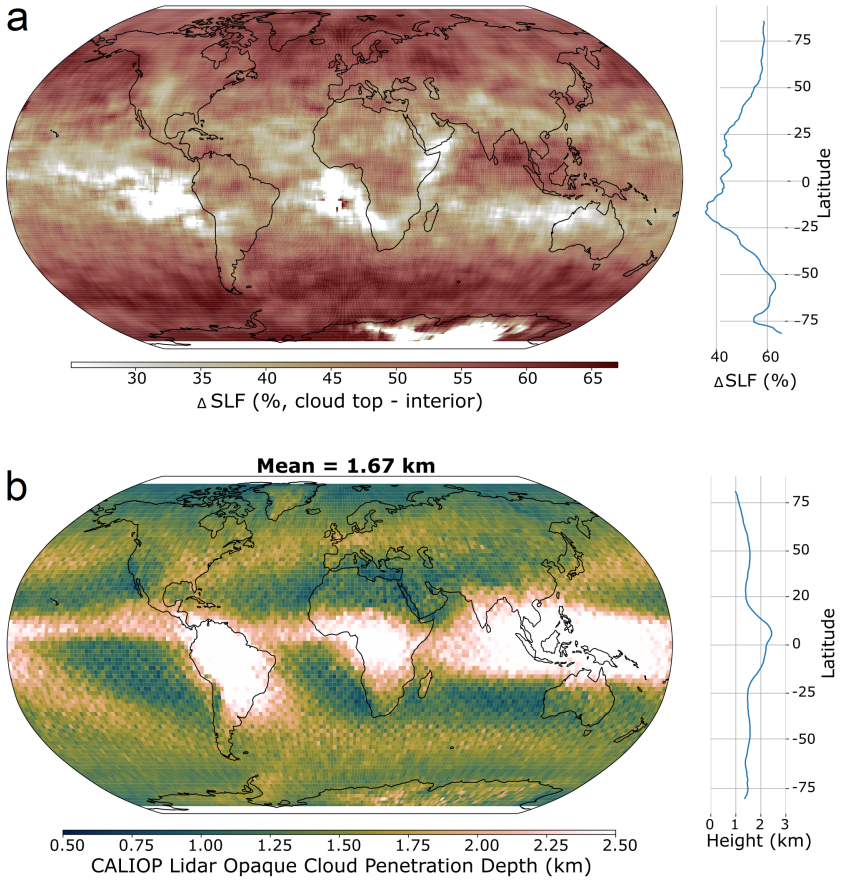

Figure 7.3.: a) Geographical differences of SLF (%) for cloud top and cloud bulk on the isotherm at -20°C (left) and its mean cross-section over latitude (right). b) Global CALIOP penetration depth (km) for opaque clouds (left, detailed data description in Guzman et al. (2017)) and its mean cross-section over latitude (right). This figure is adapted from Hofer et al. (2023)

Figure 7.4 shows that the SLF from observations is higher at the cloud top than for cloud interiors globally (a), constrained to the northern (b) and southern (c) extratropical regions, and independent of season. The largest SLF difference between cloud tops and interiors is found between $-30°C$ and $-10°C$. This temperature range may be defined by the limited INP activation at temperatures above $-10°C$ and the more and more increasing probability of homogeneous ice crystals nucleation for temperatures below $-30°C$.

The extratropical regions (Fig. 7.4(b)) show more seasonal variability in the NH than in the SH. The continental region from 30° latitude in the NH is much more heterogeneous than in the SH: more vegetation and more human activity in the NH can influence the aerosol composition with seasonal variability more than in the SH.

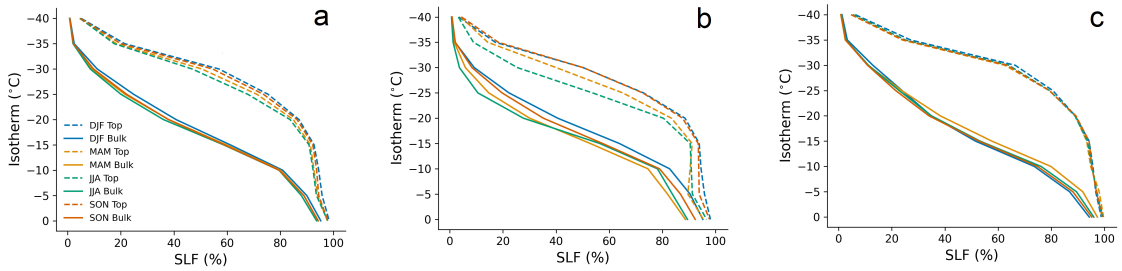

Figure 7.4.: CALIOP seasonal supercooled liquid fraction (SLF) at the cloud top (dashed line) and cloud bulk (solid line) on isotherms for: a) the whole globe; b) northern extratropical regions; c) southern extratropical regions. Different colours represent different seasons. This figure is adapted from Hofer et al. (2023)

Compared to CALIOP observations, NorESM2 produces too much ice (Fig. 7.5, left). After model tuning, NorESM2 reproduces the temperature dependence of the SLF well for all constrained regions (global, N-ET, S-ET), with consequences for the net radiation at the top of the atmosphere (Fig. 7.5, right). In particular, forcing mixed-phase clouds to have globally higher SLF (Fig. 7.5(a)) leads to a cooling effect at the top of the atmosphere in the extratropical regions, where supercooled liquid water is mostly observed (Hu et al., 2010). A possible reason for this result may be that a higher supercooled liquid fraction is associated with optically thicker (Bruno et al., 2021) and brighter clouds, resulting in more shortwave radiation being reflected back to space. The cooling effect is greater in the SH than in the NH at mid-latitudes. This may be due to the greater extent of the marine region in the SH, which is a better source of water vapour than the NH, and thus produces optically thicker clouds through larger droplets (see section 6.2).

When the model is tuned with regional (N-ET and S-ET) constraints on the SLF cloud top and bulk (Fig. 7.5(b)-(c)), similar regional cooling effects are obtained as for the global constraint. However, both constraints applied separately in the N-ET and S-ET produce effects in the tropics and (less) in the other hemisphere, although the atmospheric circulation cannot respond to local changes in net energy, as the atmospheric circulation is prescribed by the ERA reanalysis when the model is tuned with the observed SLF at the cloud top and inside the cloud.

7.5. Conclusions

Mixed-phase clouds in the Arctic have been studied using three models based on the atmospheric model CAM6. These models differ in their cloud microphysics and are named: CAM6-Oslo, CAM6 and CAM6-OsloIce. The modelled supercooled liquid fraction (SLF) at the cloud top and inside the cloud was compared with CALIOP observations. Both observations and models show higher SLF at the cloud top than in the cloud interior, suggesting that model aerosol schemes and high cloud parameterisations affect the cloud phase.

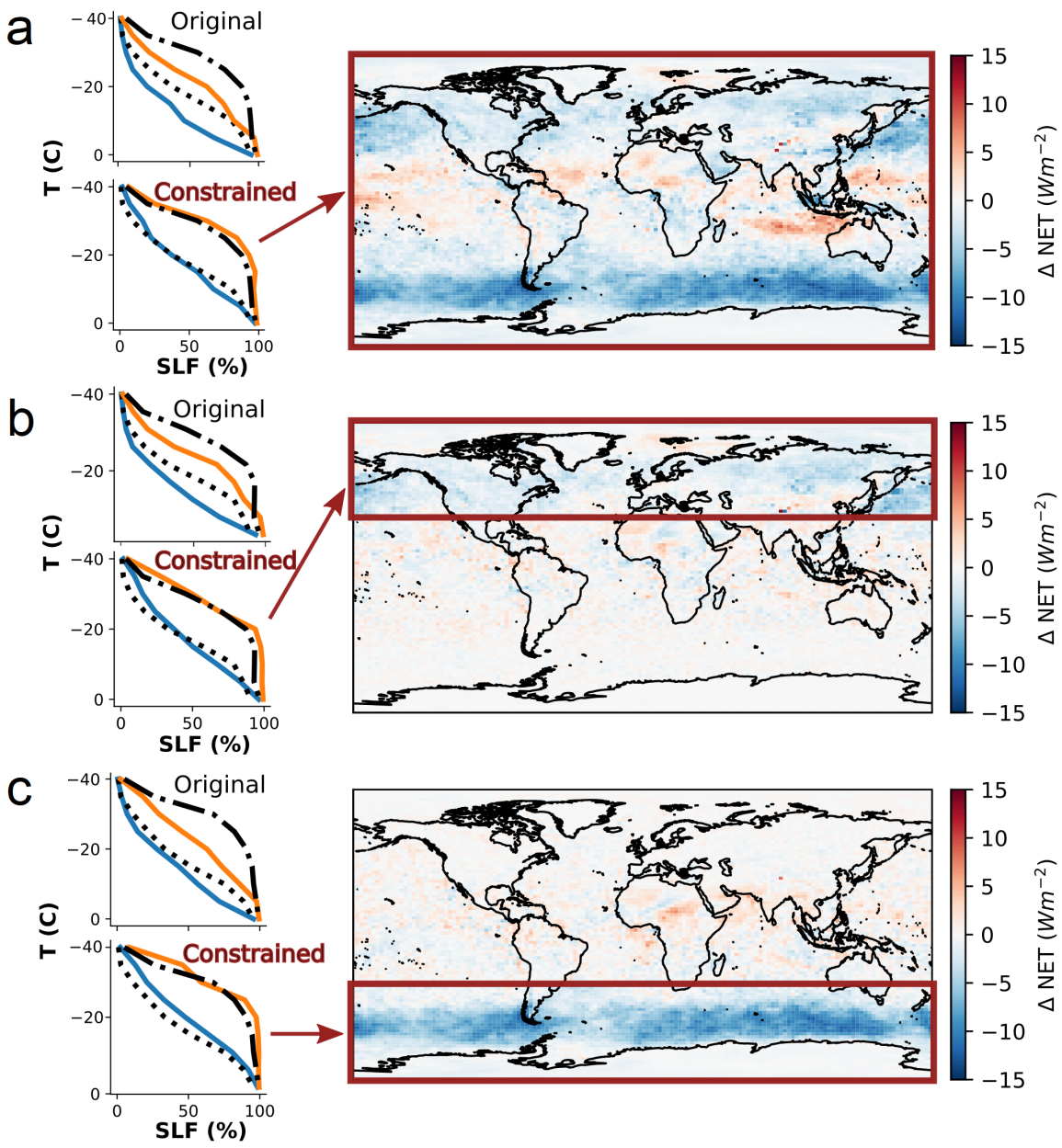

Figure 7.5.: Left: Supercooled liquid fraction (SLF) on isotherms at the cloud top (black dash-dot line for CALIOP, orange line for NorESM2) and cloud bulk (black dotted line for CALIOP, blue line for NorESM2) for: a) NorESM2 Global; b) NorESM2 N-ET; c) NorESM2 S-ET. Right: Geographical distribution of the difference in the net radiation at the top of the atmosphere of NorESM2 after and before tuning. This figure is adapted from Hofer et al. (2023)

This result is present both before and after the application of multipliers aimed at modifying the WBF process efficiency and the INP number concentration. The use of multipliers to modify the WBF process efficiency and the INP number concentration to tune the models leads to a representation of the SLF at the cloud top and inside the cloud closer to the observations.

Mixed-phase clouds have also been studied globally with NorESM2-MM, based on the CAM6 atmospheric model. CALIOP observations and NorESM2-MM show that mixed-phase clouds have a higher

SLF at the top than in the interior. This result is confirmed for CALIOP regardless of the season. Modifications of the INP concentration, the WBF process efficiency, and the temperature above which particles detrained from convection are liquid were applied to reduce the difference between the observed and simulated SLF at the top and in the interior. Both these SLF metrics were used for the first time to tune the model globally and to constrain the northern and southern extratropical regions. After model tuning, the simulated SLF at the top and in the interior of mixed-phase clouds are in good agreement with observations. As the model tuning produces higher SLF, it leads to a cooling effect on the net energy at the top of the atmosphere in the northern and southern mid-latitudes. This cooling effect is also obtained locally when the constraint is restricted to the northern and southern extratropical regions separately.

8. Conclusions and Outlook

8.1. Conclusions

This thesis provides a statistical analysis of the phase partitioning in mixed-phase clouds using observational and model data. The first part of the work focuses on cloud tops from 60°N to 60°S, in extratropical Northern and Southern Hemispheres separately, in the Tropics, over continental and oceanic areas, and for different cloud types. To classify the cloud types, cloud top pressure and cloud optical thickness thresholds were used (Rossow and Schiffer, 1999). This part of the work contains a validation study comparing datasets derived from active and passive satellite sensors and a comparison between observational and model datasets. The second part of the work focuses on improving the representation of mixed-phase clouds in global climate models (GCMs) in the Arctic and globally by tuning them.
The validation study consists of a 4-year (from 1 June 2009 to 31 May 2013) statistical analysis to better understand the relationship between cloud top phase and temperature in the mixed-phase temperature range (-40°C to 0°C). This study is based on four datasets (Cloud_cci v2, Cloud_cci v3, CLARA-A2, and CALIOP v4.20) and consists of comparing the retrieved cloud top phase and cloud top temperature in terms of the supercooled liquid fraction (SLF, see section 3.3.2.1) for specific isotherms on pixel-base. The study included collocated data, to determine the inconsistencies between retrievals, and non-collocated data for the main study. The results of the validation study are:

- The collocated data show a warm bias of the AVHRR CTT compared to CALIOP and a phase mismatch (liquid cloud tops in CALIOP are retrieved as ice in the AVHRR-based datasets). This may be due to the different sensors, one passive and one active.

- For clouds with temperatures from -40°C to 0°C at the same height level, the supercooled liquid fraction increases with cloud optical thickness.

- The supercooled liquid fraction is higher in the Southern Hemisphere than in the Northern Hemisphere between 30° and 60°, except for low-level continental clouds, for which the opposite occurs.

Despite the phase and temperature mismatch, the passive and active satellite retrievals are in qualitative agreement, showing that the passive satellite sensor AVHRR can contribute to cloud-phase research once its limitations are taken into account.
In the comparison between observations (Cloud_cci v3 and MCD06COSP) and model datasets (CAM6-Oslo, ICON, SCREAM, and GEOS), not all datasets provide 4-year data. For this reason, only February

2020 was analysed for ICON, SCREAM and GEOS from the DYAMOND Winter project. The use of different metrics for SLF was necessary to deal with the different quantities provided by the datasets for the cloud thermodynamic phase; SLF based on cloud fraction and mass fraction were used in addition to the pixel-based SLF defined and used in the validation study. Since gridboxes were introduced to compute SLF in Cloud_cci v3, a clustering algorithm (i.e., k-means clustering) was used to classify different cloud types, optimised with prescribed initial seeds. This optimisation allowed not only the most common cloud types to be captured, but also the very rare cloud types that would otherwise have been missed by the k-means algorithm. In order to compare the datasets as consistently as possible, the spatial resolution of the DYAMOND datasets was coarsened to 0.5°. In addition, the spatial resolution of the DYAMOND datasets was coarsened to 1° and 2° for a sensitivity study comparing the results at the three coarsened resolutions. To classify cloud types in the DYAMOND datasets at coarser resolutions again, the k-means algorithm with prescribed initial seeds was used. To mimic the satellite view, the cloud top had to be defined in the models. The comparison between the observational and model datasets was based on the results of the validation study, in which a NH-SH and a land-ocean dependence was found for all analysed datasets associated with cloud types: therefore, the comparison focuses on latitudes between 30° and 60° North and South, separating clouds in continental and marine areas. For the analysis, SLF-CTT (cloud top temperature) joint histograms were used. To verify if the datasets reproduced the results obtained in the validation study, CFAD (Contoured frequency by altitude diagram)-like histograms were used. The results of this comparison between the observational and model datasets are:

- There is a large variability in the amount of clouds retrieved by the datasets (from 52.9% in Cloud_cci v3 to 87.3% in MCD06COSP for all clouds and from 55.0% in MCD06COSP to 3.1% in SCREAM for mixed-phase clouds). CAM6-Oslo, ICON, SCREAM-CS, and GEOS simulate a near-global amount of "all" clouds ranging from 65.5% to 72.1% close to previous studies based on satellite measurements (ranging from 61.9% to 67.2% for afternoon satellites; PVIR, 2020). This thesis provides new references for the amount of mixed-phase clouds, which includes all cloud types up to 60° latitude.

- Comparing all or mixed-phase clouds at high, middle, and low heights, the datasets mostly agree on the geographical location of the clouds. For all clouds, the observations are dominated by low-level clouds while in the model datasets high-level clouds dominate. The model datasets show a systematic decrease in the cloud top height for high-level mixed-phase clouds compared to all clouds. Most of the datasets agree in placing mixed-phase clouds at low heights and at mid-latitudes.

- A validation study comparing ICON, SCREAM, and GEOS at coarsened spatial resolutions (0.5°, 1° and 2°) shows that the amount of all clouds decreases when the original spatial resolution is coarsened to 0.5°. This is probably due to the temperature filter applied to each gridbox, which discards cloudy boxes with a difference of less than 20°C between the warmest and coldest cloudy

pixels. This occurs only in GEOS and SCREAM when only the mixed-phase clouds are considered. The amount of clouds decreases systematically in all datasets when the spatial resolution is coarsened to 1° and 2°. Comparing the datasets at coarsened spatial resolutions using SLF on isotherms, ICON for all cloud heights at CTT < −11°C and GEOS for only high-level clouds show an increase in SLF with coarser and coarser resolutions.

- The analysed datasets show different SLF-CTT joint histograms for each cloud type (or cloud height), indicating individual cloud processes, and SLF decreasing with CTT. In ICON and SCREAM, low- and mid-level clouds do not show intermediate SLF values. However, intermediate SLF values are obtained by coarsening the spatial resolution to 0.5°. Cloud_cci v3 and CAM6-Oslo show SLF > 0.4 for stratus clouds, while similar results are shown by GEOS 0.5° for low-level clouds and SCREAM 0.5° for marine stratus clouds only in the SH.

- In Cloud_cci v3, CAM6-Oslo, GEOS, and GEOS 0.5°, high-level clouds are warmer over land than over the ocean in both hemispheres, while low-level clouds are warmer over land than over the ocean only in the SH and colder over land than over the ocean in the NH. This result is reproduced by ICON 0.5° and MCD06COSP only in the SH.

- In Cloud_cci v3, high-level clouds have higher SLF in the SH than in the NH over both land and the ocean; CAM6-Oslo and MCD06COSP agree with this result only over land, while SCREAM agrees only over the ocean. Low-level clouds in Cloud_cci v3 show higher SLF in the SH than in the NH over the ocean, but the opposite over land; CAM6-Oslo, MCD06COSP, and SCREAM agree with Cloud_cci v3 only over the ocean. Other datasets agree only over land or ocean and for limited temperature ranges.

- Cloud_cci v3 and CAM6-Oslo show evidence of enhanced ice production between −15°C and −28°C and between −11°C −15°, respectively, in marine cumulus clouds in the SH, SCREAM 0.5° at about −7°C in marine low-level and nimbostratus clouds in the SH, and GEOS 0.5° at about −15°C in both continental and marine low-level clouds in both hemispheres and in both continental and marine mid-level clouds only in the SH.

To better understand how cloud microphysics influences the results, the liquid effective radius ($\overline{r_{liq}}$) of cloud droplets was analysed using $\overline{r_{liq}}$-SLF joint histograms and $\overline{r_{liq}}$ geographical distributions for different cloud types (or heights), separated into continental and marine areas in NH and SH. For this analysis, only Cloud_cci v3, MCD06COSP, CAM6-Oslo, and GEOS were used, as they provide information on $\overline{r_{liq}}$. The results of this comparison between observational and model datasets are:

- Cloud_cci v3, MC06COSP, and CAM6-Oslo show an increase in $\overline{r_{liq}}$ with decreasing SLF.

- Some cloud types in Cloud_cci v3 (i.e., stratus, cirrus, cirrostratus, and nimbostratus) show also a decrease in $\overline{r_{liq}}$ with decreasing SLF.

- In Cloud_cci v3, low-level clouds have smaller $\overline{r_{liq}}$ over land than over the ocean in the NH, while the opposite occurs in the SH; high-level clouds show larger $\overline{r_{liq}}$ over the ocean than over land in both the NH and SH. These results are partially reproduced by MCD06COSP, for which $\overline{r_{liq}}$ is larger over the ocean than over land, regardless of the hemisphere and the cloud type, and CAM6-Oslo, which does not agree with Cloud_cci v3 for high-level clouds in the NH; GEOS 0.5° shows agreements with Cloud_cci v3 but the small variability of $\overline{r_{liq}}$ makes interpretation of the results difficult.

- In Cloud_cci v3, low-level clouds show smaller $\overline{r_{liq}}$ in the NH than in the SH over both land and the ocean, while high- (and mid-level) clouds reproduce this result only over land and show the opposite over the ocean. These land/ocean contrasts are not uniformly reproduced by all datasets: MCD06COSP shows agreement with Cloud_cci v3 for low- and high-level clouds over the ocean and only partially over land; CAM6-Oslo shows good agreement with Cloud_cci v3 over land, but opposite results over the ocean; GEOS shows agreement with Cloud_cci v3 for low-level clouds; GEOS 0.5° shows agreements with Cloud_cci v3 for both high- and low-level clouds.

- Geographical distributions of $\overline{r_{liq}}$ from observations show: smaller $\overline{r_{liq}}$ over land than over the ocean in low-level clouds; larger values of $\overline{r_{liq}}$ over land in mountainous regions; low-level mixed-phase clouds in the tropics in MCD06COSP. Furthermore, models do not show a clear land/ocean contrast like observations.

The results presented in Chapter 4 (SLF higher in the SH than in the NH, except for continental low-level clouds) were confirmed in Chapter 5 from Cloud_cci v3 using a different method. Further investigations based on $\overline{r_{liq}}$ in Chapter 6 show that the lower SLF in the SH than in the NH for low-level clouds may be due to orographic updrafts, as the continental areas between 30° and 60° latitude are mainly mountainous. To improve the ability of global climate models to reproduce observations more realistically, different versions of CAM6-Oslo were used for global GCMs simulations tuned locally (i.e., in the Arctic region, in the NH, and in the SH) and globally. These models were tuned to bring the cloud top and interior SLF from the models into agreement with the CALIOP observations by modifying the efficiency of the WBF process, the INP concentration, and the glaciation temperature of the cloud particles detrained from convective updrafts. After model tuning, the simulated SLF at the top and in the interior of mixed-phase clouds were in good agreement with observations.
In the introduction of this thesis, four questions were addressed. The answers to these questions are given here, based on the results listed above.

1. **Do mixed-phase clouds exhibit regional dependencies? And if they do, are the regional dependencies cloud-type-dependent? Can the passive satellite sensor AVHRR contribute to cloud-phase research?**

Mixed-phase clouds between 30° and 60° latitude show hemispheric dependencies, with higher SLF in the SH than in the NH. This occurs for all cloud types over land and the ocean, with the exception of continental low-level clouds, for which SLF is higher in the NH than in the SH. Therefore, the regional dependencies are cloud-type-dependent. The study presented in Chapter 4 shows that datasets derived from the active satellite sensor CALIOP and the passive satellite sensor AVHRR are qualitatively consistent with these results, making the passive satellite sensor AVHRR a good candidate for cloud-phase research.

2. **How well do global climate and global storm-resolving models represent the amount, the geographical distribution, and the phase partitioning in mixed-phase clouds globally and regionally?**

   Global climate and global storm-resolving models retrieve from 3.1% (SCREAM) to 16.7% (CAM6-Oslo) of mixed-phase clouds between 60° North and 60° South, closer to the observations of Cloud_cci v3 (6.2%) than MCD06COSP (55.0%). Despite the large variability in the amount of mixed-phase clouds retrieved, the models mostly agree with the observations in representing the geographical distribution of mixed-phase clouds, mainly at mid-latitudes. Models show a reduction of SLF with CTT, in agreement with the observations, regardless of the cloud types and regions. The hemispheric and continental-marine contrasts revealed by the observations (Cloud_cci v3) were only captured by a few models. Reasons for this may be an overly efficient WBF process in the models, a still not realistic representation of the aerosol influence (especially for low-level clouds), and an insufficient representation of the cloud macro- and microphysics.

3. **How well do global climate and global storm-resolving models represent the microphysics involved in mixed-phase clouds?**

   Global climate and global storm-resolving models were not always able to reproduce the results from the observations. Regardless of the cloud types and the regions, observations show an increase in $\overline{r_{liq}}$ with decreasing SLF. Only CAM6-Oslo agreed with this result, while GEOS failed to reproduce it. The hemispheric and continental/marine contrasts revealed by the observations were only partially reproduced by the models. Possible reasons for the disagreement between observations and models are the representation of cloud micro- and macrophysics as well as of aerosol influence (especially in the continental/marine contrast).

4. **What can be done to improve models with respect to the simulation of mixed-phase clouds?**

   Globally and regionally, it is possible to tune models to better mimic observations: by using metrics based on phase partitioning in clouds (e.g., SLF) at different cloud heights (e.g., at the top and inside clouds), microphysical-related parameters of GCMs have been adjusted (i.e., the time scale of the WBF process, the number of aerosols active as INPs, and the temperature-based lin-

ear relationship driving the phase of cloud particles detained from convective updrafts) to more realistically represent mixed-phase clouds.

8.2. Outlook

Research is an ongoing process, and although many questions have been answered, many others remain open or have arisen. Some of these are listed here as inspirational starting points for future research:

- The validation study based on datasets derived from passive and active satellite sensors shows qualitative agreement between the datasets. However, a mismatch in cloud top temperature and phase was retrieved, which was responsible for differences up to about 25°C or SLF of about 80% at a fixed temperature between active and passive retrievals. Further investigation of the clouds responsible for this discrepancy may clarify whether it is, e.g., regional or cloud-type related.

- The cloud amount of the DYAMOND datasets decreases by coarsening the spatial resolution. The application of the temperature filter to the gridboxes may cause this result. However, this hypothesis would be confirmed (or rejected) by using different temperature differences between the coldest and the warmest cloudy pixels in a gridbox as a threshold for including the gridboxes in the analysis.

- An increase in $\overline{r_{liq}}$ was found in cirriform clouds in Cloud_cci v3. Further investigation of this phenomenon using other sensors and/or regional scales may help to better understand the physics driving this result, whether it is due to misclassification of cloud type, and whether it is due to limitations of the AVHRR sensor.

- Further investigation of the increase and/or decrease in $\overline{r_{liq}}$ with decreasing SLF, distinguishing the cloud types and regions where these processes occur separately, may help to clarify the conditions under which one or both processes occur and whether they are regionally dependent.

- MCD06COSP is a combination of MODIS-Aqua and MODIS-Terra, but was created to make comparisons with models implementing the MODIS-COSP simulator more consistent, and therefore only one channel was used to retrieve the cloud optical properties. This adds uncertainties to the comparisons presented in this thesis. These uncertainties would be reduced by adding MODIS-Terra to the analysed datasets. In this way it would be clear how the MCD06COSP algorithm affected the MODIS observations, making the interpretation of the results more robust.

- In this work, neither the physical nor the chemical nature of the aerosol acting as CCN or INP was investigated. The research presented here would benefit from datasets that provide this information.

A. Appendix

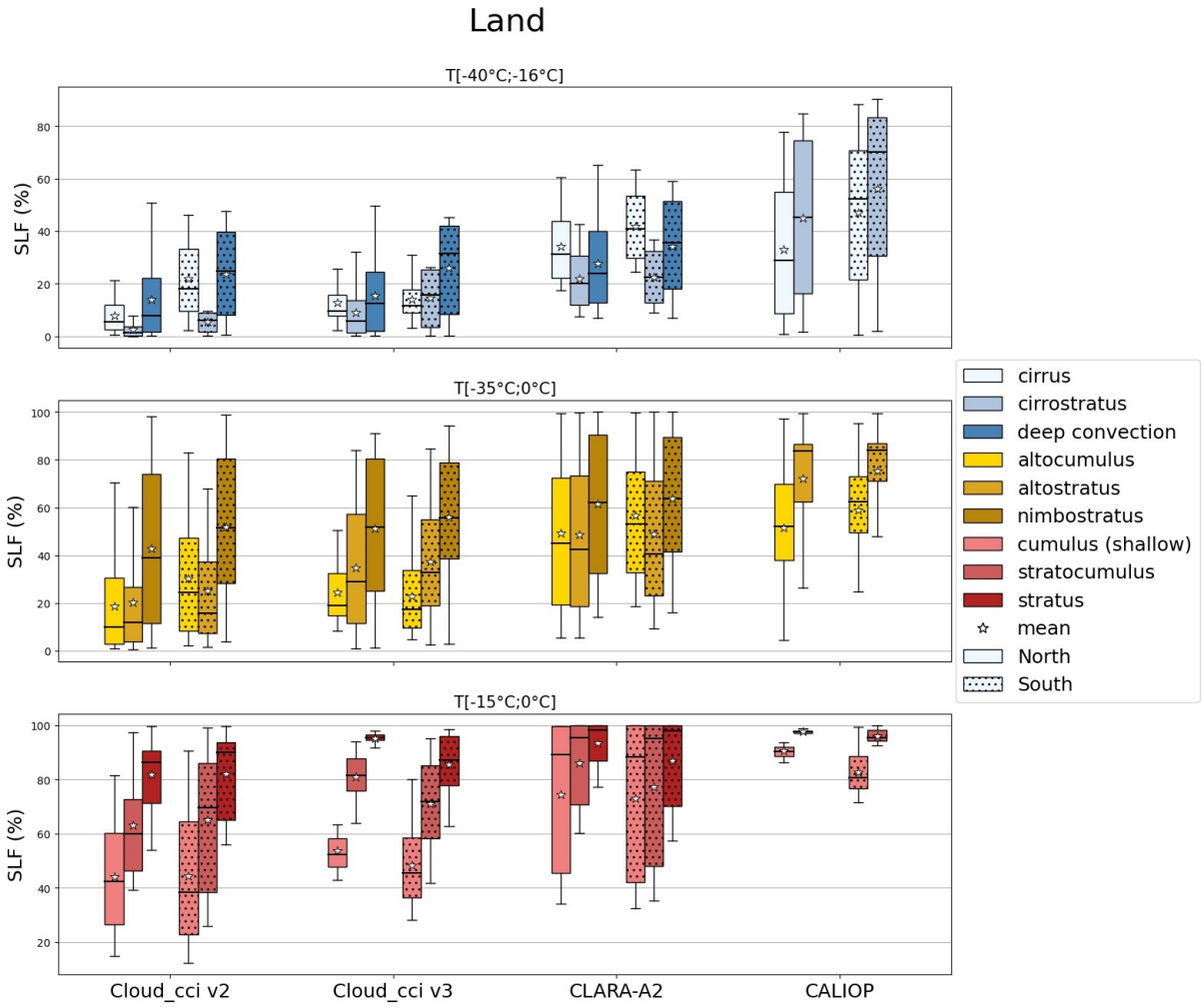

Figure A.1.: Boxplot of the supercooled liquid fraction (SLF) for different cloud types over land sorted in three height levels. Clouds at the same height-level share the same cloud top temperature range, specified at the top of each panel. The different datasets are separated by columns and every color corresponds to one cloud type. The boxes extend from the lower to upper quartile values of the data, whereas the whiskers show the entire range of the data. The horizontal lines within the boxes represent the median of the distributions, while the stars represent their mean values. For each dataset, data from the northern (without dots) and southern (with dots) hemispheres can be distinguished.

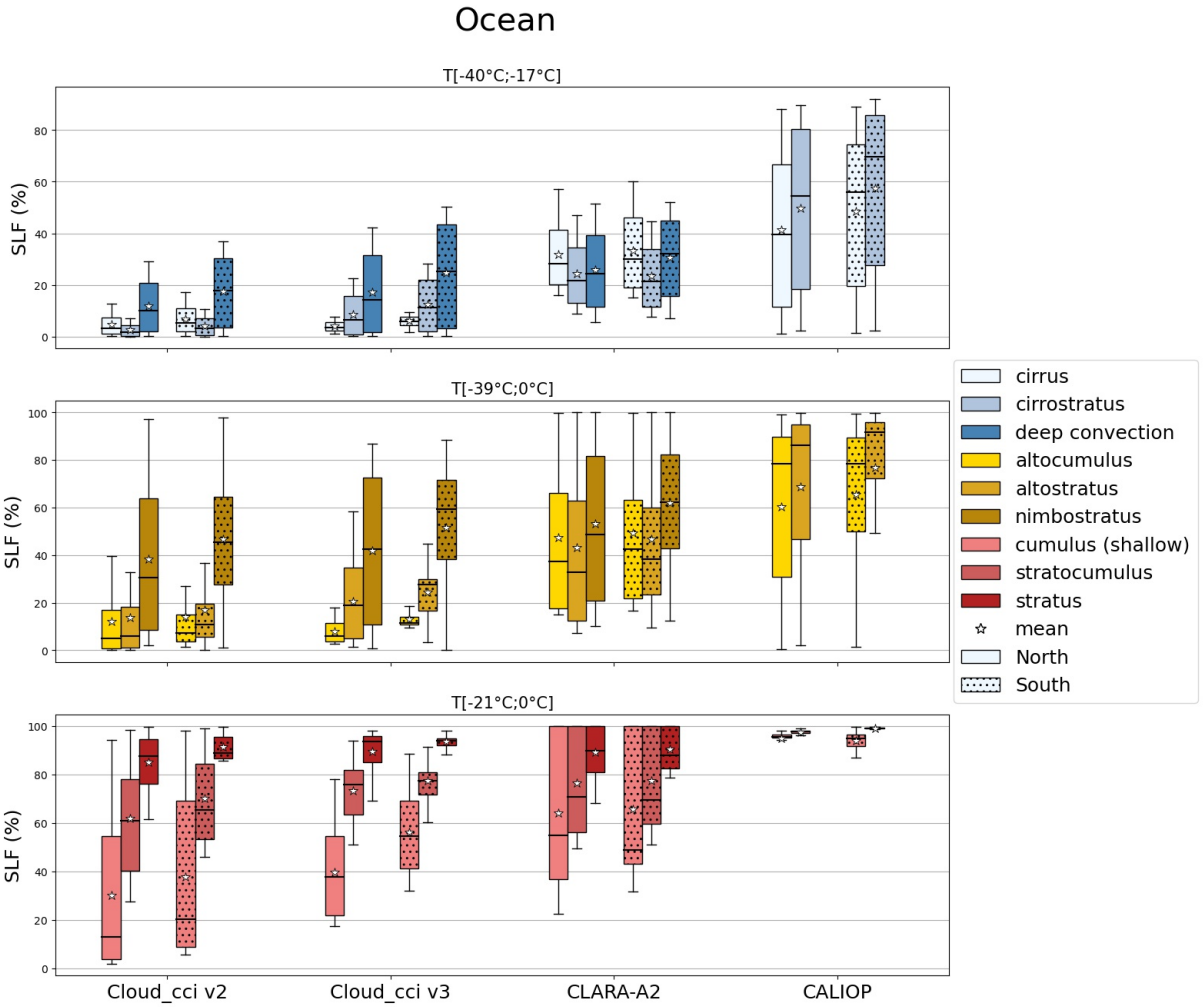

Figure A.2.: Boxplot of the supercooled liquid fraction (SLF) for different cloud types over the ocean sorted in three height levels. Clouds at the same height-level share the same cloud top temperature range, specified at the top of each panel. The different datasets are separated by columns and every color corresponds to one cloud type. The boxes extend from the lower to upper quartile values of the data, whereas the whiskers show the entire range of the data. The horizontal lines within the boxes represent the median of the distributions, while the stars represent their mean values. For each dataset, data from the northern (without dots) and southern (with dots) hemispheres can be distinguished.

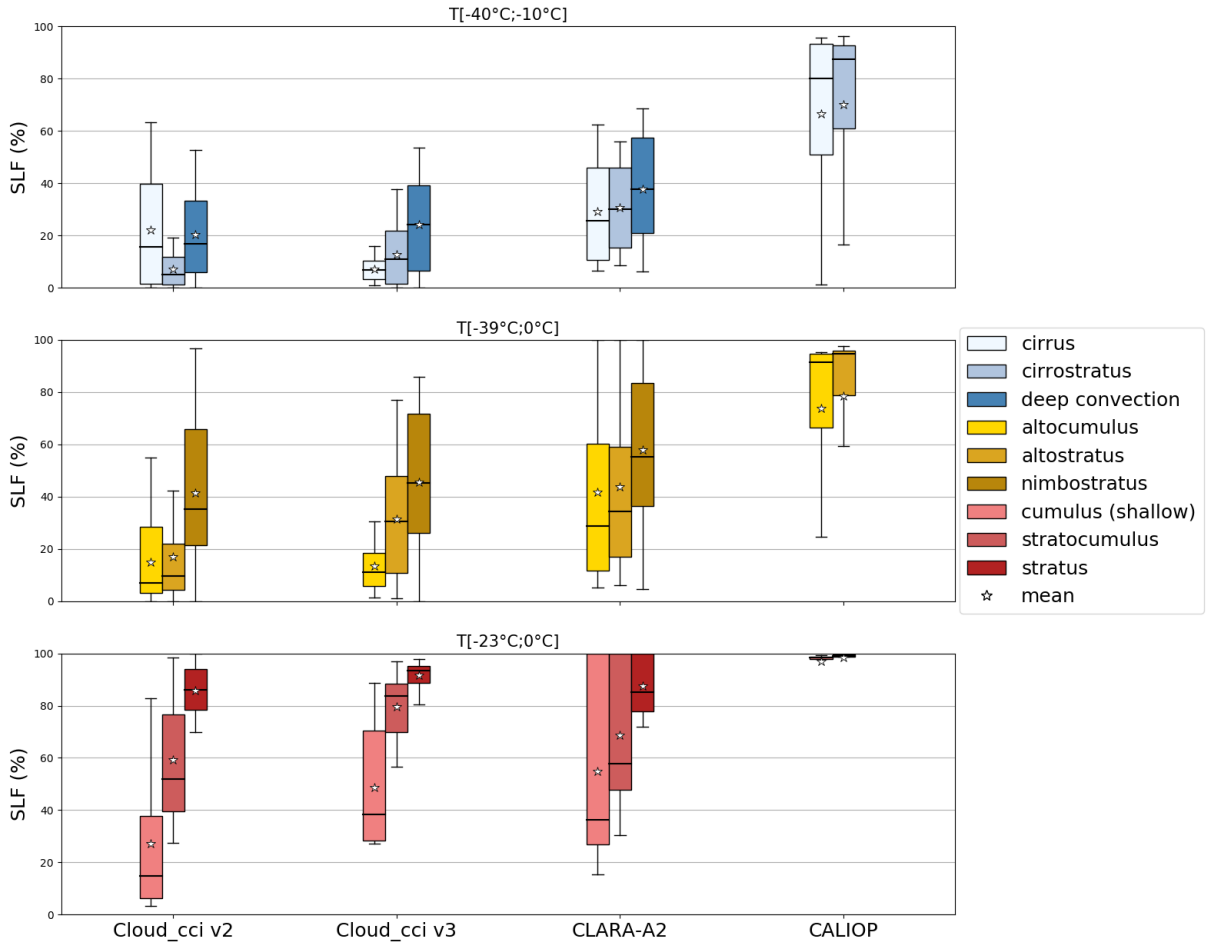

Figure A.3.: Boxplot of the supercooled liquid fraction (SLF) for different cloud types from collocated data sorted in three height levels. Clouds at the same height-level share the same cloud top temperature range, specified at the top of each panel. The different datasets are separated by columns and every color corresponds to one cloud type. The boxes extend from the lower to upper quartile values of the data, whereas the whiskers show the entire range of the data. The horizontal lines within the boxes represent the median of the distributions, while the stars represent their mean values.

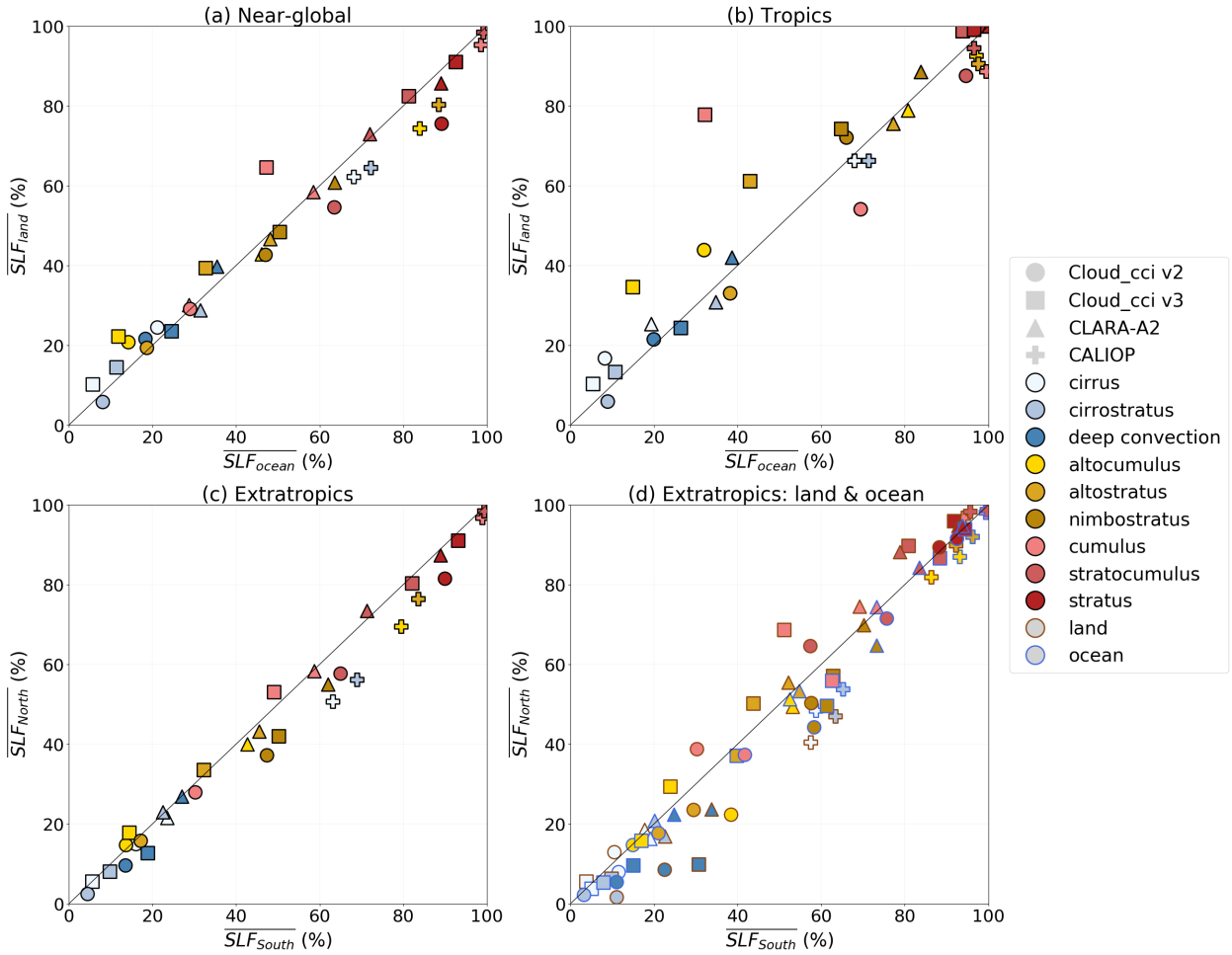

Figure A.4.: Comparison of mean SLF for different cloud types from collocated data, considered in the temperature ranges they have in common at the same height-level and for each subplot individually, for near-global marine and continental pixels (a), tropical marine and continental pixels (b), and mid-latitudes in NH and SH (c), with the further separation of marine and continental regions (d). Different markers identify different datasets, filling colors distinguish the cloud types, while edge colors refer to continental or marine pixels in (d).

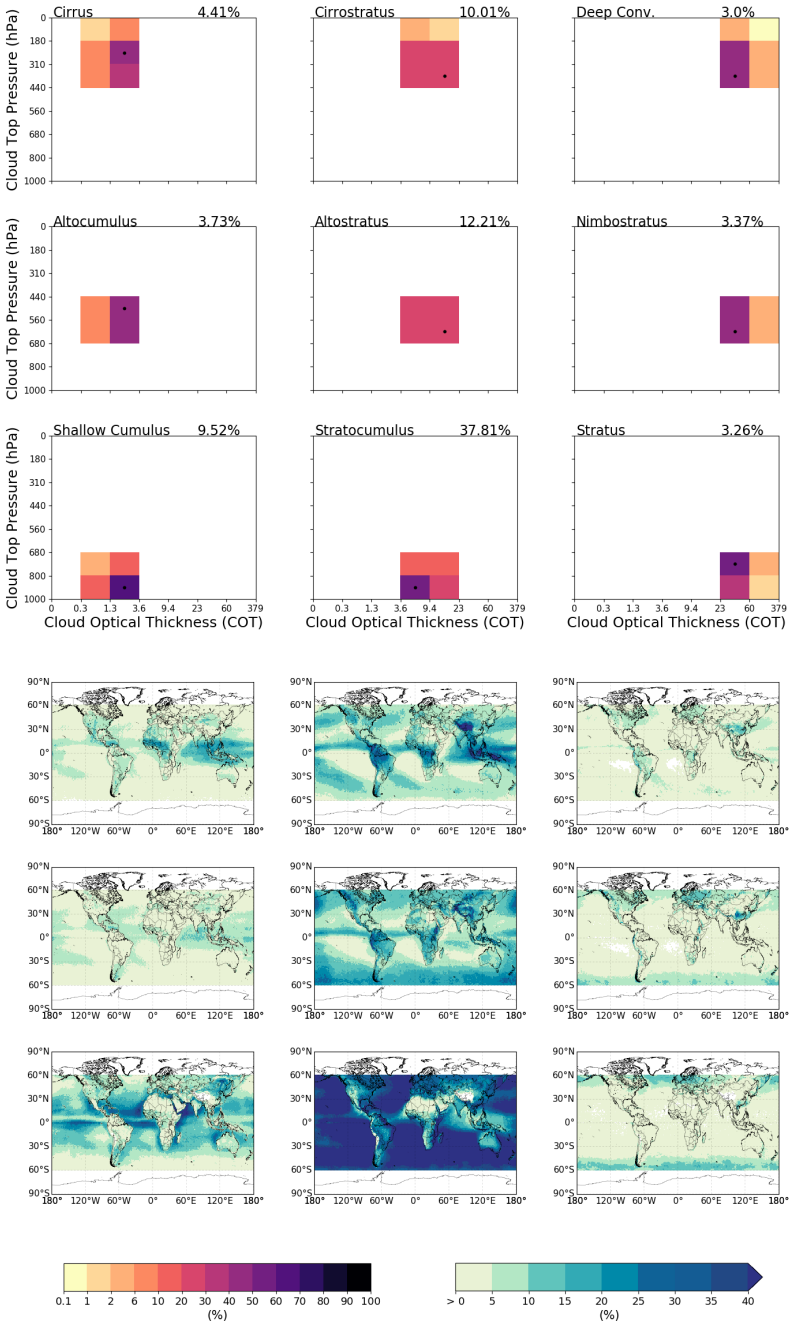

Figure A.5.: This figure is composed of two parts. The upper panel shows COT-CTP joint-histograms for different cloud types in a 3x3 matrix for MCD06COSP. Each cloud category is identified by a name in the upper left corner of the histogram, while its Relative Frequency of Occurrence (RFO) is shown in the upper right corner. Each histogram is normalized to 1 and the single dot indicates the most frequent COT-CTP combination. The lower panel shows the geographical distribution of the different cloud types in the upper panel and is also shown in a 3x3 matrix for one-to-one correspondence with the upper panel. Each pixel in the geographical distributions is normalized to the total number of measurements taken in that pixel.

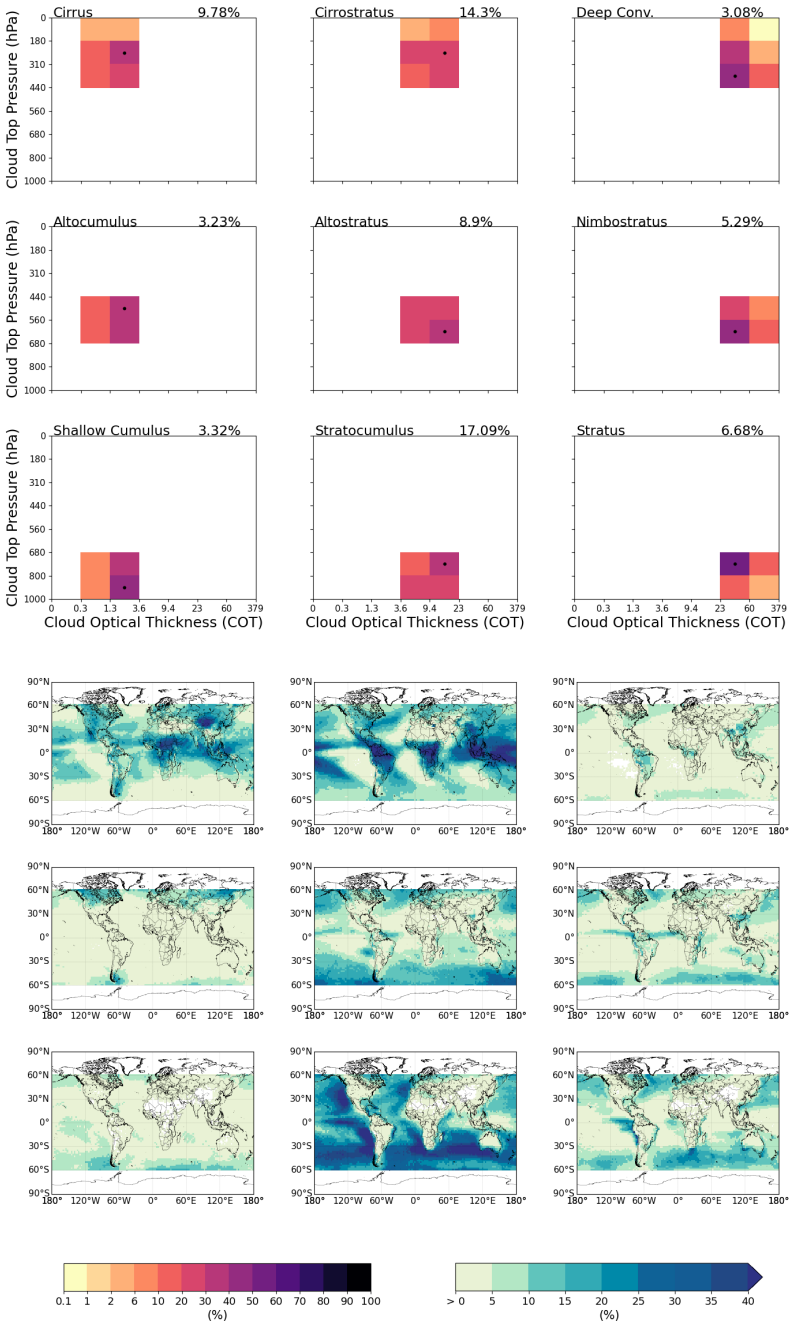

Figure A.6.: This figure is composed of two parts. The upper panel shows COT-CTP joint-histograms for different cloud types in a 3x3 matrix for CAM6-Oslo. Each cloud category is identified by a name in the upper left corner of the histogram, while its Relative Frequency of Occurrence (RFO) is shown in the upper right corner. Each histogram is normalized to 1 and the single dot indicates the most frequent COT-CTP combination. The lower panel shows the geographical distribution of the different cloud types in the upper panel and is also shown in a 3x3 matrix for one-to-one correspondence with the upper panel. Each pixel in the geographical distributions is normalized to the total number of measurements taken in that pixel.

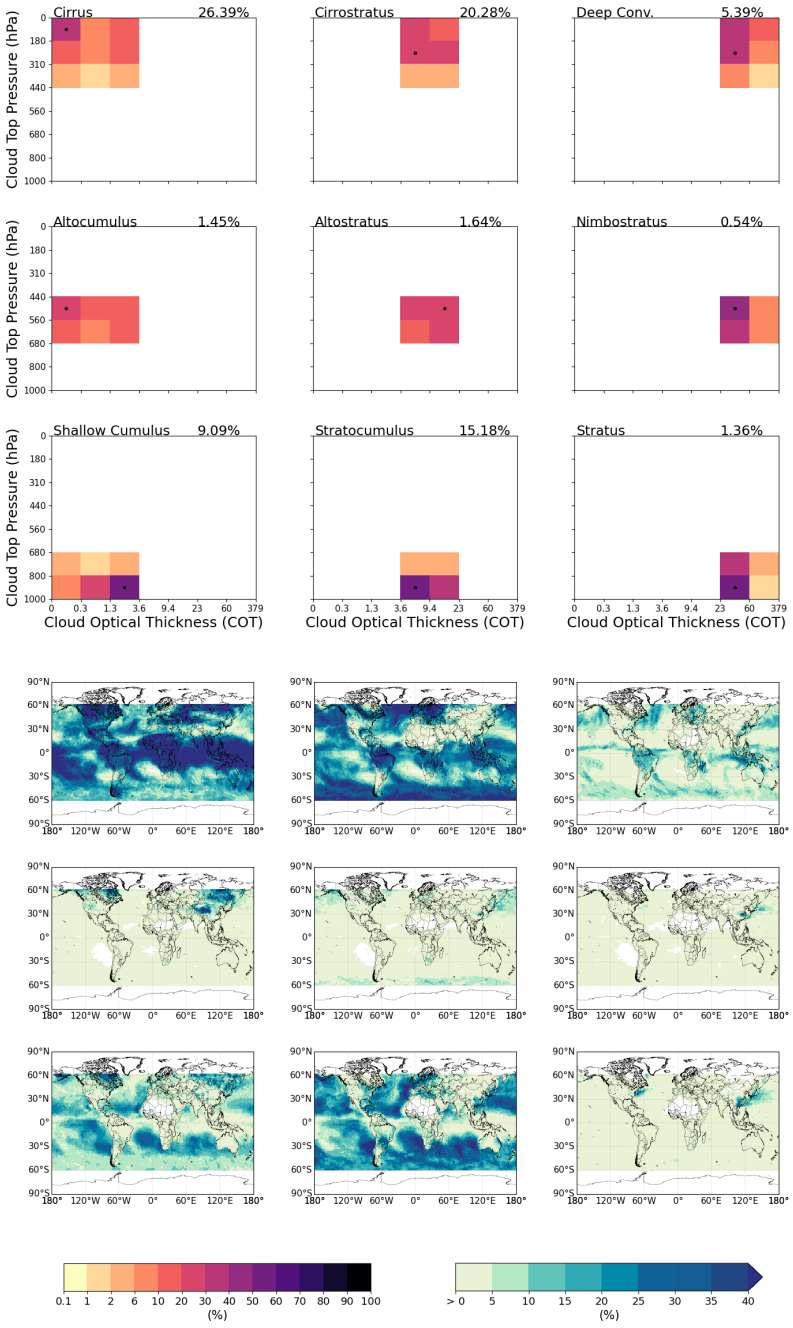

Figure A.7.: This figure is composed of two parts. The upper panel shows COT-CTP joint-histograms for different cloud types in a 3x3 matrix for SCREAM. Each cloud category is identified by a name in the upper left corner of the histogram, while its Relative Frequency of Occurrence (RFO) is shown in the upper right corner. Each histogram is normalized to 1 and the single dot indicates the most frequent COT-CTP combination. The lower panel shows the geographical distribution of the different cloud types in the upper panel and is also shown in a 3x3 matrix for one-to-one correspondence with the upper panel. Each pixel in the geographical distributions is normalized to the total number of measurements taken in that pixel.

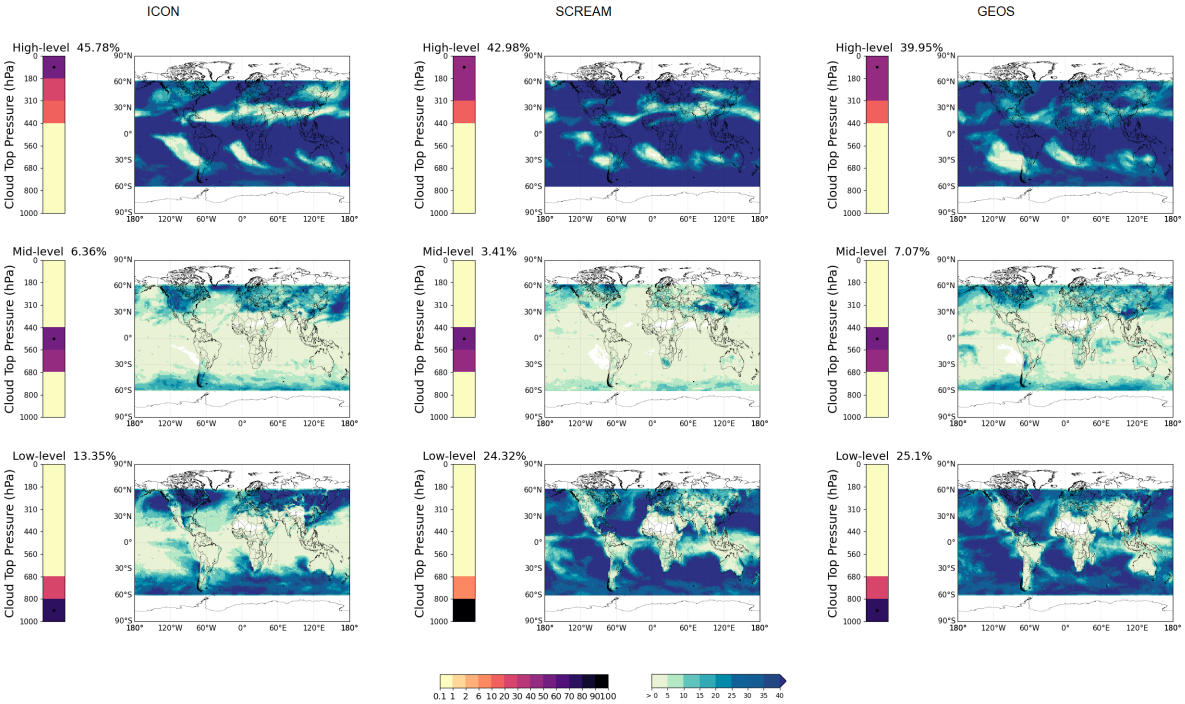

Figure A.8.: CTP histograms and geographical distributions of all clouds in high, mid and low levels for ICON (left), SCREAM-CS (middle), and GEOS (right). At the top of each histogram, the Relative Frequency of Occurrence (RFO) is shown.

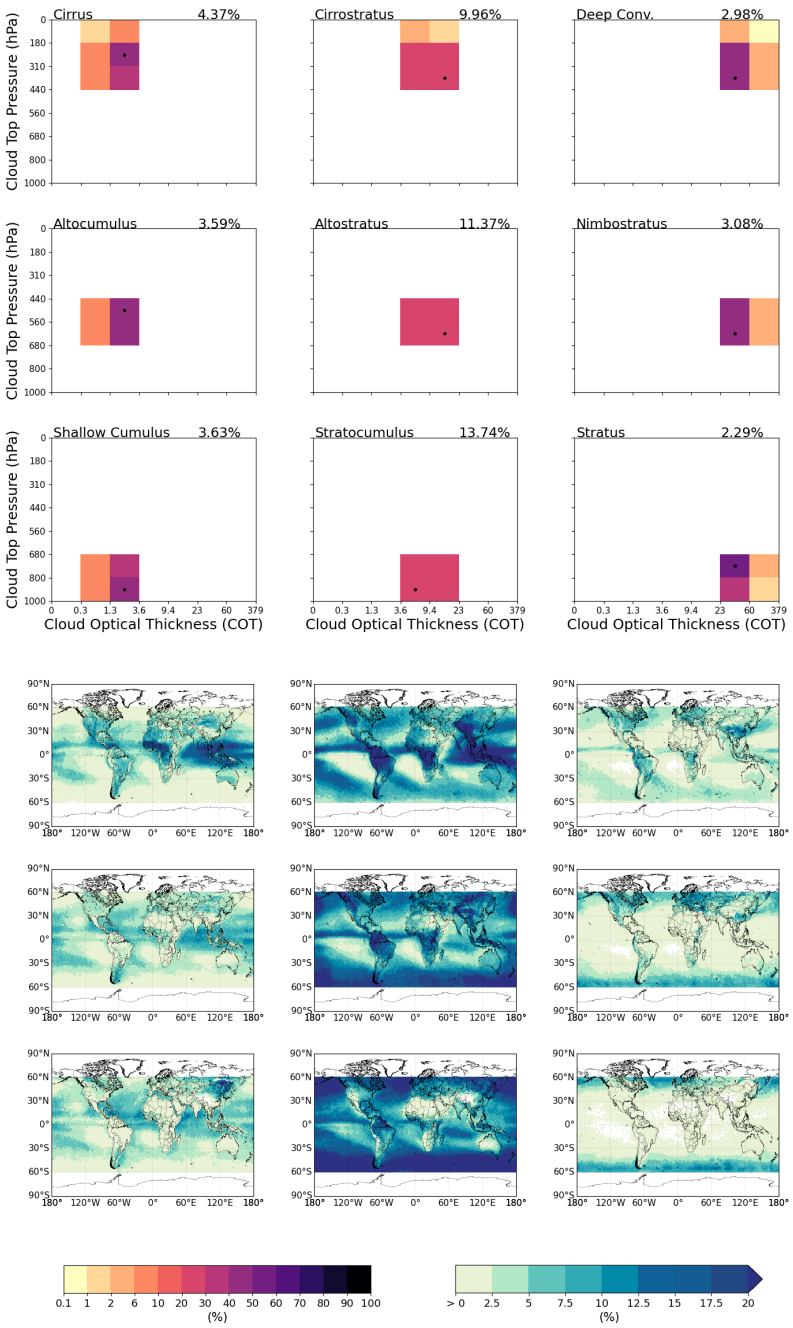

Figure A.9.: This figure is composed of two parts. The upper panel shows COT-CTP joint-histograms for different cloud types in the mixed phase in a 3x3 matrix for MCD06COSP. Each cloud category is identified by a name in the upper left corner of the histogram, while its Relative Frequency of Occurrence (RFO) is shown in the upper right corner. Each histogram is normalized to 1 and the single dot indicates the most frequent COT-CTP combination. The lower panel shows the geographical distribution of the different cloud types in the upper panel and is also shown in a 3x3 matrix for one-to-one correspondence with the upper panel. Each pixel in the geographical distributions is normalized to the total number of measurements taken in that pixel. The colorbar for the geographical distribution has a smaller range than in figures including all clouds.

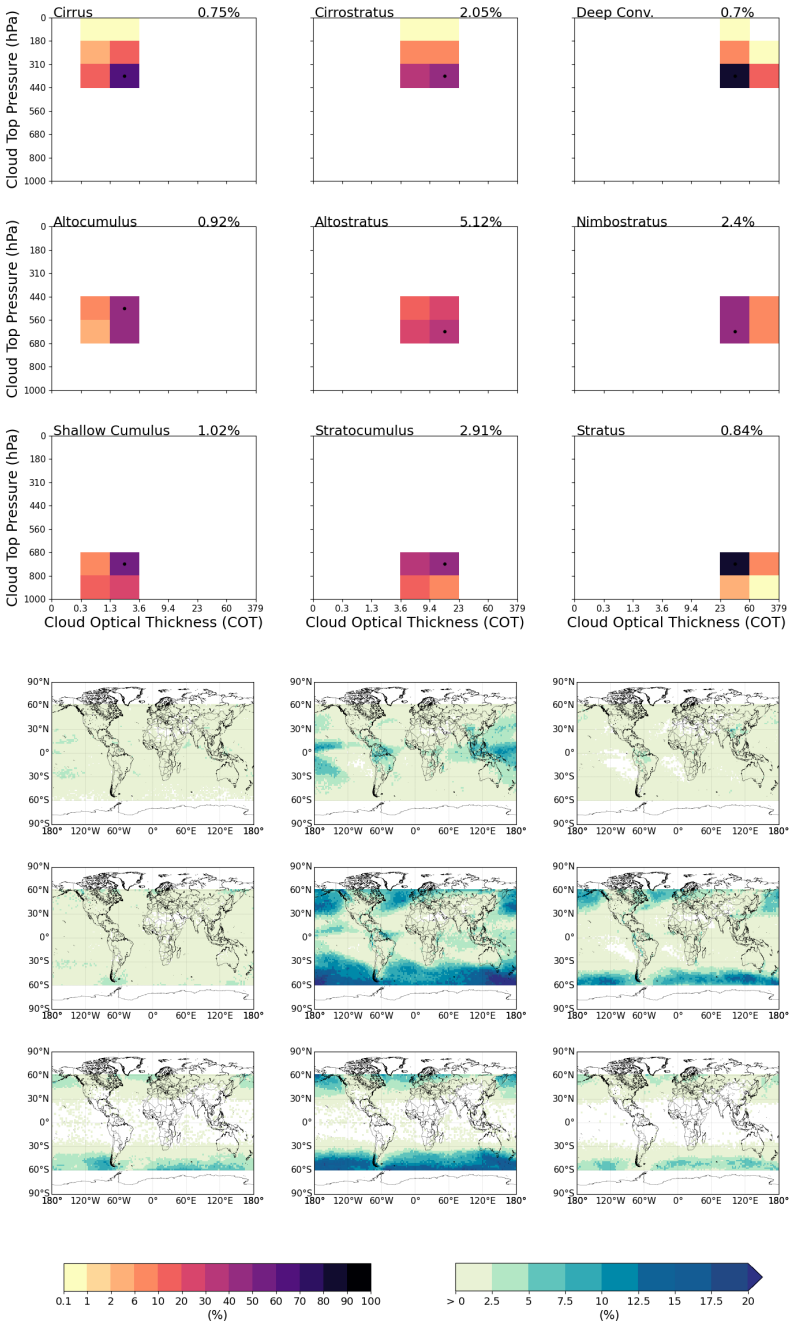

Figure A.10.: This figure is composed of two parts. The upper panel shows COT-CTP joint-histograms for different cloud types in the mixed phase in a 3x3 matrix for CAM6-Oslo. Each cloud category is identified by a name in the upper left corner of the histogram, while its Relative Frequency of Occurrence (RFO) is shown in the upper right corner. Each histogram is normalized to 1 and the single dot indicates the most frequent COT-CTP combination. The lower panel shows the geographical distribution of the different cloud types in the upper panel and is also shown in a 3x3 matrix for one-to-one correspondence with the upper panel. Each pixel in the geographical distributions is normalized to the total number of measurements taken in that pixel. The colorbar for the geographical distribution has a smaller range than in figures including all clouds.

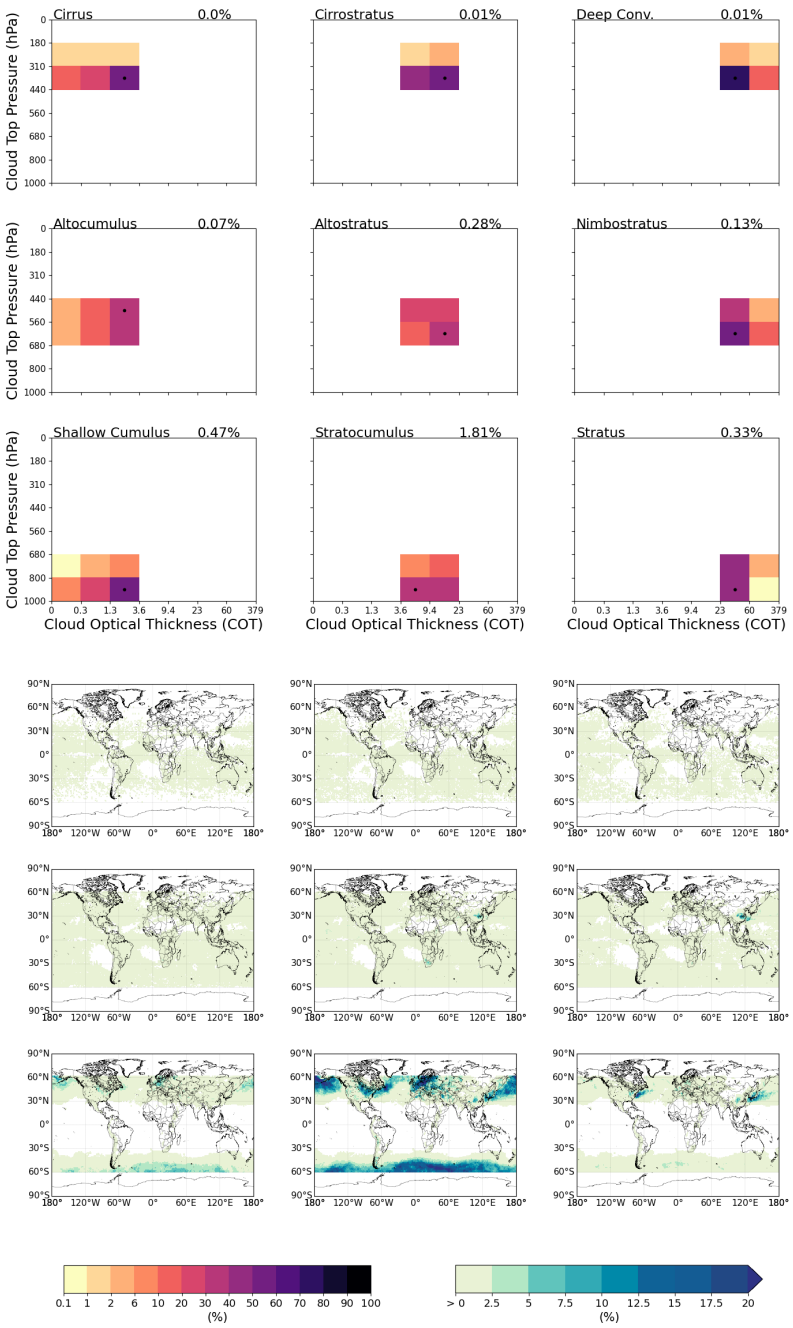

Figure A.11.: This figure is composed of two parts. The upper panel shows COT-CTP joint-histograms for different cloud types in the mixed phase in a 3x3 matrix for SCREAM. Each cloud category is identified by a name in the upper left corner of the histogram, while its Relative Frequency of Occurrence (RFO) is shown in the upper right corner. Each histogram is normalized to 1 and the single dot indicates the most frequent COT-CTP combination. The lower panel shows the geographical distribution of the different cloud types in the upper panel and is also shown in a 3x3 matrix for one-to-one correspondence with the upper panel. Each pixel in the geographical distributions is normalized to the total number of measurements taken in that pixel. The colorbar for the geographical distribution has a smaller range than in figures including all clouds.

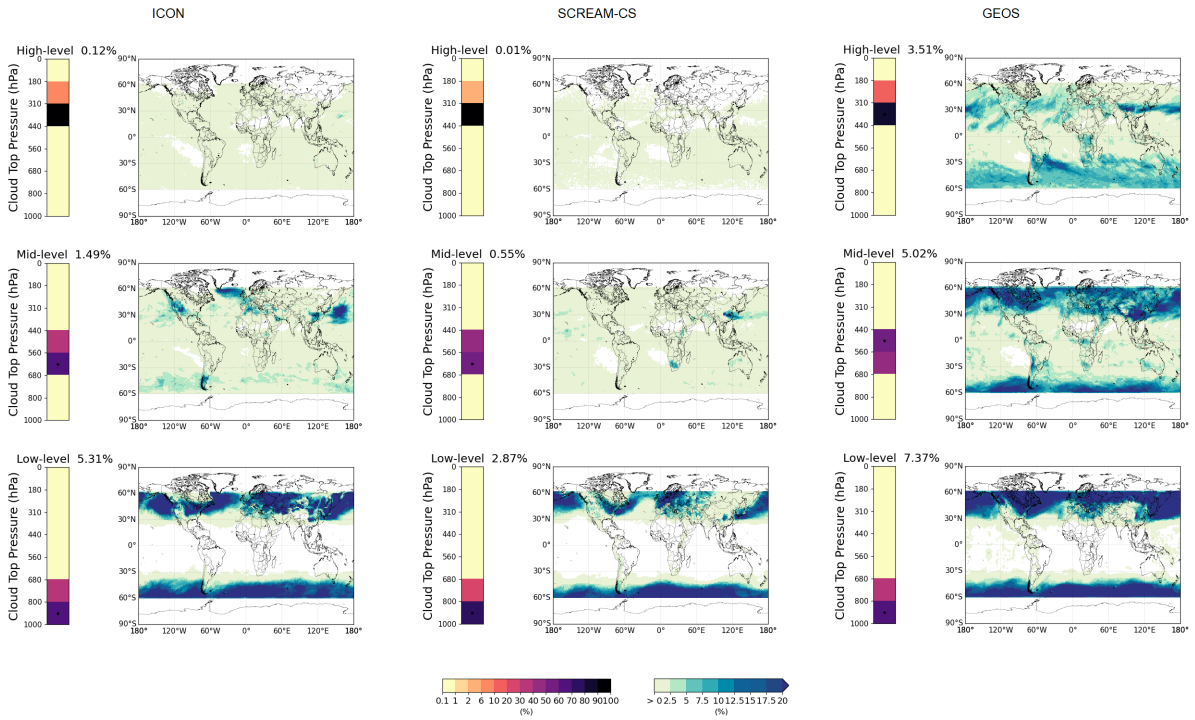

Figure A.12.: CTP histograms and geographical distributions of mixed-phase clouds in high, mid and low levels for ICON (left), SCREAM-CS (middle), and GEOS (right). At the top of each histogram, the Relative Frequency of Occurrence (RFO) is shown. The colorbar for the geographical distribution has a smaller range than in figures including all clouds.

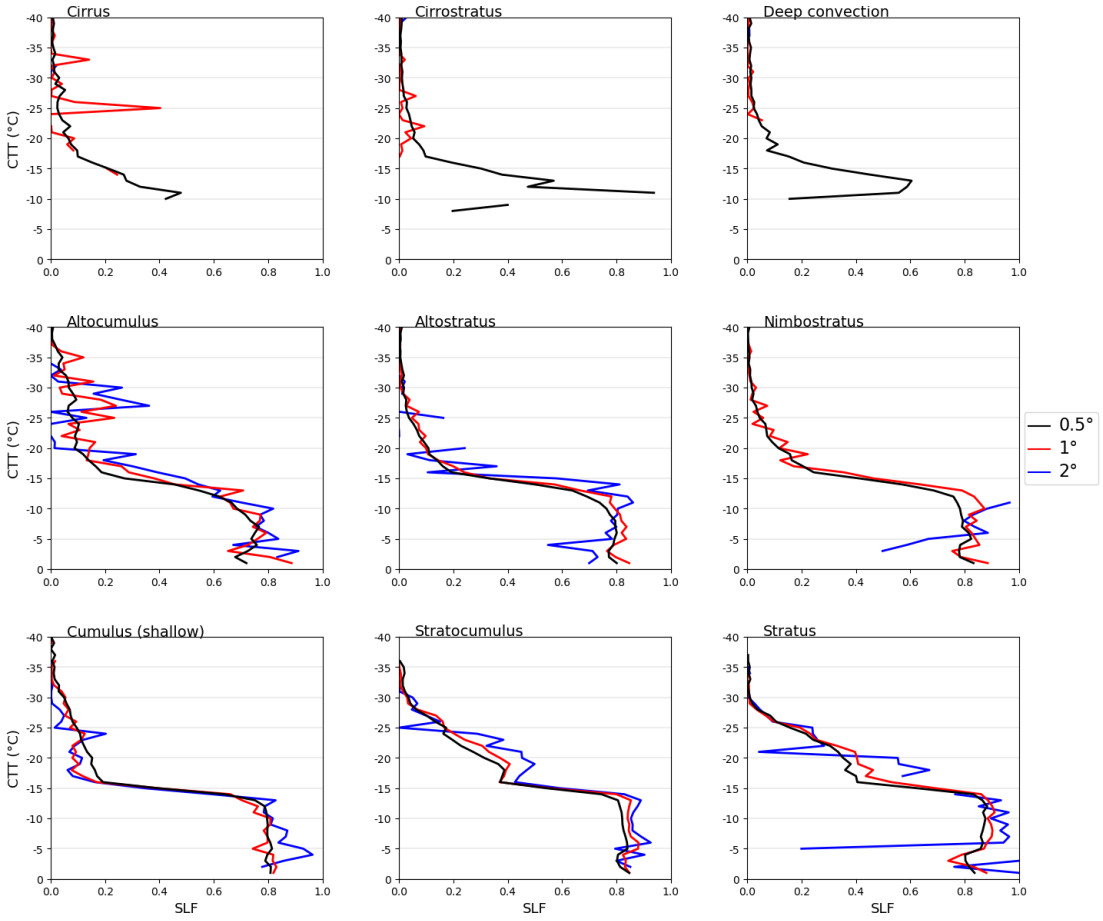

Figure A.13.: SLF profiles on isotherms of different cloud types for SCREAM at the spatial resolutions of 0.5°, 1°, and 2°.

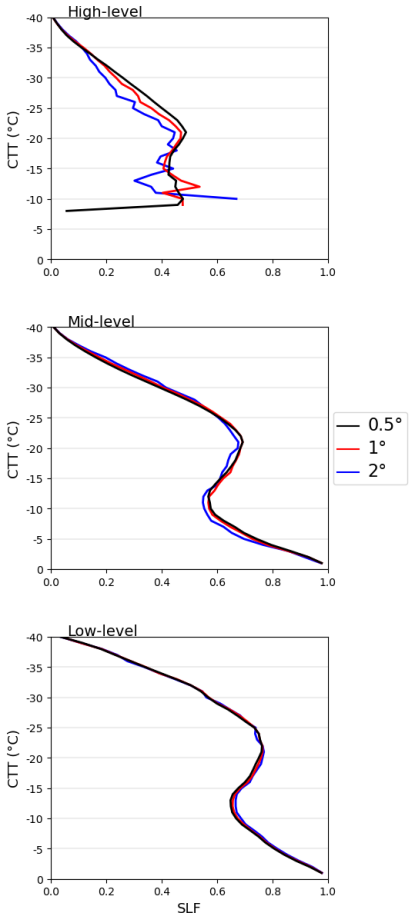

Figure A.14.: SLF profiles on isotherms of high-, mid-, and low-level clouds for GEOS at the spatial resolutions of 0.5°, 1°, and 2°.

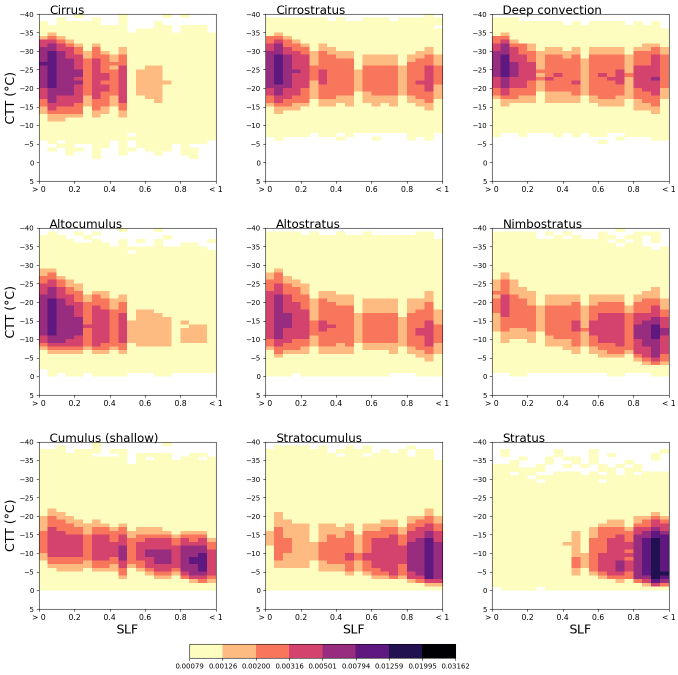

Figure A.15.: SLF-CTT joint histograms for marine mixed-phase clouds in the NH for Cloud_cci v3. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

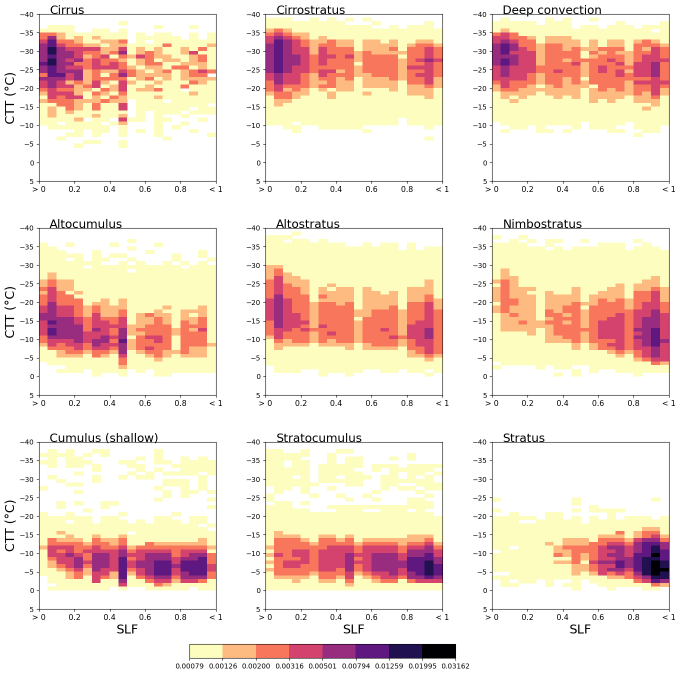

Figure A.16.: SLF-CTT joint histograms for continental mixed-phase clouds in the SH for Cloud_cci v3. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

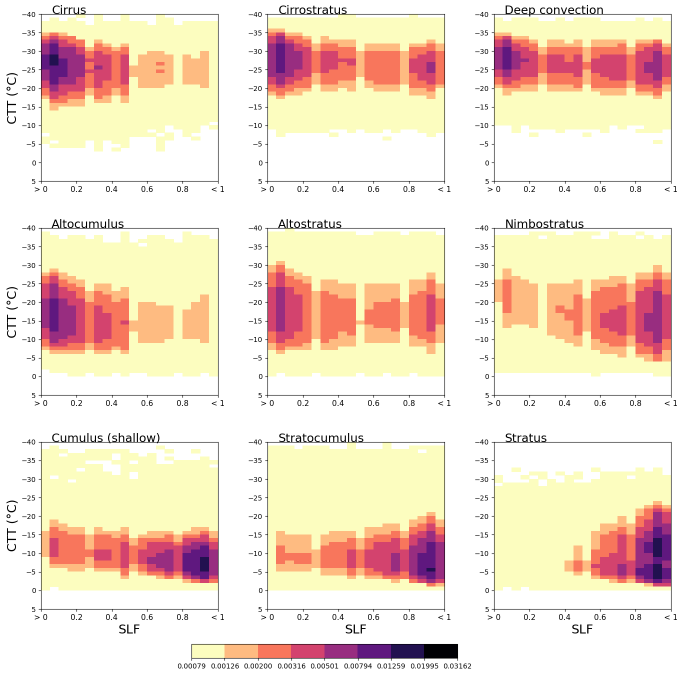

Figure A.17.: SLF-CTT joint histograms for marine mixed-phase clouds in the SH for Cloud_cci v3. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

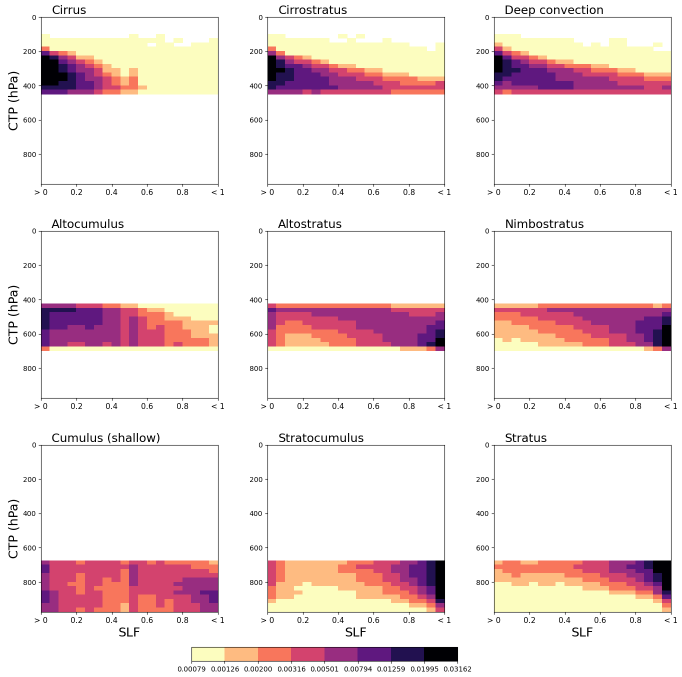

Figure A.18.: CTP-SLF joint histograms for continental mixed-phase clouds in the NH for MCD06COSP. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin.

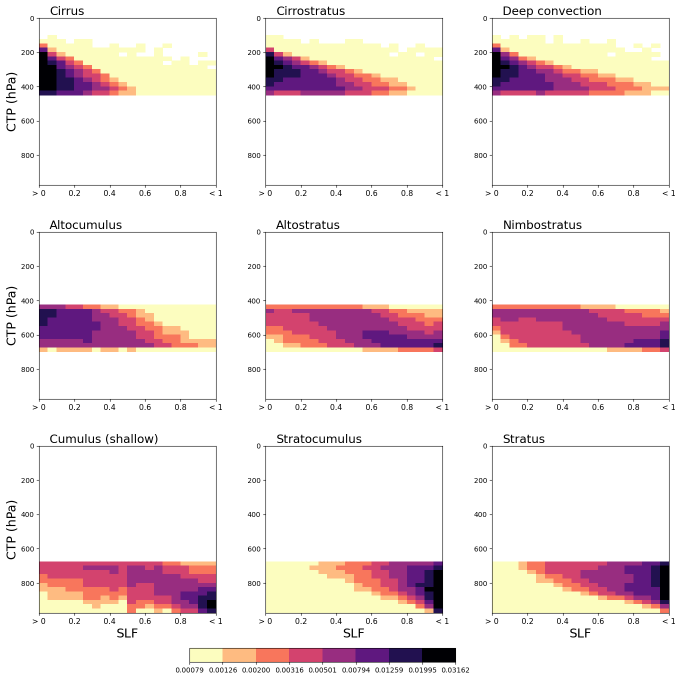

Figure A.19.: CTP-SLF joint histograms for marine mixed-phase clouds in the NH for MCD06COSP. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin.

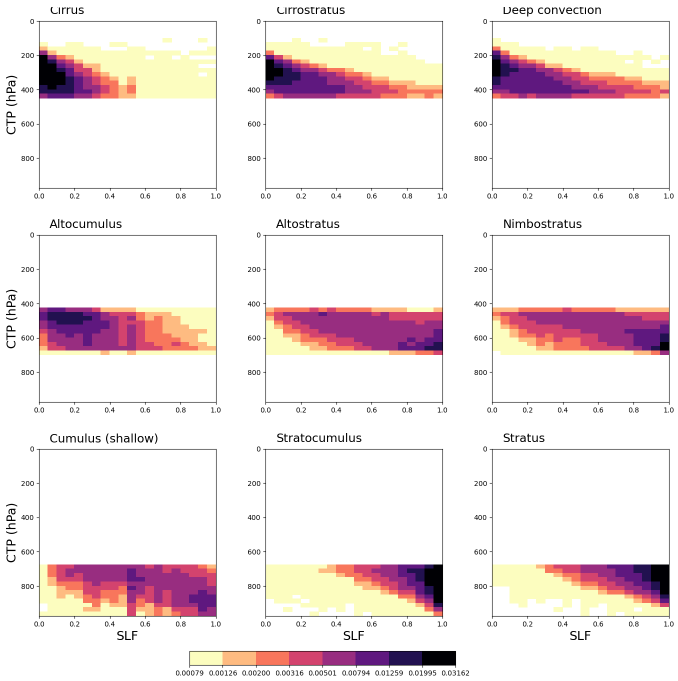

Figure A.20.: CTP-SLF joint histograms for continental mixed-phase clouds in the SH for MCD06COSP. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin.

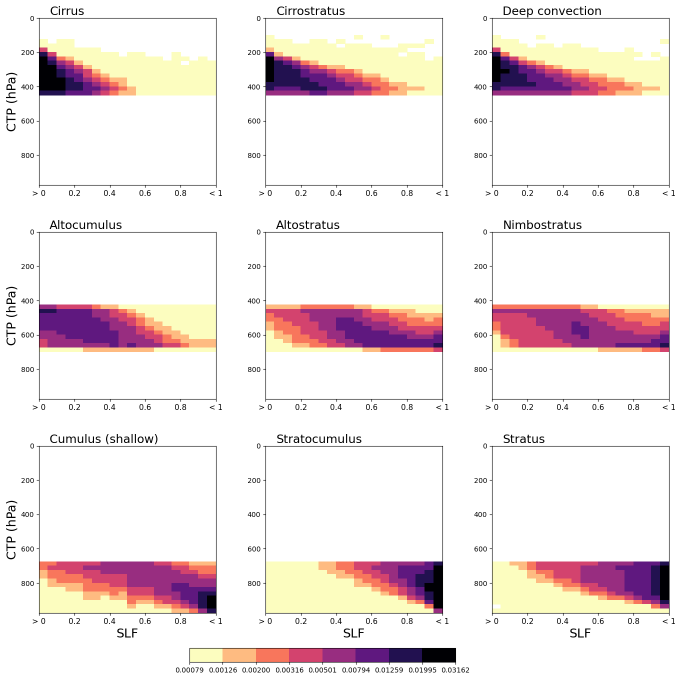

Figure A.21.: CTP-SLF joint histograms for marine mixed-phase clouds in the SH for MCD06COSP. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin.

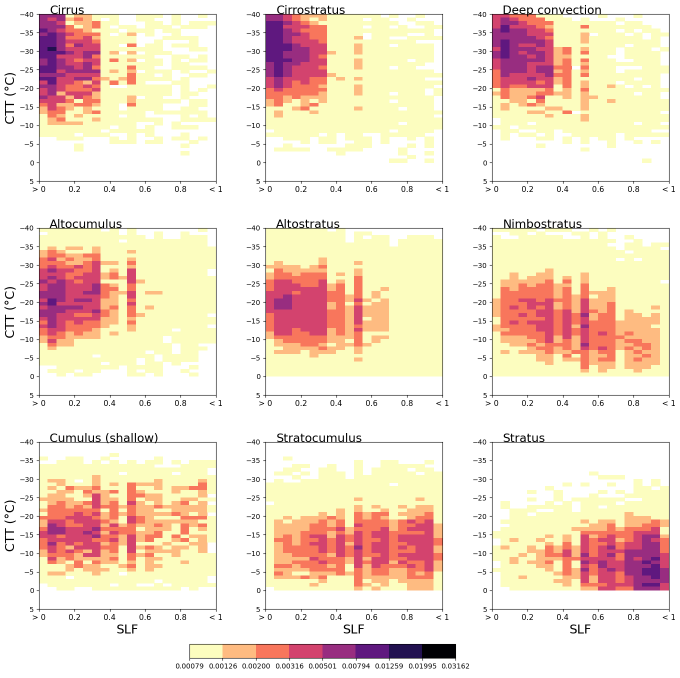

Figure A.22.: SLF-CTT joint histograms for continental mixed-phase clouds in the NH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

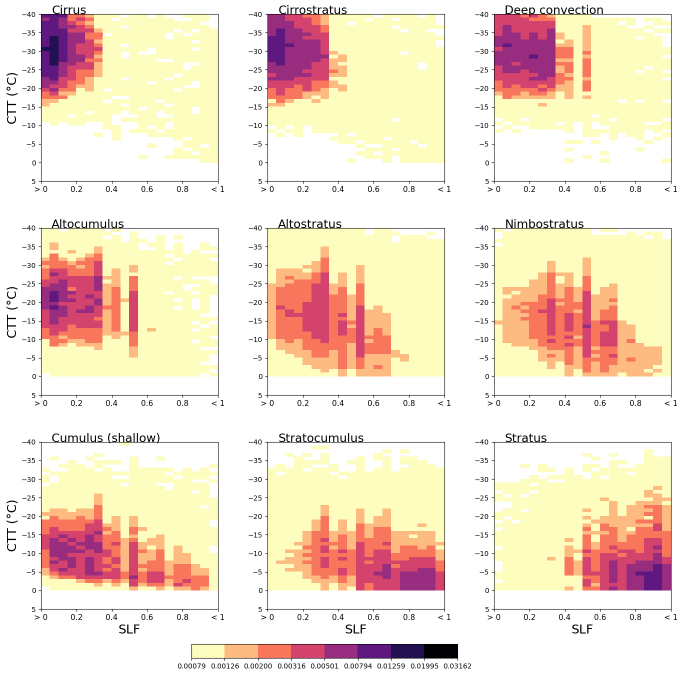

Figure A.23.: SLF-CTT joint histograms for martime mixed-phase clouds in the NH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

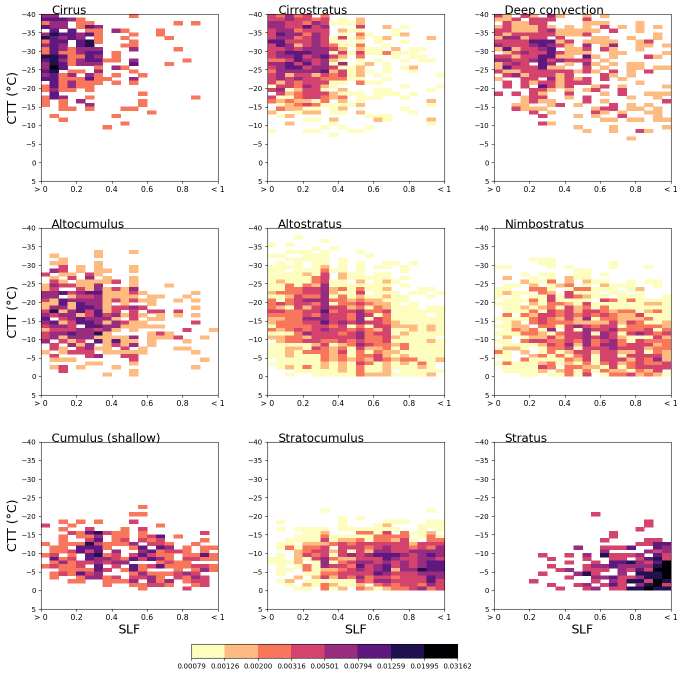

Figure A.24.: SLF-CTT joint histograms for continental mixed-phase clouds in the SH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

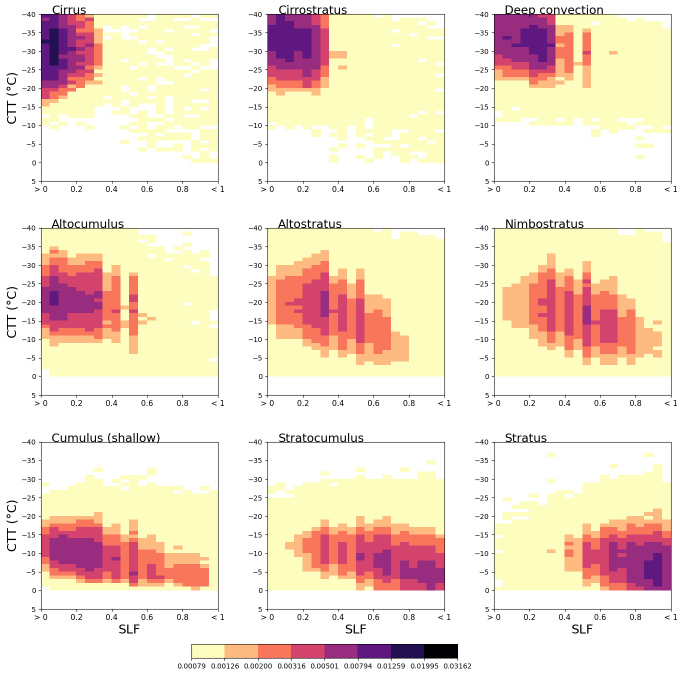

Figure A.25.: SLF-CTT joint histograms for marine mixed-phase clouds in the SH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

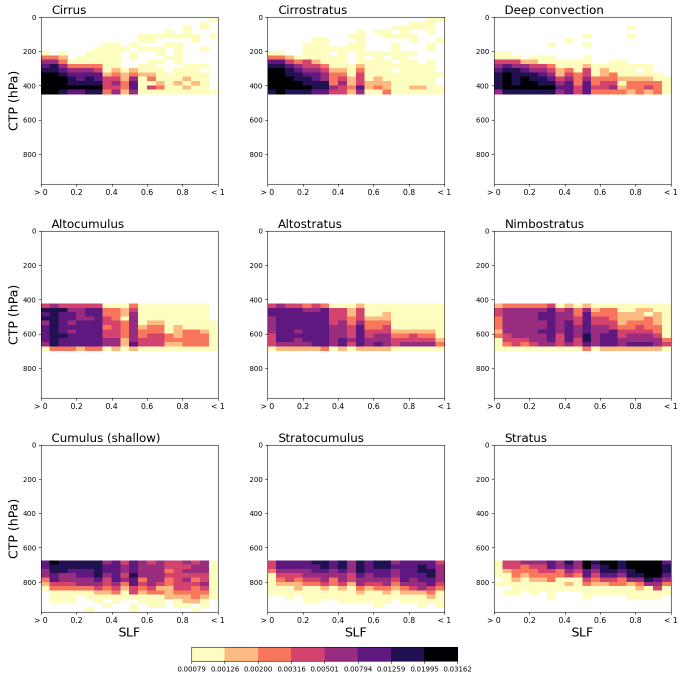

Figure A.26.: CTP-SLF joint histograms for continental mixed-phase clouds in the NH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin.

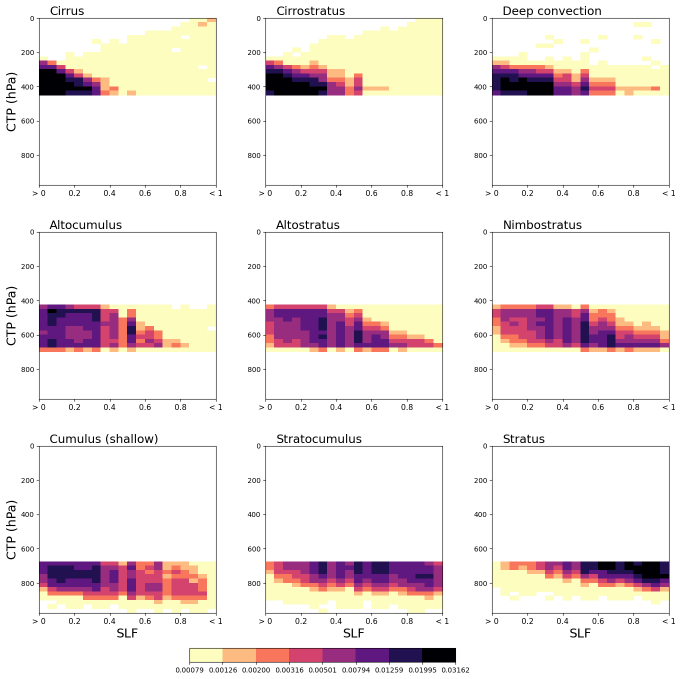

Figure A.27.: CTP-SLF joint histograms for marine mixed-phase clouds in the NH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin.

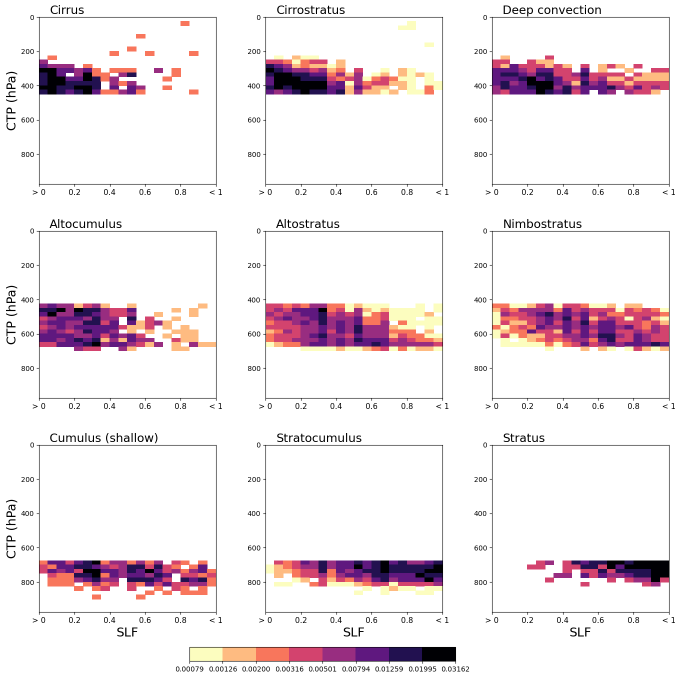

Figure A.28.: CTP-SLF joint histograms for continental mixed-phase clouds in the SH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin.

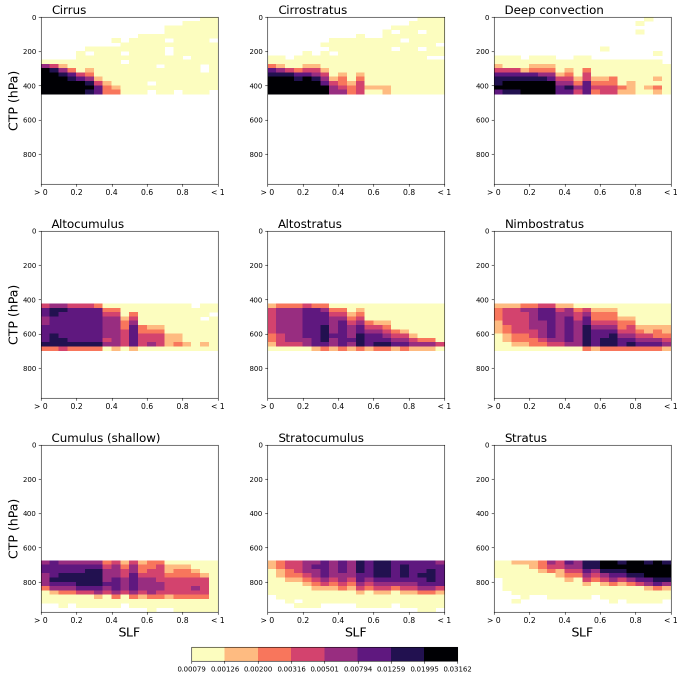

Figure A.29.: CTP-SLF joint histograms for marine mixed-phase clouds in the SH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin.

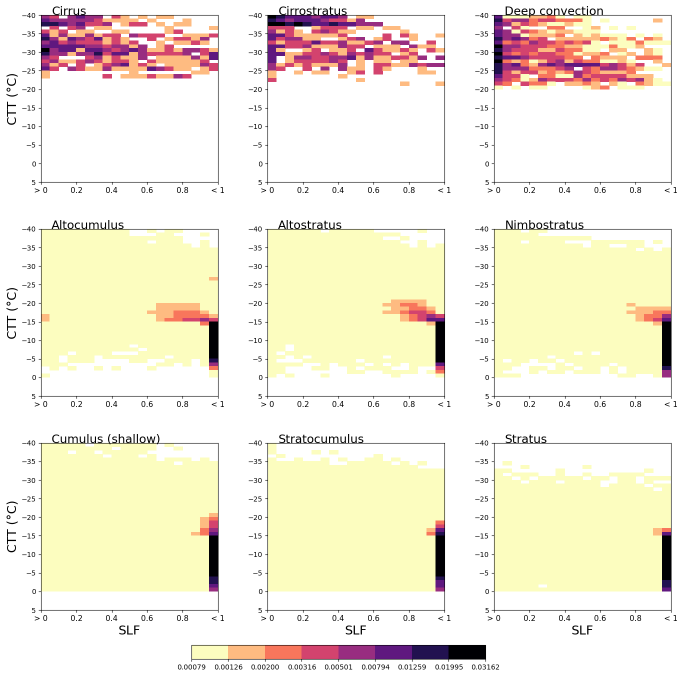

Figure A.30.: SLF-CTT joint histograms for continental mixed-phase clouds in the NH for SCREAM. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

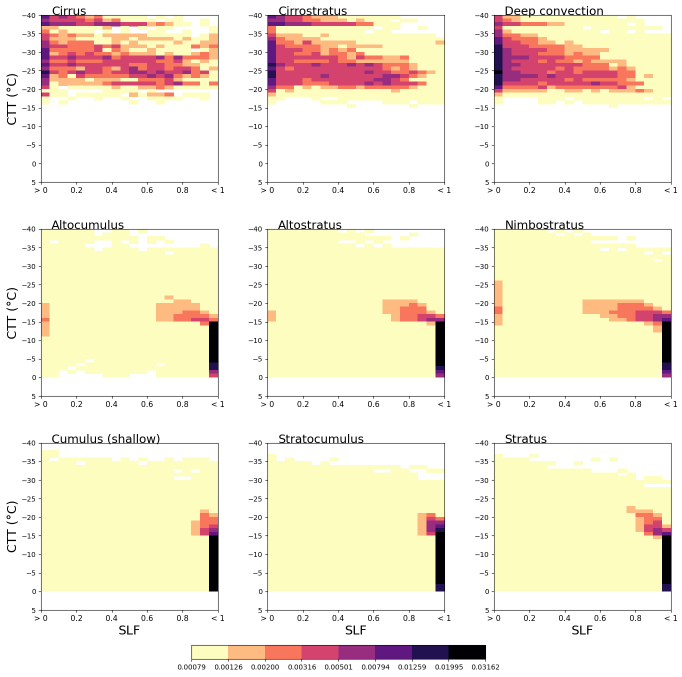

Figure A.31.: SLF-CTT joint histograms for marine mixed-phase clouds in the NH for SCREAM. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

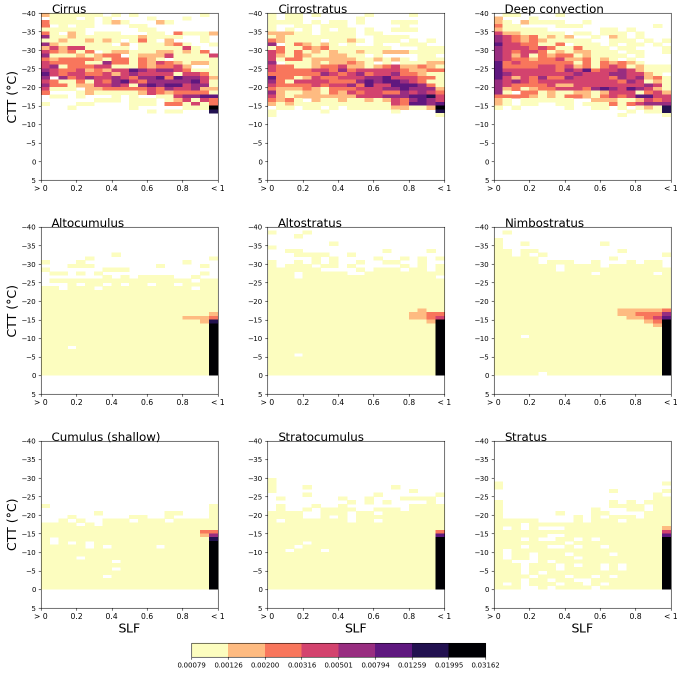

Figure A.32.: SLF-CTT joint histograms for continental mixed-phase clouds in the SH for SCREAM. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

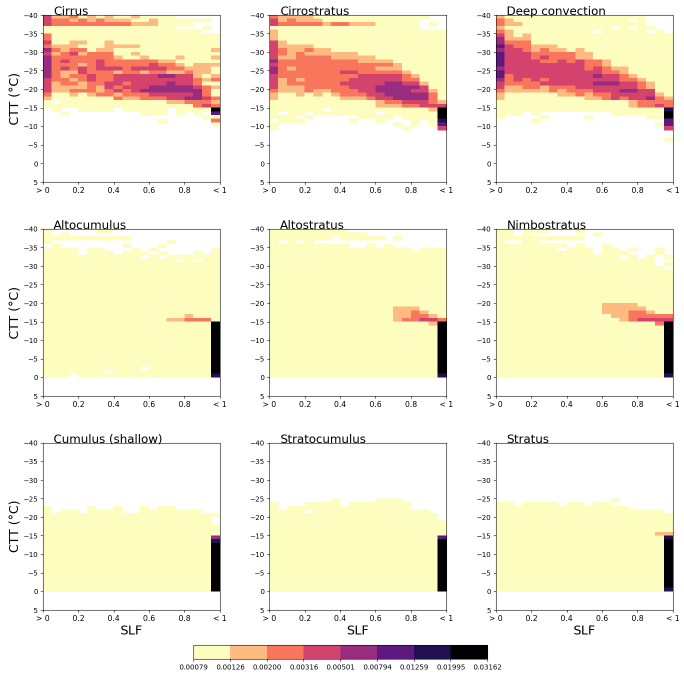

Figure A.33.: SLF-CTT joint histograms for marine mixed-phase clouds in the SH for SCREAM. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

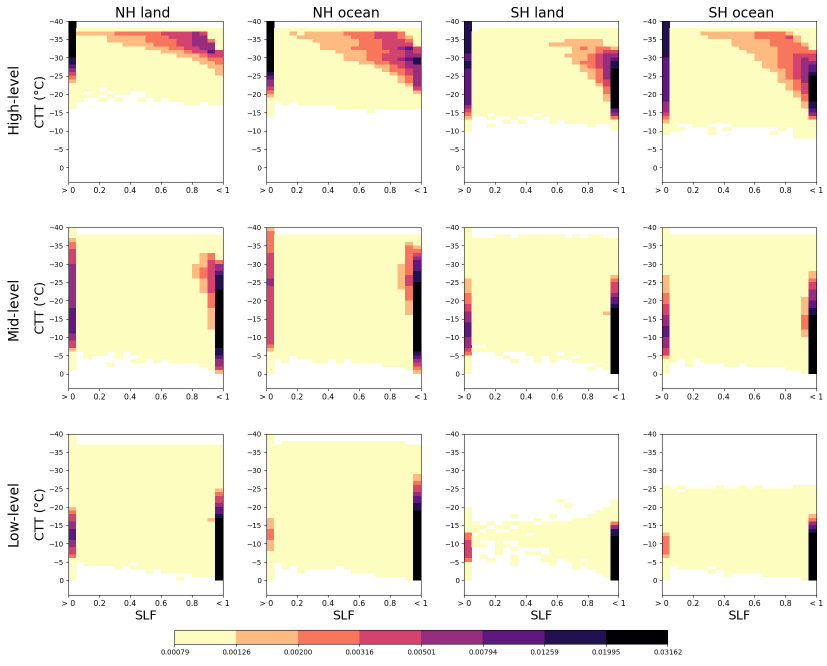

Figure A.34.: SLF-CTT joint histograms for continental and marine mixed-phase clouds in the NH and SH for ICON. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

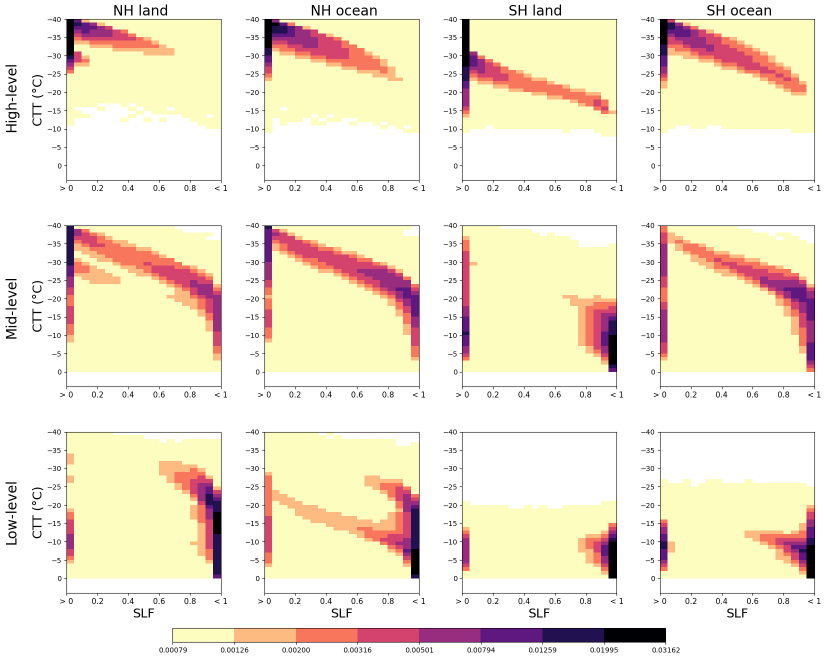

Figure A.35.: SLF-CTT joint histograms for continental and marine mixed-phase clouds in the NH and SH for GEOS. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

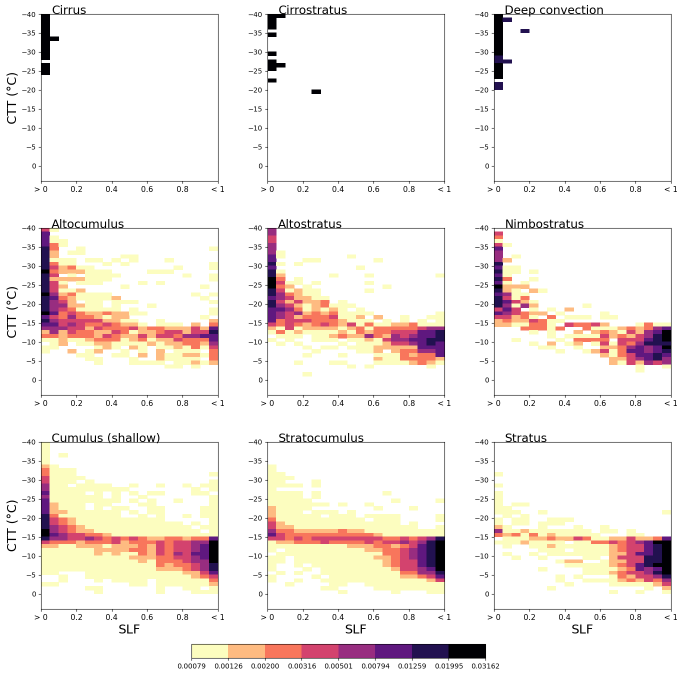

Figure A.36.: SLF-CTT joint histograms for continental mixed-phase clouds in the NH for SCREAM 0.5°. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

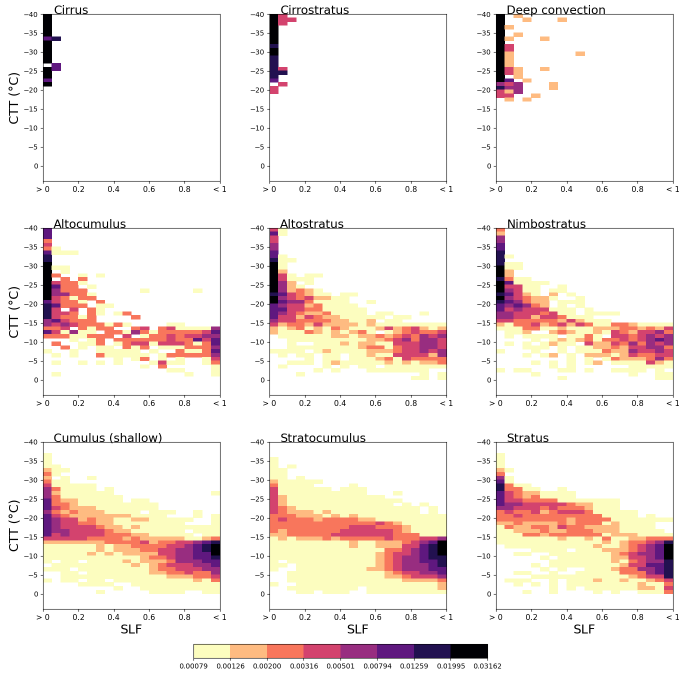

Figure A.37.: SLF-CTT joint histograms for marine mixed-phase clouds in the NH for SCREAM 0.5°. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

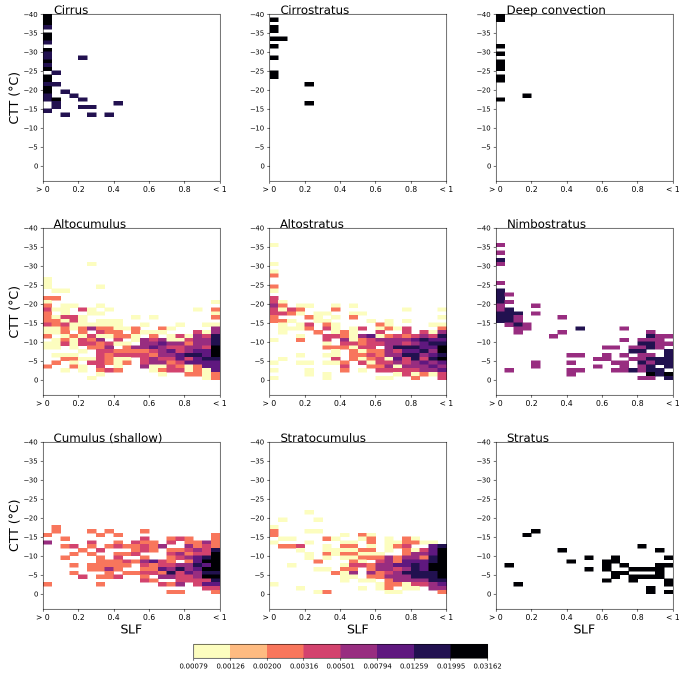

Figure A.38.: SLF-CTT joint histograms for continental mixed-phase clouds in the SH for SCREAM 0.5°. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

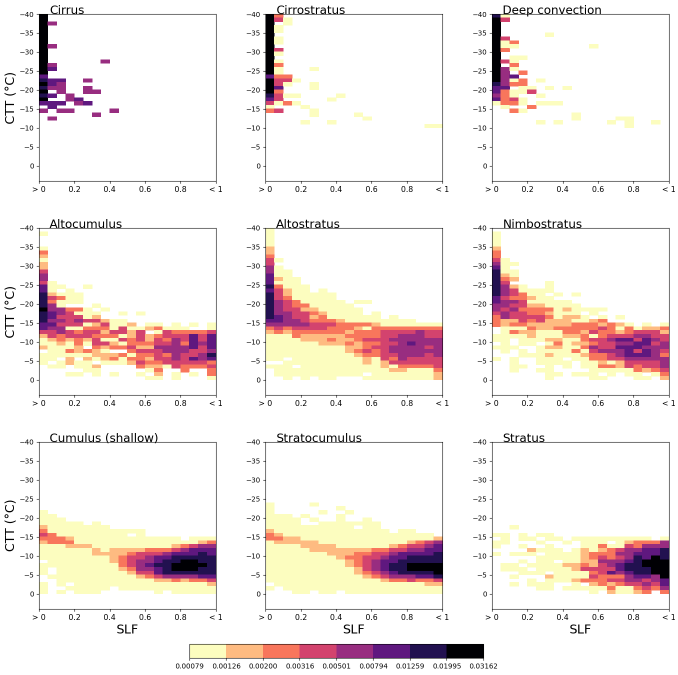

Figure A.39.: SLF-CTT joint histograms for marine mixed-phase clouds in the SH for SCREAM 0.5°. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

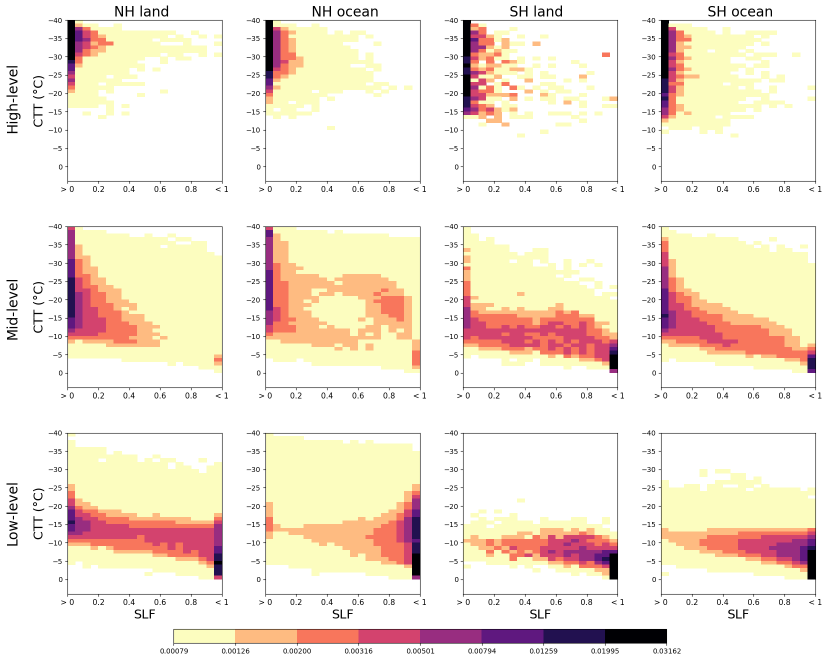

Figure A.40.: SLF-CTT joint histograms for continental and marine mixed-phase clouds in the NH and SH for ICON 0.5°. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

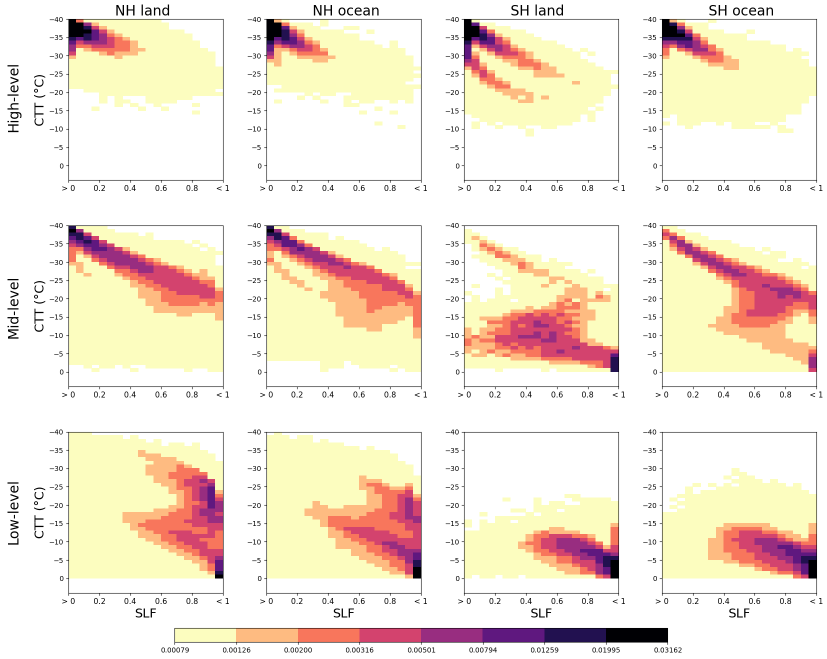

Figure A.41.: SLF-CTT joint histograms for continental and marine mixed-phase clouds in the NH and SH for GEOS 0.5°. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin.

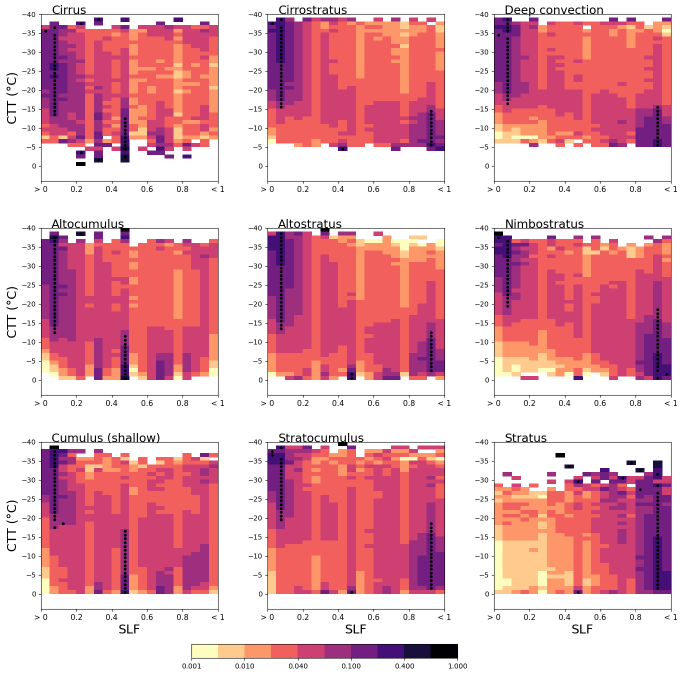

Figure A.42.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the NH for Cloud_cci v3. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

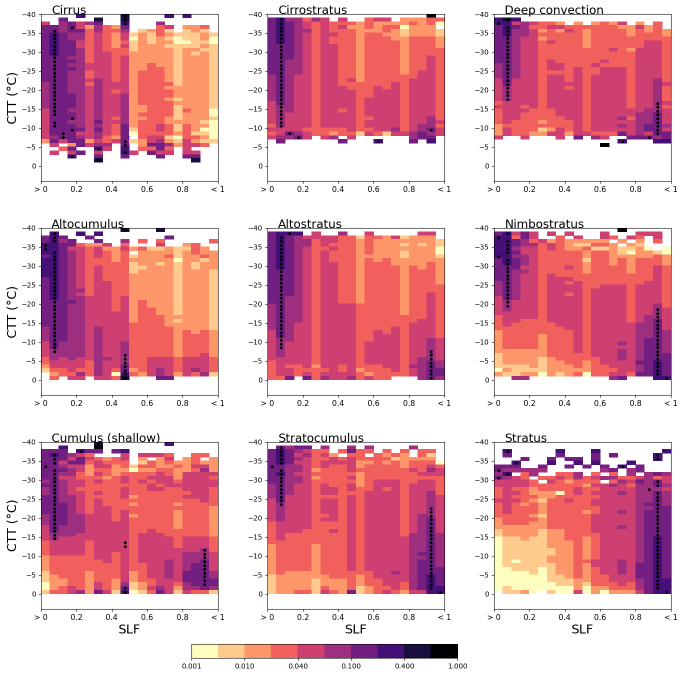

Figure A.43.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the NH for Cloud_cci v3. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

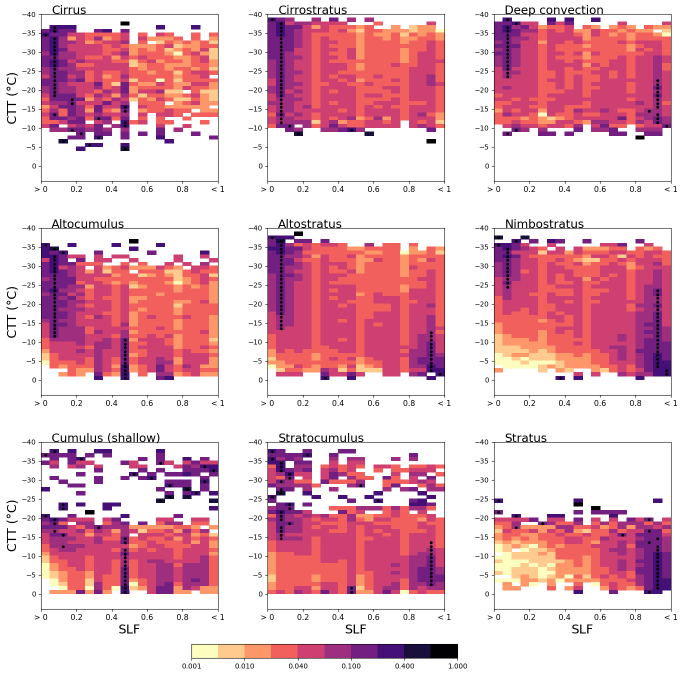

Figure A.44.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the SH for Cloud_cci v3. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

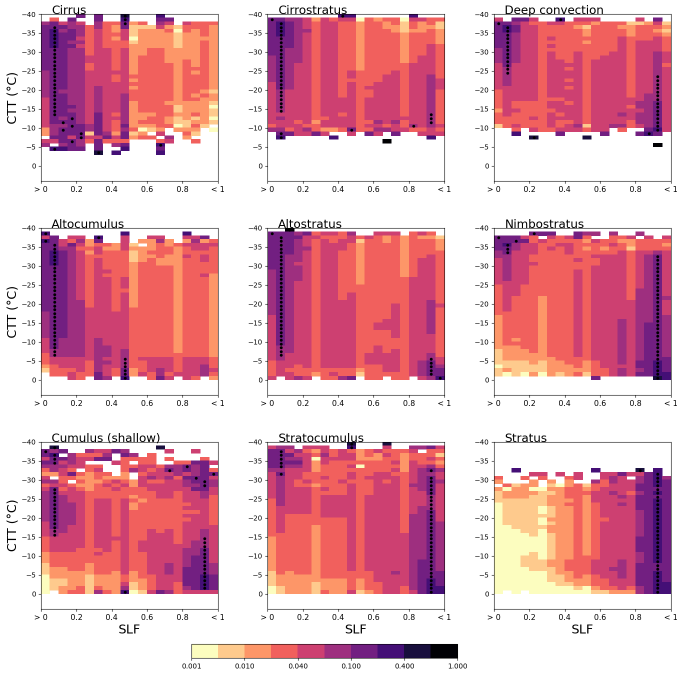

Figure A.45.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the SH for Cloud_cci v3. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

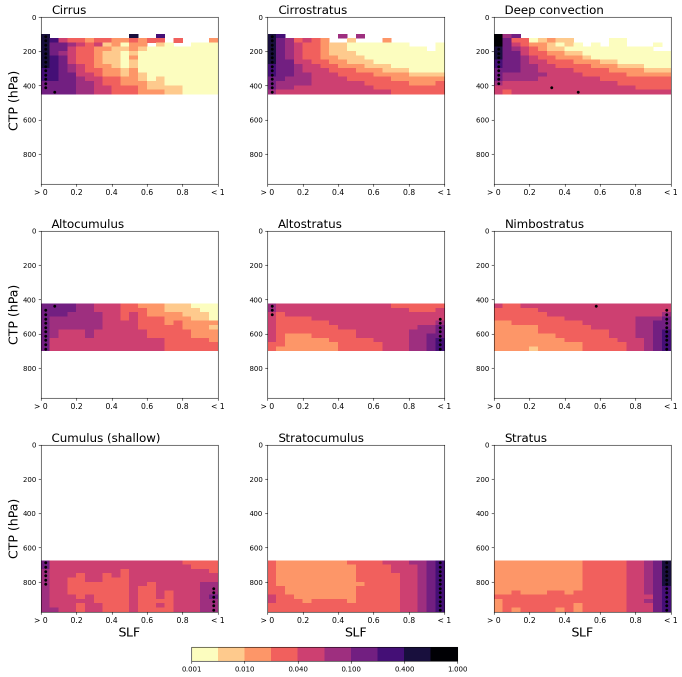

Figure A.46.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the NH for MCD06COSP. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin relative to a single isobar.

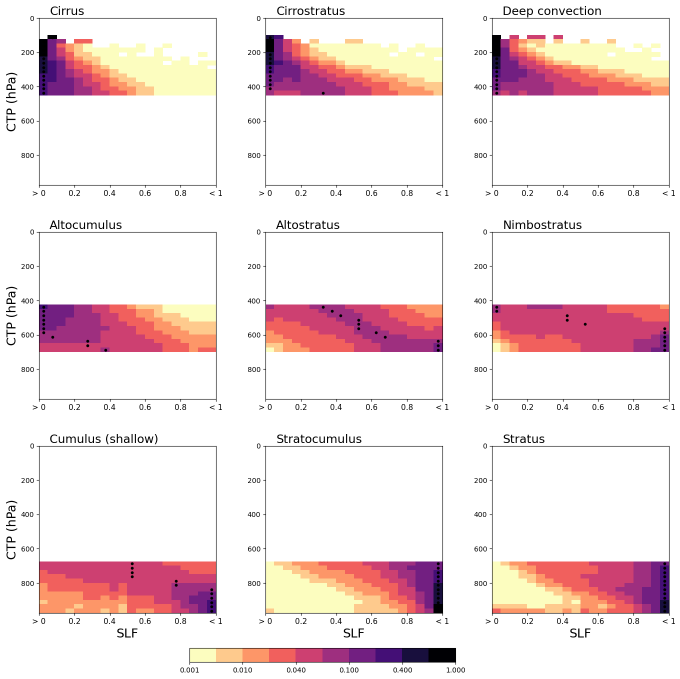

Figure A.47.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the NH for MCD06COSP. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin relative to a single isobar.

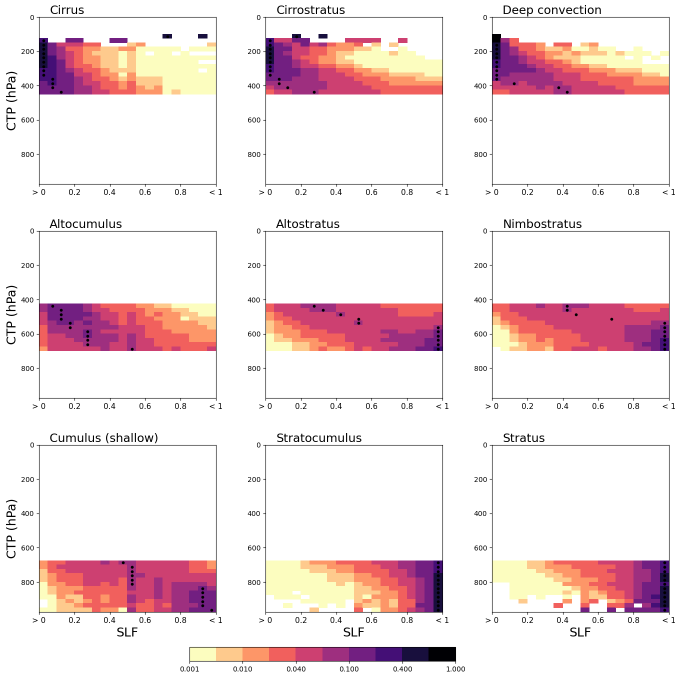

Figure A.48.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the SH for MCD06COSP. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin relative to a single isobar.

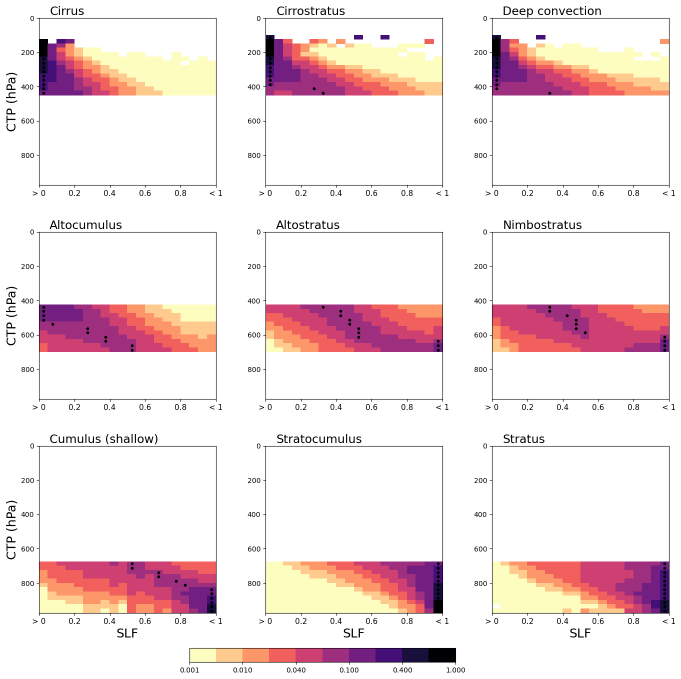

Figure A.49.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the SH for MCD06COSP. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin relative to a single isobar.

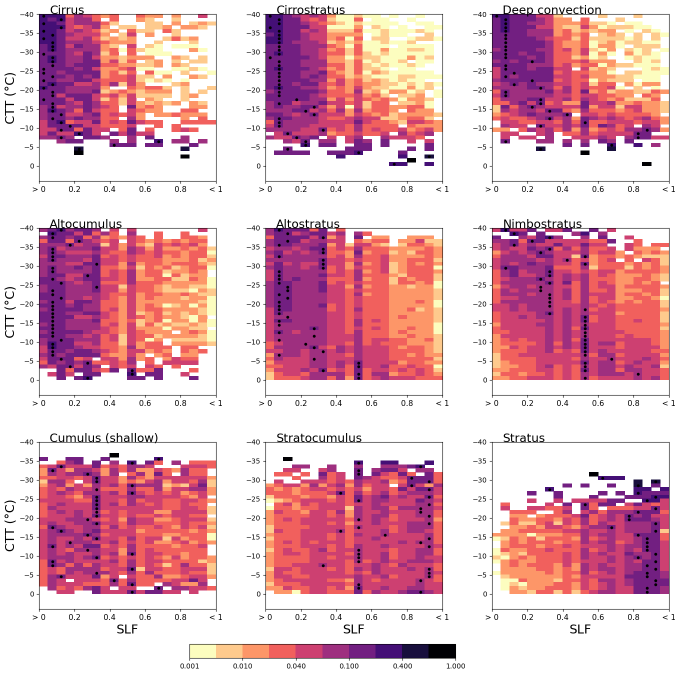

Figure A.50.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the NH for CAM6-Oslo. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

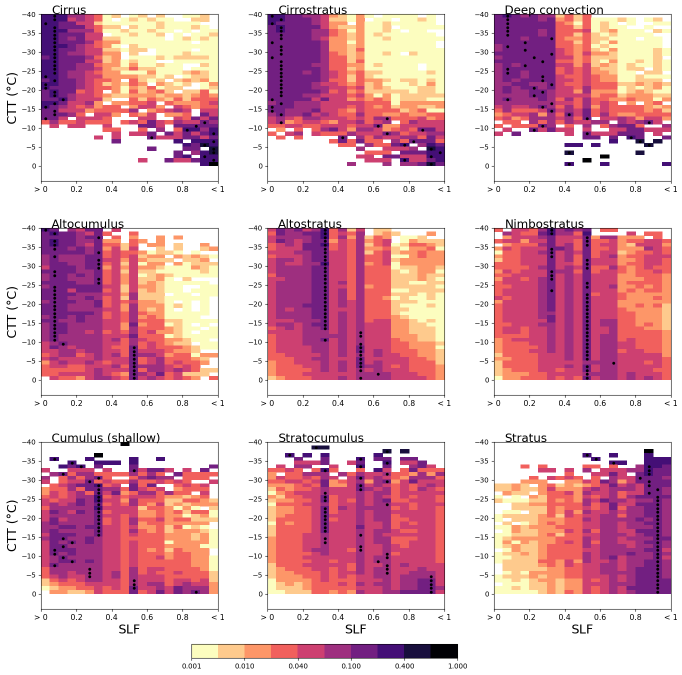

Figure A.51.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the NH for CAM6-Oslo. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

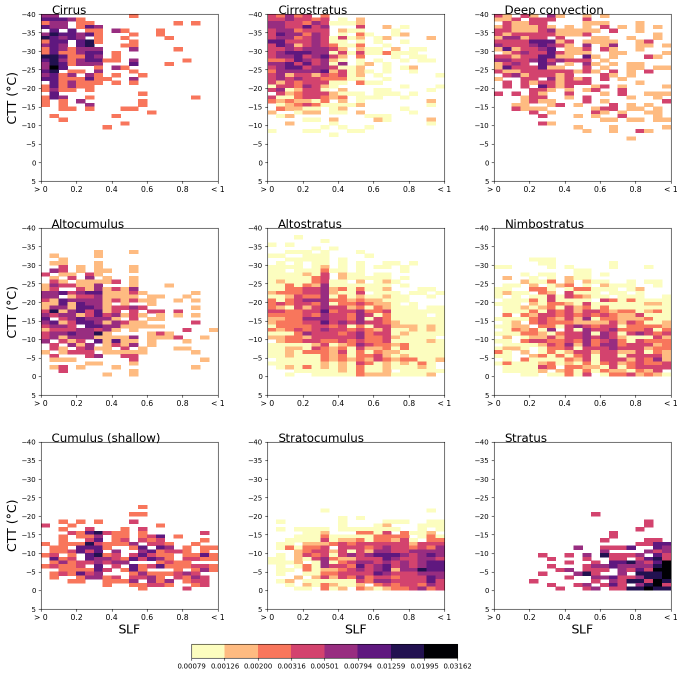

Figure A.52.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the SH for CAM6-Oslo. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

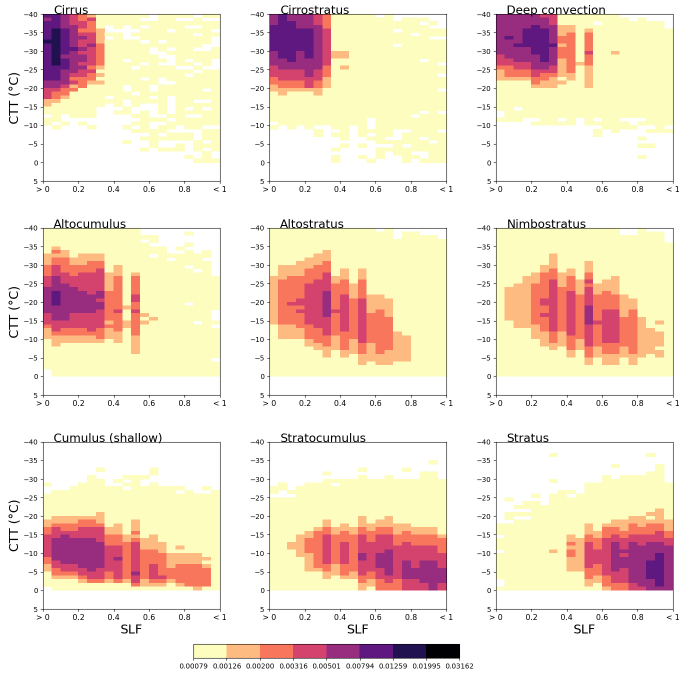

Figure A.53.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the SH for CAM6-Oslo. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

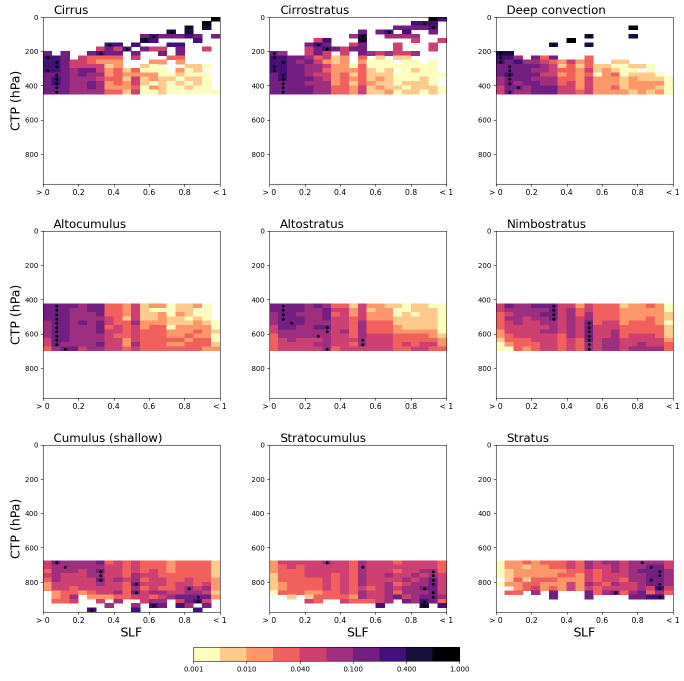

Figure A.54.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the NH for CAM6-Oslo. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin relative to a single isobar.

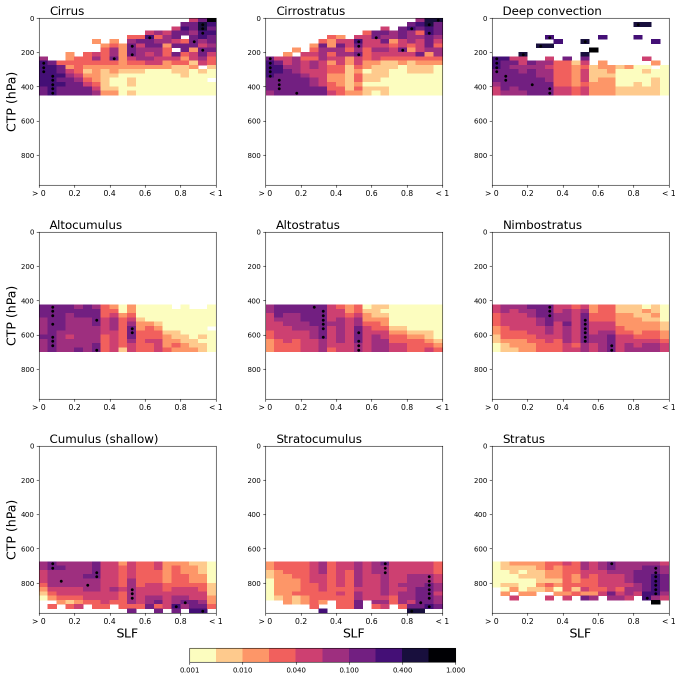

Figure A.55.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the NH for CAM6-Oslo. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin relative to a single isobar.

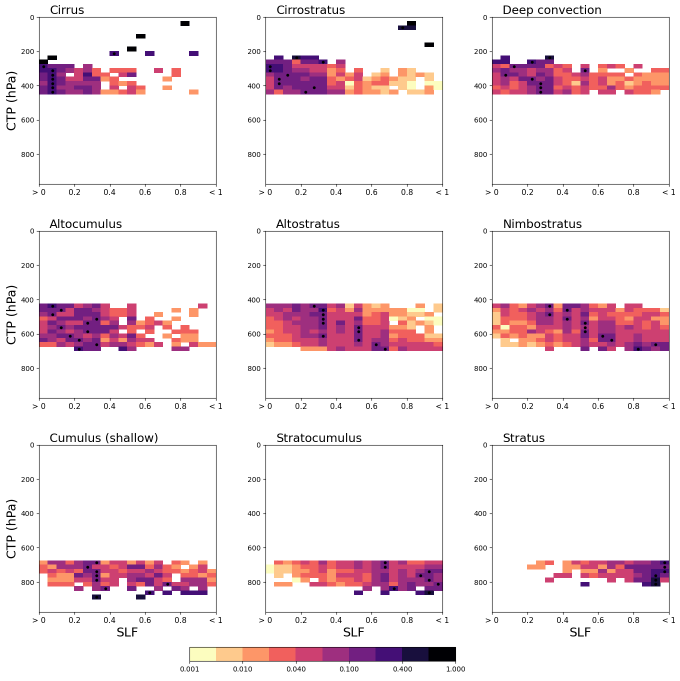

Figure A.56.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the SH for CAM6-Oslo. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given CTP-SLF bin relative to a single isobar.

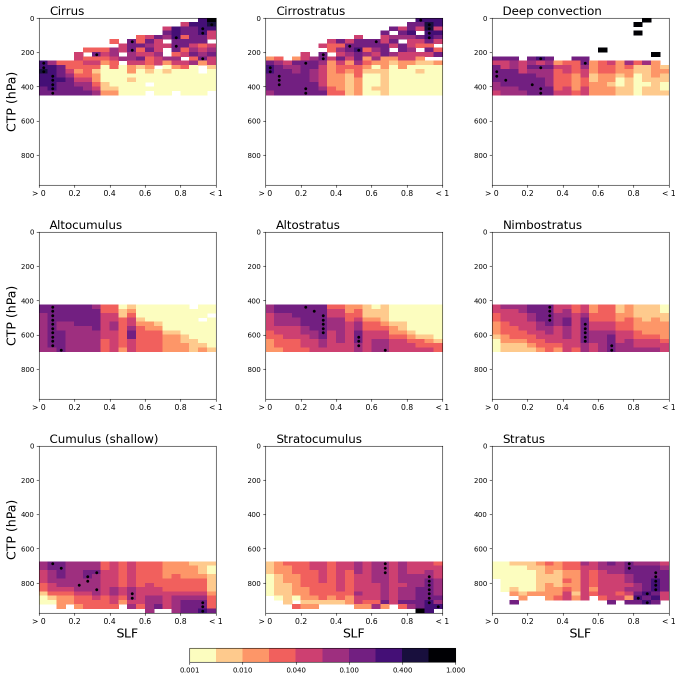

Figure A.57.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the SH for CAM6-Oslo. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isobar.

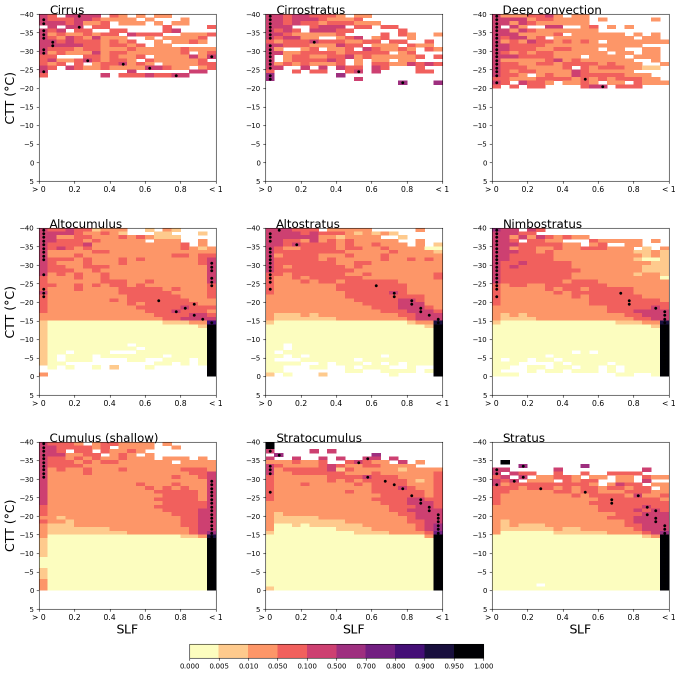

Figure A.58.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the NH for SCREAM. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

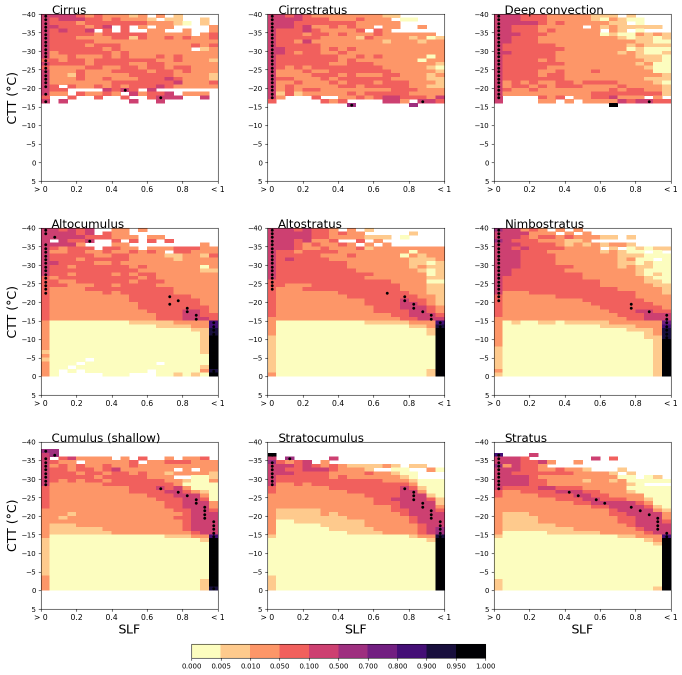

Figure A.59.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the NH for SCREAM. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

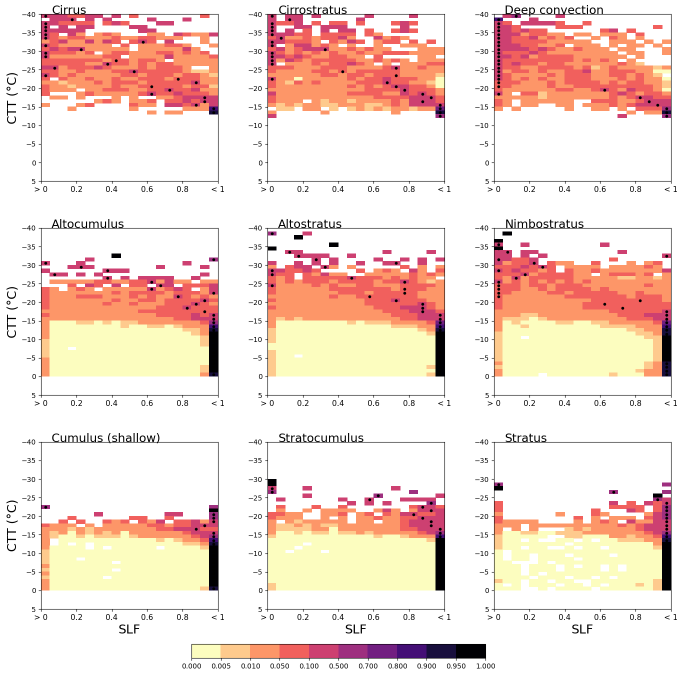

Figure A.60.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the SH for SCREAM. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

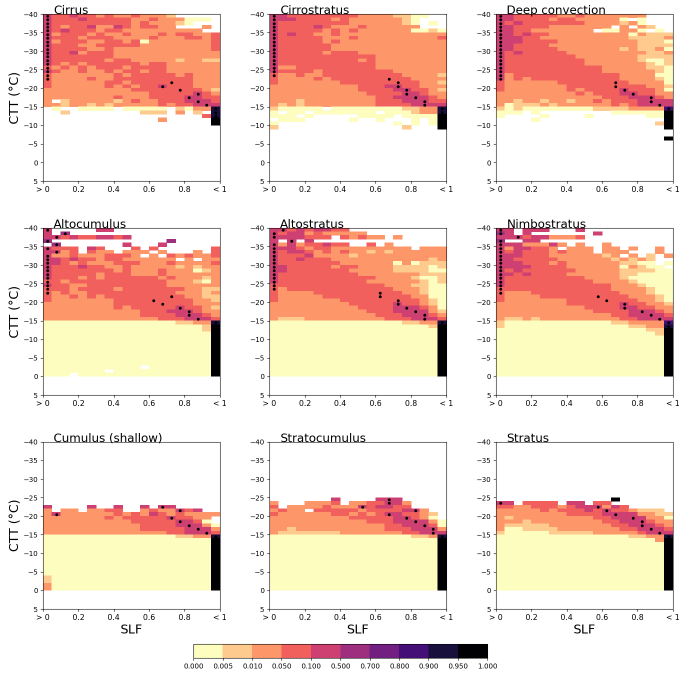

Figure A.61.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the SH for SCREAM. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

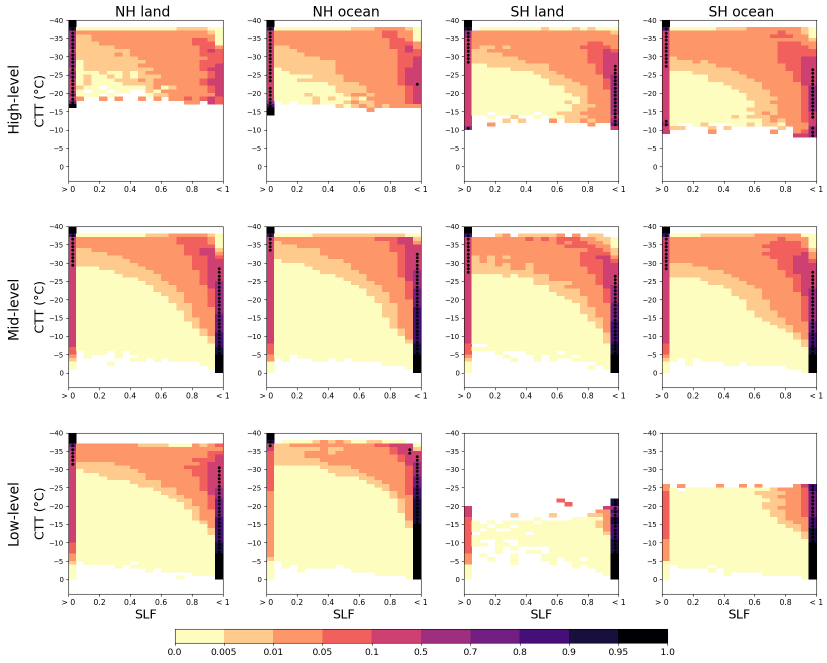

Figure A.62.: CFAD-like histograms for the SLF of continental and marine mixed-phase clouds in the NH and SH for ICON. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

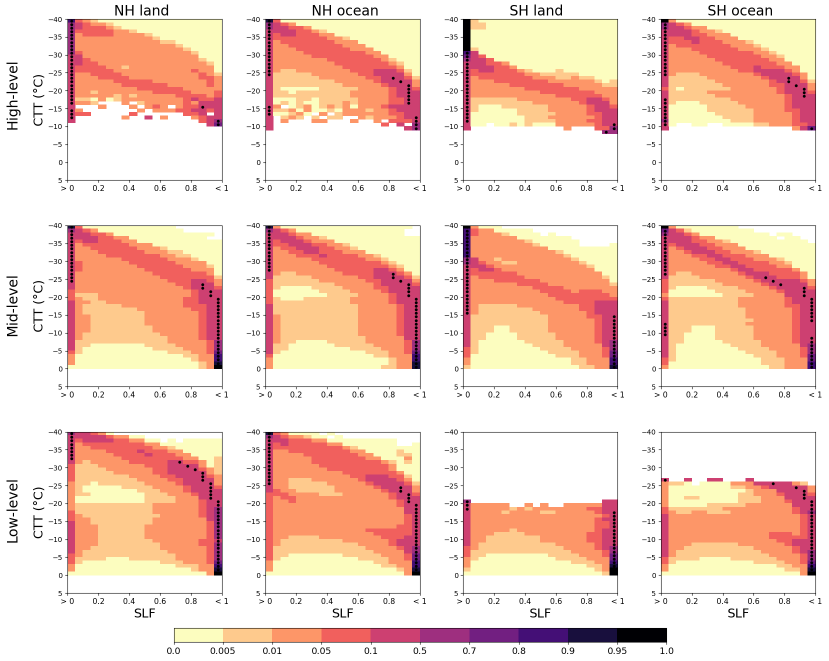

Figure A.63.: CFAD-like histograms for the SLF of continental and marine mixed-phase clouds in the NH and SH for GEOS. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

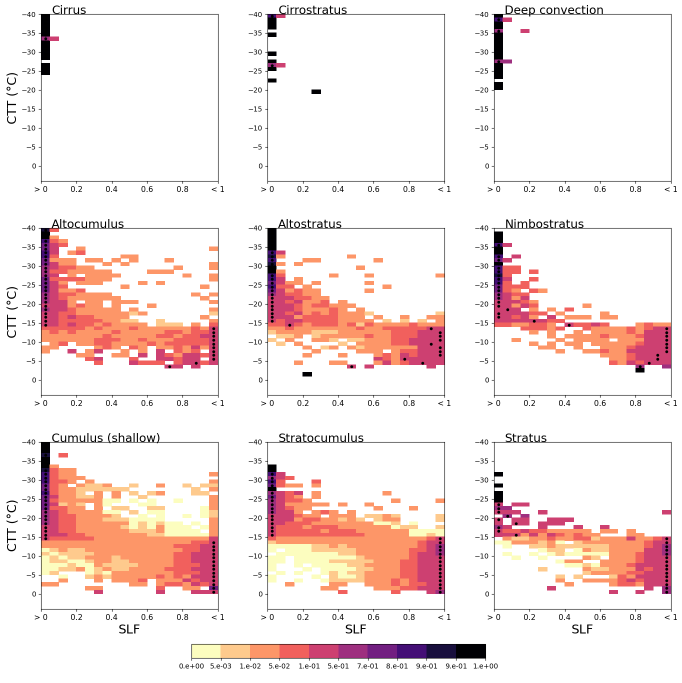

Figure A.64.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the NH for SCREAM 0.5°. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

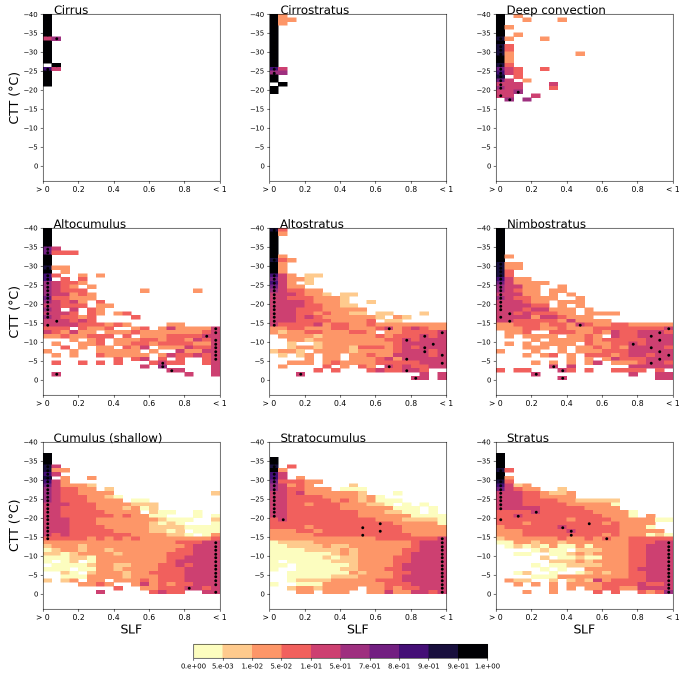

Figure A.65.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the NH for SCREAM 0.5°. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

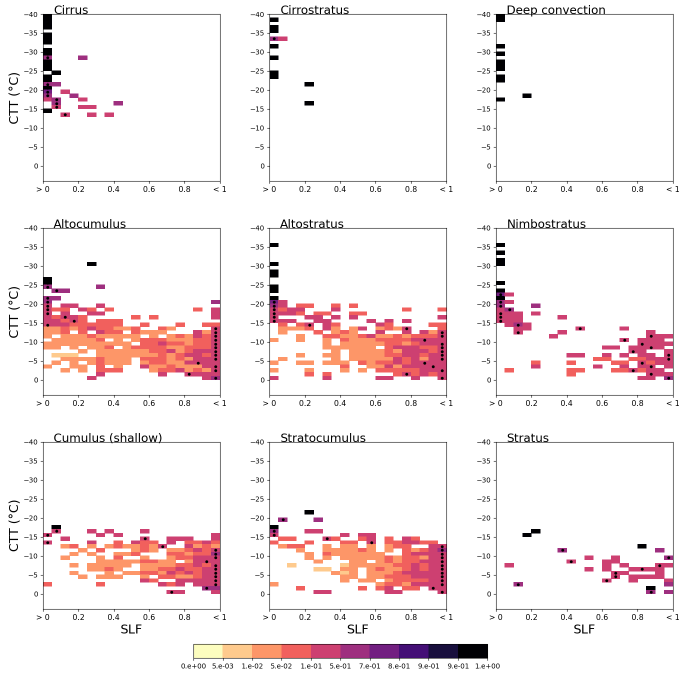

Figure A.66.: CFAD-like histograms for the SLF of continental mixed-phase clouds in the SH for SCREAM 0.5°. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

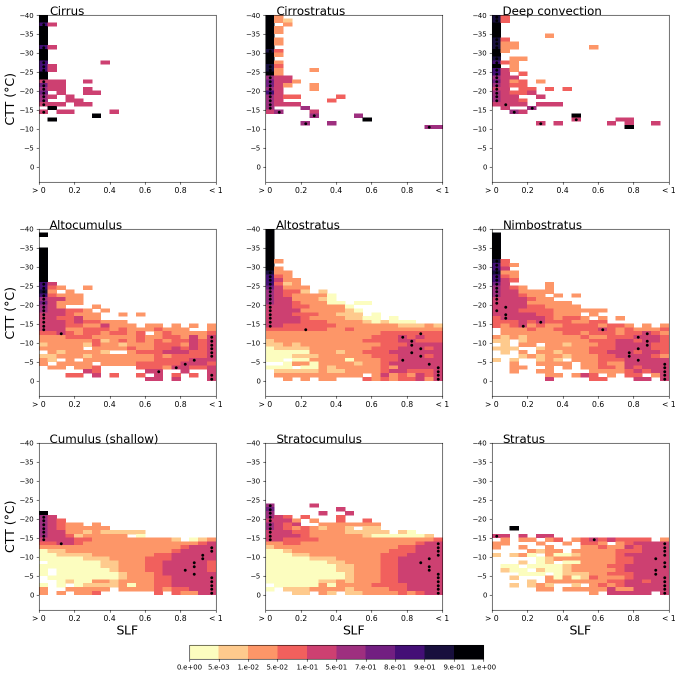

Figure A.67.: CFAD-like histograms for the SLF of marine mixed-phase clouds in the SH for SCREAM 0.5°. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

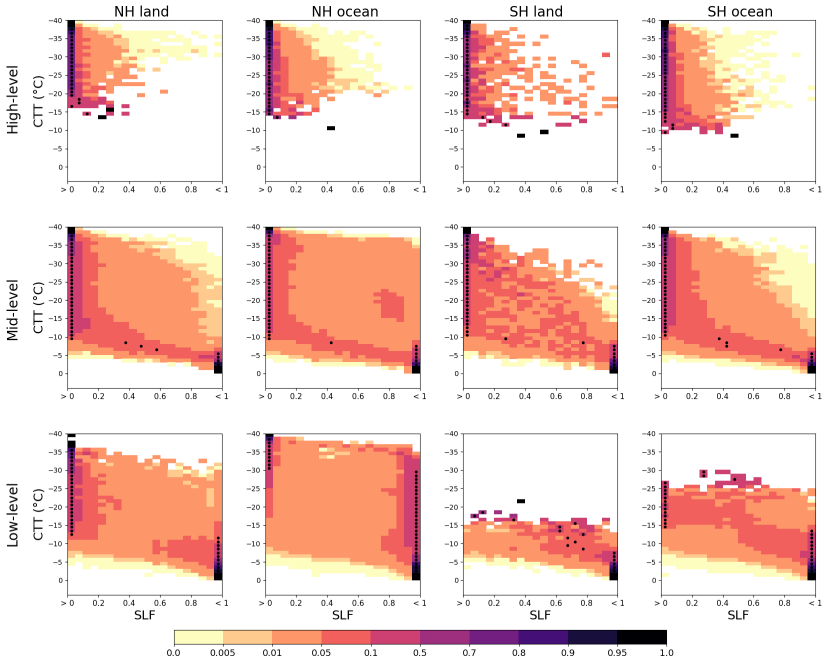

Figure A.68.: CFAD-like histograms for the SLF of continental and marine mixed-phase clouds in the NH and SH for ICON 0.5°. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

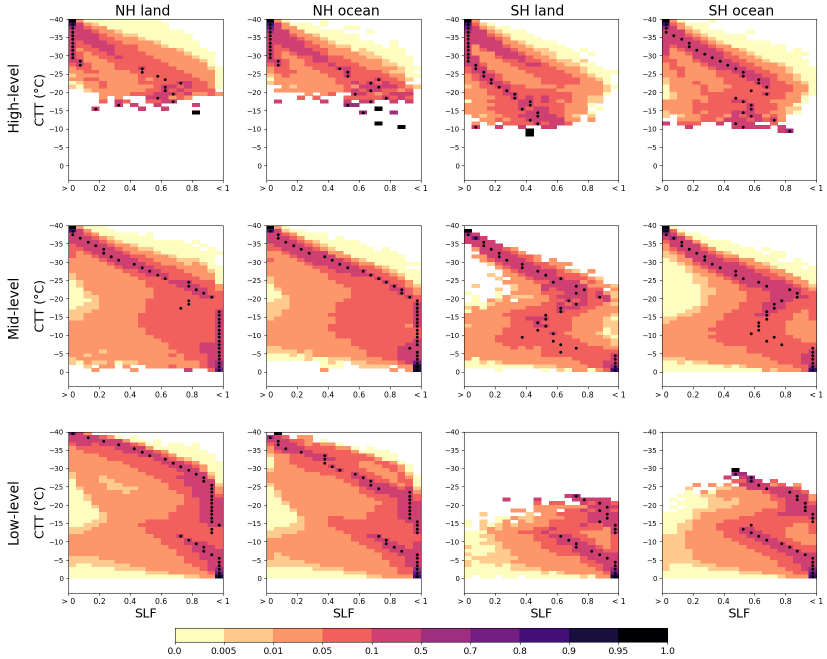

Figure A.69.: CFAD-like histograms for the SLF of continental and marine mixed-phase clouds in the NH and SH for GEOS 0.5°. Each line of the histogram is normalised individually and has a dot where the maxim of SLF is located. The colorbar indicates the relative frequency of occurrence for a given SLF-CTT bin relative to a single isotherm.

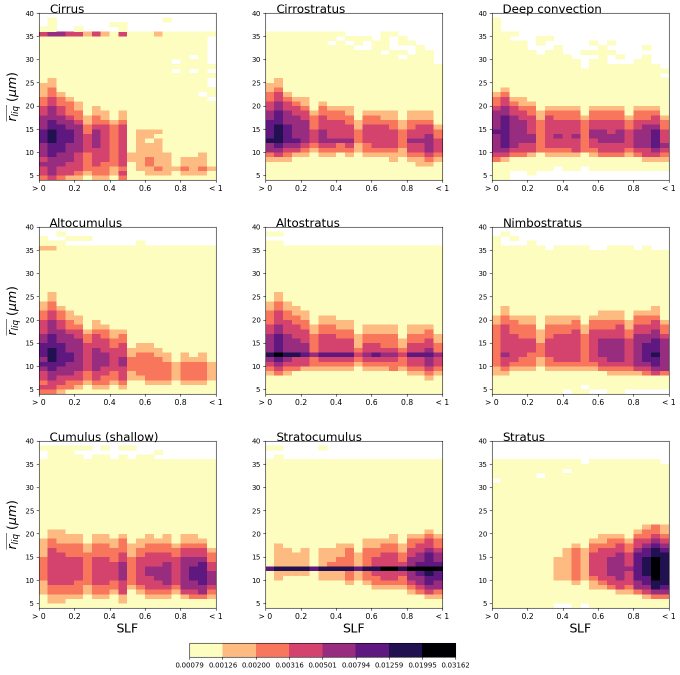

Figure A.70.: SLF-$\overline{r_{liq}}$ joint histograms for marine mixed-phase clouds in the NH for Cloud_cci v3. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

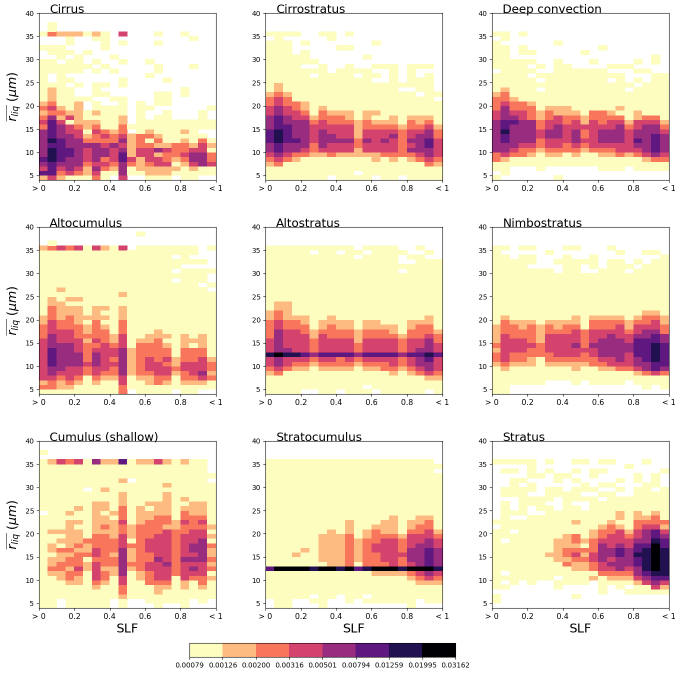

Figure A.71.: SLF-$\overline{r_{liq}}$ joint histograms for continental mixed-phase clouds in the SH for Cloud_cci v3. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

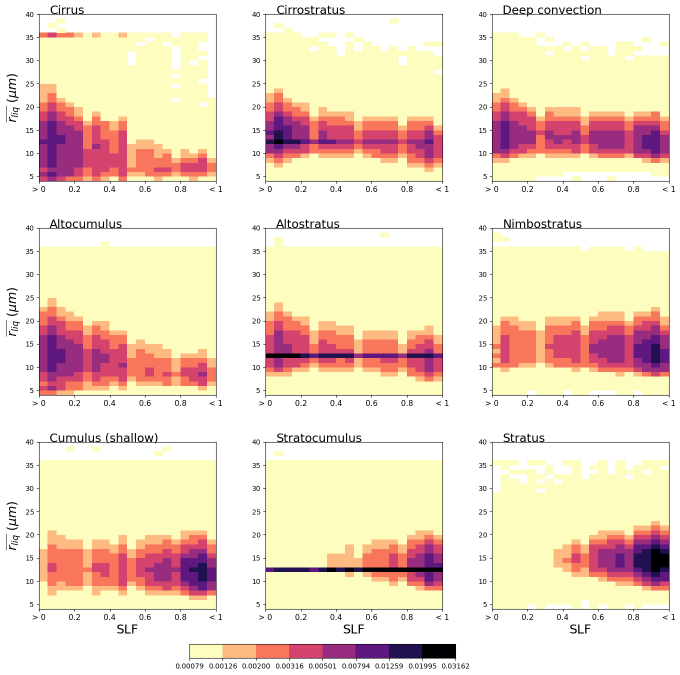

Figure A.72.: SLF-$\overline{r_{liq}}$ joint histograms for marine mixed-phase clouds in the SH for Cloud_cci v3. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

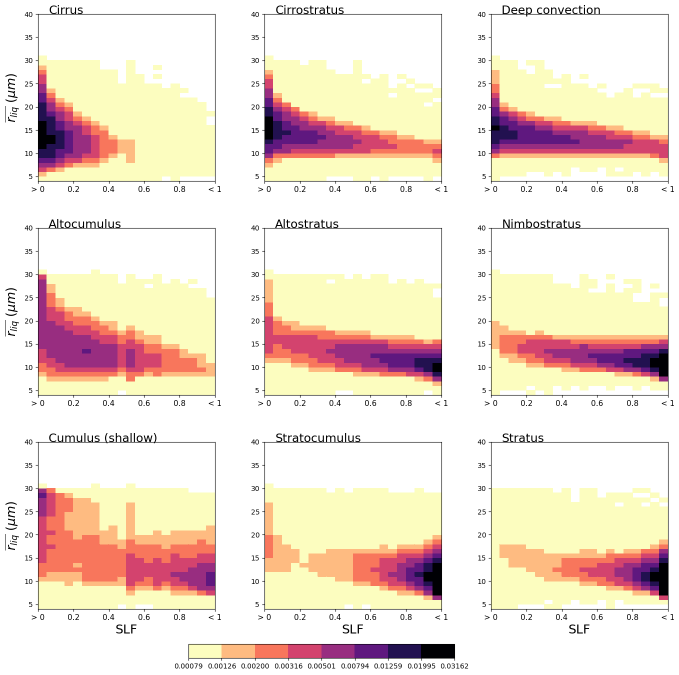

Figure A.73.: SLF-$\overline{r_{liq}}$ joint histograms for continental mixed-phase clouds in the NH for MCD06COSP. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

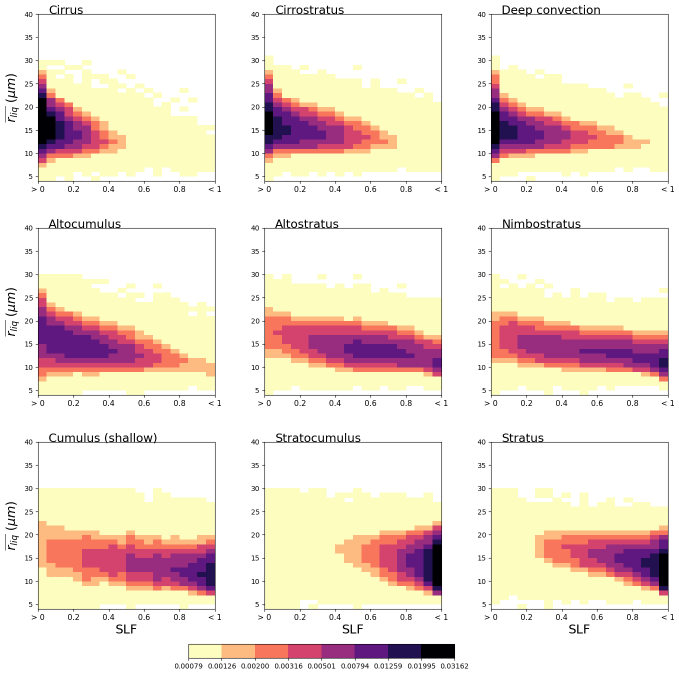

Figure A.74.: SLF-$\overline{r_{liq}}$ joint histograms for marine mixed-phase clouds in the NH for MCD06COSP. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

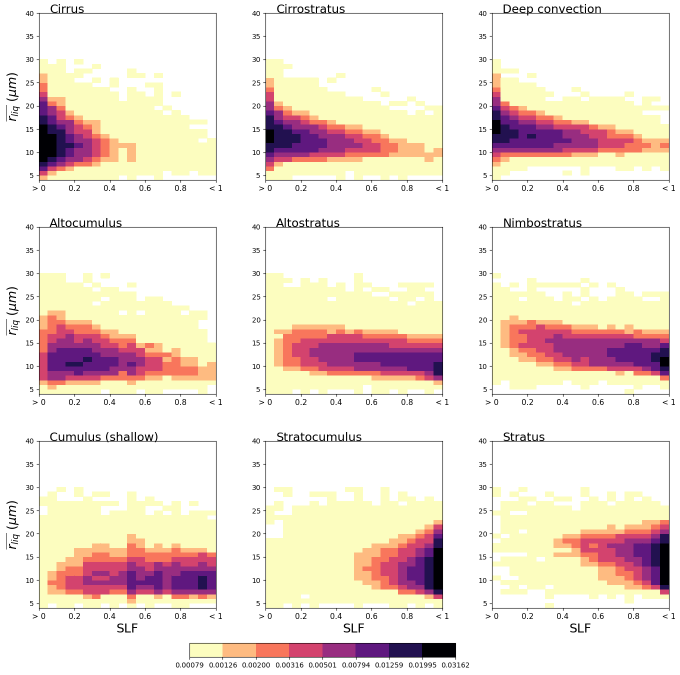

Figure A.75.: SLF-$\overline{r_{liq}}$ joint histograms for continental mixed-phase clouds in the SH for MCD06COSP. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

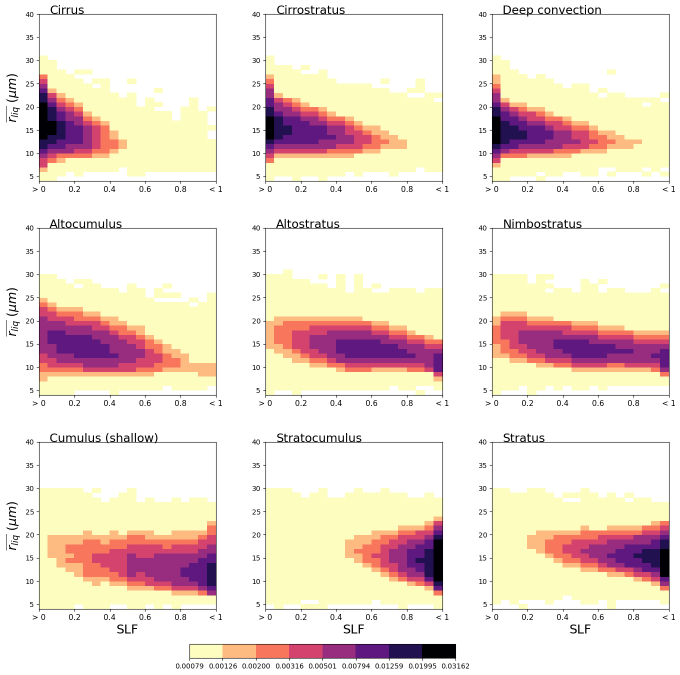

Figure A.76.: SLF-$\overline{r_{liq}}$ joint histograms for marine mixed-phase clouds in the SH for MCD06COSP. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

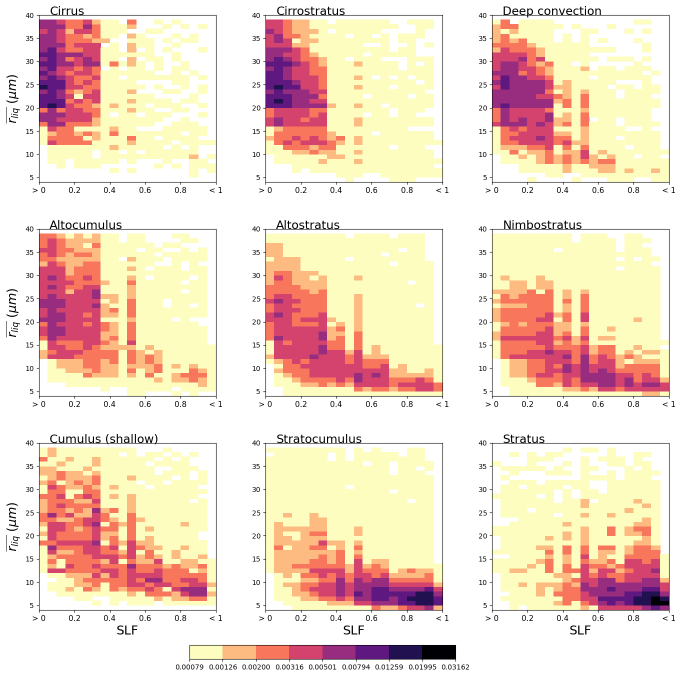

Figure A.77.: SLF-$\overline{r_{liq}}$ joint histograms for continental mixed-phase clouds in the NH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

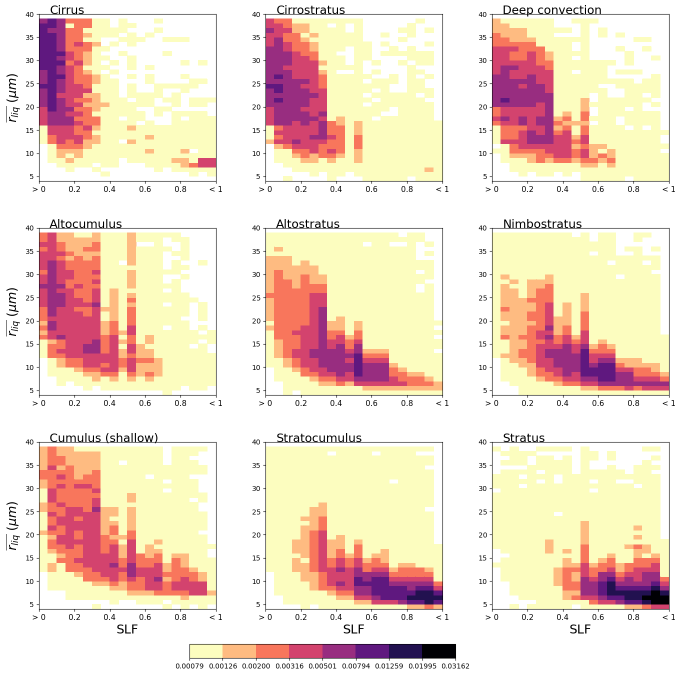

Figure A.78.: SLF-$\overline{r_{liq}}$ joint histograms for marine mixed-phase clouds in the NH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

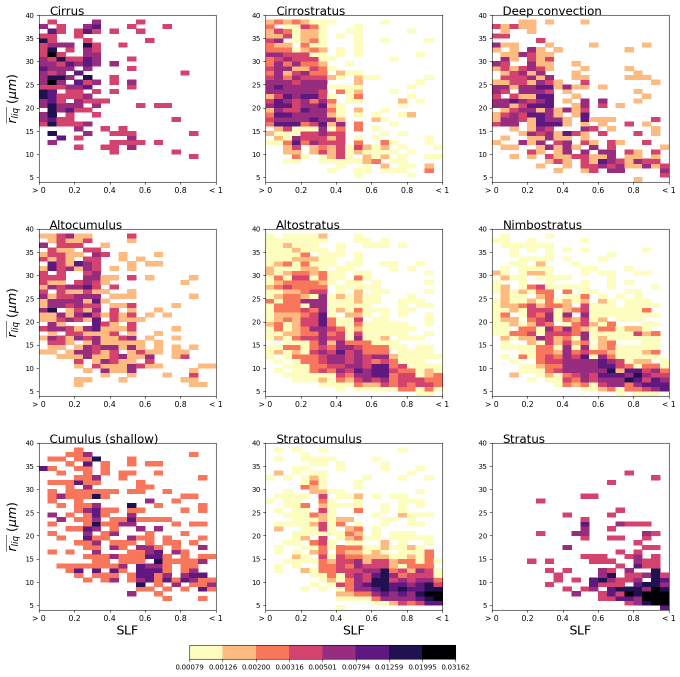

Figure A.79.: SLF-$\overline{r_{liq}}$ joint histograms for continental mixed-phase clouds in the SH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

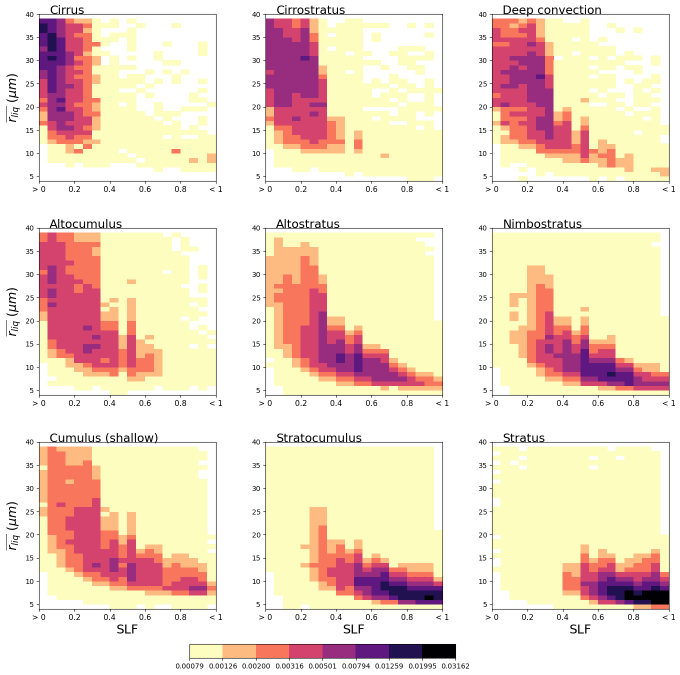

Figure A.80.: SLF-$\overline{r_{liq}}$ joint histograms for marine mixed-phase clouds in the SH for CAM6-Oslo. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

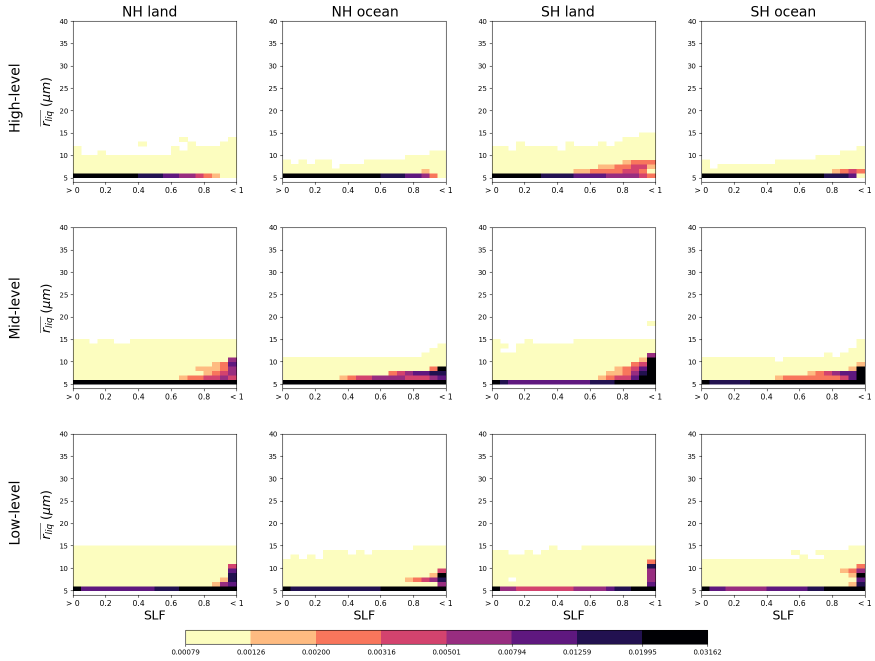

Figure A.81.: SLF-$\overline{r_{liq}}$ joint histograms for continental and marine mixed-phase clouds in the NH and SH for GEOS. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

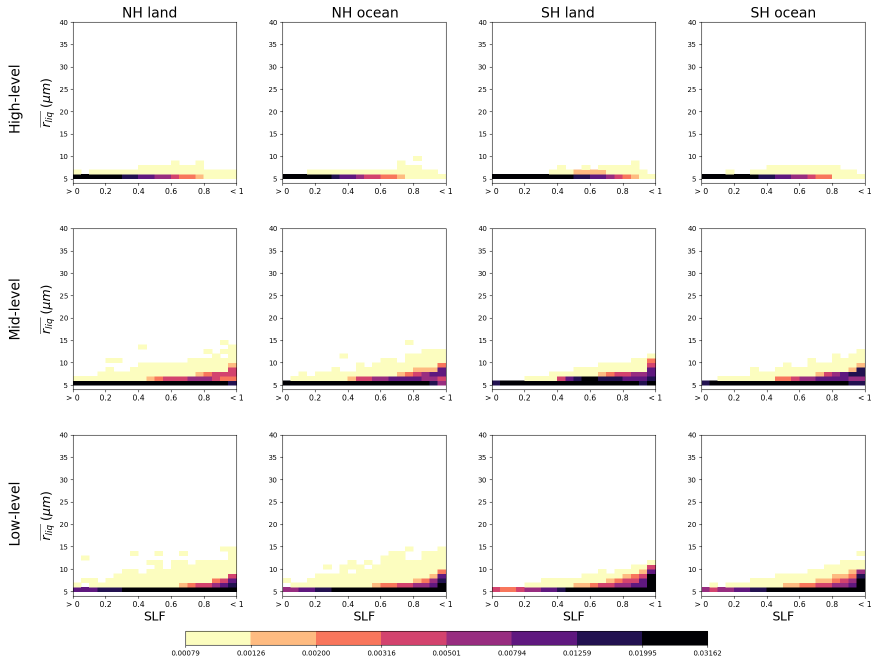

Figure A.82.: SLF-$\overline{r_{liq}}$ joint histograms for continental and marine mixed-phase clouds in the NH and SH for GEOS 0.5°. Each histogram is normalised individually. The colorbar indicates the relative frequency of occurrence for a given SLF-$\overline{r_{liq}}$ bin.

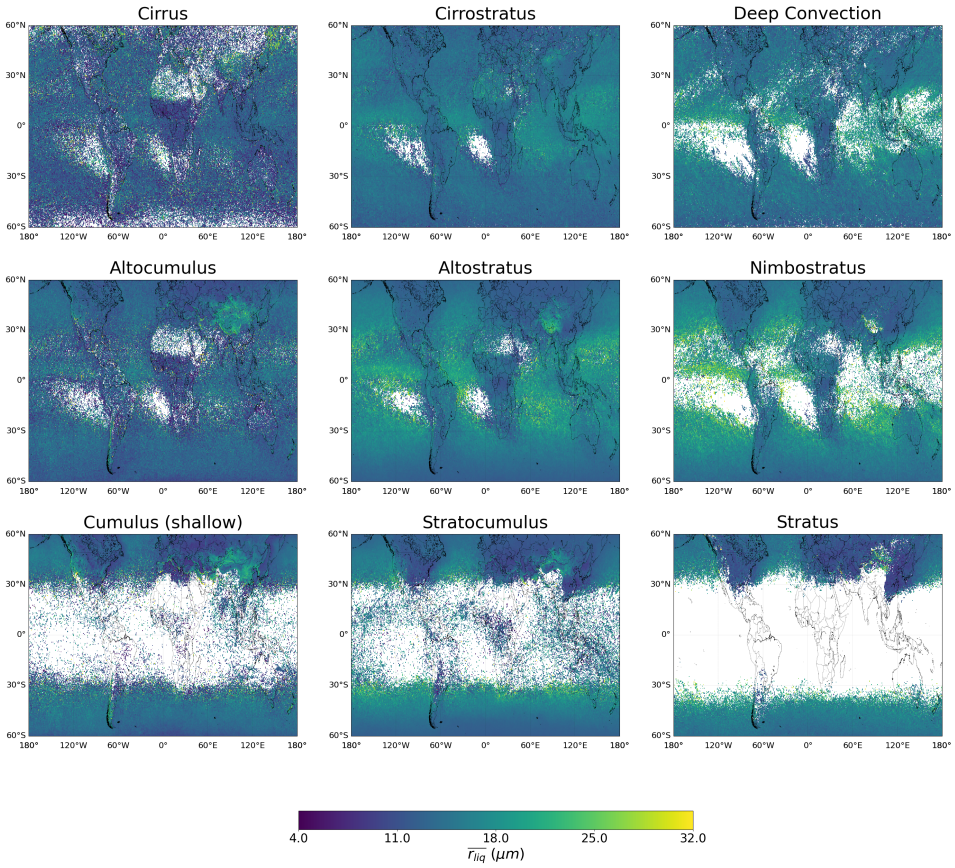

Figure A.83.: Geographical distribution of $\overline{r_{liq}}$ in different cloud types in the mixed phase for Cloud_cci v3. The colorbar indicates the size of $\overline{r_{liq}}$ in µm.

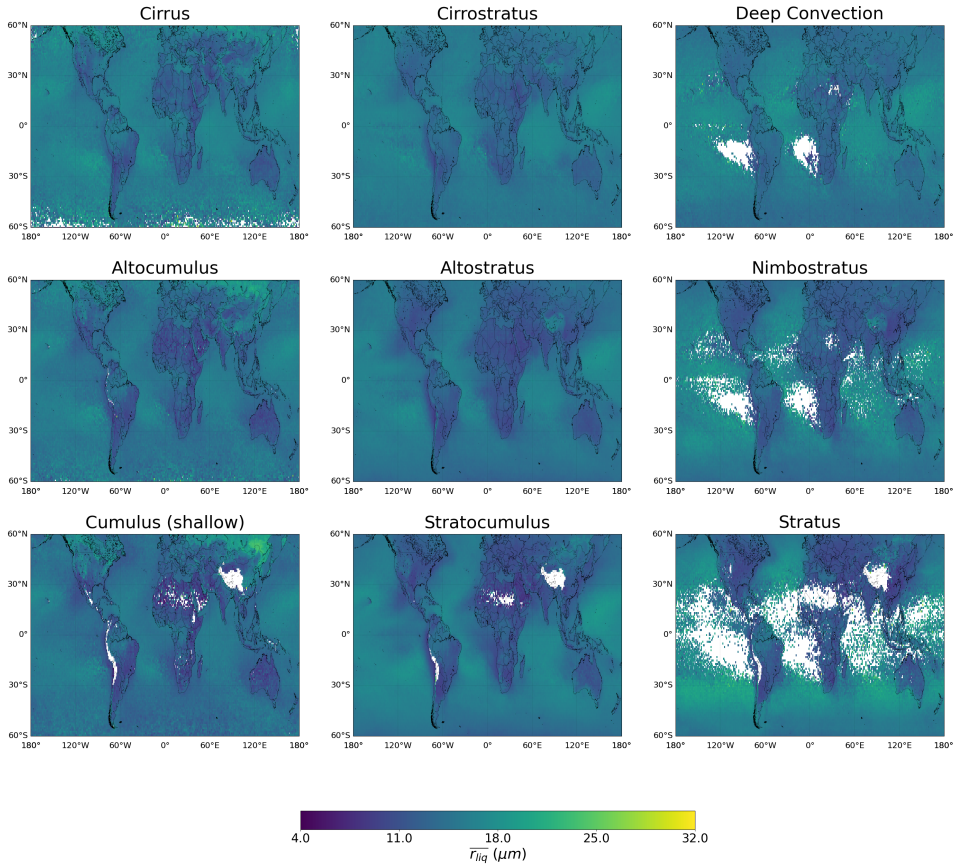

Figure A.84.: Geographical distribution of $\overline{r_{liq}}$ in different cloud types in the mixed phase for MCD06COSP. The colorbar indicates the size of $\overline{r_{liq}}$ in µm.

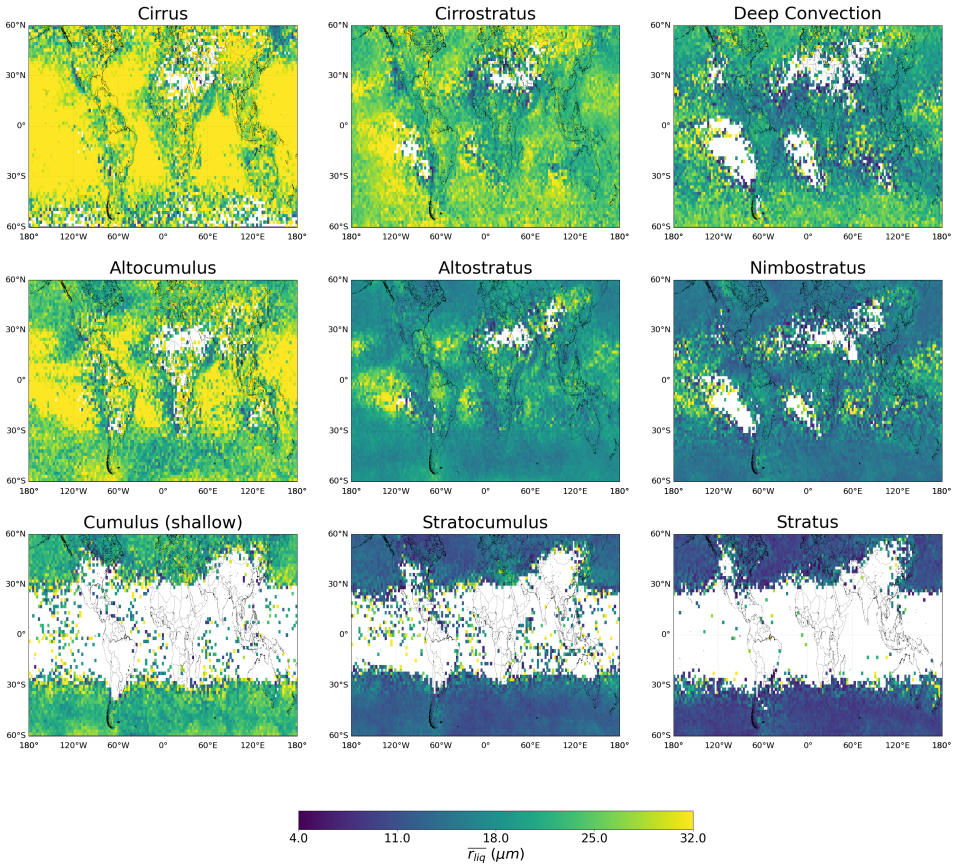

Figure A.85.: Geographical distribution of $\overline{r_{liq}}$ in different cloud types in the mixed phase for CAM6-Oslo. The colorbar indicates the size of $\overline{r_{liq}}$ in µm.

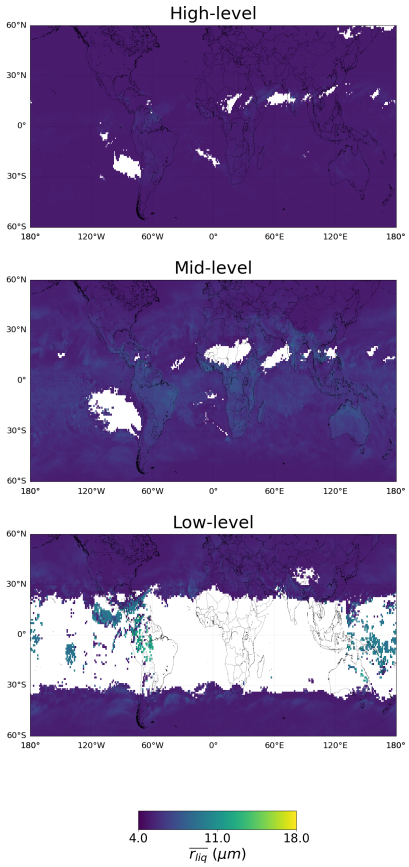

Figure A.86.: Geographical distribution of $\overline{r_{liq}}$ in different cloud heights in the mixed phase for GEOS. The colorbar indicates the size of $\overline{r_{liq}}$ in µm.

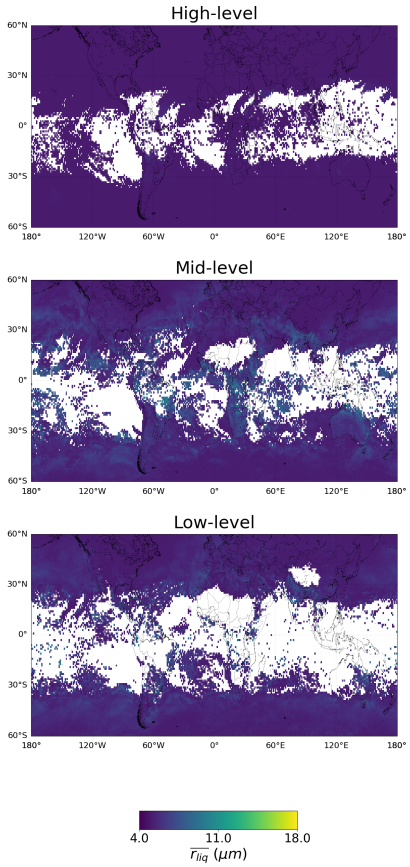

Figure A.87.: Geographical distribution of $\overline{r_{liq}}$ in different cloud heights in the mixed phase for GEOS 0.5°. The colorbar indicates the size of $\overline{r_{liq}}$ in µm.

B. Bibliography

Amrhein, V., S. Greenland, and B. McShane, 2019: Retire statistical significance. *Nature*, **567**, 305–307, URL https://doi.org/10.1038/d41586-019-00857-9.

AMS, 2012: Glossary of meteorology: Asimmetry parameter. URL https://glossary.ametsoc.org/wiki/Asymmetry_factor#:~:text=Larger%20particles%2C%20such%20as%20cloud,American%20Meteorological%20Society%20(AMS).

Anderberg, M. R., 1973: *Cluster analysis for applications*. Academic Press New York, 359 pp., URL https://doi.org/10.1137/1017065.

Ansmann, A., M. Tesche, P. Seifert, D. Althausen, R. Engelmann, J. Fruntke, U. Wandinger, I. Mattis, and D. Müller, 2009: Evolution of the ice phase in tropical altocumulus: Samum lidar observations over cape verde. *Journal of Geophysical Research*, **114**, URL https://doi.org/10.1029/2008JD011659.

ATBD-CC4CLv5.1, 2019: *Algorithm Theoretical Baseline Document ATBD CC4CL - ESA Cloud_cci*. URL https://climate.esa.int/media/documents/Cloud_Algorithm-Theoretical-Baseline-Document-ATBD-CC4CL_v5.1.pdf, issue 6, Revision: 2, date of issue: 18/10/2019.

ATBD-CC4CLv6.2, 2019: *Algorithm Theoretical Baseline Document ATBD CC4CL - ESA Cloud_cci*. URL https://climate.esa.int/media/documents/Cloud_Algorithm-Theoretical-Baseline-Document-ATBD-CC4CL_v6.2.pdf, issue 5, Revision: 1, date of issue: 28/08/2017.

ATBD-CPP_AVHRR, 2016: *Algorithm Theoretical Baseline Document - CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 2 (CLARA-A2) - Cloud Physical products*. URL http://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V002, issue/Revision Index: 2.0, date of issue: 19/08/2016.

Avery, M. A., R. A. Ryan, B. J. Getzewich, M. A. Vaughan, D. M. Winker, Y. Hu, A. Garnier, J. Pelon, and C. A. Verhappen, 2020: Caliop v4 cloud thermodynamic phase assignment and the impact of near-nadir viewing angles. *Atmospheric Measurement Techniques*, **13 (8)**, 4539–4563, URL https://doi.org/10.5194/amt-13-4539-2020.

Barrett, P. A., A. Blyth, P. R. A. Brown, and S. J. Abel, 2020: The structure of turbulence and mixed-phase cloud microphysics in a highly supercooled altocumulus cloud. *Atmospheric Chemistry and Physics*, **20 (4)**, 1921–1939, URL https://doi.org/10.5194/acp-20-1921-2020.

Bergeron, T., 1928: Über die dreidimensional verknüpfende wetteranalyse. erster teil. prinzipielle einführung in das problem der luftmassen- und frontenbildung. *Geofys. Publikasjoner*, **5 (6)**, 1–118, URL http://www.ngfweb.no/docs/NGF_GP_Vol05_no6.pdf.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Monthly Weather Review*, **124 (6)**, 1256 – 1271, URL https://doi.org/10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2.

Bodas-Salcedo, A., M. J. Webb, S. Bony, H. Chepfer, J.-L. Dufresne, S. A. Klein, Y. Zhang, R. Marchand, J. M. Haynes, R. Pincus, and V. O. John, 2011: Cosp: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, **92 (8)**, 1023 – 1043, URL https://doi.org/10.1175/2011BAMS2856.1.

Bony, S., J.-L. Dufresne, H. Le Treut, J.-J. Morcrette, and C. Senior, 2004: On dynamic and thermodynamic components of cloud changes. *Climate Dynamics*, **22**, 71–86, URL https://doi.org/10.1007/s00382-003-0369-6.

Boucher, O., 2013: Clouds and Aerosols. *Climate Change 2013 - The Physical Science Basis*, Intergovernmental Panel on Climate Change, Ed., Cambridge University Press, Cambridge, 571–658, URL https://doi.org/10.1017/CBO9781107415324.016.

Boucher, O. and U. Lohmann, 1995: The sulfate-CCN-cloud albedo effect. *Tellus B: Chemical and Physical Meteorology*, **47 (3)**, 281–300, URL https://doi.org/10.3402/tellusb.v47i3.16048.

Bruno, O., C. Hoose, T. Storelvmo, Q. Coopman, and M. Stengel, 2021: Exploring the cloud top phase partitioning in different cloud types using active and passive satellite sensors. *Geophysical Research Letters*, **48 (2)**, e2020GL089 863, URL https://doi.org/10.1029/2020GL089863.

Bureau, R., 1946: Measures en météorologie. *Meteorologie*, **3**, 292.

Caldwell, P. M., C. R. Terai, B. Hillman, N. D. Keen, P. Bogenschutz, W. Lin, H. Beydoun, M. Taylor, L. Bertagna, A. M. Bradley, T. C. Clevenger, A. S. Donahue, C. Eldred, J. Foucar, J.-C. Golaz, O. Guba, R. Jacob, J. Johnson, J. Krishna, W. Liu, K. Pressel, A. G. Salinger, B. Singh, A. Steyer, P. Ullrich, D. Wu, X. Yuan, J. Shpund, H.-Y. Ma, and C. S. Zender, 2021: Convection-permitting simulations with the e3sm global atmosphere model. *Journal of Advances in Modeling Earth Systems*, **13 (11)**, e2021MS002 544, URL https://doi.org/10.1029/2021MS002544, e2021MS0025.

Carey, L. D., J. Niu, P. Yang, J. A. Kankiewicz, V. E. Larson, and T. H. V. Haar, 2008: The vertical profile of liquid and ice water content in midlatitude mixed-phase altocumulus clouds. *Journal*

of Applied Meteorology and Climatology, **47 (9)**, 2487 – 2495, URL https://doi.org/10.1175/2008JAMC1885.1.

Cesana, G., A. D. Del Genio, and H. Chepfer, 2019: The Cumulus And Stratocumulus CloudSat-CALIPSO Dataset (CASCCAD). *Earth System Science Data Discussions*, **11**, 1–33, URL https://doi.org/10.5194/essd-2019-73.

Cesana, G. and T. Storelvmo, 2017: Improving climate projections by understanding how cloud phase affects radiation. *Journal of Geophysical Research*, **122 (8)**, 4594–4599, URL https://doi.org/10.1002/2017JD026927.

Cesana, G., D. E. Waliser, X. Jiang, and J.-L. F. Li, 2015: Multimodel evaluation of cloud phase transition using satellite and reanalysis data. *Journal of Geophysical Research: Atmospheres*, **120 (15)**, 7871–7892, URL https://doi.org/10.1002/2014JD022932.

Chandrasekhar, S., 1960: *Radiative transfer*.

Chen, T., J. Guo, Z. Li, C. Zhao, H. Liu, M. Cribb, F. Wang, and J. He, 2016: A cloudsat perspective on the cloud climatology and its association with aerosol perturbations in the vertical over eastern china. *Journal of the Atmospheric Sciences*, **73 (9)**, 3599 – 3616, URL https://doi.org/10.1175/JAS-D-15-0309.1.

Coopman, Q., C. Hoose, and M. Stengel, 2019: Detection of Mixed-Phase Convective Clouds by a Binary Phase Information From the Passive Geostationary Instrument SEVIRI. *Journal of Geophysical Research: Atmospheres*, **124 (9)**, 5045–5057, URL https://doi.org/10.1029/2018JD029772.

———, 2021: Analyzing the thermodynamic phase partitioning of mixed phase clouds over the southern ocean using passive satellite observations. *Geophysical Research Letters*, **48 (7)**, e2021GL093 225, URL https://doi.org/10.1029/2021GL093225.

Coopman, Q., J. Riedi, S. Zeng, and T. J. Garrett, 2020: Space-based analysis of the cloud thermodynamic phase transition for varying microphysical and meteorological regimes. *Geophysical Research Letters*, **47 (6)**, e2020GL087 122, URL https://doi.org/10.1029/2020GL087122.

Costa, A., J. Meyer, A. Afchine, A. Luebke, G. Günther, J. R. Dorsey, M. W. Gallagher, A. Ehrlich, M. Wendisch, D. Baumgardner, H. Wex, and M. Krämer, 2017: Classification of Arctic, midlatitude and tropical clouds in the mixed-phase temperature regime. *Atmospheric Chemistry and Physics*, **17 (19)**, 12 219–12 238, URL https://doi.org/10.5194/acp-17-12219-2017.

Costa-Surós, M., O. Sourdeval, C. Acquistapace, H. Baars, C. Carbajal Henken, C. Genz, J. Hesemann, C. Jimenez, M. König, J. Kretzschmar, N. Madenach, C. I. Meyer, R. Schrödner, P. Seifert, F. Senf, M. Brueck, G. Cioni, J. F. Engels, K. Fieg, K. Gorges, R. Heinze, P. K. Siligam,

U. Burkhardt, S. Crewell, C. Hoose, A. Seifert, I. Tegen, and J. Quaas, 2020: Detection and attribution of aerosol–cloud interactions in large-domain large-eddy simulations with the icosahedral non-hydrostatic model. *Atmospheric Chemistry and Physics*, **20 (9)**, 5657–5678, URL https://doi.org/10.5194/acp-20-5657-2020.

Cotton, W. R. and R. A. Anthes, 1992: *Storm and Cloud Dynamics*. Academic Press New York, 883 pp.

Cubasch, U., D. Wuebbles, D. Chen, M. Facchini, D. Frame, N. Mahowald, and J.-G. Winther, 2013: *Introduction*, chap. 1, 119—158. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, URL https://10.1017/CBO9781107415324.007.

Curry, J. A., P. V. Hobbs, M. D. King, D. A. Randall, P. Minnis, G. A. Isaac, J. O. Pinto, T. Uttal, A. Bucholtz, D. G. Cripe, H. Gerber, C. W. Fairall, T. J. Garrett, J. Hudson, J. M. Intrieri, C. Jakob, T. Jensen, P. Lawson, D. Marcotte, L. Nguyen, P. Pilewskie, A. Rangno, D. C. Rogers, K. B. Strawbridge, F. P. J. Valero, A. G. Williams, and D. Wylie, 2000: Fire arctic clouds experiment. *Bulletin of the American Meteorological Society*, **81 (1)**, 5 – 30, URL https://doi.org/10.1175/1520-0477(2000)081<0005:FACE>2.3.CO;2.

Dahiru, T., 2008: P-value, a true test of statistical significance? a cautionary note. *Annals of Ibadan postgraduate medicine*, **6**, 21–26, URL https://doi.org/10.4314/aipm.v6i1.64038.

Danabasoglu, G., J.-F. Lamarque, J. Bacmeister, D. A. Bailey, A. K. DuVivier, J. Edwards, L. K. Emmons, J. Fasullo, R. Garcia, A. Gettelman, C. Hannay, M. M. Holland, W. G. Large, P. H. Lauritzen, D. M. Lawrence, J. T. M. Lenaerts, K. Lindsay, W. H. Lipscomb, M. J. Mills, R. Neale, K. W. Oleson, B. Otto-Bliesner, A. S. Phillips, W. Sacks, S. Tilmes, L. van Kampenhout, M. Vertenstein, A. Bertini, J. Dennis, C. Deser, C. Fischer, B. Fox-Kemper, J. E. Kay, D. Kinnison, P. J. Kushner, V. E. Larson, M. C. Long, S. Mickelson, J. K. Moore, E. Nienhouse, L. Polvani, P. J. Rasch, and W. G. Strand, 2020: The community earth system model version 2 (cesm2). *Journal of Advances in Modeling Earth Systems*, **12 (2)**, e2019MS001 916, URL https://doi.org/10.1029/2019MS001916.

Danker, J., O. Sourdeval, I. L. McCoy, R. Wood, and A. Possner, 2022: Exploring relations between cloud morphology, cloud phase, and cloud radiative properties in southern ocean's stratocumulus clouds. *Atmospheric Chemistry and Physics*, **22 (15)**, 10 247–10 265, URL https://doi.org/10.5194/acp-22-10247-2022.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, 2011: The era-interim

reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137 (656)**, 553–597, URL https://doi.org/10.1002/qj.828.

Duras, J., F. Ziemen, and D. Klocke, 2021: The dyamond winter data collection. URL https://doi.org/10.5194/egusphere-egu21-4687.

D'Alessandro, J. J., M. Diao, C. Wu, X. Liu, J. B. Jensen, and B. B. Stephens, 2019: Cloud phase and relative humidity distributions over the southern ocean in austral summer based on in situ observations and cam5 simulations. *Journal of Climate*, **32 (10)**, 2781 – 2805, URL https://doi.org/10.1175/JCLI-D-18-0232.1.

Efremenko, D. and A. Kokhanovsky, 2021: *Foundations of Atmospheric Remote Sensing*. Springer International Publishing, URL https://books.google.de/books?id=iT8vEAAAQBAJ.

Ewald, F., T. Zinner, T. Kölling, and B. Mayer, 2019: Remote sensing of cloud droplet radius profiles using solar reflectance from cloud sides – part 1: Retrieval development and characterization. *Atmospheric Measurement Techniques*, **12 (2)**, 1183–1206, URL https://doi.org/10.5194/amt-12-1183-2019.

Findeisen, W., 1931: *Messungen der Größe und Anzahl der Nebeltropfen zum Studium der Koagulation-inhomogenen Nebels*. Hamburgische Universitat.

Fisher, R., 1925: *Statistical Methods for Research Workers*. Biological monographs and manuals, Oliver and Boyd, URL https://books.google.de/books?id=I0NBAAAAIAAJ.

Forbes, R. M. and M. Ahlgrimm, 2014: On the Representation of High-Latitude Boundary Layer Mixed-Phase Cloud in the ECMWF Global Model. *Monthly Weather Review*, **142 (9)**, 3425–3445, URL https://doi.org/10.1175/MWR-D-13-00325.1.

Fränti, P. and S. Sieranoja, 2019: How much k-means can be improved by using better initialization and repeats? *Pattern Recognition*, **93**, URL https://doi.org/10.1016/j.patcog.2019.04.014.

Fu, D., L. Di Girolamo, R. M. Rauber, G. M. McFarquhar, S. W. Nesbitt, J. Loveridge, Y. Hong, B. van Diedenhoven, B. Cairns, M. D. Alexandrov, P. Lawson, S. Woods, S. Tanelli, S. Schmidt, C. Hostetler, and A. J. Scarino, 2022: An evaluation of the liquid cloud droplet effective radius derived from modis, airborne remote sensing, and in situ measurements from camp^{2}ex. *Atmospheric Chemistry and Physics*, **22 (12)**, 8259–8285, URL https://doi.org/10.5194/acp-22-8259-2022.

Fu, Y., Y. Lin, G. Liu, and Q. Wang, 2003: Seasonal characteristics of precipitation in 1998 over East Asia as derived from TRMM PR. *Advances in Atmospheric Sciences*, **20 (4)**, 511–529, URL https://doi.org/10.1007/BF02915495.

Garcia, R. R., 2003: Meteorology, dynamic (stratosphere). *Encyclopedia of Physical Science and Technology (Third Edition)*, Meyers, R. A., Ed., Academic Press, New York, third edition ed., 603–627, URL https://doi.org/10.1016/B0-12-227410-5/00919-4.

Gettelman, A. and H. Morrison, 2015: Advanced two-moment bulk microphysics for global models. part i: Off-line tests and comparison with other schemes. *Journal of Climate*, **28 (3)**, 1268 – 1287, URL https://doi.org/10.1175/JCLI-D-14-00102.1.

Gierens, R., S. Kneifel, M. D. Shupe, K. Ebell, M. Maturilli, and U. Löhnert, 2020: Low-level mixed-phase clouds in a complex arctic environment. *Atmospheric Chemistry and Physics*, **20 (6)**, 3459–3481, URL https://doi.org/10.5194/acp-20-3459-2020.

Grosvenor, D. P. and R. Wood, 2014: The effect of solar zenith angle on modis cloud optical and microphysical retrievals within marine liquid water clouds. *Atmospheric Chemistry and Physics*, **14 (14)**, 7291–7321, URL https://doi.org/10.5194/acp-14-7291-2014.

Guzman, R., H. Chepfer, V. Noel, T. Vaillant de Guélis, J. E. Kay, P. Raberanto, G. Cesana, M. A. Vaughan, and D. M. Winker, 2017: Direct atmosphere opacity observations from calipso provide new constraints on cloud-radiation interactions. *Journal of Geophysical Research: Atmospheres*, **122 (2)**, 1066–1085, URL https://doi.org/10.1002/2016JD025946.

Hahn, C. J., W. B. Rossow, and S. G. Warren, 2001: ISCCP Cloud Properties Associated with Standard Cloud Types Identified in Individual Surface Observations. *Journal of Climate*, **14 (1)**, 11–28, URL https://doi.org/10.1175/1520-0442(2001)014<0011:ICPAWS>2.0.CO;2.

Hamilton, D. S., L. A. Lee, K. J. Pringle, C. L. Reddington, D. V. Spracklen, and K. S. Carslaw, 2014: Occurrence of pristine aerosol environments on a polluted planet. *Proceedings of the National Academy of Sciences*, **111 (52)**, 18 466–18 471, URL https://doi.org/10.1073/pnas.1415440111.

Henneberger, J., J. P. Fugal, O. Stetzer, and U. Lohmann, 2013: HOLIMO II: A digital holographic instrument for ground-based in situ observations of microphysical properties of mixed-phase clouds. *Atmospheric Measurement Techniques*, **6 (11)**, 2975–2987, URL https://doi.org/10.5194/amt-6-2975-2013.

Herman, B. M. and S. R. Browning, 1965: A numerical solution to the equation of radiative transfer. *Journal of Atmospheric Sciences*, **22 (5)**, 559 – 566, URL https://doi.org/10.1175/1520-0469(1965)022<0559:ANSTTE>2.0.CO;2.

Heymsfield, A. J., 1977: Precipitation development in stratiform ice clouds: A microphysical and dynamical study. *Journal of Atmospheric Sciences*, **34 (2)**, 367 – 381, URL https://doi.org/10.1175/1520-0469(1977)034<0367:PDISIC>2.0.CO;2.

Hobbs, P. V. and A. L. Rangno, 1998: Microstructures of low and middle-level clouds over the beaufort sea. *Quarterly Journal of the Royal Meteorological Society*, **124 (550)**, 2035–2071, URL https://doi.org/10.1002/qj.49712455012.

Hofer, S., L. Hahn, J. Shaw, Z. McGraw, O. Bruno, F. Hellmuth, M. Pietschnig, I. Mostue, R. David, T. Carlsen, and T. Storelvmo, 2023: Realistic representation of mixed-phase clouds increases future climate warming. *Nature Communications Earth and Environment*, preprint, URL https://doi.org/10.21203/rs.3.rs-2981113/v1.

Hoose, C., M. Karrer, and C. Barthlott, 2018: Cloud top phase distributions of simulated deep convective clouds. *Journal of Geophysical Research: Atmospheres*, **123 (18)**, 10,464–10,476, URL https://doi.org/10.1029/2018JD028381.

Hoose, C. and O. Möhler, 2012: Heterogeneous ice nucleation on atmospheric aerosols: A review of results from laboratory experiments. *Atmospheric Chemistry and Physics*, **12 (20)**, 9817–9854, URL https://doi.org/10.5194/acp-12-9817-2012.

Hostetler, C. A., Z. Liu, J. Reagan, M. Vaughan, D. Winker, M. Osborn, W. H. Hunt, K. A. Powell, and C. Trepte, 2006: CALIOP algorithm theoretical basis document: Calibration and level 1 data products. URL https://www-calipso.larc.nasa.gov/resources/pdfs/PC-SCI-201v1.0.pdf.

Houze, R., 2014: *Cloud Dynamics - Second Edition*.

Howard, L., 1803a: I. on the modifications of clouds, and on the principles of their production, suspension, and destruction; being the substance of an essay read before the askesian society in the session 1802–3. *The Philosophical Magazine*, **17 (65)**, 5–11, URL https://doi.org/10.1080/14786440308676365.

———, 1803b: Lxiv. on the modifications of clouds, and on the principles of their production, suspension, and destruction; being the substance of an essay read before the askesian society in the session 1802–3. *The Philosophical Magazine*, **16 (64)**, 344–357, URL https://doi.org/10.1080/14786440308676358.

———, 1803c: Xviii. on the modifications of clouds, and on the principles of their production, suspension, and destruction; being the substance of an essay read before the askesian society in the session 1802–3. *The Philosophical Magazine*, **16 (62)**, 97–107, URL https://doi.org/10.1080/14786440308676310.

Hu, Y., S. Rodier, K.-m. Xu, W. Sun, J. Huang, B. Lin, P. Zhai, and D. Josset, 2010: Occurrence, liquid water content, and fraction of supercooled water clouds from combined caliop/iir/modis measurements. *Journal of Geophysical Research: Atmospheres*, **115 (D4)**, URL https://doi.org/10.1029/2009JD012384.

Hu, Y., D. Winker, M. Vaughan, B. Lin, A. Omar, C. Trepte, D. Flittner, P. Yang, S. L. Nasiri, B. Baum, R. Holz, W. Sun, Z. Liu, Z. Wang, S. Young, K. Stamnes, J. Huang, and R. Kuehn, 2009: CALIPSO/CALIOP Cloud Phase Discrimination Algorithm. *Journal of Atmospheric and Oceanic Technology*, **26 (11)**, 2293–2309, URL https://doi.org/10.1175/2009JTECHA1280.1.

Hussain, S. and A. Haji-Akbari, 2021: Role of nanoscale interfacial proximity in contact freezing in water. *Journal of the American Chemical Society*, **143 (5)**, 2272–2284, URL https://doi.org/10.1021/jacs.0c10663, pMID: 33507741.

Jackson, J., 1998: *Classical Electrodynamics*.

Jakob, C., 2010: Accelerating progress in global atmospheric model development through improved parameterizations: Challenges, opportunities, and strategies. *Bulletin of the American Meteorological Society*, **91 (7)**, 869 – 876, URL https://doi.org/10.1175/2009BAMS2898.1.

Jakob, C. and G. Tselioudis, 2003: Objective identification of cloud regimes in the Tropical Western Pacific. *Geophysical Research Letters*, **30 (21)**, 1–4, URL https://doi.org/10.1029/2003GL018367.

Jia, H., X. Ma, J. Quaas, Y. Yin, and T. Qiu, 2019: Is positive correlation between cloud droplet effective radius and aerosol optical depth over land due to retrieval artifacts or real physical processes? *Atmospheric Chemistry and Physics*, **19 (13)**, 8879–8896, URL https://doi.org/10.5194/acp-19-8879-2019.

Karlsson, K. G., K. Anttila, J. Trentmann, M. Stengel, J. Fokke Meirink, A. Devasthale, T. Hanschmann, S. Kothe, E. Jaäskelaïnen, J. Sedlar, N. Benas, G. J. Van Zadelhoff, C. Schlundt, D. Stein, S. Finkensieper, N. Häkansson, and R. Hollmann, 2017: CLARA-A2: The second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data. *Atmospheric Chemistry and Physics*, **17 (9)**, 5809–5828, URL https://doi.org/10.5194/acp-17-5809-2017.

Karlsson, K.-G. and N. Håkansson, 2018: Characterization of avhrr global cloud detection sensitivity based on calipso-caliop cloud optical thickness information: demonstration of results based on the cm saf clara-a2 climate data record. *Atmospheric Measurement Techniques*, **11 (1)**, 633–649, URL https://doi.org/10.5194/amt-11-633-2018.

Karlsson, K.-G., E. Johansson, and A. Devasthale, 2015: Advancing the uncertainty characterisation of cloud masking in passive satellite imagery: Probabilistic formulations for noaa avhrr data. *Remote Sensing of Environment*, **158**, 126–139, URL https://doi.org/10.1016/j.rse.2014.10.028.

Kelkar, R., 2007: *Satellite Meteorology*. BS Publications, URL https://books.google.de/books?id=htN3oAEACAAJ.

King, M. D., S. Platnick, W. P. Menzel, S. A. Ackerman, and P. A. Hubanks, 2013: Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites. *IEEE Transactions on Geoscience and Remote Sensing*, **51 (7)**, 3826–3852, URL https://doi.org/10.1109/TGRS.2012.2227333.

King, M. D., S. Platnick, P. Yang, G. T. Arnold, M. A. Gray, J. C. Riedi, S. A. Ackerman, and K. N. Liou, 2004: Remote sensing of liquid water and ice cloud optical thickness and effective radius in the Arctic: Application of airborne multispectral MAS data. *Journal of Atmospheric and Oceanic Technology*, **21 (6)**, 857–875, URL https://doi.org/10.1175/1520-0426(2004)021<0857:RSOLWA>2.0.CO;2.

Kirkevåg, A., A. Grini, D. Olivié, Ø. Seland, K. Alterskjær, M. Hummel, I. H. H. Karset, A. Lewinschal, X. Liu, R. Makkonen, I. Bethke, J. Griesfeller, M. Schulz, and T. Iversen, 2018: A production-tagged aerosol module for earth system models, osloaero5.3 – extensions and updates for cam5.3-oslo. *Geoscientific Model Development*, **11 (10)**, 3945–3982, URL https://doi.org/10.5194/gmd-11-3945-2018.

Kirkevåg, A., T. Iversen, Ø. Seland, C. Hoose, J. E. Kristjánsson, H. Struthers, A. M. L. Ekman, S. Ghan, J. Griesfeller, E. D. Nilsson, and M. Schulz, 2013: Aerosol–climate interactions in the norwegian earth system model – noresm1-m. *Geoscientific Model Development*, **6 (1)**, 207–244, URL https://doi.org/10.5194/gmd-6-207-2013.

Klaes, K., 2018: 1.07 - the eumetsat polar system. *Comprehensive Remote Sensing*, Liang, S., Ed., Elsevier, Oxford, 192–219, URL https://doi.org/10.1016/B978-0-12-409548-9.10318-5.

Komurcu, M. and T. Storelvmo, 2014: Intercomparison of the cloud water phase among global climate models. *Journal of . . .*, URL https://doi.org/10.1002/2013JD021119.Received.

Korolev, A., P. J. DeMott, I. Heckman, M. Wolde, E. Williams, D. J. Smalley, and M. F. Donovan, 2022: Observation of secondary ice production in clouds at low temperatures. *Atmospheric Chemistry and Physics*, **22 (19)**, 13 103–13 113, URL https://doi.org/10.5194/acp-22-13103-2022.

Korolev, A. and T. Leisner, 2020: Review of experimental studies of secondary ice production. *Atmospheric Chemistry and Physics*, **20 (20)**, 11 767–11 797, URL https://acp.copernicus.org/articles/20/11767/2020/.

Korolev, A. V., 2007: Limitations of the Wegener–Bergeron–Findeisen Mechanism in the Evolution of Mixed-Phase Clouds. *Journal of the Atmospheric Sciences*, **64 (9)**, 3372–3375, URL https://doi.org/10.1175/JAS4035.1.

———, 2008: Rates of phase transformations in mixed-phase clouds. *Quarterly Journal of the Royal Meteorological Society*, **134 (632)**, 595–608, URL https://doi.org/10.1002/qj.230.

Korolev, A. V. and I. P. Mazin, 2003: Supersaturation of water vapor in clouds. *Journal of the Atmospheric Sciences*, **60 (24)**, 2957 – 2974, URL https://doi.org/10.1175/1520-0469(2003)060<2957:SOWVIC>2.0.CO;2.

Korolev, A. V., G. McFarquhar, P. R. Field, C. Franklin, P. Lawson, Z. Wang, E. Williams, S. J. Abel, D. Axisa, S. Borrmann, J. Crosier, J. Fugal, M. Krämer, U. Lohmann, O. Schlenczek, M. Schnaiter, and M. Wendisch, 2017: Mixed-Phase Clouds: Progress and Challenges. *Meteorological Monographs*, **58**, URL https://doi.org/10.1175/AMSMONOGRAPHS-D-17-0001.1.

Krishnananda, 1984: *The Chhandogya Upanishad*. Divine Life Society, Shivanandanagar, Distt. Tehri-Garhwal, U.P., Himalayas, India.

Lamb, D. and J. Verlinde, 2011: *Physics and Chemistry of Clouds*.

Libbrecht, K. G., 2005: The physics of snow crystals. *Reports on Progress in Physics*, **68 (4)**, 855, URL https://dx.doi.org/10.1088/0034-4885/68/4/R03.

Liu, F., F. Mao, D. Rosenfeld, Z. Pan, L. Zang, Y. Zhu, J. Yin, and W. Gong, 2022: Opposing comparable large effects of fine aerosols and coarse sea spray on marine warm clouds. *Communications Earth Environment*, **3**, URL https://doi.org/10.1038/s43247-022-00562-y.

Liu, X., P.-L. Ma, H. Wang, S. Tilmes, B. Singh, R. C. Easter, S. J. Ghan, and P. J. Rasch, 2016: Description and evaluation of a new four-mode version of the modal aerosol module (mam4) within version 5.3 of the community atmosphere model. *Geoscientific Model Development*, **9 (2)**, 505–522, URL https://doi.org/10.5194/gmd-9-505-2016.

Liu, X., J. E. Penner, and M. Wang, 2009: Influence of anthropogenic sulfate and black carbon on upper tropospheric clouds in the NCAR CAM3 model coupled to the IMPACT global aerosol model. *Journal of Geophysical Research: Atmospheres*, **114 (D3)**, URL https://doi.org/10.1029/2008JD010492.

Liu, Z., J. Kar, S. Zeng, J. Tackett, M. Vaughan, M. Avery, J. Pelon, B. Getzewich, K. P. Lee, B. Magill, A. Omar, P. Lucker, C. Trepte, and D. Winker, 2019: Discriminating between clouds and aerosols in the CALIOP version 4.1 data products. *Atmospheric Measurement Techniques*, **12 (1)**, 703–734, URL https://doi.org/10.5194/amt-12-703-2019.

Marchant, B., S. Platnick, K. Meyer, G. T. Arnold, and J. Riedi, 2016: Modis collection 6 shortwave-derived cloud phase classification algorithm and comparisons with caliop. *Atmospheric Measurement Techniques*, **9 (4)**, 1587–1599, URL https://doi.org/10.5194/amt-9-1587-2016.

Martins, J. V., A. Marshak, L. A. Remer, D. Rosenfeld, Y. J. Kaufman, R. Fernandez-Borda, I. Koren, A. L. Correia, V. Zubko, and P. Artaxo, 2011: Remote sensing the vertical profile of cloud droplet

effective radius, thermodynamic phase, and temperature. *Atmospheric Chemistry and Physics*, **11 (18)**, 9485–9501, URL https://doi.org/10.5194/acp-11-9485-2011.

Mason, B., 1957: *The Physics of Clouds*.

Mauritsen, T., B. Stevens, E. Roeckner, T. Crueger, M. Esch, M. Giorgetta, H. Haak, J. Jungclaus, D. Klocke, D. Matei, U. Mikolajewicz, D. Notz, R. Pincus, H. Schmidt, and L. Tomassini, 2012: Tuning the climate of a global model. *Journal of Advances in Modeling Earth Systems*, **4 (3)**, URL https://doi.org/10.1029/2012MS000154.

Mazin, I. P., 2006: Cloud phase structure: Experimental data analysis and parameterization. *Journal of the Atmospheric Sciences*, **63 (2)**, 667 – 681, URL https://doi.org/10.1175/JAS3660.1.

McCoy, D. T., I. Tan, D. L. Hartmann, M. D. Zelinka, and T. Storelvmo, 2016: On the relationships among cloud cover, mixed-phase partitioning, and planetary albedo in GCMs. *Journal of Advances in Modeling Earth Systems*, **8 (2)**, 650–668, URL https://doi.org/10.1002/2015MS000589.

McIlhattan, E. A., T. S. L'Ecuyer, and N. B. Miller, 2017: Observational evidence linking arctic supercooled liquid cloud biases in cesm to snowfall processes. *Journal of Climate*, **30 (12)**, 4477 – 4495, URL https://doi.org/10.1175/JCLI-D-16-0666.1.

Mignani, C., J. M. Creamean, L. Zimmermann, C. Alewell, and F. Conen, 2019: New type of evidence for secondary ice formation at around -15 c in mixed-phase clouds. *Atmospheric Chemistry and Physics*, **19 (2)**, 877–886, URL https://doi.org/10.5194/acp-19-877-2019.

Minikin, A., A. Petzold, J. Ström, R. Krejci, M. Seifert, P. Velthoven, H. Schlager, and U. Schumann, 2003: Aircraft observations of the upper tropospheric fine particle aerosol in the northern and southern hemispheres at midlatitudes. *Geophys. Res. Lett*, **30**, URL https://doi.org/10.1029/2002GL016458.

Mittaz, J., C. Merchant, and E. Woolliams, 2019: Applying principles of metrology to historical earth observations from satellites. *Metrologia*, **56**, 032 002, URL https://doi.org/10.1088/1681-7575/ab1705.

Morrison, H. and J. A. Milbrandt, 2015: Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. part i: Scheme description and idealized tests. *Journal of the Atmospheric Sciences*, **72 (1)**, 287 – 311, URL https://doi.org/10.1175/JAS-D-14-0065.1.

Mossop, S. C., 1980: The mechanism of ice splinter production during riming. *Geophysical Research Letters*, **7 (2)**, 167–169, URL https://doi.org/10.1029/GL007i002p00167.

Mossop, S. C. and J. Hallett, 1974: Ice crystal concentration in cumulus clouds: Influence of the drop spectrum. *Science*, **186 (4164)**, 632–634, URL https://doi.org/10.1126/science.186.4164.632.

Myhre, G., D. Shindell, F.-M. Breon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura, and H. Zhang, 2013: *Anthropogenic and Natural Radiative Forcing*, chap. 8, 659—740. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, URL https://doi.org/10.1017/CBO9781107415324.018.

Nagao, T. M. and K. Suzuki, 2022: Characterizing vertical stratification of the cloud thermodynamic phase with a combined use of calipso lidar and modis swir measurements. *Journal of Geophysical Research: Atmospheres*, **127 (21)**, e2022JD036 826, URL https://doi.org/10.1029/2022JD036826, e2022JD036826 2022JD036826.

Nakajima, T. and M. D. King, 1990: Determination of the Optical Thickness and Effective Particle Radius of Clouds from Reflected Solar Radiation Measurements. Part I: Theory. URL https://doi.org/10.1175/1520-0469(1990)047<1878:DOTOTA>2.0.CO;2.

NASA, 2022: *MCD06COSP_M3_MODIS - MODIS (Aqua/Terra) Cloud Properties Level 3 Monthly, 1x1 Degree Grid*. URL http://doi.org/10.5067/MODIS/MCD06COSP_D3_MODIS.062, version 2, date of issue: 27/10/2022.

Nishihama, M., R. Wolfe, D. Solomon, F. Patt, J. Blanchette, A. Fleig, and E. Masuoka, 1997: MODIS Level IA Earth Location: Algorithm Theoretical Basis document. URL https://modis.gsfc.nasa.gov/data/atbd/.

Noh, Y.-J., S. D. Miller, A. K. Heidinger, G. G. Mace, A. Protat, and S. P. Alexander, 2019: Satellite-based detection of daytime supercooled liquid-topped mixed-phase clouds over the southern ocean using the advanced himawari imager. *Journal of Geophysical Research: Atmospheres*, **124 (5)**, 2677–2701, URL https://doi.org/10.1029/2018JD029524.

Norris, J. R. and C. P. Weaver, 2001: Improved techniques for evaluating gcm cloudiness applied to the ncar ccm3. *Journal of Climate*, **14 (12)**, 2540 – 2550, URL https://doi.org/10.1175/1520-0442(2001)014<2540:ITFEGC>2.0.CO;2.

Oreopoulos, L., N. Cho, and D. Lee, 2017: Using MODIS cloud regimes to sort diagnostic signals of aerosol-cloud-precipitation interactions. *Journal of Geophysical Research Atmospheres*, **122**, 5416–5440, URL https://doi.org/10.1002/2016JD026120.

Ostwald, W., 1897: Studien über die bildung und umwandlung fester körper. *Zeitschrift für Physikalische Chemie*, **22U (1)**, 289–330, URL https://doi.org/10.1515/zpch-1897-2233.

Pavolonis, M. J. and A. K. Heidinger, 2004: Daytime Cloud Overlap Detection from AVHRR and VIIRS. *Journal of Applied Meteorology*, **43**, 762–778, URL https://doi.org/10.1175/2099.1.

Pavolonis, M. J., A. K. Heidinger, and T. Uttal, 2005: Daytime Global Cloud Typing from AVHRR and VIIRS: Algorithm Description, Validation, and Comparisons. *Journal of Applied Meteorology*, **44 (6)**, 804–826, URL https://doi.org/10.1175/JAM2236.1.

Petty, G. D., 2004: *A First Course In Atmospheric Radiation*.

Pincus, R., P. A. Hubanks, S. A. Platnick, K. Meyer, R. E. Holz, D. Botambekov, and C. J. Wall, 2022: Updated observations of clouds by modis for global model assessment. *Earth System Science Data Discussions*, **pre-print**, 1–20, URL https://doi.org/10.5194/essd-2022-282.

Pincus, R., S. Platnick, S. A. Ackerman, R. S. Hemler, and R. J. P. Hofmann, 2012: Reconciling simulated and observed views of clouds: Modis, isccp, and the limits of instrument simulators. *Journal of Climate*, **25 (13)**, 4699 – 4720, URL https://doi.org/10.1175/JCLI-D-11-00267.1.

Pinto, J. O., 1998: Autumnal mixed-phase cloudy boundary layers in the arctic. *Journal of the Atmospheric Sciences*, **55 (11)**, 2016 – 2038, URL https://doi.org/10.1175/1520-0469(1998)055<2016:AMPCBL>2.0.CO;2.

Platt, C. M. R., 1977: Lidar observation of a mixed-phase altostratus cloud. *Journal of Applied Meteorology (1962-1982)*, **16 (4)**, 339–345, URL http://www.jstor.org/stable/26178122.

Plummer, D. M., G. M. Mcfarquhar, R. M. Rauber, B. F. Jewett, and D. C. Leon, 2014: Structure and statistical analysis of the microphysical properties of generating cells in the comma head region of continental winter cyclones. *Journal of the Atmospheric Sciences*, **71 (11)**, 4181–4203, URL https://doi.org/10.1175/JAS-D-14-0100.1.

Prenni, A. J., P. J. Demott, D. C. Rogers, S. M. Kreidenweis, G. M. McFarquhar, G. Zhang, and M. R. Poellot, 2009: Ice nuclei characteristics from m-pace and their relation to ice formation in clouds. *Tellus B*, **61 (2)**, 436–448, URL https://doi.org/10.1111/j.1600-0889.2009.00415.x.

Pruppacher, H. and J. Klett, 1997: *Microphysics of Clouds and Precipitation*.

Putman, W. M. and M. Suarez, 2011: Cloud-system resolving simulations with the nasa goddard earth observing system global atmospheric model (geos-5). *Geophysical Research Letters*, **38 (16)**, URL https://doi.org/10.1029/2011GL048438.

PVIR, 2020: *Product Validation and Intercomparison Report (PVIR) - ESA Cloud_cci*. URL https://climate.esa.int/media/documents/Cloud_Product-Validation-and-Intercomparison-Report-PVIR_v6.0.pdf, issue 0, Revision: 0, date of issue: 03/02/2020.

Pörtge, V., T. Kölling, A. Weber, L. Volkmer, C. Emde, T. Zinner, L. Forster, and B. Mayer, 2023: High-spatial-resolution retrieval of cloud droplet size distribution from polarized observations of the cloudbow. *Atmospheric Measurement Techniques*, **16 (3)**, 645–667, URL https://doi.org/10.5194/amt-16-645-2023.

Qin, F., X. Tao, and Y. Fu, 2018: Cloud-precipitation parameters and radiative forcing of warm precipitating cloud over the tropical pacific ocean based on trmm datasets and radiative transfer model. *Atmosphere*, **9**, 206, URL https://doi.org/10.3390/atmos9060206.

Rangno, A., 2003: Clouds | classification. *Encyclopedia of Atmospheric Sciences*, Holton, J. R., Ed., Academic Press, Oxford, 467–475, URL https://doi.org/10.1016/B0-12-227090-8/00112-3.

Rangno, A., 2015: Clouds and fog | classification of clouds. *Encyclopedia of Atmospheric Sciences (Second Edition)*, North, G. R., J. Pyle, and F. Zhang, Eds., Academic Press, Oxford, second edition ed., 141–160, URL https://doi.org/10.1016/B978-0-12-382225-3.00112-2.

Reagan, J. A., X. Wang, and M. T. Osborn, 2002: Spaceborne lidar calibration from cirrus and molecular backscatter returns. *IEEE Trans. Geosci. Remote. Sens.*, **40 (10)**, 2285–2290, URL https://doi.org/10.1109/TGRS.2002.802464.

Reinert, D., F. Prill, H. Frank, M. Denhard, M. Baldauf, C. Schraff, C. Gebhardt, C. Marsigli, and G. Zängl, 2023: *DWD Database Reference for the Global and Regional ICON and ICON-EPS Forecasting System*. URL https://www.dwd.de/SharedDocs/downloads/DE/modelldokumentationen/nwv/icon/icon_dbbeschr_aktuell.pdf?view=nasPublication&nn=13934, version 2.2.1, date of issue: 2023.

Riley, E. M. and B. E. Mapes, 2009: Unexpected peak near 15°c in cloudsat echo top climatology. *Geophysical Research Letters*, **36 (9)**, URL https://doi.org/10.1029/2009GL037558.

Ringer, M. A. and R. P. Allan, 2004: Evaluating climate model simulations of tropical cloud. *Tellus A*, **56 (4)**, 308–327, URL https://doi.org/10.1111/j.1600-0870.2004.00061.x.

Rossow, W. B. and R. A. Schiffer, 1999: Advances in understanding clouds from isccp. *Bulletin of the American Meteorological Society*, **80 (11)**, 2261 – 2288, URL https://doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2.

Rossow, W. B., G. Tselioudis, A. Polak, and C. Jakob, 2005: Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud property mixtures. *Geophysical Research Letters*, **32 (21)**, URL https://doi.org/10.1029/2005GL024584.

Sassen, K., 1992: Evidence for liquid-phase cirrus cloud formation from volcanic aerosols: Climatic implications. *Science*, **257 (5069)**, 516–519, URL https://doi.org/10.1126/science.257.5069.516.

Schuddeboom, A., A. J. McDonald, O. Morgenstern, M. Harvey, and S. Parsons, 2018: Regional regime-based evaluation of present-day general circulation model cloud simulations using self-organizing maps. *Journal of Geophysical Research: Atmospheres*, **123 (8)**, 4259–4272, URL https://doi.org/10.1002/2017JD028196.

Schuster, A., 1905: Radiation through a foggy atmosphere. *Astrophysical Journal*, **21 (1)**, 1–22, URL https://doi.org/10.1086/141186.

Seland, Ø., M. Bentsen, D. Olivié, T. Toniazzo, A. Gjermundsen, L. S. Graff, J. B. Debernard, A. K. Gupta, Y.-C. He, A. Kirkevåg, J. Schwinger, J. Tjiputra, K. S. Aas, I. Bethke, Y. Fan, J. Griesfeller, A. Grini, C. Guo, M. Ilicak, I. H. H. Karset, O. Landgren, J. Liakka, K. O. Moseid, A. Nummelin, C. Spensberger, H. Tang, Z. Zhang, C. Heinze, T. Iversen, and M. Schulz, 2020: Overview of the norwegian earth system model (noresm2) and key climate response of cmip6 deck, historical, and scenario simulations. *Geoscientific Model Development*, **13 (12)**, 6165–6200, URL https://doi.org/10.5194/gmd-13-6165-2020.

Shaw, J., Z. McGraw, O. Bruno, T. Storelvmo, and S. Hofer, 2022: Using satellite observations to evaluate model microphysical representation of arctic mixed-phase clouds. *Geophysical Research Letters*, **49 (3)**, e2021GL096 191, URL https://doi.org/10.1029/2021GL096191, e2021GL096191 2021GL096191.

Shen, C., G. Li, and Y. Dong, 2022: Vertical structures associated with orographic precipitation during warm season in the sichuan basin and its surrounding areas at different altitudes from 8-year gpm dpr observations. *Remote Sensing*, **14**, 4222, URL https://doi.org/10.3390/rs14174222.

Shupe, M. D. and J. M. Intrieri, 2004: Cloud radiative forcing of the arctic surface: The influence of cloud properties, surface albedo, and solar zenith angle. *Journal of Climate*, **17 (3)**, 616 – 628, URL https://doi.org/10.1175/1520-0442(2004)017<0616:CRFOTA>2.0.CO;2.

Shupe, M. D., S. Y. Matrosov, and T. Uttal, 2006: Arctic mixed-phase cloud properties derived from surface-based sensors at SHEBA. *Journal of the Atmospheric Sciences*, **63 (2)**, 697–711, URL https://doi.org/10.1175/JAS3659.1.

Siebesma, A. P., S. Bony, C. Jakob, and B. Stevens, 2020: *Clouds and Climate*.

Silber, I., P. S. McGlynn, J. Y. Harrington, and J. Verlinde, 2021: Habit-dependent vapor growth modulates arctic supercooled water occurrence. *Geophysical Research Letters*, **48 (10)**, e2021GL092 767, URL https://doi.org/10.1029/2021GL092767.

Sivananda, S., 1985: *The Brihadaranyaka Upanishad: Sanskrit Text*. Divine Life Society, Shivanandanagar, Distt. Tehri-Garhwal, U.P., Himalayas, India.

Sokol, A. and T. Storelvmo, 2022: The spatial heterogeneity of cloud phase observed by satellite. *Geophysical Research Letter (Under Review)*, URL https://doi.org/10.1002/essoar.10511313.1.

Stengel, M., S. Mieruch, M. Jerg, K. G. Karlsson, R. Scheirer, B. Maddux, J. F. Meirink, C. Poulsen, R. Siddans, A. Walther, and R. Hollmann, 2015: The Clouds Climate Change Initiative: Assessment of state-of-the-art cloud property retrieval schemes applied to AVHRR heritage measurements. *Remote Sensing of Environment*, **162**, 363–379, URL https://doi.org/10.1016/j.rse.2013.10.035.

Stengel, M., S. Stapelberg, O. Sus, S. Finkensieper, B. Würzler, D. Philipp, R. Hollmann, C. Poulsen, M. Christensen, and G. McGarragh, 2020: Cloud_cci advanced very high resolution radiometer post meridiem (avhrr-pm) dataset version 3: 35-year climatology of global cloud and radiation properties. *Earth System Science Data*, **12 (1)**, 41–60, URL https://doi.org/10.5194/essd-12-41-2020.

Stengel, M., S. Stapelberg, O. Sus, C. Schlundt, C. Poulsen, G. Thomas, M. Christensen, C. C. Henken, R. Preusker, J. Fischer, A. Devasthale, U. Willén, K. G. Karlsson, G. R. McGarragh, S. Proud, A. C. Povey, R. G. Grainger, J. F. Meirink, A. Feofilov, R. Bennartz, J. S. Bojanowski, and R. Hollmann, 2017: Cloud property datasets retrieved from AVHRR, MODIS, AATSR and MERIS in the framework of the Cloud-cci project. *Earth System Science Data*, **9 (2)**, 881–904, URL https://doi.org/10.5194/essd-9-881-2017.

Stephens, G. L., 1994: *Remote Sensing of the Lower Atmosphere: An Introduction.*

Storelvmo, T., I. Tan, and A. V. Korolev, 2015: Cloud Phase Changes Induced by CO2 Warming—a Powerful yet Poorly Constrained Cloud-Climate Feedback. *Current Climate Change Reports*, **1 (4)**, 288–296, URL https://doi.org/10.1007/s40641-015-0026-2.

Swales, D. J., R. Pincus, and A. Bodas-Salcedo, 2018: The cloud feedback model intercomparison project observational simulator package: Version 2. *Geoscientific Model Development*, **11 (1)**, 77–81, URL https://doi.org/10.5194/gmd-11-77-2018.

Tan, I. and T. Storelvmo, 2019: Evidence of strong contributions from mixed-phase clouds to arctic climate change. *Geophysical Research Letters*, **46 (5)**, 2894–2902, URL https://doi.org/10.1029/2018GL081871.

Tan, I., T. Storelvmo, and Y. S. Choi, 2014: Spaceborne lidar observations of the ice-nucleating potential of dust, polluted dust, and smoke aerosols in mixed-phase clouds. *Journal of Geophysical Research*, **119 (11)**, 6653–6665, URL https://doi.org/10.1002/2013JD021333.

Tan, J., L. Oreopoulos, C. Jakob, and D. Jin, 2018: Evaluating rainfall errors in global climate models through cloud regimes. *Climate Dynamics*, **50**, 3301–3314, URL https://doi.org/10.1007/s00382-017-3806-7.

Tselioudis, G., Y. Zhang, and W. B. Rossow, 2000: Cloud and radiation variations associated with northern midlatitude low and high sea level pressure regimes. *Journal of Climate*, **13 (2)**, 312 – 327, URL https://doi.org/10.1175/1520-0442(2000)013<0312:CARVAW>2.0.CO;2.

Vali, G., P. J. DeMott, O. Möhler, and T. F. Whale, 2015: Technical note: A proposal for ice nucleation terminology. *Atmospheric Chemistry and Physics*, **15 (18)**, 10 263–10 270, URL https://doi.org/10.5194/acp-15-10263-2015.

van de Hulst, H. C. and W. M. Irvine, 1963: General Report on Radiation Transfer in Planets Scattering in Model Planetary Atmospheres. **11**, 78–98.

Verlinde, J., J. Y. Harrington, G. M. McFarquhar, V. T. Yannuzzi, A. Avramov, S. Greenberg, N. Johnson, G. Zhang, M. R. Poellot, J. H. Mather, D. D. Turner, E. W. Eloranta, B. D. Zak, A. J. Prenni, J. S. Daniel, G. L. Kok, D. C. Tobin, R. Holz, K. Sassen, D. Spangenberg, P. Minnis, T. P. Tooman, M. D. Ivey, S. J. Richardson, C. P. Bahrmann, M. Shupe, P. J. DeMott, A. J. Heymsfield, and R. Schofield, 2007: The mixed-phase arctic cloud experiment. *Bulletin of the American Meteorological Society*, **88 (2)**, 205 – 222, URL https://doi.org/10.1175/BAMS-88-2-205.

Villanueva, D., B. Heinold, P. Seifert, H. Deneke, M. Radenz, and I. Tegen, 2020: The day-to-day co-variability between mineral dust and cloud glaciation: a proxy for heterogeneous freezing. *Atmospheric Chemistry and Physics*, **20 (4)**, 2177–2199, URL https://doi.org/10.5194/acp-20-2177-2020.

Villanueva, D., F. Senf, and I. Tegen, 2021: Hemispheric and seasonal contrast in cloud thermodynamic phase from a-train spaceborne instruments. *Journal of Geophysical Research: Atmospheres*, **126 (6)**, e2020JD034 322, URL https://doi.org/10.1029/2020JD034322.

Wang, Z., J. French, G. Vali, P. Wechsler, S. Haimov, A. Rodi, M. Deng, D. Leon, J. Snider, L. Peng, and A. L. Pazmany, 2012: Single aircraft integration of remote sensing and in situ sampling for the study of cloud microphysics and dynamics. *Bulletin of the American Meteorological Society*, **93 (5)**, 653–668, URL https://doi.org/10.1175/BAMS-D-11-00044.1.

Wegener, A., 1926: *Thermodynamik der Atmosphäre*, 156–189. Springer Berlin Heidelberg, Berlin, Heidelberg, URL https://doi.org/10.1007/978-3-642-90779-1_3.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*.

Williams, K., M. Ringer, and C. Senior, 2003: Evaluating the cloud response to climate change and current climate variability. *Climate Dynamics*, **20**, 705—-721, URL https://doi.org/10.1007/s00382-002-0303-3.

Williams, K. D., M. A. Ringer, C. A. Senior, M. J. Webb, B. J. McAvaney, N. Andronova, S. Bony, J. L. Dufresne, S. Emori, R. Gudgel, T. Knutson, B. Li, K. Lo, I. Musat, J. Wegner, A. Slingo, and

J. F. B. Mitchell, 2006: Evaluation of a component of the cloud response to climate change in an intercomparison of climate models. *Climate Dynamics*, **26 (2-3)**, 145–165, URL https://doi.org/10.1007/s00382-005-0067-7.

Williams, K. D., C. A. Senior, A. Slingo, and J. F. B. Mitchell, 2005: Towards evaluating cloud response to climate change using clustering technique identification of cloud regimes. *Climate Dynamics*, **24**, 701–719, URL https://doi.org/10.1007/s00382-004-0512-z.

Williams, K. D. and G. Tselioudis, 2007: Gcm intercomparison of global cloud regimes: Present-day evaluation and climate change response. *Clim. Dyn.*, **29**, 231–250, URL https://doi.org/10.1007/s00382-007-0232-2.

Williamson, C. J., A. Kupc, A. Rollins, J. Kazil, K. D. Froyd, E. A. Ray, D. M. Murphy, G. P. Schill, J. Peischl, C. Thompson, I. Bourgeois, T. B. Ryerson, G. S. Diskin, J. P. DiGangi, D. R. Blake, T. P. V. Bui, M. Dollner, B. Weinzierl, and C. A. Brock, 2021: Large hemispheric difference in nucleation mode aerosol concentrations in the lowermost stratosphere at mid- and high latitudes. *Atmospheric Chemistry and Physics*, **21 (11)**, 9065–9088, URL https://doi.org/10.5194/acp-21-9065-2021.

Yang, J., Y. Zhang, Z. Wang, and D. Zhang, 2022: Cloud type and life stage dependency of liquid-ndash;ice mass partitioning in mixed-phase clouds. *Remote Sensing*, **14 (6)**, URL https://doi.org/10.3390/rs14061431.

Young, S. A., M. A. Vaughan, A. Garnier, J. L. Tackett, J. D. Lambeth, and K. A. Powell, 2018: Extinction and optical depth retrievals for calipso's version 4 data release. *Atmospheric Measurement Techniques*, **11 (10)**, 5701–5727, URL https://doi.org/10.5194/amt-11-5701-2018.

Yu, G., J. Verlinde, E. E. Clothiaux, and Y.-S. Chen, 2014: Mixed-phase cloud phase partitioning using millimeter wavelength cloud radar doppler velocity spectra. *Journal of Geophysical Research: Atmospheres*, **119 (12)**, 7556–7576, URL https://doi.org/10.1002/2013JD021182.

Yuan, T. and Z. Li, 2010: General macro- and microphysical properties of deep convective clouds as observed by modis. *Journal of Climate*, **23 (13)**, 3457 – 3473, URL https://doi.org/10.1175/2009JCLI3136.1.

Yuter, S. E. and R. A. Houze, 1995: Three-dimensional kinematic and microphysical evolution of florida cumulonimbus. part ii: Frequency distributions of vertical velocity, reflectivity, and differential reflectivity. *Monthly Weather Review*, **123 (7)**, 1941 – 1963, URL https://doi.org/10.1175/1520-0493(1995)123<1941:TDKAME>2.0.CO;2.

Zhang, D., A. Vogelmann, P. Kollias, E. Luke, F. Yang, D. Lubin, and Z. Wang, 2019a: Comparison of antarctic and arctic single-layer stratiform mixed-phase cloud properties using ground-based remote

sensing measurements. *Journal of Geophysical Research: Atmospheres*, **124 (17-18)**, 10 186–10 204, URL https://doi.org/10.1029/2019JD030673.

Zhang, D., Z. Wang, and D. Liu, 2010: A global view of midlevel liquid-layer topped stratiform cloud distribution and phase partition from calipso and cloudsat measurements. *Journal of Geophysical Research: Atmospheres*, **115 (D4)**, URL https://doi.org/10.1029/2009JD012143.

Zhang, M., X. Liu, M. Diao, J. J. D'Alessandro, Y. Wang, C. Wu, D. Zhang, Z. Wang, and S. Xie, 2019b: Impacts of representing heterogeneous distribution of cloud liquid and ice on phase partitioning of arctic mixed-phase clouds with ncar cam5. *Journal of Geophysical Research: Atmospheres*, **124 (23)**, 13 071–13 090, URL https://doi.org/10.1029/2019JD030502.

Zhang, W., J. Wang, D. Jin, L. Oreopoulos, and Z. Zhang, 2018: A Deterministic Self-Organizing Map Approach and its Application on Satellite Data based Cloud Type Classification. *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, 2027–2034, URL https://doi.org/10.1109/BigData.2018.8622558.

Zängl, G., D. Reinert, P. Rípodas, and M. Baldauf, 2015: The icon (icosahedral non-hydrostatic) modelling framework of dwd and mpi-m: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, **141 (687)**, 563–579, URL https://doi.org/10.1002/qj.2378.

C. List of Figures

D. List of Tables

E. Acknowledgments

First of all I would like to thank my supervisor Prof. Dr. Corinna Hoose. She gave me the opportunity to work in my favourite field and to actively contribute to cloud physics research. She allowed me to improve and grow as a scientist and as a person, her door was always open and she supported me whenever I needed it. She was a compass, a role model, a mentor. I will always remember the years working together with joy and gratitude.

I would also like to thank Prof. Dr. Jan Cermak, who agreed to serve as a referee for this thesis and accompanied me throughout my Ph.D. with interesting and stimulating discussions that improved my work.

I thank the Graduate School for Climate and Environment (GRACE) for funding a conference and several courses I attended, as well as my research stay at the University of Oslo. I thank my GRACE steering group, consisting of Prof. Dr. Corinna Hoose, Prof. Dr. Jan Cermak and Dr. Martin Stengel, for the joint discussions, which gave me the opportunity to look at my work from other perspectives.

I thank Prof. Dr. Trude Storelvmo for supervising me during my time at the University of Oslo, improving my knowledge and inspiring me to take a different approach to cloud physics research. I would also like to thank Jonah Shaw and Stefan Hofer for the fruitful collaboration, the interesting discussions, and for always being available for questions and virtual meetings, even with different jet-leg.

I would like to thank the colleagues who shared the office with me during my PhD for the sometimes scientific and sometimes funny discussions, the relaxed atmosphere during working hours and the honest support we gave each other.

I would like to thank Gerhard Brückel, Doris Stenschke, Friederike Schönbein, Alexandra Beideck, Jan-Hendrik Daub and Roswitha Marioth for their administrative, organisational and IT support.

Furthermore, I would like to thank all my colleagues at IMK-TRO and especially on the 13th floor for the nice conversations, the coffee breaks, the time spent together after work and the work-related discussions.

I would also like to thank my high school teachers Filadelfio Di Leo, who made me discover the beauty of physics and inspired me to start my career in this wonderful field, where every apparently magical phenomenon has a fascinating explanation, and Sanni Prestipino Giarritta, who discovered my interest in atmospheric physics and meteorology. In addition, I would like to thank Prof. Dr. Vincenzo Levizzani, who made me develop my passion for cloud physics.

A special thanks goes to my family and friends who have always supported me during my studies and my career.

Most of all, I would like to thank my husband, Demetrio Vilardi, who tolerated and supported me during this wonderful, crazy, hard, demanding, frustrating, but always inspiring time.
Finally, I would like to thank all the resources listed here that made this thesis possible:

- This project has received funding from the European Research Council (ERC) under the European Union‘s Horizon 2020 research and innovation programme under grant agreement No 714062 (ERC Starting Grant "C2Phase").

- This work was performed on the supercomputer ForHLR funded by the Ministry of Science, Research and the Arts Baden-Württemberg and by the Federal Ministry of Education and Research.

- The authors acknowledge support by the state of Baden-Württemberg through bwHPC.

- DYAMOND data management was provided by the German Climate Computing Center (DKRZ) and supported through the projects ESiWACE and ESiWACE2. The projects ESiWACE and ESiWACE2 have received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreements No 675191 and 823988. This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project IDs bk1040 and bb1153.