Tropical Cirrus in Global Storm-Resolving Models. Part II: Cirrus Life Cycle and Top-of-Atmosphere Radiative Fluxes

Samantha M Turbeville[1,1], Jacqueline M Nugent[1,1], Thomas P Ackerman[1,1], Christopher S. Bretherton[1,1], and Peter N. Blossey[1,1]

[1]University of Washington

November 30, 2022

Abstract

Cirrus clouds of various thicknesses and radiative characteristics extend over much of the tropics, especially around deep convection. They can be difficult to observe due to their high altitude and sometimes small optical depths. They are also difficult to simulate in conventional global climate models, which have coarse grid spacings and simplified parameterizations of deep convection and cirrus formation. We investigate the representation of tropical cirrus in global storm-resolving models (GSRMs), which have higher spatial resolution and explicit convection and could more accurately represent cirrus cloud processes. This study uses GSRMs from the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) project. The aggregate life cycle of tropical cirrus is analyzed using joint albedo and outgoing longwave radiation (OLR) histograms to assess the fidelity of models in capturing the observed cirrus cloud populations over representative tropical ocean and land regions. The proportions of optically-thick deep convection, anvils, and cirrus vary across models and are reflected in the vertical distribution of cloud cover and top-of-atmosphere radiative fluxes. Model differences in cirrus populations, likely driven by subgrid processes such as ice microphysics, dominate over regional differences between convectively-active tropical land and ocean locations.

Tropical Cirrus in Global Storm-Resolving Models. Part II: Cirrus Life Cycle and Top-of-Atmosphere Radiative Fluxes

S. M. Turbeville[1], J. M. Nugent[1], T. P. Ackerman[1,2], C. S. Bretherton[1,3], and P. N. Blossey[1]

[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA
[2]Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA
[3]Vulcan, Inc., Seattle, WA

Key Points:

- Statistics of tropical cirrus in 40-day high-resolution model simulations scatter around observational estimates
- The joint albedo-OLR histogram is a good observationally-testable diagnostic of cirrus life cycle
- Large differences between models are driven by ice microphysics and model dynamics

Corresponding author: S. M. Turbeville, smturbev@uw.edu

Abstract

Cirrus clouds of various thicknesses and radiative characteristics extend over much of the tropics, especially around deep convection. They can be difficult to observe due to their high altitude and sometimes small optical depths. They are also difficult to simulate in conventional global climate models, which have coarse grid spacings and simplified parameterizations of deep convection and cirrus formation. We investigate the representation of tropical cirrus in global storm-resolving models (GSRMs), which have higher spatial resolution and explicit convection and could more accurately represent cirrus cloud processes. This study uses GSRMs from the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) project. The aggregate life cycle of tropical cirrus is analyzed using joint albedo and outgoing longwave radiation (OLR) histograms to assess the fidelity of models in capturing the observed cirrus cloud populations over representative tropical ocean and land regions. The proportions of optically-thick deep convection, anvils, and cirrus vary across models and are reflected in the vertical distribution of cloud cover and top-of-atmosphere radiative fluxes. Model differences in cirrus populations, likely driven by subgrid processes such as ice microphysics, dominate over regional differences between convectively-active tropical land and ocean locations.

Plain Language Summary

Cirrus (ice) clouds vary in thickness and so have a wide range of impacts on Earth's energy budget. Unlike other clouds, thin cirrus reduce the amount of energy escaping to space, slightly warming the Earth. It is important to understand the differences in tropical cirrus cloud life cycles between models because tropical cirrus are a major source of uncertainty in the prediction of future climates. Cirrus clouds cover a large area in space and can last up to several days, yet they are difficult to measure with satellites and ground-based instruments. We instead use computer models to simulate tropical cirrus, specifically global storm-resolving models (GSRMs) which are able provide a level of detail not possible through observations. Unfortunately, most models have large biases in cloud properties. These differences arise from the imperfect representation of ice in the models. Our goal is to understand the model differences in representation of ice clouds using statistical analysis of the life cycle of cirrus in each model.

1 Introduction

High clouds in the tropics have been difficult to reproduce in global climate models (GCMs) because of their complex microphysics and radiative properties (e.g., Del Genio, 2012; Stephens, 2005). Proper representation of the properties of tropical cirrus, especially cloud amount and hydrometeor distribution, is a key issue for improving GCMs (Inoue et al., 2010). GCMs generally have a low spatial resolution, which is unable to explicitly represent convection and the subsequent tropical cloud life cycle. Diverse convective and ice microphysical parameterizations lead to large differences between GCMs in the ice cloud population (Del Genio, 2012).

This is the second of two papers comparing the formation and properties of tropical cirrus in relation to deep convection in several high-resolution global storm-resolving models (GSRMs). Nugent et al. (2021) (hereafter Part I) focuses on deep convection and its role as a source of ice and vapor for cirrus formation. Here, we compare the simulated ice cloud populations with satellite observations, interpreting them in terms of an aggregate cirrus life cycle.

As noted in Part I, GSRMs are attractive for modeling tropical cirrus. Unlike conventional climate models with horizontal grid spacings of 25–200 km, GSRMs have sub-5 km grid spacing that enables them to explicitly simulate deep convection and its detrainment

of ice into the upper troposphere and better represent the mesoscale gravity wave spectrum and orographic features. Therefore, GSRMs have the potential to act as a better surrogate for the real atmosphere to study scientific questions about tropical cloud populations and the aggregate life cycle of tropical cirrus, including how they might respond to climate change. However, before we can use GSRMs to try to answer such questions, we must evaluate their performance compared to available observations.

We focus on three contrasting $10^\circ \times 10^\circ$ regions in or near frequent deep convection during the simulated period of 1 August–10 September 2016. The first two regions were also used in Part I. Our main focus of analysis is a tropical West Pacific (TWP) region centered near Manus, Papua New Guinea. The second region, which sampled continental convection, is centered over the Sahel (SHL) in west Africa. We added a third region (NAU) centered near Nauru Island on the edge of the Pacific warm pool, where cirrus are often thinner and have originated from remote convection.

1.1 Archetypal Life Cycle of Tropical Cirrus

Convection plays a significant role in the formation and composition of tropical cirrus clouds (Fueglistaler et al., 2009). Houze (1981) describes the idealized life cycle of a typical tropical mesoscale convective system, which begins with a narrow core of deep convection in the formation stage that then matures and spreads into a system with stratiform precipitation and cirrus outflow. As the system weakens, the stratiform precipitation persists but tends not to reach the ground. Finally, the upper-level cirrus clouds break up and thin in the dissipating stage (Houze, 1981). During this evolution, the cloud system has notably different effects on the reflected shortwave (RSW) and outgoing longwave radiation (OLR) at the top of atmosphere (TOA) due to the changing amount of liquid and ice in the vertical column and the change in the associated optical depth.

We focus on the radiative impacts of the life cycle through the populations of thick, medium, and thin clouds and their relative cloud radiative effects. Following Hartmann and Short (1980), we use joint albedo-OLR histograms to characterize and compare cloud populations and their TOA radiative effects in models and satellite observations. These histograms do not follow individual cloud systems, but can be interpreted as proxies for the cirrus life cycle. As tropical cirrus clouds evolve from thick to thin, they tend to follow a characteristic path in the albedo-OLR plane associated with cooling to warming cloud radiative effects (CRE), switching signs for an ice water path (IWP) around 200 g m^{-2}, (Berry & Mace, 2014) that can be qualitatively reproduced in cloud-resolving simulations (Gasparini et al., 2019).

The overall radiative properties of tropical cloud fields integrate over all phases of the cirrus life cycle to be approximately neutral in the TWP. This observed cancellation of longwave (LW) and shortwave (SW) CRE in the TWP has been well documented, but whether it is coincidental or reflects some deeper physical regulating principle is still controversial (e.g., Ramanathan et al., 1989; Hartmann et al., 2001; Kiehl, 1994). Ackerman et al. (1988) showed that cirrus with an optical depth less than $\sim$5 have a net warming effect since they absorb LW radiation more than they reflect SW solar radiation at TOA. Conversely, thick anvils and deep convective clouds reflect SW radiation more than they absorb LW radiation, resulting in a net cooling effect. In the TWP and NAU, the large negative CRE from relatively infrequent deep convective events is nearly offset by the widespread, thin cirrus clouds with slightly positive CRE (Wall et al., 2018). In this study, we analyze whether GSRMs capture this observed neutrality in the TWP and NAU regions.

We also investigate the representation of cirrus in the tropical tropopause layer (TTL), a transitional layer from troposphere to stratosphere defined as in Part I to extend from 14–18 km altitude. TTL cirrus are the highest clouds in the tropics and are thought to play an active role in the thermodynamic structure of this layer and in setting lower-stratospheric water vapor concentrations.

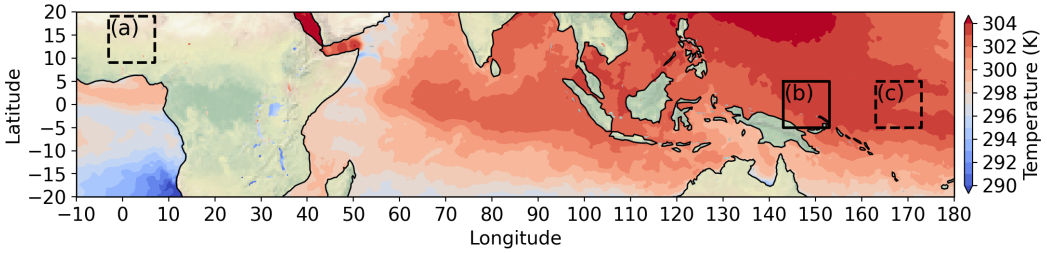

Figure 1. The three analysis regions: (a) the Sahel (SHL; 3°W–7°E, 9°N–19°N), (b) Manus, Papua New Guinea (TWP; 143°E–153°E, 5°N–5°S), and (c) Nauru Island (NAU; 163°E–173°E, 5°N–5°S). Shading shows ERA5 sea surface temperatures on 1 August 2016.

In section 2, we review data sources reused from Part I and introduce new data and analysis methods. Results for the TWP region are presented in section 3 and compared across regions in section 4. We analyze TTL cirrus simulated by the GSRMs in section 5. Section 6 presents our conclusions.

2 Methods and Data

The focus of this paper is comparing the population and implied life cycle of cirrus in GSRMs vs. satellite observations, in different parts of the tropics, with a focus on their TOA radiative properties.

The GSRM radiative transfer schemes operate on individual grid columns for computational efficiency. Consequently, our analysis largely focuses on column-by-column statistics, not horizontal relationships between columns. Properties of GSRM grid columns that we will compare with observations include OLR, albedo, ice water content, and frozen water path.

In this section, we describe the study domains, models, and data used. The models and some data are described in more detail in Part I, so we only elaborate on our new data sources and analysis methods.

2.1 Study Domains

The global simulations and observational data are studied in three contrasting $10° \times 10°$ latitude-longitude regions across the tropics shown in Figure 1. These three regions each contain a substantial but manageable volume of model output (at least 66,000 grid columns). Two of these regions are used in Part I: (a) the Sahel (SHL) in west Africa, representative of continental deep convection, and (b) the TWP region near Manus Island, Papua New Guinea, representative of warm oceanic deep convection, which will be our primary focus. We add (c) a secondary region in the west Pacific near Nauru Island (NAU), in which there is less local deep convection, so more of the cirrus has aged. Each region at some time included an observational Department of Energy Atmospheric Radiation Measurement site offering a comprehensive set of ground-based observations for future work.

The three study regions experienced typical amounts of precipitation over the simulated period in August and early September 2016, with a weak MJO in mid-August and slight La Niña conditions. This justifies comparing model output with climatological observations from the same season in other years.

2.2 The DYAMOND GSRMs

Like Part I, this study uses outputs from GSRMs that participated in the DYnamics of the Atmosphere general circulation Modeled On Non-hydrostatic Domains (DYAMOND) intercomparison project. The DYAMOND project compared free-running GSRMs over the 40-day period from 00 UTC 1 August to 00 UTC 11 September 2016, initialized from ERA5 reanalysis, with the sea surface temperature evolution also based on ERA5. As noted in Part I, the DYAMOND GSRMs use diverse choices of ice microphysics, advection schemes, etc. However, because of their fine grid, they do not use deep convection parameterizations, allowing the detrainment of ice into the tropical upper troposphere to be explicitly simulated. Thus, it is interesting to compare their simulated populations of tropical cirrus clouds. Because the GSRMs are free-running, we compare the model output statistically rather than comparing specific weather systems. We disregard the first two days of model output for all models as a spin-up period for cloud processes.

We focus on four models: NICAM, FV3, ICON, and SAM (collectively referred to as NFIS) because those models provide sufficient output to compute the total IWP and also include 3D output for cloud ice (see Table 1). For more information, see section 2.1 of Part I.

Table 1. List of DYAMOND GSRMs used in this study by their acronym, horizontal grid spacing, and availability of ice/snow/graupel for 3D water content (WC) and 2D integrated water path (WP).

	Grid	3D WC			2D WP		
Model	spacing	I	S	G	I	S	G
NICAM	3.5 km	X	X	X	X	X	X
FV3	3.25 km	X			X	X	X
ICON	2.5 km	X			X	X	X
SAM	4 km	X			X	X	X
ARPNH*	2.5 km				X		
IFS*	4 km				X	X	-
MPAS*	3.5 km				X	X	X
UM*	7.8 km				X		

* Coarsened model output with 0.1° resolution used
\- Graupel not included in microphysics scheme

The DYAMOND models output 2D fields, such as precipitation, available integrated water paths (listed in Table 1), and SW and LW fluxes at TOA, at 15 minute intervals. Some 2D comparisons are based on output coarsened to a $0.1^\circ \times 0.1^\circ$ grid in all models. We also use 3-hourly 3D output of temperature, hydrometeor mixing ratios (available frozen hydrometeors listed in Table 1), specific humidity, vertical velocity, pressure, and height. The vertical grid spacings are similar between models. For example, NICAM has 10 vertical levels within the 14–18 km TTL, while FV3, ICON, and SAM all have 8 vertical levels in the TTL.

Figure 2 shows a typical snapshot of total water content from FV3 in a $10^\circ \times 10^\circ$ latitude-longitude area for the TWP. Darker shading indicates more ice mass in a particular grid cell. Figure 2 demonstrates the detail and complexity of simulated cloud systems afforded by the high resolution. Some convective cores are surrounded by anvil outflow and thinner cirrus with underlying low clouds.

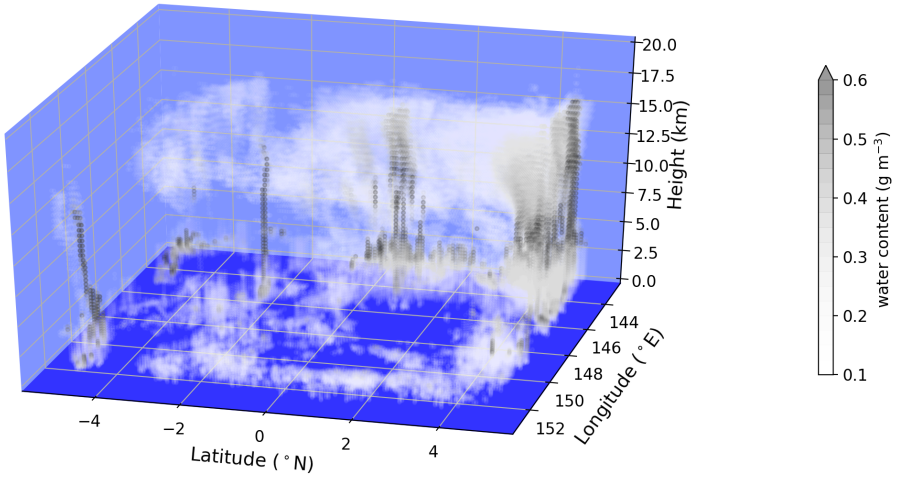

Figure 2. Snapshot of cloud ice water content (g m^{-3}) in FV3 at 6:00 UTC on 3 August 2016 for the 10° × 10° TWP study region.

2.3 Data

The CALIPSO-CloudSat-CERES-MODIS (CCCM) merged product and the raDAR-liDAR (DARDAR) product are reused from Part I. In addition, this paper uses TOA radiative fluxes from the CERES SYN1 deg product.

CCCM provides collocated cloud microphysics and TOA radiative fluxes with a horizontal footprint of roughly 30 km (Kato et al., 2011). We use daytime values from this dataset (13:30 local time) when albedo can be measured and seasonally subset to July–September 2007–2010.

DARDAR combines CloudSat, CALIPSO, and MODIS satellite measurements to derive microphysical properties of cirrus clouds over a broad range of optical thicknesses (Deng et al., 2013; Sokol & Hartmann, 2020) with a fine horizontal spacing of ∼1.1 km and a vertical resolution of 60 m. DARDAR has a higher retrieval uncertainty for optically thin cirrus layers, which are visible with only the lidar (Cazenave et al., 2019). We use both daytime and nighttime values from the 2009 July-August-September data. For both TWP and NAU, there was approximately 50% more precipitation in August 2009 compared to August 2016 according to ECMWF reanalysis; this should be kept in mind when comparing DARDAR to the DYAMOND output in these regions.

We use OLR and RSW radiation from the CERES SYN1 product, available from 2000–2019. CERES SYN1 data have a 1° × 1° horizontal resolution with an approximate uncertainty of 3 W m^{-2} (Kato et al., 2018). We seasonally subset this data to 1 August to 10 September of each year.

For comparisons between GSRM output and these observations, we coarsen the model output to 0.3° × 0.3° to match CCCM or 1° × 1° to match CERES. The GSRM output is not coarsened for comparisons with DARDAR, which has a similar footprint to the model grid spacings.

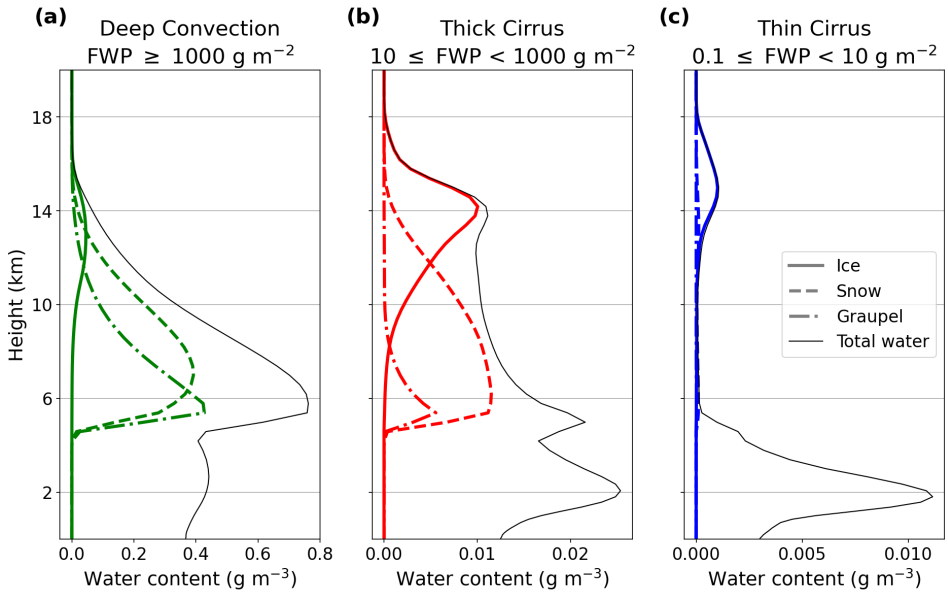

Figure 3. Vertical profiles of water content (g m^{-3}) for NICAM showing the breakdown between ice (solid), snow (dashed), and graupel (dot-dashed) in areas of (a) deep convection (green), (b) thick cirrus (red), and (c) thin cirrus (blue). The total water content, including liquid, is shown in the thin black line. The x-axis is different for each plot.

2.4 FWP Categorization

We classify each grid column according to its column-integrated frozen water path (FWP), defined as the sum of ice, snow, and graupel water paths. As in section 2.5 of Part I, we define three thickness categories that approximately correspond to (1) deep convection, (2) thick cirrus, and (3) thin cirrus:

FWP Categories:				
CAT 1	Deep convection		FWP	≥ 1000 g m^{-2}
CAT 2	Thick cirrus	$10 \leq$	FWP	< 1000 g m^{-2}
CAT 3	Thin cirrus	$0.1 \leq$	FWP	< 10 g m^{-2}
	Cirrus-free		FWP	< 0.1 g m^{-2}

For comparisons to observational data and previous studies, we use the model output of hydrometeor mixing ratios to calculate IWC and frozen water content (FWC) where possible as well as liquid water content (LWC). For FV3, ICON, and SAM, only cloud ice and cloud water were saved for the 3D hydrometeors, so we can only calculate the FWC for NICAM. We show vertical profiles of cloud ice, snow, and graupel in NICAM in Figure 3 to demonstrate how FWC differs between cloud types. Snow and graupel in NICAM make up the bulk of frozen mass below the TTL and a small, but non-negligible, portion of the total frozen water in the upper troposphere in the rare columns with vigorous deep convection (Figure 3a). Columns with thick cirrus in NICAM have snow and graupel dominant below and ice dominant above 12.5 km (Figure 3b). Thin cirrus are made mostly of cloud ice with a negligible amount of snow or graupel. We therefore assume the IWC is a decent representation of FWC above 12.5 km for thin cirrus, but underestimates FWC in deep convection or anvils in models that do not output 3D snow or graupel.

2.5 Clear Sky Thresholds

The DYAMOND models produce few columns that are truly ice-free. The model microphysics schemes typically retain small amounts of FWC. FWCs below 5×10^{-4} g m^{-3} are difficult to detect from aircraft or satellite data (Deng et al., 2013). Therefore, we adopt this value as a 'clear-sky' or 'cirrus-free' threshold for simulated IWC. The frequency of thin cirrus clouds is sensitive to this threshold in all of the models.

Given this threshold and assuming a GSRM vertical grid spacing of 200 m, the column-integrated cloud ice water path must exceed 0.1 g m^{-2} for the volume to be considered cloudy, so this is chosen as our cirrus-free FWP threshold. DARDAR does not retrieve FWP values less than 0.1 g m^{-2}.

3 The Aggregate Life Cycle of Tropical Cirrus in the TWP

In this section, we determine the cloud populations from the albedo-OLR histograms and compare the radiative climatologies in each model to satellite observations. Then, the life cycle of tropical cirrus is quantified using FWP categories for the NFIS models in the TWP. The vertical structure and ice properties of clouds in the TWP are used to explain the differences in cloud populations between models. Our analysis relies on statistical aggregates from snapshots of cloud properties and their corresponding radiative fluxes.

3.1 Cloud Populations

A joint albedo-OLR histogram gives a useful summary of how the cloud populations relate to the TOA radiation budget (Hartmann & Short, 1980). Over ocean, the albedo can be regarded as a proxy for cloud optical depth. The OLR is an indicator of cloud-top height; thus, we can differentiate between cloud types with these two 2D variables. Using albedo restricts our analysis of the models to mid-day (10 am–2 pm local time) to eliminate any anomalies occurring near dawn and dusk. The model outputs are coarsened to match the 30 km CCCM footprint.

Figure 4 compares simulated albedo-OLR histograms for the TWP during 1 August–10 September 2016 (DYAMOND GSRMs) and July-August-September 2007–2010 (CCCM observations). The observations are shown in Figure 4a. The distribution also encompasses a banana-shaped curve extending from a maximum probability density at the clear-sky values in the lower right up to a secondary deep convective peak at high albedo and low OLR. There is also a slight shoulder of moderate albedo and high OLR occurrences.

Figure 4f interprets this joint histogram in terms of component cloud types. A typical tropical cirrus life cycle roughly aligns with the black arrow in Figure 4f, starting with deep convection then progressing through thick and then thin cirrus. The light blue ellipse indicates a population of low, liquid-dominated clouds with tops below 8 km. These can induce relatively strong albedo perturbations with less impact on OLR than deep convective cloud systems. The orange circle demarcates clear sky. These cloud types may shift somewhat in albedo-OLR space in other convectively-active regions of the tropics with different surface albedo and humidity/temperature profiles.

In comparing the TWP albedo-OLR histograms for each model (Figure 4b–e,g–j), we look for three characteristic properties of the observations: (1) a clear sky peak, (2) a secondary peak in deep convection, and (3) a banana-shaped distribution. All of the models successfully capture the high OLR-low albedo clear-sky peak, though FV3, ICON, and IFS have slightly higher clear-sky OLR values. FV3, MPAS, and UM also capture the local maximum corresponding to observed deep convective radiative characteristics. NICAM and IFS match the general banana shape of the histogram very well but do not capture the secondary deep convective peak. ICON, SAM, and ARPNH all fail to simulate the secondary deep convective peak and do not match the banana-shaped distribution of the

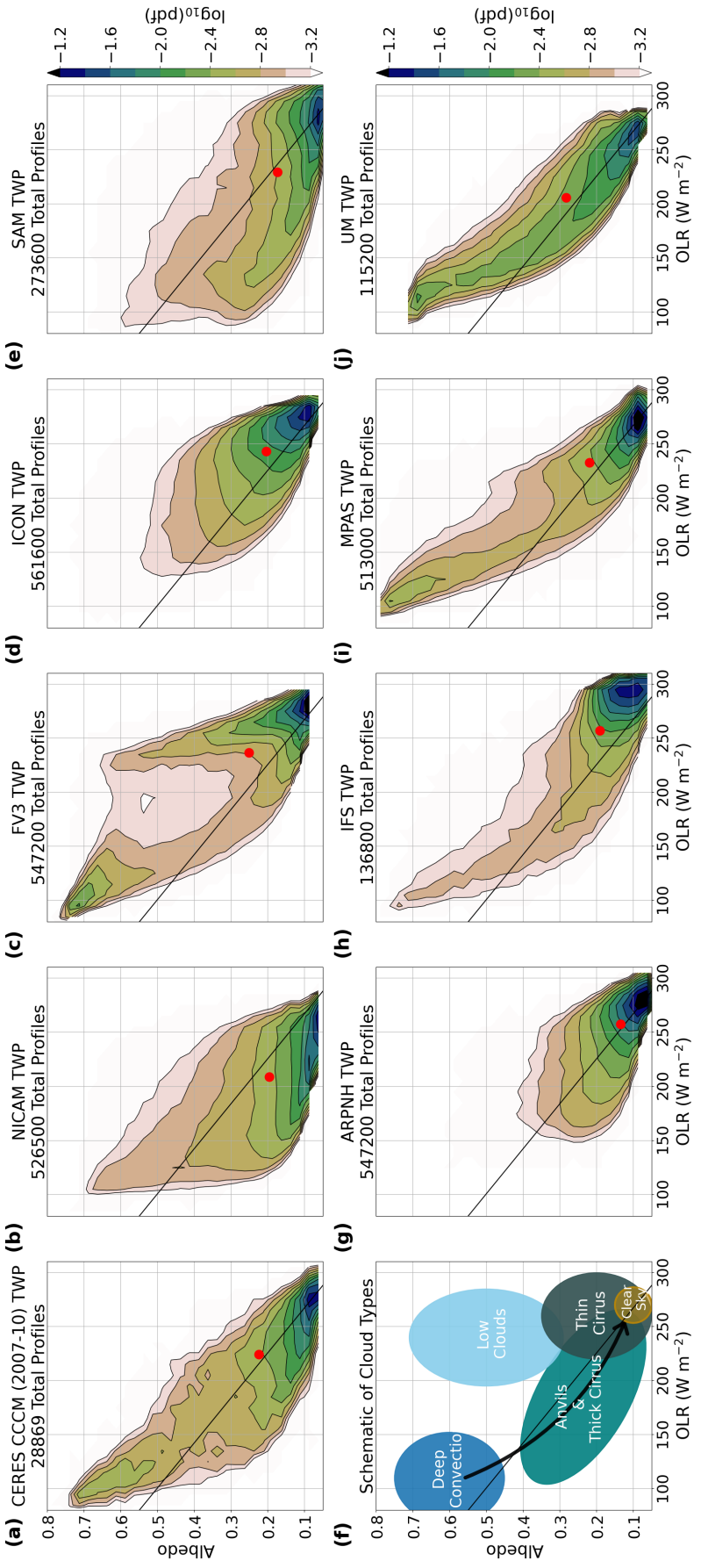

Figure 4. Density plot of albedo and OLR for (a) CCCM and (b–e,g–j) the DYAMOND models for days 3–40. The schematic of cloud types is reproduced in (f). Model output averaged over 0.3° to match the CCCM footprint. There are 32 and 27 evenly spaced bins for the OLR and albedo, respectively (as in Gasparini et al., 2019). The red dot indicates the mean albedo-OLR pairing for the distribution. Neutral CRE is calculated from observation as the black line.

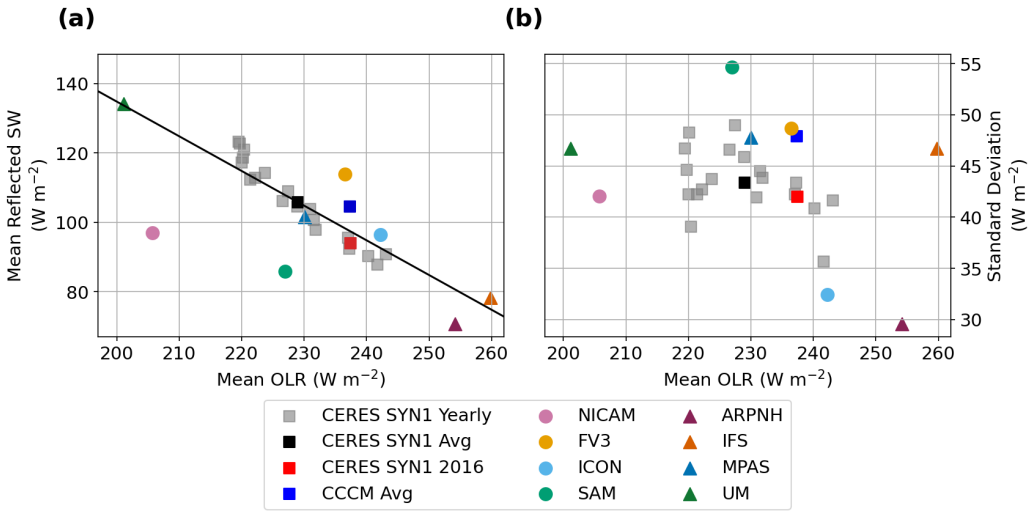

Figure 5. (a) Mean OLR plotted against mean reflected SW radiation at TOA from days 3–40. The linear fit of CERES SYN1 is plotted as the black line. (b) Daily mean versus standard deviation of OLR. Before computing the standard deviation, the OLR from each model was coarsened to match the hourly and $1^\circ \times 1^\circ$ horizontal resolution of CERES SYN1.

observations. ICON and ARPNH do not have albedo-OLR values characteristic of deep convection because their deep convective cores are not surrounded by sufficient thick, cold anvil within the CCCM footprint. SAM has convection reaching the upper troposphere, shown by the low OLR values, but the associated cloud is optically too thin to match the observed deep convective characteristics. FV3 has a distinct population of liquid-dominated clouds reaching high albedo values despite relatively high OLR.

In summary, MPAS and UM capture all of the key characteristics observed in the TWP. NICAM, FV3, and IFS also capture most key traits of cloud populations in the TWP, while ICON, SAM, and ARPNH deviate greatly from the observed cloud populations.

3.2 TOA Radiation Balance and CRE

The TWP-mean albedo and OLR (the red dot in each panel of Figure 4) can be inferred by averaging over the albedo-OLR histogram. The GSRMs scatter around the observed albedo of 0.22, with albedos ranging from 0.14 (ARPNH) to 0.3 (UM). They similarly scatter about the observed OLR of 225 W m^{-2}, from 205 W m^{-2} (NICAM) to 255 W m^{-2} (ARPNH). Some of this scatter comes from different characteristic clear-sky albedo and OLR, but most of it is due to the different simulated cloud populations.

Hartmann et al. (2001) noted a characteristic of the TWP cloud population–near zero net cloud radiative effect (CRE, defined as the cloud-induced change in the net TOA radiation). This is due to a balance between thick convective and cirrus clouds, which increase RSW more than they decrease OLR, and thin cirrus, which do the opposite. The black neutral-CRE line in all of the panels of Figure 4 shows the albedos and OLRs for which this balance would occur in the observations. It is computed by locating the clear-sky occurrence peaks of OLR and albedo in the CCCM data. The longwave CRE is calculated as the difference between the estimated clear-sky OLR and the actual OLR. while the shortwave CRE (which is negative) is calculated as the difference between the clear-sky albedo and

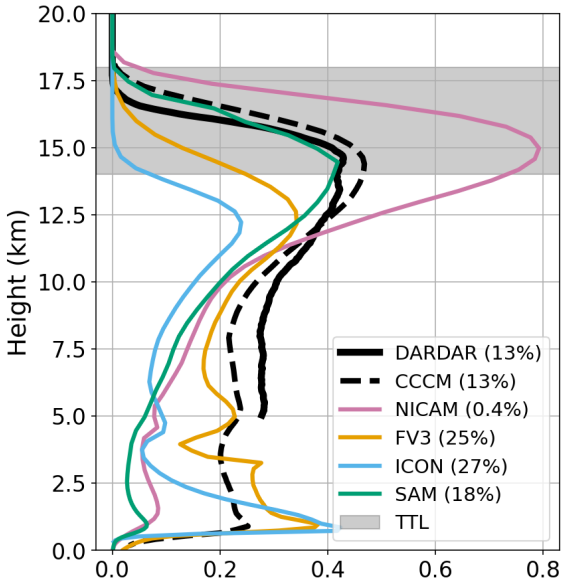

Figure 6. Vertical occurrence of clouds for each model from days 3–40 compared to observations from DARDAR and CCCM. For the models, only cloud ice is used to calculate cloud fraction; the observations implicitly include all frozen hydrometeors.

the actual albedo, multiplied by the daily-mean insolation at the given location and season (413.2 W m^{-2} in the TWP and NAU regions and 435.3 W m^{-2} in the SHL region).

The red dot in the CCCM panel almost overlays this black line, showing the net neutrality of CRE in the TWP. For the model panels, a red dot above (below) the black line signifies the clouds contribute an overall radiative cooling (warming) effect. The models approximate the observed CRE neutrality to within 30 W m^{-2}. Only NICAM simulates substantial positive net CRE (cloud-radiative warming). The black line corresponds to a constant total outgoing radiative flux (LW+SW) equal to the 3-year mean. The model variability mostly scatters $\pm$10 W m^{-2} around the 3-year mean along the black line, which also is approximately a line of CRE neutrality.

We include CCCM and CERES SYN1 satellite retrievals of TOA radiative fluxes in Figure 5. Since the CERES SYN1 observations span 20 years of 40-day periods and the GSRM simulations span 1 year for 40 days, we can assess whether the GSRM results are within the range of interannual variability of TOA radiative fluxes. The model spread should ideally be contained within this range of natural annual variability. That is true for MPAS, FV3, and ICON, but not for the other models (Figure 5a). Some of those models near CRE neutrality (SAM and ICON in Figure 4) lie outside the observed interannual range (Figure 5). Extreme examples of this are UM and IFS, which each had biases of $\sim$30 W m^{-2} in OLR (high in IFS, low in UM) with offsetting biases in reflected SW. These two models match the banana shape of the observed albedo-OLR distribution in Figure 4, but their frequency of occurrence at the two ends of the banana are biased, with IFS tending to produce near clear-sky values and UM tending towards thick anvils. Other models, especially NICAM, lie far from CRE neutrality so as to be inconsistent with observations from any year.

3.3 Sources of GSRM TOA Radiation Bias

As a group, the models capture key observed radiative characteristics of the cloud populations. However, the albedo-OLR histograms show that individual GSRMs simulate clouds differently from each other and the observations. How do these differences arise? Two potential contributors are the vertical profile of fractional coverage and the thickness of clouds, as measured by their frozen water path.

Observed vertical profiles of cloud fraction in convectively active regions of the tropics have their largest peak in the upper troposphere (Hollars et al., 2004), where it is associated with cirrus clouds. Different GSRMs simulate different cirrus cloud-top heights, likely due to their formulations of ice microphysics and dynamics. Virts et al. (2010) show that the fractional coverage of upper troposphere cirrus maximizes in the 14–15 km layer, which corresponds to the base of the TTL. Figure 6 compares observational estimates of vertically-resolved cloud fraction from DARDAR (solid) and CCCM (dashed) with the NFIS GSRMs in the TWP. The cloud fraction peaks near 15 km in NICAM and SAM, consistent with these observations, but peaks well below the TTL in FV3 and ICON at $\sim$12.5 km.

NICAM has too much cloud cover in the upper troposphere and TTL but too few low clouds when compared to DARDAR and CCCM. This contributes to its low bias in OLR. FV3 has too few upper-level clouds and too many low clouds; this leads to a high bias in RSW. ICON has too many low clouds but too few ice clouds everywhere above 2 km altitude, also leading to a high bias in OLR. SAM is reasonably close in frequency and shape to the observed cirrus cloud occurrence, but has too little low cloud, leading to a low bias in RSW. Part I shows that the upper level cloud occurrence in the NFIS models approximately tracks their relative humidity profiles (Figure 2b, d in Part I).

3.4 Frozen Water Path Distribution

We investigate the characteristics of ice clouds in each model more closely through the distribution of frozen mass in each column over the August–September period. We further break down the ice into three main categories that roughly align with the deep convection, thick cirrus, and thin cirrus regimes in the albedo-OLR diagram shown in Figure 4f (schematic). Because the FWP also influences both the SW and LW radiative characteristics of the cirrus, this analysis can also shed further light on GSRM cloud-radiative biases; that will be addressed in the following section. We ignore LWP in this categorization since our focus is on cirrus clouds. This categorization is done at the full model resolution, since that is comparable to that of the DARDAR FWP observations.

Figure 7 shows FWP histograms for DARDAR and the four NFIS models. The FWP categories are plotted as different colors corresponding to deep convection (CAT 1; green), thick cirrus (CAT 2; red), thin cirrus (CAT 3; blue), and cirrus-free (gray).

Deep convection occupies 11% of all columns in space and time in the DARDAR observations, and 3–5% in the models. The DARDAR measurements were taken during an August with nearly 50% more convection than August 2016. This suggests an observationally-based estimate for 2016 of 7% occurrence of CAT 1. The DARDAR methodology may not accurately retrieve FWP in deep convective columns dominated by snow and graupel, so the remaining difference in CAT 1 frequency between the models and observations may not be meaningful.

Thick cirrus occupy the most area in space and time, 48% for DARDAR and 30–67% for the models. Thin cirrus occurs slightly less often, 25% for DARDAR and 28–38% for the GSRMs. Lastly, the cirrus-free category, occupying 16% fractional area in DARDAR, has a broad range of 1–32% across the models. Berry and Mace (2014) find that the CRE transitions from cooling to warming as FWP decreases to ∼200 g m^{-2}. This transition occurs in CAT 2; subsequently, CAT 2 has the strongest influence on the overall cloud radiative effect in the region. The models range from cooling in FV3 to warming in NICAM.

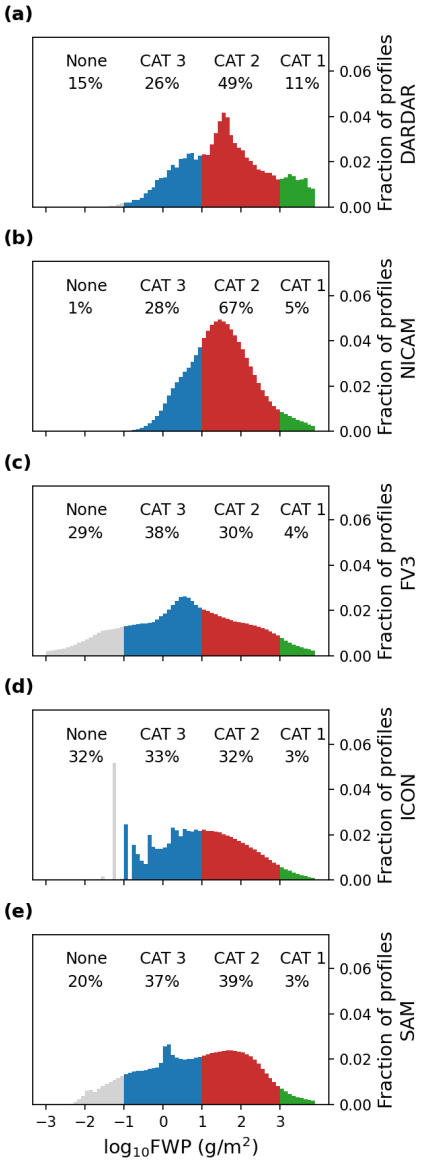

Figure 7. Total column FWP for observations from (a) DARDAR and model output from days 3–40 for (b) NICAM, (c) FV3, (d) ICON, and (e) SAM colored by the FWP categories: (green) deep convection, (red) thick cirrus, and (blue) thin cirrus.

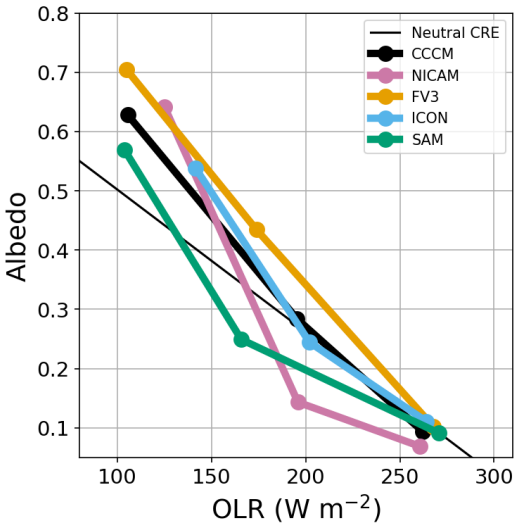

Figure 8. The median albedo-OLR pairing for each category is plotted as closed circles for the data and models. Models are not coarsened to match the CCCM footprint. CAT 1 is the upper-leftmost point and is connected to CAT 2 and then CAT 3 by straight lines, signifying the aggregate life cycle of tropical cirrus.

There is a large range of FWP histogram shapes among the models. NICAM most closely matches the shape of DARDAR including the peak FWP near 30 g m^{-2}, but has almost no cirrus-free columns. FV3 and ICON peak at a smaller FWP while SAM peaks at a larger FWP.

Our results confirm the spatial dominance of the maturing and dissipating stages of the ideal life cycle from Houze (1981). The model FWP distributions are diverse, presumably reflecting differences in their ice microphysics schemes and perhaps also their model dynamics. As a group, the FWP distributions tend to scatter around the DARDAR observations in both modal value and category frequency.

3.5 FWP as a Control on Albedo and OLR

To what extent do intermodel differences in the FWP histogram correlate with their differences in albedo-OLR histograms and in TOA radiative fluxes seen in Figure 4? Binning the albedo-OLR diagrams into FWP categories helps address this question, while allowing us to construct a simple aggregate radiative description of the cirrus life cycle for each NFIS model and the observations.

In Figure 8, the median value of OLR and albedo for each FWP category is plotted, with CAT 1 deep convection in the upper left corner progressing through anvils to CAT 2 thick cirrus and ending with CAT 3 thin cirrus in the lower right. Each model traces a unique path from one category to the next in the albedo-OLR plane. The median values are used because Berry and Mace (2014) found that the mean value of IWP is a poor diagnostic of radiative properties of tropical cirrus, and because the mean values of albedo are dominated by high-albedo liquid clouds, while our focus is on ice clouds. The GSRMs are analyzed at their fine native grid resolution to allow better FWP sorting, so the observations have a footprint about 6X larger in each direction than the models. To partly compensate, the CCCM FWP is calculated only for the cloudy fraction of the footprint.

Category 1 covers only a small fractional area, especially in the models, but is of particular interest since deep convection is important for bringing ice into the TTL (Part I) that may also affect the thick and thin cirrus categories. Figure 8 shows that FV3 and SAM have the lowest median OLR for CAT 1, comparable to the CCCM observations, meaning that they produce tall convective towers with the coldest cloud tops, even though NICAM has more high cold cirrus overall. Despite its low OLR, SAM has a slightly lower albedo than FV3, NICAM, or the observations, indicating that its convective clouds are not as optically thick. ICON has the highest OLR and lowest albedo, meaning it does not have convective towers that reach as high or are as reflective as the other models.

The largest differences between models occur in CAT 2, with OLR ranging from 156–202 W m^{-2} and albedo from 0.15–0.44. This category covers much of the TWP area, so these differences are predictive of area-mean net CRE. The net CRE in CAT 2 is most positive (warming) for NICAM, followed by SAM, with slight negative CAT 2 CRE for ICON and more negative CAT 2 CRE for FV3. This is exactly the same ordering seen for area-mean CRE in Figure 5a.

The NFIS models have much more similar median albedo and OLR for Category 3 (thin cirrus), which has relatively weak radiative effects. Thus, thin cirrus are comparatively unimportant to intermodel differences in area-mean net CRE over the TWP.

Even with FWP fixed, a major factor affecting albedo is the presence of liquid clouds, because they are more reflective than ice clouds. FV3 tends to have some liquid clouds underlying CAT 2 and 3 cirrus. ICON tends to have more liquid clouds underlying CAT 3 cirrus clouds.

4 Regional Comparison

Here, we compare results from the TWP with our two other tropical study regions described in section 2.1, SHL (land convection) and NAU (on the edge of an active ocean convection region), to see how the NFIS models simulate cirrus in these different yet climatically significant regions.

4.1 Cloud Populations

The SHL has a much brighter surface than the dark ocean. The northern half of the SHL region is the high-albedo Sahara while the southern half has more vegetation and a lower surface albedo. This causes a larger variation within the region in clear-sky albedo and OLR values and a higher albedo overall than the ocean, both in the CCCM observations and in the four NFIS GSRMs (see Figure 9a–e). Most clouds lie below the CCCM-based line of neutral CRE in the SHL (Figure 9a), implying a stronger warming effect of clouds over the SHL and/or a GSRM with a lower surface albedo than that inferred from the CCCM observations. CCCM data also have fewer deep convective columns with high albedo and low OLR in SHL than the TWP. This cloud distribution reflects more sporadic deep convection in a drier environment. Qualitatively, all four GSRMs also show this pattern.

Compared to the TWP, NAU has fewer deep convective columns and no secondary peak at low OLR-high albedo values. The mean albedo-OLR pair (red dot in Figure 9k) in NAU is slightly above the line of neutral CRE, signifying a small net cloud cooling effect. This cooling is likely due to the larger population of low clouds, which are more reflective than extended cirrus clouds. The four GSRMs all simulate the weakening of the deep convective peak, but only SAM shows an enhanced shallow-cloud peak of higher albedo and high OLR at NAU compared to TWP (Figure 9l–o).

Many biases of individual GSRMs identified in the TWP persist across regions. For example, NICAM has frequent thin cirrus, FV3 has a large liquid cloud population, and ICON and SAM do not have sufficient deep convection (Figure 9).

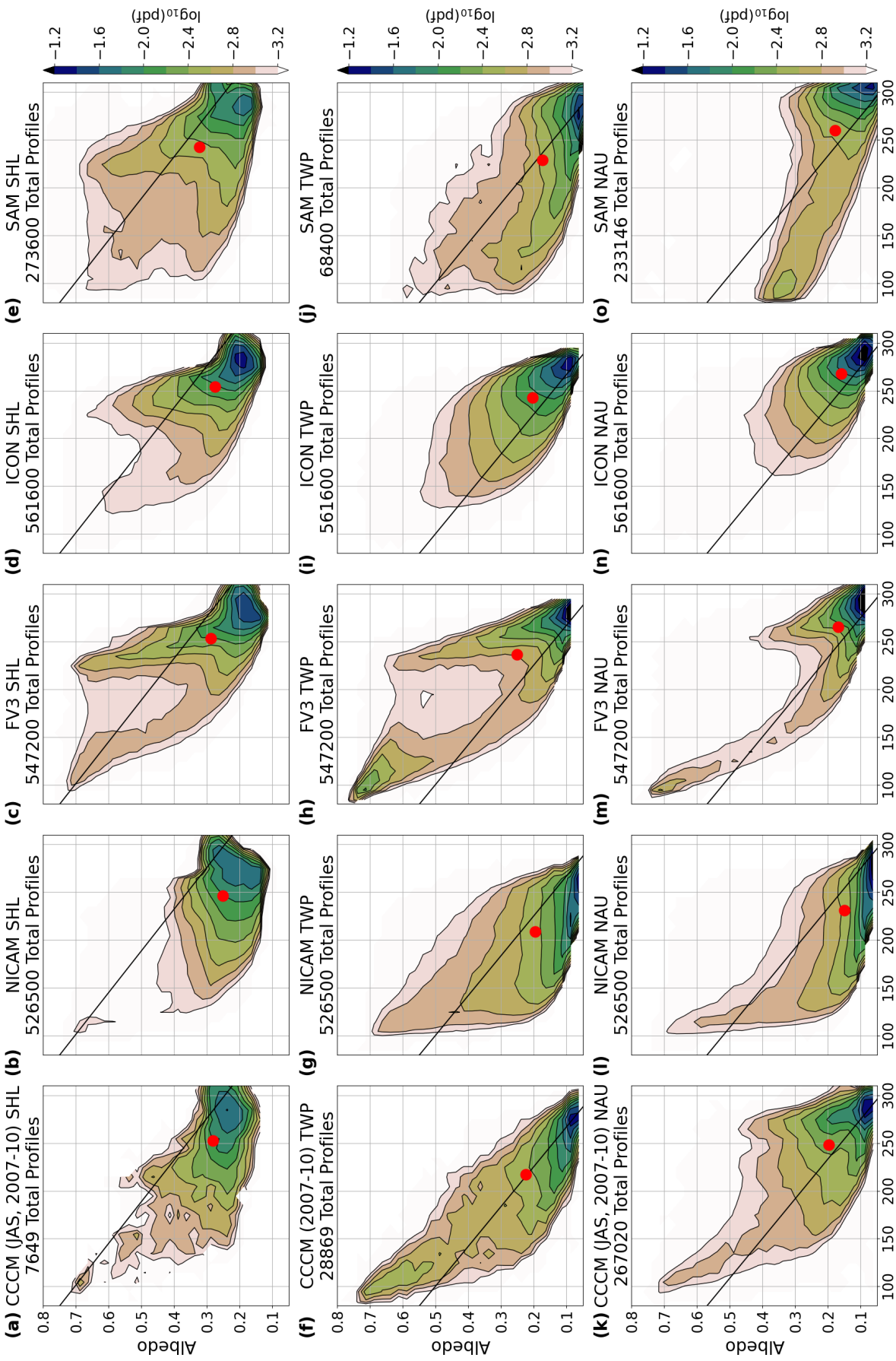

Figure 9. Joint albedo-OLR histograms for SHL (a–e), TWP (f-j), and NAU (k–o) for (a, f, k) the observations from CCCM and the model output from (b, g, l) NICAM, (c, h, m) FV3, (d, i, n) ICON, and (e, j, o) SAM. The black line of neutral CRE was calculated for each region from the observations.

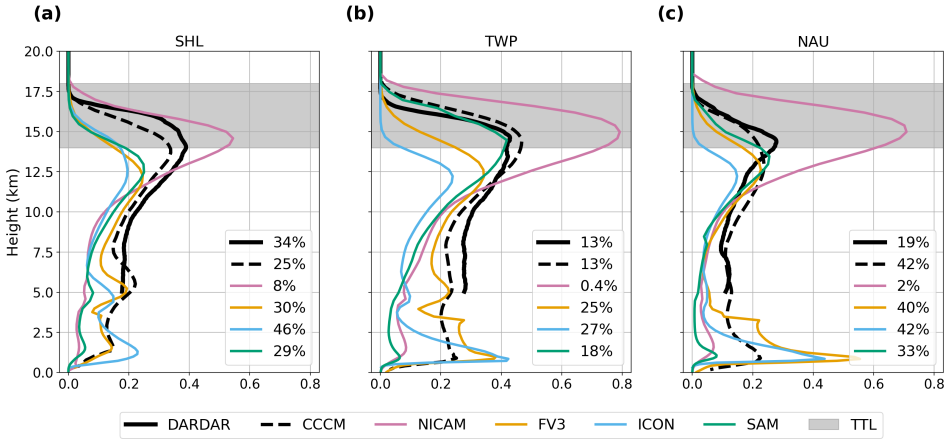

Figure 10. Same as Figure 6 but for the (a) SHL, (b) TWP, and (c) NAU. The radar-lidar product from CCCM shows the combined liquid and ice-phase cloud occurrence. The clear-sky percentages are noted in the legend for SHL, TWP, and NAU, respectively.

4.2 Vertical Structure of Clouds

In the CCCM and DARDAR observations, all regions have a primary peak in cloud occurrence near the base of the TTL (Figure 10b). The TWP has the highest cloud occurrence at all heights (Figure 10b) and the most cloud cover (time-area percentage of cloudy columns) at nearly 87%. The SHL (Figure 10a) has comparable vertical cloud occurrence in the TTL but much less column cloud cover (66% in DARDAR and 75% in CCCM). DARDAR also shows a stronger trimodal peak in cloud occurrence in the SHL, as noted by Johnson et al. (1999). Fewer TTL and mid-troposphere clouds as well as a weaker peak in cloud occurrence at the freezing level in NAU indicate weaker or less frequent deep convection (Figure 10c). FV3, ICON, and SAM reproduce the trimodal distribution of clouds in the SHL, but fail to simulate enough TTL clouds. In all regions, NICAM overestimates the cloud occurrence in the TTL and underestimates clouds in the lower troposphere.

The models have characteristic biases in the vertical structure of clouds that are fairly similar in the SHL and NAU to the TWP and correlate well with the model differences in the joint albedo-OLR histogram. Compared to the TWP, each model simulates fewer TTL and mid-troposphere clouds in NAU consistent with observations and the TOA radiative fluxes shown in the joint albedo-OLR histograms (Figure 4). ICON has the lowest cloud-top heights and correspondingly few instances of low OLR values. The population of low clouds in FV3 shows up in the vertical occurrence as a strong peak in boundary layer and intermediary clouds near the freezing level. The joint histogram of NICAM does not visually capture the high occurrence of clouds in the TTL. The low occurrence of deep convection in SAM is not apparent in the vertical occurrence; hence, the properties of clouds in SAM need to be further examined.

4.3 FWP Categorization

The FWP histograms for the TWP for the DARDAR observations and the four NFIS models from Figure 7a–e are reproduced as the green lines in Figure 11a–e. The FWP histograms for the SHL and NAU are in the orange and purple lines, respectively. The

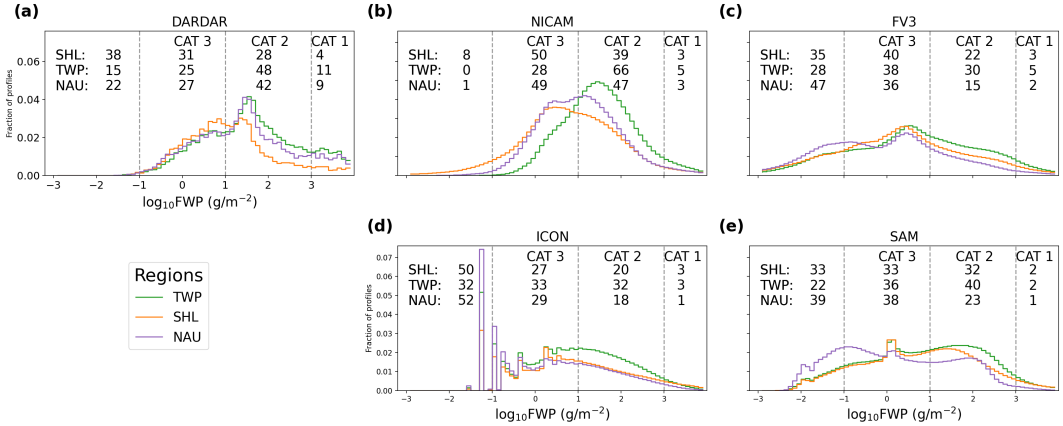

Figure 11. Normalized histograms of total column FWP for the SHL (orange), NAU (purple), and TWP histogram (green; same as black line in Figure 7 histograms) for (a) DARDAR, (b) NICAM, (c) FV3, (d) ICON, and (e) SAM. The categories are divided by vertical dashed lines. Columns with FWP below 0.1 g m^{-2} are considered to be cirrus-free. Listed at the top of each subplot is the percentage of columns from days 3–40 in each region for each category.

fraction of columns in each category is listed above the histograms in each plot. Recall that DARDAR observations are from 2009, which had more precipitation in the TWP and NAU than 2016. This difference should be considered when comparing the observed and simulated FWP histograms.

The DARDAR FWP histogram for the SHL has a similar mode (20–30 g m^{-2}) as for the TWP, but FWPs exceeding this mode are much rarer in the SHL while smaller FWPs are slightly more common. The DARDAR FWP histogram for NAU is similar to the TWP, but with slightly less occurrence at FWPs above the modal value (Figure 11a). Qualitatively similar regional differences are seen in the GSRMs; however, for NICAM and SAM, their modal FWPs also shift downward at the SHL and NAU relative to TWP (Figure 11b, e). The FWP histogram changes more from TWP to NAU in all GSRMs than in DARDAR. The differences between GSRMs are broadly similar in the SHL and NAU as in the TWP.

4.4 Regional Albedo and OLR vs. FWP category

Figure 12 compares the median values of albedo and OLR plane in each FWP category across models and regions. It builds on Figure 8, which presented similar information for the TWP in a complementary form. The models are not coarsened to match CCCM in order to capitalize on the fine resolution of the native grids of the GSRMs. The black line of neutral CRE is only drawn for the TWP region; the line for NAU is nearly identical but the SHL line is at a much higher albedo.

Looking first at Figure 12a, the CCCM data suggest that the OLR and albedo of the thickest (CAT 1) tropical deep convective clouds are similar across regions. Similarly, for the thickest clouds in all NFIS models except SAM, there is generally more intermodel variability in than intra-model regional variability. As we progress to categories with successively smaller FWPs, the opposite occurs, with the CCCM observations and the models generally agreeing on the regional differences in clear-sky albedo and OLR. These regional differences are driven by the higher surface albedo of the SHL and the drier conditions in the SHL and NAU than in the TWP (Figure 12d).

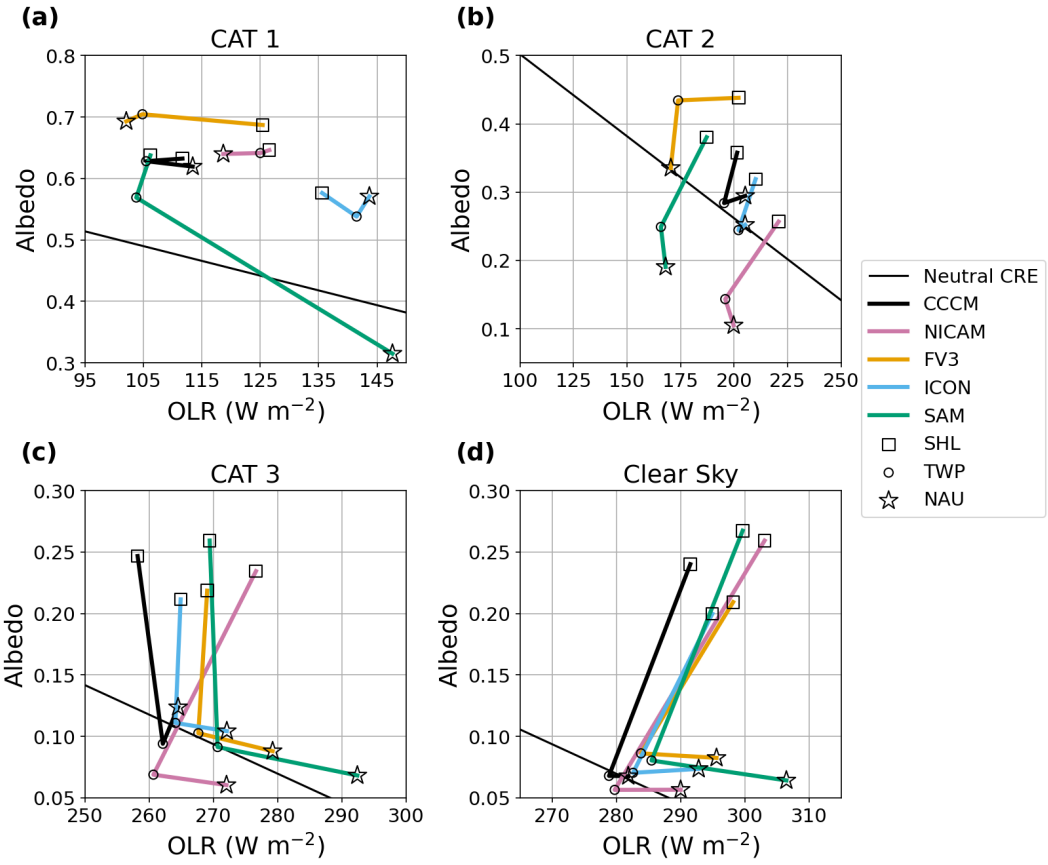

Figure 12. (a) CAT 1, (b) CAT 2, (c) CAT 3, and (d) clear-sky median values for the SHL (square), TWP (circle), and NAU (star) in each model and CCCM. The scale and axis limits are different for CATs 1 and 2. Compared to clear sky, CAT 3 is shifted to lower OLR values by 15 W m^{-2} but has the same x and y scales. The line of neutral CRE calculated from CCCM is shown for the TWP.

In CAT 1 (deep convection, shown in Figure 12a), the models have a diversity of small regional shifts in OLR indicating disagreement on which region has the coldest cloud tops (SHL for ICON; NAU for FV3 and NICAM; and TWP for SAM and the observations). However, consistent with Figure 11 and 9, the differences between models in this category are mostly similar across regions, except for SAM in NAU and FV3 in the SHL.

In CAT 2 (anvil and thick cirrus, shown in Figure 12b), systematic regional differences are seen in both the observations and the models. This FWP category is associated with a higher albedo over the bright underlying surface of the SHL than in the two other ocean regions. The OLR of thick cirrus, through which little longwave radiation can penetrate from below, varies less between regions because its cloud-top temperature is similar across regions.

In CAT 3 (thin cirrus, shown in Figure 12c), regional differences start to dominate model differences. Both in models and observations, the albedos are only slightly higher than the clear-sky values shown in Figure 12d, but the OLRs are reduced about 20 W m^{-2} over the ocean regions and 30 W m^{-2} over the drier SHL region. The bright surface of the SHL dominates the regional albedo variations. Over the ocean, the thin cirrus in ICON often have underlying liquid clouds which give it the largest CAT 3 albedo among the models

and bring it closest to the observed albedo in NAU. Except for NICAM in the TWP, the observed OLR is overestimated in CAT 3 and the clear sky category by all models in all regions.

5 TTL Cirrus

In this section, we investigate a particular subset of cirrus with especially high cloud tops: TTL cirrus. The TTL is the layer of the atmosphere that has properties transitioning between the troposphere and stratosphere, approximately 14–18 km above sea level in the tropics. TTL cirrus are interesting to study because they have a significant impact on water vapor in the TTL and stratosphere (e.g., Virts & Houze, 2015; Holton & Gettelman, 2001). Some TTL cirrus are thick enough to significantly reduce OLR (Fueglistaler et al., 2009).

5.1 Categorization

The TTL IWP was calculated by integrating the 3D cloud ice within the 14–18 km layer for NFIS. Even though NICAM has TTL FWP, we use TTL IWP for consistency between models; the results differ slightly in CAT 1. We use the total frozen condensate (FWP) in the TTL for DARDAR because DARDAR does not differentiate the different hydrometeors. The NICAM hydrometeor profiles in Figure 3 suggest that the missing snow and graupel will mostly affect CAT 1 (deep convection); there is likely minimal impact on CAT 2 or 3.

In Figure 13, the TTL IWP histogram for CAT 1 (green) is plotted on the bottom, with the TTL IWP histograms for CAT 2 (red) and CAT 3 (blue) stacked on top. The black line shows the histogram of the total TTL IWP (FWP for DARDAR). The time-mean frequency of columns in each category that contain TTL cirrus is listed in the legend. A TTL IWP (FWP for DARDAR) threshold of 0.1 g m^{-2} is used to determine the presence of TTL cloud. Columns below this threshold are shaded in gray.

Of the models, NICAM's TTL IWP distribution best matches the shape of DARDAR TTL FWP in each region (Figure 13ii) but is shifted to larger FWPs by an order of magnitude for CATs 2 and 3. That is, NICAM has thicker cirrus in the TTL than observations suggest.

On the other hand, FV3, ICON, and SAM appear to have too little ice in TTL cirrus. These models all have roughly equal proportions of TTL cirrus in CATs 2 and 3. The TTL IWP is smallest in ICON, for which it rarely exceeds 5 g m^{-2} in the TWP or 1 g m^{-2} in the SHL and NAU. Based on Part I, this is likely due to a lack of convection in ICON that is deep enough to reach the TTL. These three models also often generate nonzero ice concentrations in the TTL that are too small to classify as a cloud. Again, the NFIS models look very similar across regions but have different characteristics of ice in the TTL from each other. Like the NFIS FWP modes and category frequencies scatter around observations, so do the NFIS TTL IWP characteristics. We have shown that ICON consistently has too little ice in the TTL and NICAM consistently has too much. SAM matches the category frequencies and the range in FWP and TTL IWP values well.

5.2 Conditional Occurrence of TTL Cirrus vs. Overall FWP Category

The percentages of columns with TTL cirrus are listed in Table 2 for the models and observations in each region. In CAT 1 columns, TTL cirrus typically overlay deep convection that has penetrated the TTL (see Part I). In CAT 2 columns, TTL cirrus may form the top of a thicker ice cloud or overlay other ice cloud layers. CAT 3 columns typically contain one cirrus layer that may lie partly or wholly in the TTL. Table 2 also shows the fraction of columns in which all cloud ice is in TTL cirrus alone; these are mainly a subset of CAT 3.

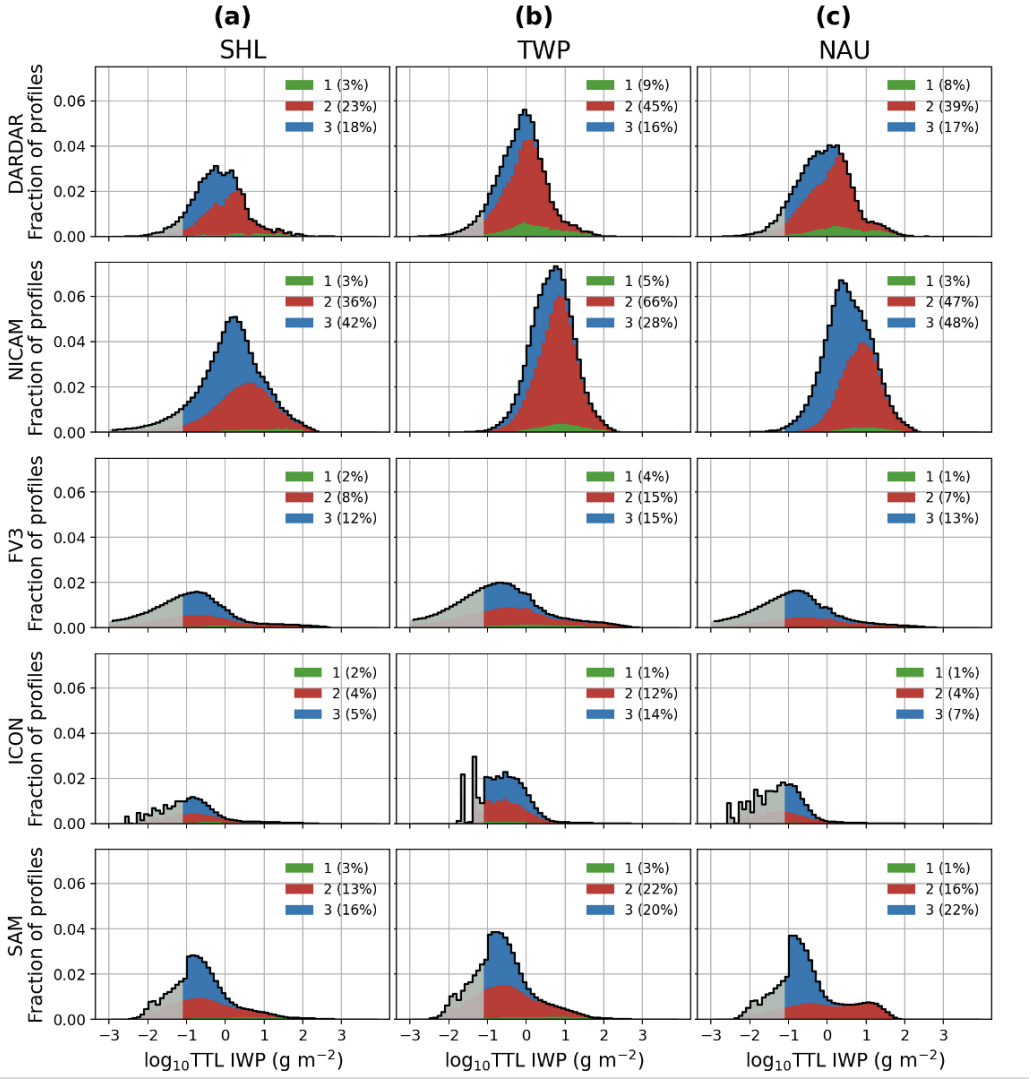

Figure 13. Stacked histograms of the TTL IWP (g m^{-2}) integrated from 14–18km for each total-column FWP category for the models and TTL FWP for DARDAR. CAT 1 (green) indicates the TTL IWP over deep convection; CAT 2 (red) is TTL cirrus above or connected to anvil and is stacked on top of CAT 1; CAT 3 (blue) is stacked on top of CAT 2. The black line shows the total histogram of TTL FWP/IWP.

Taking the TWP as a representative example, CAT 1 has the highest conditional occurrence of ice in the TTL for both observational data (84–88%) and all models (41–100%). This is consistent with the findings from Part I that nearly all of the mass flux of frozen water into the TTL is associated with deep convection. CAT 2 and CAT 3 columns have a somewhat lower conditional occurrence frequency of TTL cirrus. The large disparity between the two observational data, especially for CAT 3, may reflect the more restrictive cloud mask used by DARDAR (Cazenave et al., 2019).

NICAM has too much TTL cirrus conditioned on all three FWP categories, while ICON has too little. FV3 and SAM have plausible TTL cirrus frequency in CAT 1, but too low a TTL cirrus frequency in CATs 2 and 3.

Table 2. Frequency of TTL cirrus for the SHL, TWP, and NAU is shown for each category as a percentage of columns in that category. The frequency of isolated TTL cirrus (i.e., clouds in the TTL but no clouds below the TTL) is shown as a percentage of all columns.

	SHL				TWP				NAU			
	CAT 1	CAT 2	CAT 3	Iso TTL Ci	CAT 1	CAT 2	CAT 3	Iso TTL Ci	CAT 1	CAT 2	CAT 3	Iso TTL Ci
DARDAR	81	74	67	1	84	79	71	1	88	81	73	1
CCCM	71	32	8	3	88	49	22	6	66	24	8	2
NICAM	97	94	86	9	100	100	99	3	100	100	99	11
FV3	80	41	32	2	89	53	40	1	92	59	43	1
ICON	37	21	12	1	41	34	35	3	20	15	15	1
SAM	79	40	46	5	84	53	53	4	85	67	56	7

Note. Iso TTL Ci = Isolated TTL cirrus.

In Table 2, the fourth column for each region shows that isolated TTL cirrus are relatively rare, both in the observations and the GSRMs (except NICAM). The IWP of isolated TTL cirrus is similarly small in FV3, ICON, and SAM (0.27 g m^{-2}, 0.43 g m^{-2}, and 0.51 g m^{-2}, respectively), but NICAM has a mean TTL IWP an order of magnitude larger (3.4 g m^{-2}).

6 Summary and Discussion

We have compared the population of tropical cirrus clouds simulated by the DYAMOND GSRMs with a suite of satellite observations. We focus especially on albedo-OLR joint histograms, since they directly connect to the longwave and shortwave radiative effects of the cloud population. GSRMs explicitly resolve deep convection, the most important cirrus-initiating process, and they also resolve the resulting cirrus clouds, so it is plausible that they will simulate tropical cirrus much more realistically than conventional GCMs.

We use a statistical approach to examine the life cycle of modeled convective systems and cirrus in the DYAMOND GSRMs. Three $10^\circ \times 10^\circ$ latitude-longitude regions in the tropics are used to represent different deep convective regimes: local and remote ocean convection (TWP and NAU) and land convection (SHL).

We interpret the albedo-OLR histograms in terms of an aggregate life cycle of tropical cirrus that begins with convective systems, then transitions through anvils to cirrus with correspondingly smaller frozen water paths. The effect of deep convection on cirrus was investigated in Part I; we pick up from there. After convective injection, each model simulates a path to thin cirrus in the albedo-OLR plane. These paths are different between models in ways that are robust across the three study regions. This is due to inherent model differences in the formulation of microphysics and fine-scale dynamics that influence the typical cirrus height, optical depth, and cloud fractional coverage. As a group, the GSRMs also capture the more subtle differences in the cirrus population between the three study regions.

Many model biases in the albedo-OLR histogram and their overall TOA radiative fluxes can be traced to aspects of their simulation of cirrus clouds. Four GSRMs provided the necessary outputs for this analysis. NICAM has a low mean OLR due to its excessive cirrus extent and high cloud-top heights, but reasonably reproduces the observed cloud populations and captures the regional differences quite well. FV3 has the smallest radiative biases in the 38-day climatology and simulates deep convection well in all three regions. FV3 also consistently produces a population of bright clouds with intermediate OLR that we interpret as congestus mostly composed of liquid water. These liquid-dominated clouds are observed at NAU but not at TWP or SHL. ICON consistently fails to simulate convection and cirrus that are deep enough or cirrus that are extensive enough, leading to excessive OLR. SAM

does not simulate sufficiently optically thick deep convective cores and instead tends to generate thick anvil clouds with low albedo.

The FWP histogram is another revealing and observable metric of ice cloud properties. In GSRMs, this metric is sensitive to the ice microphysics parameterization. The models as a group produce a wide range of FWP distributions with modes and amplitudes that scatter around observed values. Modest regional differences are observed in the FWP histogram. These differences are qualitatively reproduced by the GSRMs, but are much smaller than the tropics-wide differences in the FWP histograms simulated by the four models. The shape of the observed histogram is most closely matched by NICAM, even though NICAM simulates too much cirrus cloud, especially in the TTL.

In the TTL, the models retain aspects of the total column FWP distribution but with the mode shifted to small IWPs ($\sim$0.3 g m^{-2}). While most CAT 1 deep convective columns have ice in the TTL, the majority of TTL cirrus occur in CATs 2 and 3.

One motivation for using GSRMs is to expand our understanding of tropical cirrus by filling in observational gaps, but the models need further improvements to be more useful for studying the physical properties of the TTL and the life cycle of tropical clouds. Improvements for each model could include changes to the model physics, such as improved ice microphysics. In addition, analysis would be enhanced by saving model variables such as full 3D frozen hydrometeor output for radiatively active hydrometeors, optical depths, and clear-sky radiative fluxes to directly compare radiative effects to observations. Nonetheless, GSRMs in the DYAMOND simulations provide an illuminating range of model behavior and simulated TTL cirrus which will only be enhanced by further improvements in model subgrid processes.

Sensitivity studies would help determine how microphysics parameterizations affect the presence and radiative effects of TTL cirrus. Future work could include the role of cirrus in dehydrating the TTL in GSRMs. Simulations of boreal winter from DYAMOND Winter, the second phase of the DYAMOND project, may also provide an interesting contrast to the summer simulations presented here. The GSRMs in both phases of DYAMOND provide ample opportunity for studying the TTL, its properties, and cirrus clouds.

Acknowledgments

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. DARDAR data were obtained from the ICARE Data and Services Center (`http://www.icare.univlille1.fr/projects/dardar`). The satellite data from the A-Train Integrated CALIPSO, CloudSat, CERES, and MODIS Merged Product Release B1 (CCCM) and CERES Synoptic 1-degree product were obtained from the NASA Langley Research Center Atmospheric Science Data Center (`https://doi.org/10.5067/AQUA/CERES/NEWS_CCCM-FM3-MODIS-CAL-CS_L2.RELB1`). The source code for the analysis to this paper can be found on `https://github.com/smturbev/turbeville_2021_lifecycle_paper/`. We thank Adam Sokol and Blaž Gasparini for their insightful discussions. This material is based upon work supported by the National Science Foundation through the Partnerships in International Research and Education program under grant number OISE-1743753. The corresponding author, Turbeville, is an ARCS Foundation Fellow and GO-MAP scholar.

References

Ackerman, T. P., Liou, K.-N., Valero, F., & Pfister, L. (1988). Heating rates in tropical anvils. *J. Atmos. Sci.*, *45*(10), 1606-1623. doi: 10.1175/1520-0469(1988)045⟨1606:HRITA⟩2.0.CO;2

Berry, E., & Mace, G. G. (2014). Cloud properties and radiative effects of the asian summer monsoon derived from a-train data. *Journal of Geophysical Research: Atmospheres*, *119*(15), 9492-9508. doi: 10.1002/2014JD021458

Cazenave, Q., Ceccaldi, M., Delanoë, J., Pelon, J., Groß, S., & Heymsfield, A. (2019). Evolution of dardar-cloud ice cloud retrievals: new parameters and impacts on the retrieved microphysical properties. *Atmospheric Measurement Techniques*, *12*(5), 2819–2835. doi: 10.5194/amt-12-2819-2019

Del Genio, A. (2012). Representing the sensitivity of convective cloud systems to tropospheric humidity in general circulation models. *Surv Geophys*, *33*, 637–656. doi: 10.1007/s10712-011-9148-9

Deng, M., Mace, G. G., Wang, Z., & Lawson, R. P. (2013, 04). Evaluation of Several A-Train Ice Cloud Retrieval Products with In Situ Measurements Collected during the SPARTICUS Campaign. *Journal of Applied Meteorology and Climatology*, *52*(4), 1014-1030. doi: 10.1175/JAMC-D-12-054.1

Fueglistaler, S., Dessler, A. E., Dunkerton, T. J., Folkins, I., Fu, Q., & Mote, P. W. (2009). Tropical tropopause layer. *Rev. Geophys.*, *47*(RG1004). doi: 10.1029/2008RG000267

Gasparini, B., Blossey, P. N., Hartmann, D. L., Lin, G., & Fan, J. (2019). What drives the life cycle of tropical anvil clouds? *Journal of Advances in Modeling Earth Systems*, *11*(8), 2586-2605. doi: 10.1029/2019MS001736

Hartmann, D. L., Moy, L. A., & Fu, Q. (2001). Tropical convection and the energy balance at the top of the atmosphere. *Journal of Climate*, *14*, 4495-4511. doi: 10.1175/1520-0442(2001)014⟨4495:TCATEB⟩2.0.CO;2

Hartmann, D. L., & Short, D. A. (1980, 06). On the Use of Earth Radiation Budget Statistics for Studies of Clouds and Climate. *Journal of the Atmospheric Sciences*, *37*(6), 1233-1250. doi: 10.1175/1520-0469(1980)037⟨1233:OTUOER⟩2.0.CO;2

Hollars, S., Fu, Q., Comstock, J., & Ackerman, T. (2004). Comparison of cloud-top height retrievals from ground-based 35 ghz mmcr and gms-5 satellite observations at arm twp manus site. *Atmospheric Research*, *72*(1), 169-186. (Clouds and Radiation) doi: 10.1016/j.atmosres.2004.03.015

Holton, J. R., & Gettelman, A. (2001). Horizontal transport and the dehydration of the stratosphere. *Geophysical Research Letters*, *28*(14), 2799-2802. doi: 10.1029/2001GL013148

Houze, R. A. (1981). Cloud clusters in large-scale vertical motions in the tropics. *Journal of the Meteorological Society of Japan*, *60*(1), 396-410. doi: 10.2151/jmsj1965.60.1_396

Inoue, T., Satoh, M., Hagihara, Y., Miura, H., & Schmetz, J. (2010). Comparison of high-level clouds represented in a global cloud system–resolving model with calipso/cloudsat and geostationary satellite observations. *Journal of Geophysical Research: Atmospheres*, *115*(D4). doi: 10.1029/2009JD012371

Johnson, R. H., Rickenbach, T. M., Rutledge, S. A., Ciesielski, P. E., & Schubert, W. H. (1999, 08). Trimodal Characteristics of Tropical Convection. *Journal of Climate*, *12*(8), 2397-2418. doi: 10.1175/1520-0442(1999)012⟨2397:TCOTC⟩2.0.CO;2

Kato, S., Rose, F. G., Rutan, D. A., Thorsen, T. J., Loeb, N. G., Doelling, D. R., ... Ham, S.-H. (2018, 05). Surface Irradiances of Edition 4.0 Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Data Product. *Journal of Climate*, *31*(11), 4501-4527. doi: 10.1175/JCLI-D-17-0523.1

Kato, S., Rose, F. G., Sun-Mack, S., Miller, W. F., Chen, Y., Rutan, D. A., ... Collins, W. D. (2011). Improvements of top-of-atmosphere and surface irradiance computations with calipso-, cloudsat-, and modis-derived cloud and aerosol

properties. *Journal of Geophysical Research: Atmospheres*, *116*(D19). doi: https://doi.org/10.1029/2011JD016050

Kiehl, J. T. (1994). On the observed near cancellation between longwave and shortwave cloud forcing in tropical regions. *Journal of Climate*, *7*(4), 559-565. doi: 10.1175/1520-0442(1994)007⟨0559:OTONCB⟩2.0.CO;2

Nugent, J. M., Turbeville, S. M., Bretherton, C. S., Blossey, P. N., & Ackerman, T. P. (2021). *Tropical cirrus in global storm-resolving models. part i: Role of deep convection.* (Unpublished manuscript)

Ramanathan, V., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E., & Hartmann, D. (1989). Cloud-radiative forcing and climate: Results from the earth radiation budget experiment. *Science*, *243*(4887), 57–63. doi: 10.1126/science.243.4887.57

Sokol, A., & Hartmann, D. (2020). Tropical anvil clouds: Radiative driving towards a preferred state. *JGR: Atmospheres*. doi: 10.1029/2020JD033107

Stephens, G. (2005). Cloud feedbacks in the climate system: A critical review. *Journal of Climate*, *18*(2), 237-273. doi: 10.1175/JCLI-3243.1

Virts, K. S., & Houze, R. A. (2015). Clouds and water vapor in the tropical tropopause transition layer over mesoscale convective systems. *Journal of the Atmospheric Sciences*, *72*(12), 4739 - 4753. doi: 10.1175/JAS-D-15-0122.1

Virts, K. S., Wallace, J. M., Fu, Q., & Ackerman, T. P. (2010, 10). Tropical Tropopause Transition Layer Cirrus as Represented by CALIPSO Lidar Observations. *Journal of the Atmospheric Sciences*, *67*(10), 3113-3129. doi: 10.1175/2010JAS3412.1

Wall, C. J., Hartmann, D. L., Thieman, M. M., Smith, J., William L., & Minnis, P. (2018, 11). The Life Cycle of Anvil Clouds and the Top-of-Atmosphere Radiation Balance over the Tropical West Pacific. *Journal of Climate*, *31*(24), 10059-10080. doi: 10.1175/JCLI-D-18-0154.1

Supporting Information for "Tropical Cirrus in Global Storm-Resolving Models. Part I: Role of Deep Convection"

J. M. Nugent[1], S. M. Turbeville[1], C. S. Bretherton[1,2], P. N. Blossey[1], and

T. P. Ackerman[1,3]

[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA

[2]Vulcan, Inc., Seattle, WA

[3]Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA

Contents of this file

Introduction This supplement includes the figures and table for the tropical West Pacific (TWP) region (Figures S1–S3, S7–S9; Table S1) that correspond to the Sahel-only figures and table in the main article. The time series of outgoing longwave radiation, precipitation rate, total-column and tropical tropopause layer ice water paths, and fractional areas of each category for NICAM, ICON, and SAM (Figures S4–S6) are also included; these time series are only shown for FV3 in the main article. Descriptions of all supplementary figures and the supplementary table are included in the main article.

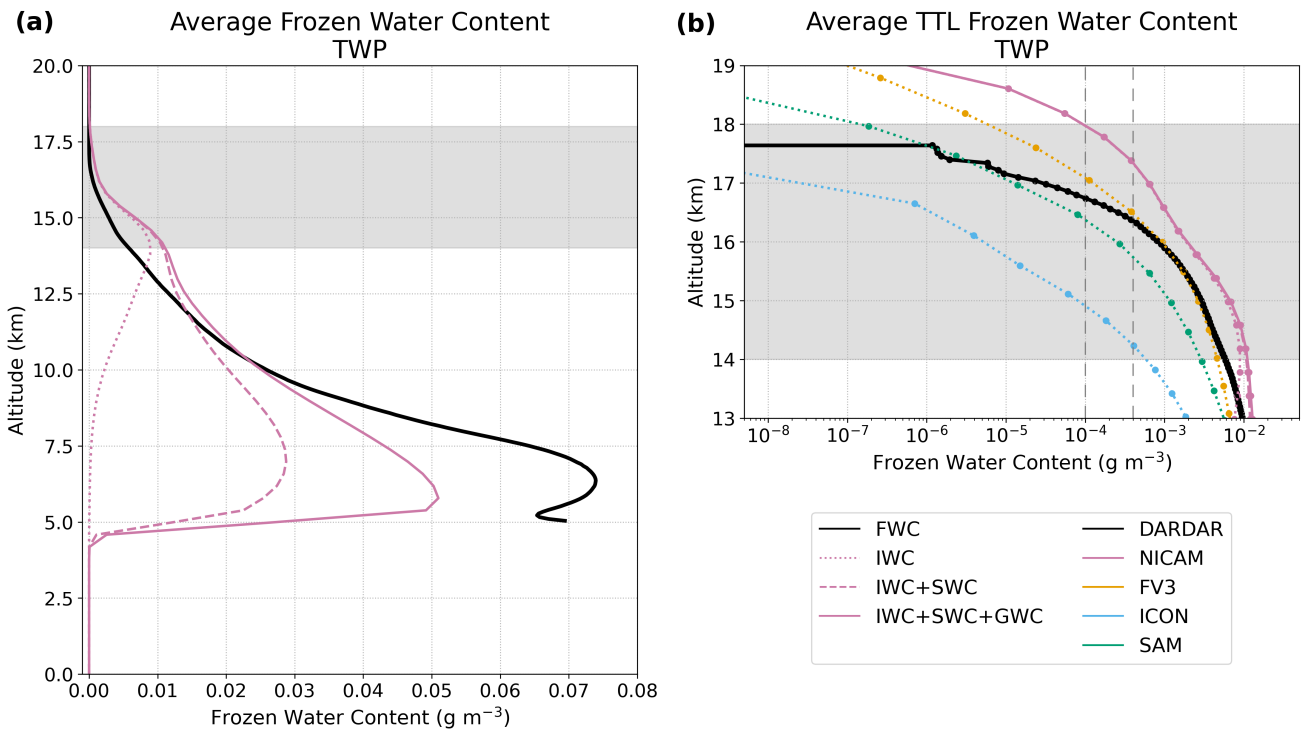

Figure S1. Same as Figure 9 in the paper, but for the TWP. The TWP region experienced anomalously high precipitation in 2009 when DARDAR measurements were taken, so the apparent underestimation of peak FWC in NICAM is likely exaggerated here.

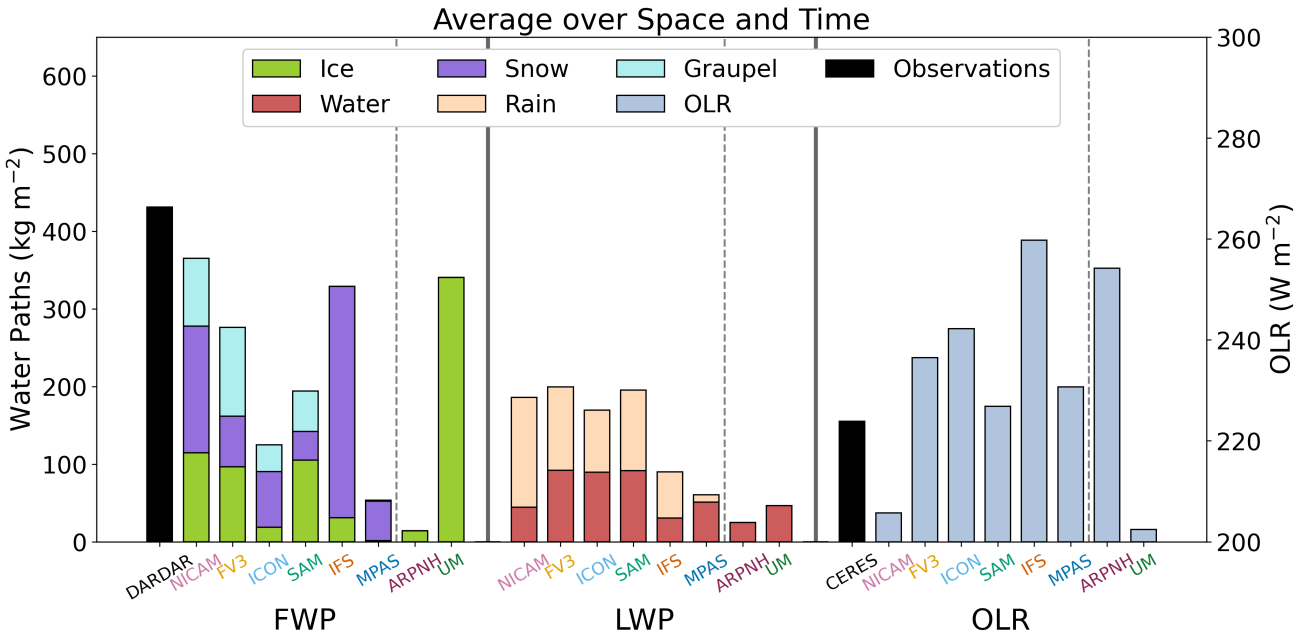

Figure S2. Same as Figure 10 in the paper, but for the TWP.

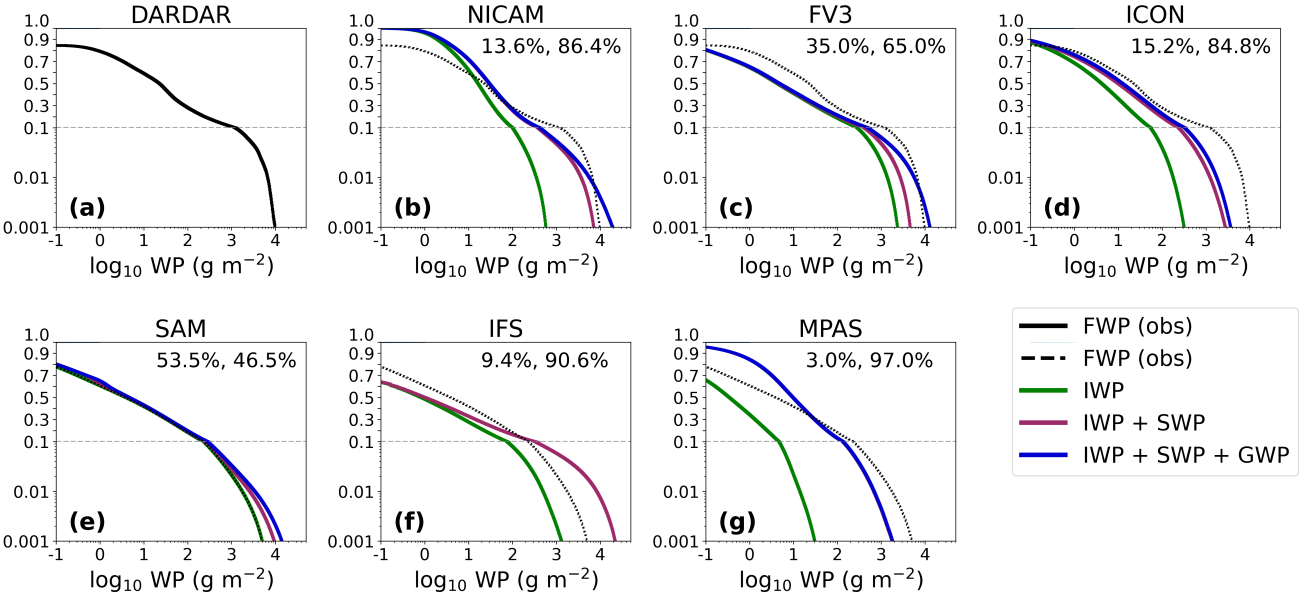

Figure S3. Same as Figure 11 in the paper, but for the TWP.

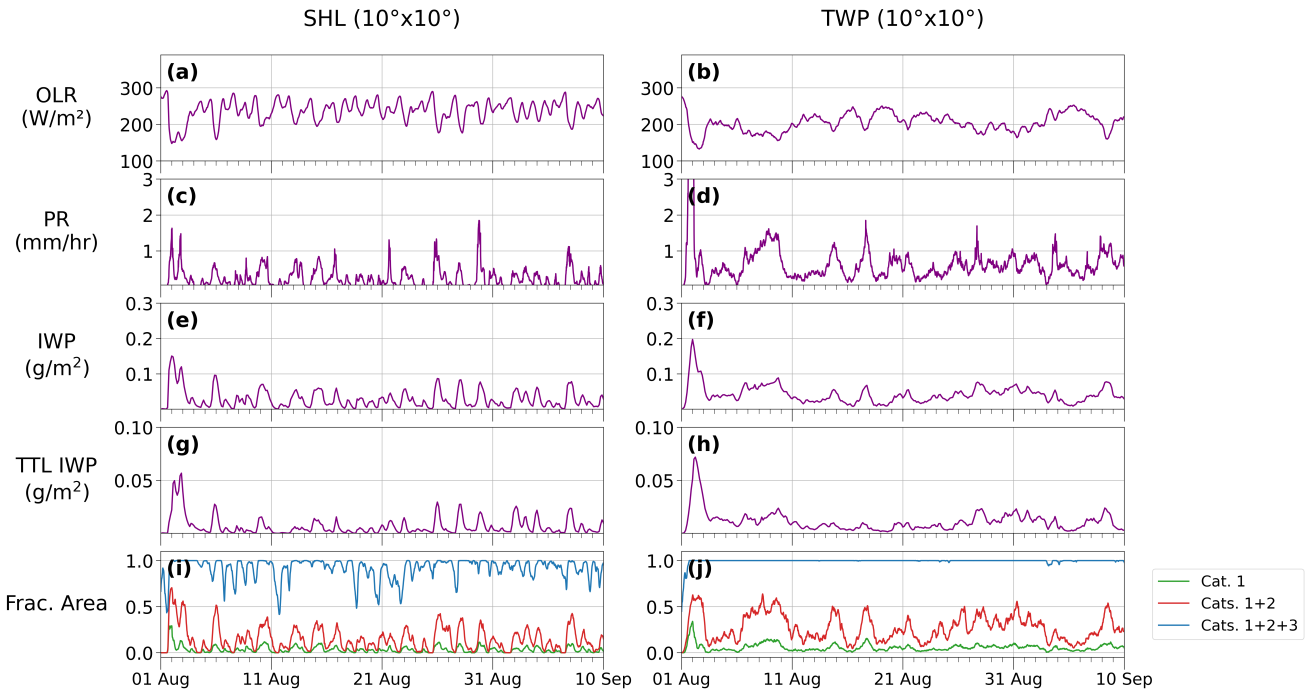

Figure S4. Same as Figure 12 in the paper, but for NICAM.

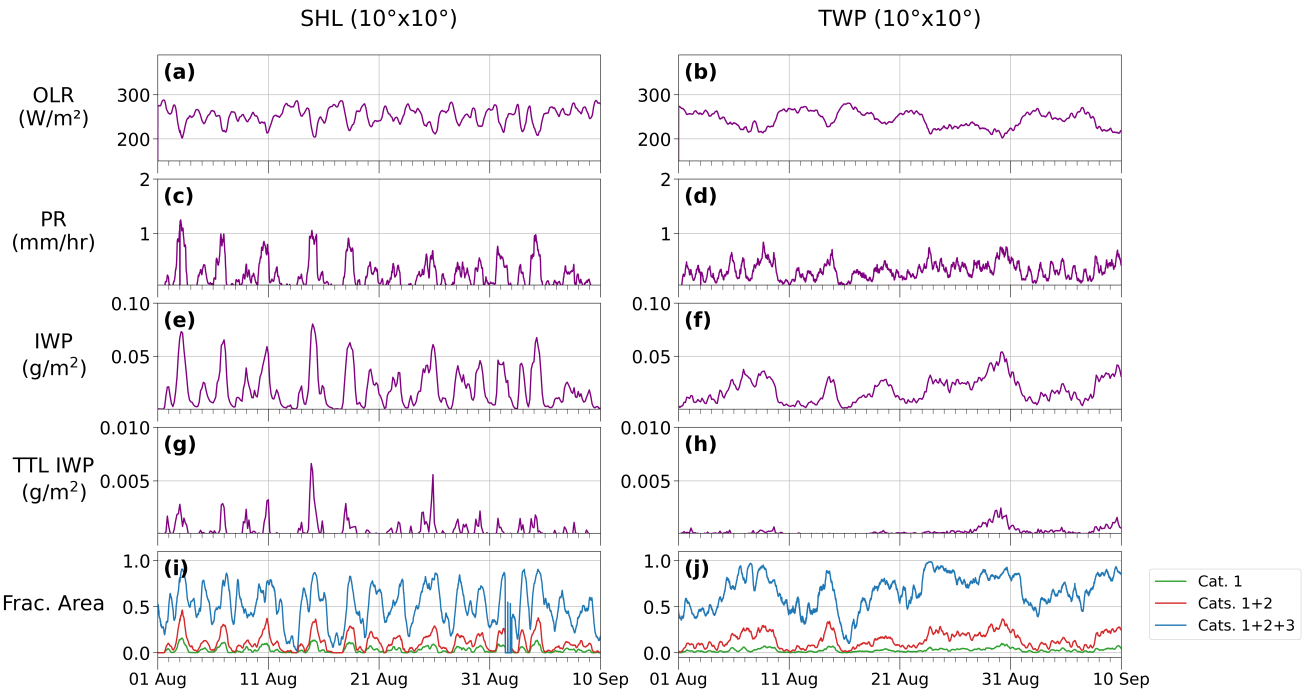

Figure S5. Same as Figure 12 in the paper, but for ICON.

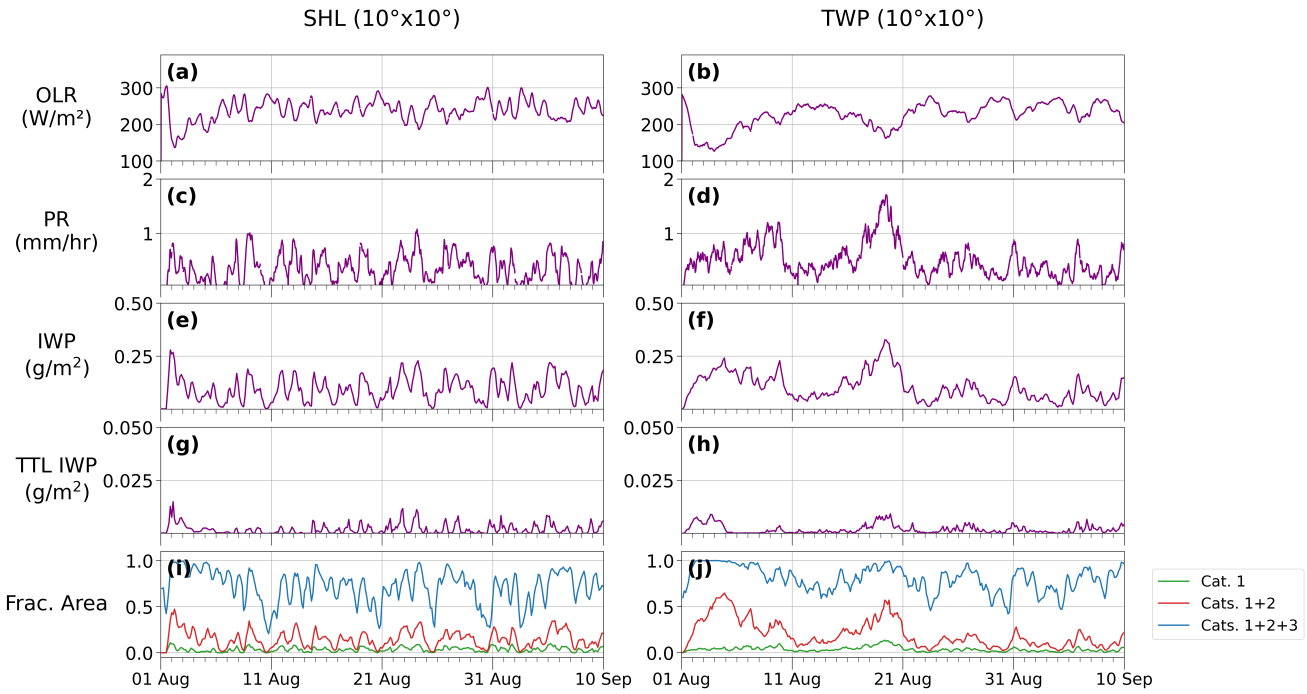

Figure S6. Same as Figure 12 in the paper, but for SAM.

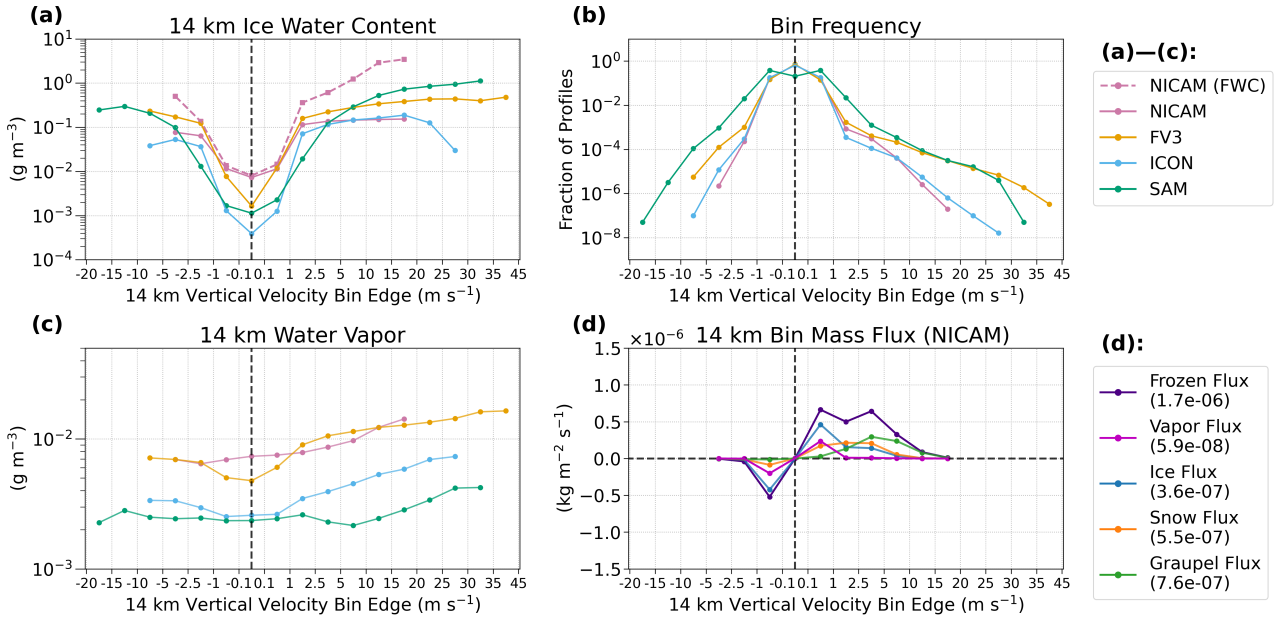

Figure S7. Same as Figure 13 in the paper, but for the TWP.

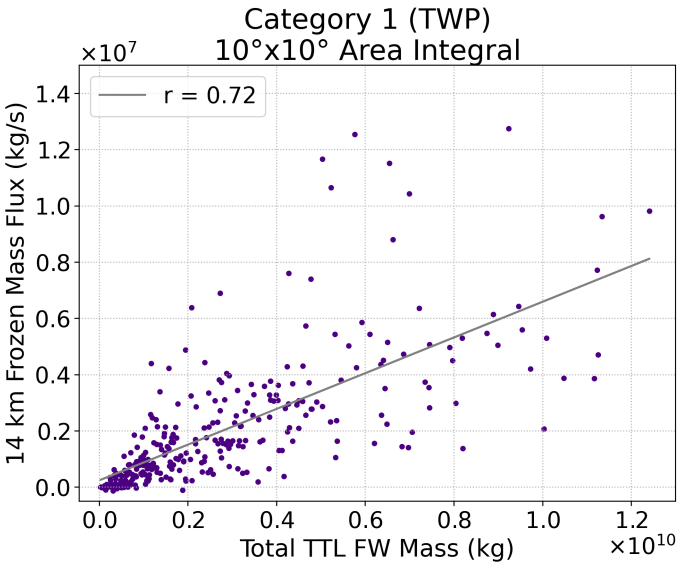

Figure S8. Same as Figure 14 in the paper, but for the TWP.

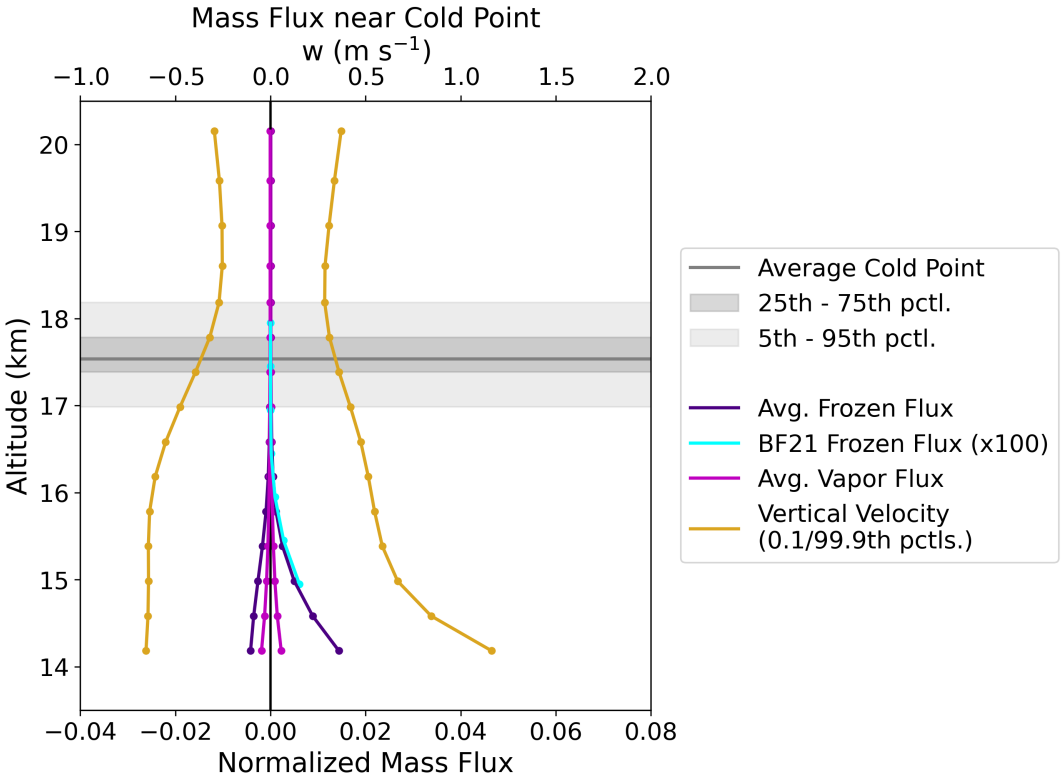

Figure S9. Same as Figure 15 in the paper, but for the TWP.

Table S1. Results from Frozen Water Path Categorization in the TWP

Model	Cat.	Freq.	Freq. of $\lvert w\rvert \geq 2.5$ m s^{-1}	Avg. vapor flux (kg m^{-2} s^{-1})	Avg. frozen flux (kg m^{-2} s^{-1})	Avg. ice flux (kg m^{-2} s^{-1})	Avg. snow flux (kg m^{-2} s^{-1})	Avg. graupel flux (kg m^{-2} s^{-1})
NICAM	1	5.4 %	0.010 %	7.1×10^{-10}	1.6×10^{-6}	3.2×10^{-7}	5.4×10^{-7}	7.6×10^{-7}
	2	66.4 %	0.022 %	4.9×10^{-9}	5.5×10^{-8}	3.9×10^{-8}	1.5×10^{-8}	4.5×10^{-11}
	3	28.1 %	0.003 %	3.2×10^{-9}	5.0×10^{-10}	3.9×10^{-10}	1.1×10^{-10}	2.2×10^{-16}
FV3	1	4.8 %	0.039 %	7.3×10^{-8}		1.8×10^{-6}		
	2	29.2 %	0.039 %	1.4×10^{-8}		6.6×10^{-8}		
	3	37.2 %	0.007 %	1.1×10^{-8}		3.1×10^{-9}		
ICON	1	2.8 %	0.001 %	1.4×10^{-9}		1.1×10^{-7}		
	2	32.1 %	0.009 %	-3.9×10^{-9}		9.5×10^{-9}		
	3	33.5 %	0.005 %	4.8×10^{-9}		4.7×10^{-10}		
SAM	1	3.7 %	0.213 %	9.9×10^{-9}		2.3×10^{-6}		
	2	39.5 %	0.055 %	5.8×10^{-9}		1.9×10^{-7}		
	3	37.0 %	0.008 %	2.8×10^{-8}		1.5×10^{-8}		

Note. Same as Table 2 in the paper, but for the TWP.

Tropical Cirrus in Global Storm-Resolving Models. Part I: Role of Deep Convection

J. M. Nugent[1], S. M. Turbeville[1], C. S. Bretherton[1,2], P. N. Blossey[1], and T. P. Ackerman[1,3]

[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA
[2]Vulcan, Inc., Seattle, WA
[3]Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA

Key Points:

- Characteristics of tropical cirrus over land and ocean in nine global storm-resolving models (GSRMs) scatter around observational ranges
- Most GSRMs reasonably simulate convective organization and rainfall, but with diverse vertical cloud structure through the upper troposphere
- Deep convection supplies most water to the tropical tropopause layer, with inter-model differences due to updraft speeds and microphysics

Corresponding author: J. M. Nugent, jnug@atmos.washington.edu

Abstract

Pervasive cirrus clouds in the upper troposphere and tropical tropopause layer (TTL) influence the climate by altering the top-of-atmosphere radiation balance and stratospheric water vapor budget. These cirrus are often associated with deep convection, which global climate models must parameterize and struggle to accurately simulate. By comparing high-resolution global storm-resolving models from the Dynamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) intercomparison that explicitly simulate deep convection to satellite observations, we assess how well these models simulate deep convection, convectively generated cirrus, and deep convective injection of water into the TTL over representative tropical land and ocean regions. The DYAMOND models simulate deep convective precipitation, organization, and cloud structure fairly well over land and ocean regions, but with clear intermodel differences. All models produce frequent overshooting convection whose strongest updrafts humidify the TTL and are its main source of frozen water. Inter-model differences in cloud properties and convective injection exceed differences between land and ocean regions in each model. We argue that global storm-resolving models can better represent tropical cirrus and deep convection in present and future climates than coarser-resolution climate models. To realize this potential, they must use available observations to perfect their ice microphysics and dynamical flow solvers.

Plain Language Summary

High-altitude tropical cirrus (ice) clouds influence the earth's climate by reflecting sunlight, trapping upwelling radiative energy from the earth's surface, and affecting the temperature and humidity of the upper atmosphere. These clouds are initiated by systems of strong thunderstorms, whose most vigorous updrafts loft water vapor and ice high into the atmosphere. Computer models used to study the global climate struggle to accurately represent tropical thunderstorms because their updrafts are far narrower than the width of a modeled grid cell. Models with very fine grids can better represent the air flows that form these clouds. We investigate how well several fine-grid models reproduce observed characteristics of tropical thunderstorm systems and cirrus. We find generally good agreement but also substantial differences between individual models, mainly because of their diverse ways of representing ice and snow formation and their evolution. With further observationally-motivated improvements, such fine-grid models should enable more reliable simulations of the role of tropical cirrus in our changing climate.

1 Introduction

Upper-tropospheric cirrus influence the climate through local radiative heating. These cirrus extend throughout the tropics and can be advected up to 1000 km during their long lifetimes (Luo & Rossow, 2004). Jensen et al. (1996a) found that very thin cirrus can warm the surrounding atmosphere by a few Kelvins per day. Due to the prevalence of the cirrus, this heating alters the top-of-atmosphere radiation balance (Lee et al., 2009; Haladay & Stephens, 2009).

Tropical cirrus in the upper troposphere are strongly related to deep convection. Areas with a high occurrence of cirrus clouds are often collocated with frequent convection (e.g., Lee et al., 2009; Sassen et al., 2009; Schoeberl et al., 2018). Near the tropopause, cirrus can form through convective anvil detrainment as well as in situ ice nucleation (Jensen et al., 1996b). Some of the cirrus formed in situ may also be related to convection if the ice nucleation results from cooling caused by gravity wave perturbations (Dessler et al., 2006; Jensen et al., 2016; Krämer et al., 2016). Pervasive, mostly optically thin cirrus characterize the transition region between the upper troposphere and lower stratosphere, known as the tropical tropopause layer (TTL; see review article by Fueglistaler et al., 2009). Several definitions for the TTL boundaries have been proposed in the literature;

here, we take the TTL to be the 14–18 km layer in the tropics. The formation and maintenance of TTL cirrus are of particular interest to us. Like other upper-tropospheric cirrus, TTL cirrus are associated with deep convection. Overshooting convection that reaches into the TTL is especially important for TTL cirrus as it injects water vapor and ice crystals into the layer to support cirrus formation (e.g., Massie et al., 2002; Luo & Rossow, 2004; Mace et al., 2006; Krämer et al., 2016) and can ultimately alter the TTL composition (Fierli et al., 2011; Virts & Houze, 2015). Additionally, Jensen et al. (1996a) determined that cirrus near the tropopause formed outside of convection act to dehydrate the stratosphere. Both TTL cirrus and changes in stratospheric water vapor concentrations radiatively impact large-scale upper-atmospheric circulations and the global climate (Solomon et al., 2010).

Overshooting convection drives water transport into the upper troposphere and is a major issue for global climate models (GCMs) in simulating tropical cirrus clouds (Fueglistaler et al., 2009). Deep convection is not resolved by the typical GCM horizontal grid spacings of 25–200 km and must be parameterized, which is notoriously challenging due to the complex small-scale structure of deep convection. The parameterizations inevitably lead to substantial biases and intermodel differences in simulated deep convection and related clouds. For instance, most GCMs simulate the diurnal cycle of convection over land poorly, with a maximum in precipitation often occurring too early in the day (Yang & Slingo, 2001; Khairoutdinov et al., 2005; Chao, 2013). An accurate diurnal cycle is needed to simulate tropical cirrus because the timing of anvil detrainment may affect the cirrus characteristics and thus their radiative effects (Sokol & Hartmann, 2020). Land convection is particularly important for TTL cirrus because it penetrates the TTL more often (Liu & Zipser, 2005) and is more intense than ocean convection (Yang & Slingo, 2001; Zipser et al., 2006). However, different convective parameterizations in GCMs disagree on the representation of convective fluxes and often struggle to reproduce convection that overshoots the cold point (Arteta et al., 2009).

Cloud-resolving models (CRMs) have sufficiently fine horizontal grid spacings to explicitly resolve deep convection without parameterization. An early CRM study by Bechtold et al. (2004) found that the precipitation maximum over tropical land realistically shifted to later in the day with nonparameterized deep convection. This improvement in the simulated diurnal cycle has been reproduced by numerous studies (e.g., Duda & Gallus, 2013; Berthou et al., 2019).

Increased computing power has since allowed for the development of global storm-resolving models (GSRMs), global CRMs of 1–5 km horizontal grid resolution, that can resolve individual convective storms. Recently, several GSRMs were brought together as part of the Dynamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) project (Stevens et al., 2019). The DYAMOND project provides a novel opportunity for the intercomparison of TTL cirrus and convection modeled at such high resolutions. Because of the expected improvements from explicitly simulating deep convection, the DYAMOND models have the potential to more accurately model tropical cirrus in the upper troposphere and TTL in a changing climate. But is this potential realized, even in the present climate that we can observe? Other parameterizations still needed in GSRMs, especially ice microphysics, could still be major sources of model error. GSRM fidelity in accurately simulating strong grid-scale convective updrafts and downdrafts also needs to be assessed.

The goal of this paper is to use a variety of observations, mainly from satellites, over representative land and ocean regions to assess how well tropical deep convection and convectively generated cirrus are simulated in the DYAMOND GSRMs. In Part II of this paper, Turbeville et al. (2021) analyze the full life cycle of tropical cirrus and their radiative effects in the DYAMOND GSRMs. The remainder of the paper is organized as follows: Section 2 provides an overview of the methods, including descriptions of the models, datasets, and study regions. We assess the fidelity of the DYAMOND models

in simulating both land and ocean convection in section 3. We examine the differences in the model output of microphysical variables related to cirrus in section 4. In section 5, we relate deep convection to the simulated cirrus by analyzing the convective injection of water vapor and ice into the TTL. Finally, section 6 summarizes and discusses the results and their implications for our understanding of modeling tropical cirrus.

2 Methods and Data

2.1 DYAMOND Models

The DYAMOND initiative brought together nine different GSRMs with 5 km or finer horizontal grid spacings and explicit, nonparameterized deep convection (Stevens et al., 2019). These GSRMs each have their own specific design described in detail in Stevens et al. (2019), but all models were initialized from the same meteorological analysis and run freely for a total of 40 days from 1 August to 10 September 2016. In addition to the high spatial resolution, the DYAMOND GSRMs also have high temporal resolution, with output saved every 15 minutes for 2D fields and every 3 hours for 3D fields.

Our analysis focuses primarily on the four models: NICAM, FV3, ICON, and SAM (starred models in Table 1). Hereafter we refer to these models collectively as NFIS. This limited subset of models allows for a detailed, in-depth evaluation of such high resolution simulations while maintaining a manageable output volume. To place the NFIS models into context with the rest of the intercomparison, we also evaluate the output of certain 2D fields for the other DYAMOND models: IFS, ARPNH, MPAS, and UM (Table 1). We have omitted the ninth model, the Goddard Earth Observing System (GEOS), because the saved DYAMOND run included a poor implementation of its microphysics scheme that distorted the model results (Putman, 2020).

Table 1. List of DYAMOND Models

Model name	Abbreviation	Horiz. grid spacing (km)
Non-hydrostatic Icosahedral Atmospheric Model	NICAM*	3.5
Finite-Volume Cubed-Sphere Dynamical Core	FV3*	3.3
Icosahedral Non-hydrostatic Model	ICON*	2.5
Global System for Atmospheric Modeling	SAM*	4.3
Integrated Forecast System	IFS	4.8
ARPEGE-NH	ARPNH	2.5
Model for Predicting Across Scales	MPAS	3.8
United Model	UM	7.8

Note. We focus on the first four models (starred) and use the others for comparisons of certain 2D variables.

We have analyzed the NFIS models at their native grid scales where possible; however, some interpolation was necessary for ICON and FV3. ICON output is saved on an icosahedral grid, so for spatial plots all variables were first interpolated via nearest neighbors onto the finest latitude-longitude grid permitted by the number of cells. For FV3, the 3D hydrometeor profiles, 2D integrated water paths, and pressure were regridded through first-order conservative remapping onto a regular latitude-longitude grid while all other variables were saved on the native cubed sphere grid. These regridded variables had a larger number of cells in the study regions than the native variables, so for any analysis involving both types, all variables were first aligned onto a new latitude-longitude grid through nearest neighbors interpolation. In addition, FV3 uses a hybrid pressure-sigma vertical coordinate and the altitude was only saved on a limited number of pressure levels. We calculated a regional-mean altitude of all model levels from horizontal regional

averages of pressure, temperature, and specific humidity by integrating the hypsometric equation up from the known horizontal-mean surface elevation and pressure.

Most DYAMOND models have one-moment microphysics schemes, which partition the frozen water into cloud ice, snow, and graupel. Unlike the models, satellite measurements generally cannot distinguish between the types of frozen hydrometeors. The observed "ice" is most accurately compared to the total sum of what each model defines as cloud ice, snow, and graupel, which we refer to here as the total frozen water. In addition to the 3D profiles of cloud ice, all models except ARPNH and UM saved output for the 2D column-integrated ice water path (IWP), snow water path (SWP), and graupel water path (GWP), whose sum we call the frozen water path (FWP). However, NICAM is the only model that also saved the 3D output of snow and graupel, so we can only directly compare 3D profiles of the observed frozen water content (FWC) to NICAM output. Section 4 further discusses the implications of this limitation.

Finally, we have omitted the first 48 hours of each model from any averaging or sampling to avoid contamination from the spin-up period, which includes an initial shock in precipitation in NICAM. Therefore, "time-mean" refers to the average of days 3–40 in each DYAMOND model run (i.e., 3 August to 10 September 2016).

2.2 Datasets

Since the DYAMOND models were run freely, they will not reproduce specific, observed weather systems after the first few days. Therefore, we assess model fidelity by comparing time means and distributions of their output to a set of observational and reanalysis data from the same time of year.

We take FWC measurements from DARDAR-CLOUD v2.1.1, which combines radar and lidar retrievals from CloudSat and CALIPSO (Delanoë & Hogan, 2008, 2010). These satellites, part of NASA's A-Train constellation, cross the equator at approximately 01:30 and 13:30 local time for nighttime and daytime measurements, respectively (Stephens et al., 2002). The combination of radar and lidar data in DARDAR (raDAR-liDAR) capitalizes on the detection strengths of both instruments to enable retrievals in areas of deep convection as well as thin cirrus. The DARDAR data have a 1.1 km horizontal grid spacing and 60 m vertical grid spacing (e.g., Sokol & Hartmann, 2020). We use FWC retrievals from July-August-September (JAS) 2009, which corresponds to the season of the DYAMOND simulations. The data are then integrated vertically to compute the observed FWP. Although solar noise increases the uncertainty in lidar measurements during the day, we include both daytime and nighttime retrievals to avoid diurnal bias.

For precipitation, we build an 11-year climatology during the DYAMOND period by merging estimates from 1 August to 10 September 2006–2016. We use the Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) 3B42 version 7 product, which contains 3 hourly data at a spatial resolution of 0.25° (Huffman et al., 2007, 2010). This product combines microwave and infrared rainfall estimates with precipitation gauge measurements using the algorithm described by Huffman et al. (2007) with modifications for the current version detailed in Huffman and Bolvin (2018).

We obtain retrievals of outgoing longwave radiation (OLR) from the sun-synchronous Clouds and the Earth's Radiant Energy System (CERES) satellite for JAS 2007–2010. The CERES data are contained in the CALIPSO, CloudSat, CERES, and MODIS merged data product (CCCM) described by Kato et al. (2010, 2011). CERES has a ∼20 km footprint and also belongs to the A-Train constellation.

For temperature and humidity, we use the high-resolution European Center for Medium Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) from 1 August to 10 September 2016. The ERA5 dataset contains global hourly output on 37 pressure levels at a res-

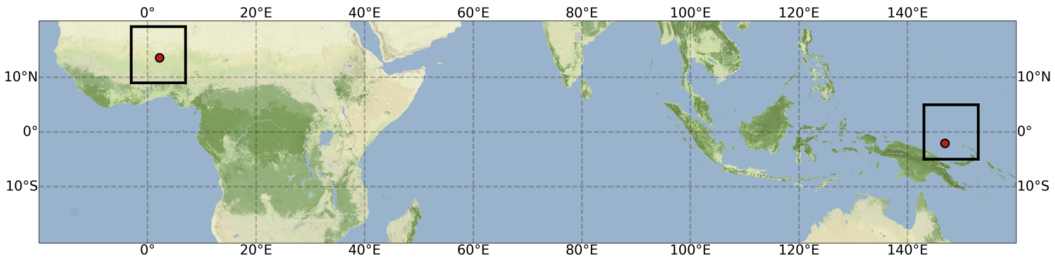

Figure 1. Locations of the analysis regions in (left) the Sahel, 3°W–7°E, 9°N–19°N; and (right) the tropical West Pacific, 143°E–153°E, 5°S–5°N. Red dots indicate locations of the (left) AMMA and (right) ARM sites.

olution of 31 km regridded to 0.25° (Hersbach et al., 2020). We calculate the altitude of each pressure level from the ERA5 geopotential variable.

2.3 Study Regions

This study examines two representative 10°×10° latitude-longitude regions in the tropics as shown in Figure 1: a continental area over the Sahel (SHL) in western Africa and an oceanic area over the tropical West Pacific (TWP).

The Sahel region was chosen to represent tropical land for several reasons. This area experiences frequent deep convection during the DYAMOND period when the West African Monsoon is active. Furthermore, a large proportion of overshooting convection that extends above 14 km in the tropics occurs over Africa (Liu & Zipser, 2005). The Sahel region encompasses several different climate regions, from the moist and convectively active savanna to the Sahara desert where detrained anvils are the main cloud type during the monsoon season (Futyan et al., 2004). Finally, this region is approximately centered over the African Monsoon Multidisciplinary Analysis (AMMA) measurement site in Niamey, Niger, where the RADAGAST field campaign took place in 2006 (Slingo et al., 2008). RADAGAST generated a rich suite of measurements of the overlying atmospheric column that may in the future be climatologically compared with the DYAMOND output.

The TWP region contains frequent ocean convection, a valuable complement to the Sahel region. The TWP box is approximately centered over the Atmospheric Radiation Measurement (ARM) site in Manus, Papua New Guinea (Long et al., 2016) and is the primary study area used in Part II.

2.4 Mass Flux Calculation

One goal of this study is to analyze how the DYAMOND models bring the ice and moisture necessary to support cirrus cloud formation and maintenance into the TTL. To this end, we examine the vertical advective mass fluxes of frozen water and water vapor at 14 km, a representative height of the TTL base. The mass fluxes (F, kg m^{-2} s^{-1}) of water vapor, frozen water, cloud ice, snow, and graupel are calculated in equation (1):

$$F = w\rho q^* \tag{1}$$

Here, w is the vertical velocity (m s^{-1}), ρ is the air density (kg m^{-3}), and q^* (kg kg^{-1}) is either the specific humidity for the water vapor flux or the specific water content (frozen, cloud ice, snow, or graupel) for the other fluxes. The w used in equation (1) does not include any ice fall velocity, which was not an output archived by the models. We calculate the water vapor and cloud ice mass fluxes for all NFIS models. Since FV3, ICON,

and SAM do not save the 3D model outputs of snow and graupel that are needed to determine the total FWC, we can only calculate the frozen water, snow, and graupel mass fluxes for NICAM.

2.5 Frozen Water Path Categorization

We will use the FWP to divide ice-containing air columns into three categories with distinct physical characteristics, following Sokol and Hartmann (2020):

Category 1: FWP $\geq$ 1000 g m^{-2} (deep convection)
Category 2: 10 $\leq$ FWP $<$ 1000 g m^{-2} (thick "anvil" cirrus)
Category 3: 0.1 $\leq$ FWP $<$ 10 g m^{-2} (thin cirrus).

The models can simulate cirrus layers so thin that lidar cannot distinguish them from clear sky. The lower threshold of Category 3 is chosen to exclude most such cirrus layers from our model-observation comparisons. It corresponds to a 200 m thick cirrus layer (narrower than the 400–600 m vertical grid spacing of the NFIS models within the TTL) with an average FWC of 5×10^{-4} g m^{-3}, which Deng et al. (2013) found to be the smallest value measurable by satellite-based lidar. Columns with FWP $<$ 0.1 g m^{-2} are considered to be cirrus-free.

The SAM 2D water paths require additional processing. These files were saved with 2-byte compression that quantizes values into integer multiples of 1/64,000th of the field maximum. While IWP and liquid water path (LWP) can be computed by integrating profiles of cloud ice and liquid in the 3D outputs, we statistically correct the errors due to quantization in the snow, graupel, and rain water paths by adding random perturbations to any values below 1 g m^{-2}. This categorization, as well as the SAM processing, is also used to evaluate TTL cirrus in Part II.

3 Convection

In this section, we evaluate the performance of the DYAMOND models in simulating both land and ocean convection.

3.1 Thermodynamic Background

Since the DYAMOND models are not nudged towards observations, they each settle into their own climates over the course of the model run. We compare the model output for temperature and relative humidity to ERA5 reanalysis data in order to understand the mean thermodynamic state of each model. The relative humidity with respect to ice in the DYAMOND models was calculated using the equation for ice-saturation vapor pressure from Murphy and Koop (2005).

Across both the Sahel and TWP study regions, the time-mean TTL temperature profiles vary between the models but all lie within 5 K of the ERA5 reanalysis (Figure 2a–b). The cold point temperatures are a few degrees lower and closer to saturation in the TWP than in the Sahel. In both regions, the cold points are located within approximately 1.5 km of ERA5 and have a temperature range of about 3.5 K (Figure 2a–b).

Despite their microphysical diversity, the models all simulate mean relative humidity profiles that scatter within ∼15% of ERA5 below the cold point and fall well within one standard deviation of the ERA5 daily mean (Figure 2c–d), which is not true of the temperature profiles. This disagreement in temperature but not relative humidity suggests that convection has a lasting influence on the TTL in the NFIS models; the mean temperature profiles alone do not explain the relative humidity profiles, so the infrequent

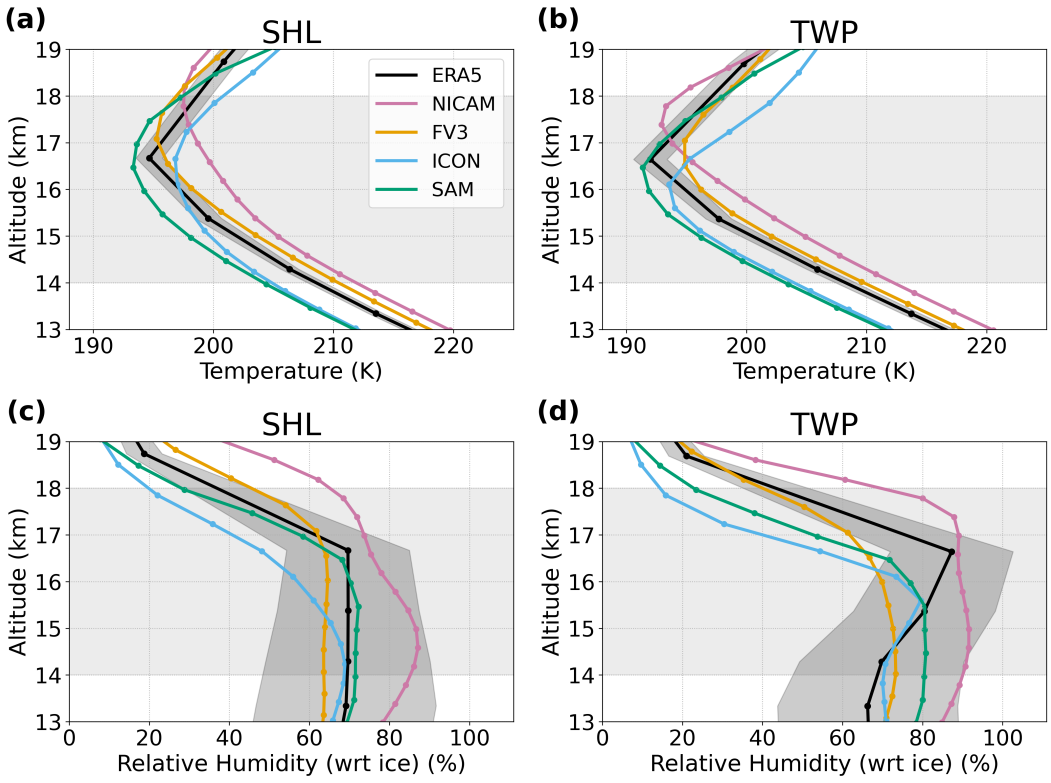

Figure 2. Time- and regional-mean vertical profiles of (top) temperature and (bottom) relative humidity with respect to ice for the (left) Sahel and (right) TWP for the NFIS models and ERA5 reanalysis data. The standard deviation of the daily mean ERA5 data is shaded in dark gray. The TTL is shaded in light gray.

periods of deep convection likely help to humidify the upper troposphere to around 70–90%, consistent with Ueyama et al. (2018).

3.2 Accumulated Precipitation

The DYAMOND models scatter around the observed accumulated precipitation in both regions, with generally higher precipitation over the TWP (Figure 3b) than the Sahel (Figure 3a). The ratio of TWP to Sahel 40-day accumulated precipitation ranges considerably from 1.0 (ARPNH, IFS) to 3.6 (UM), compared to an 11-year TRMM climatological estimate of 1.6 for the DYAMOND period in 2006–2016. Since the models are free-running and a 40-day average is short, this ratio may not be fully representative of each model's climatological behavior, but Figure 3 suggests some models systematically precipitate more than others in each region.

The DYAMOND models as a whole reasonably simulate the total precipitation in both regions with most model accumulation falling within the 11-year TRMM climatological range. Among the NFIS models, none simulate precipitation notably better than the others. SAM consistently accumulates about 50% more precipitation than the climatological rate. NICAM, FV3, and ICON are very close to the climatological rate in the Sahel, but ICON underestimates and FV3 overestimates precipitation in the TWP. The strong NICAM precipitation on 1 August in the TWP results from the initial shock (see section 2.1); thereafter, the total accumulated precipitation in NICAM tracks the top end of the TRMM range.

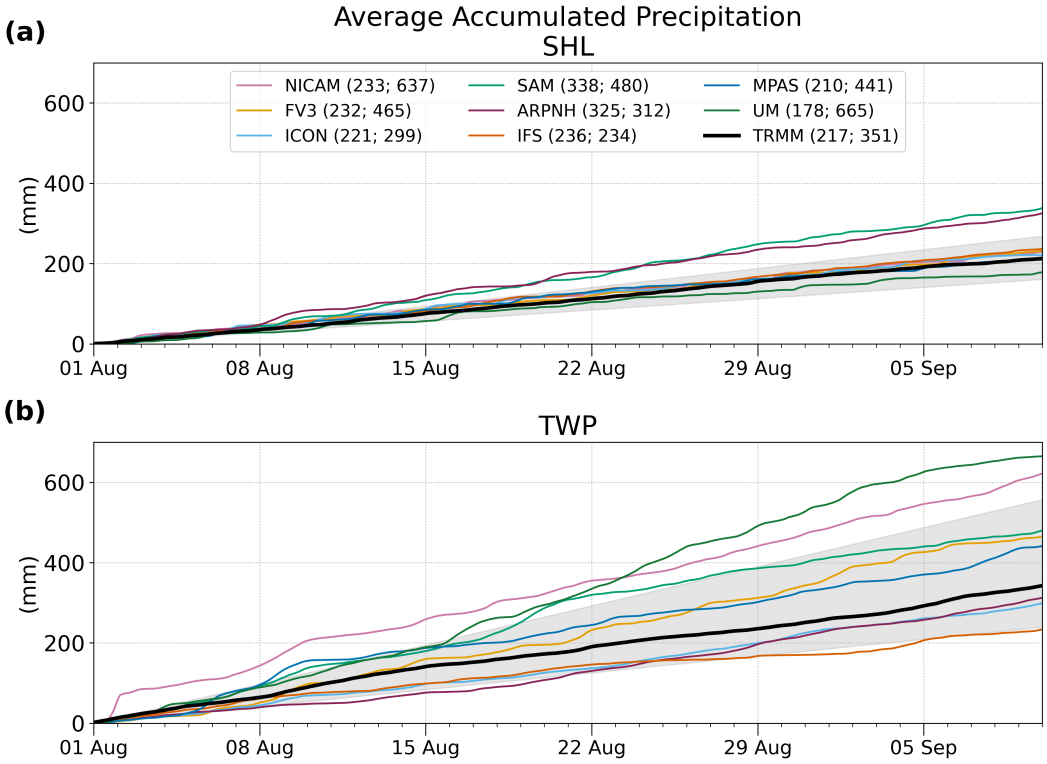

Figure 3. Regional-mean accumulated precipitation at each time step for the (a) Sahel and (b) TWP. The black line represents the 2006–2016 TRMM climatological mean accumulation during the DYAMOND time period. The gray shading shows the range in accumulation observed by TRMM over this 11-year period. The numbers in the legend are the 40-day accumulations in (left) the Sahel and (right) the TWP in mm.

3.3 Texture of Convection

The "texture" (fine-scale morphology) of precipitation is an illuminating diagnostic of how well models simulate the details of convection and differences between land and ocean regions (Inoue et al., 2008). Figures 4 and 5 show characteristic snapshots of precipitation in the NFIS models for the Sahel and TWP, respectively. For each model, we have selected the output times with the maximum regional-mean precipitation rate after the 48-hour spin-up period.

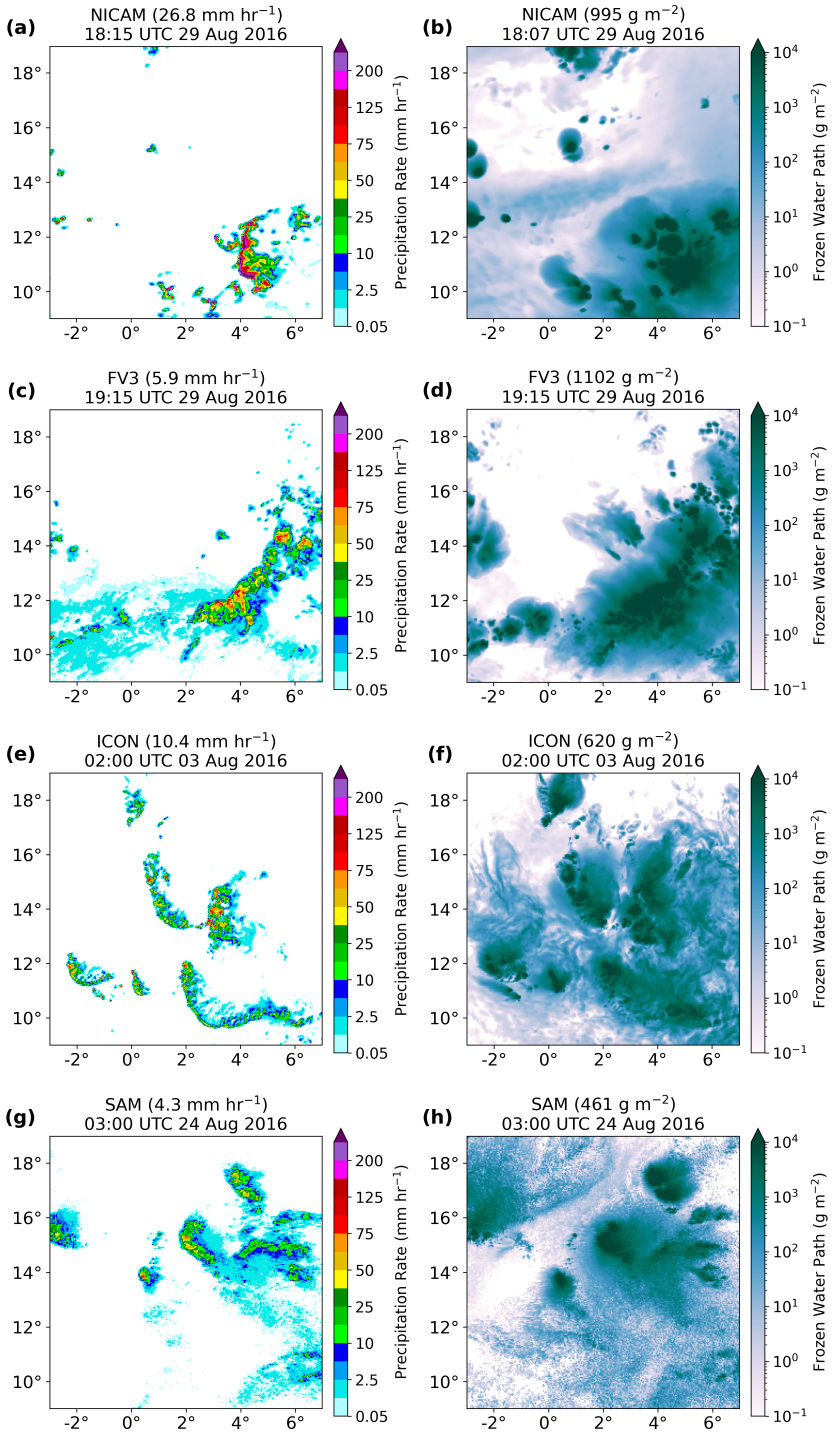

Figure 4. Snapshots of (left) precipitation rate and (right) frozen water path (FWP) over the entire 10° × 10° box in the Sahel. Rates below 0.05 mm hr^{-1} and FWPs below 0.1 g m^{-2} are masked. The numbers in each panel title indicate the regional-mean precipitation rate or FWP for each snapshot. The timing of the saved FWP output for NICAM is offset from the precipitation by 7.5 minutes.

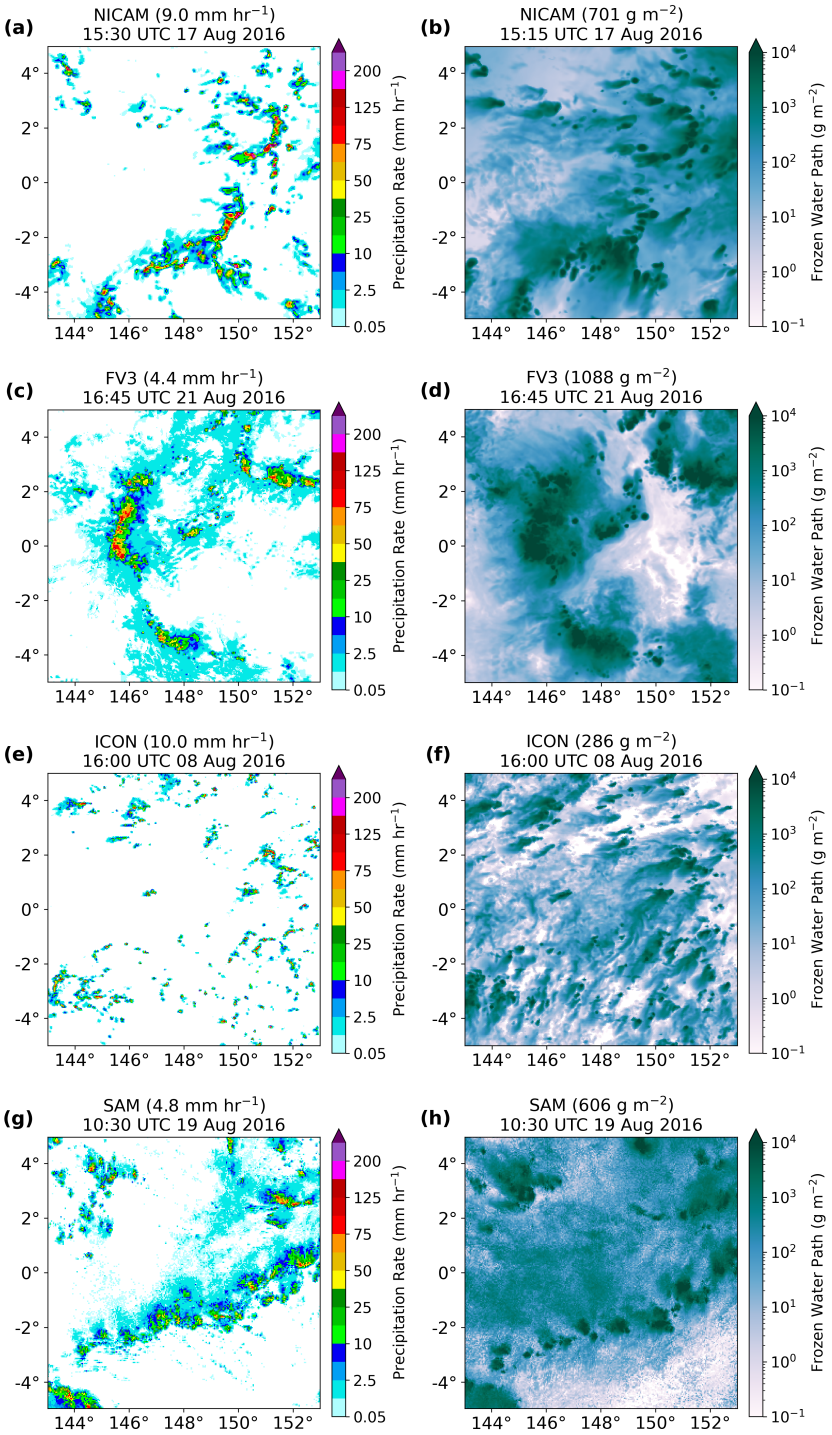

Figure 5. As in Figure 4, but for the TWP.

In the Sahel, the NFIS models all produce realistic squall lines with narrow, short-lived bands of precipitation that are characteristic of the region (Figure 4) (Redelsperger & Lafore, 1988; Redelsperger et al., 2002). However, FV3 and SAM have an excess of light precipitation that is uncharacteristic of the Sahel (Figure 4c, g). Much of this light precipitation comes from shallow cumulus clouds (parameterized in FV3) which are not

apparent in the FWP snapshot. The areal extent of thin cirrus clouds in NICAM is much larger than in the other models (Figure 4b). In all models, there is a strong collocation between the FWP and precipitation fields in which the precipitation tracks with the cloud field.

Similarly in the TWP, NICAM, FV3, and SAM reproduce the large, widespread convective systems typically observed over tropical oceans (Figure 5) (Bousquet & Chong, 2000). FV3 and SAM again have expansive regions of unrealistically light precipitation, but to a lesser extent than in the Sahel. The precipitation in ICON consists of small pockets of "popcorn" convection scattered throughout the region (Figure 5e). As in the Sahel, areas of intense precipitation collocate with higher FWPs. In NICAM, the entire $10° \times 10°$ TWP box has frozen water in the atmospheric column (Figure 5b). Although there is diversity in the details of precipitation and FWP in individual models that is generally consistent in the Sahel and TWP, the models simulate a reasonable texture of both precipitation and convection and capture the expected regional differences in land and ocean convection.

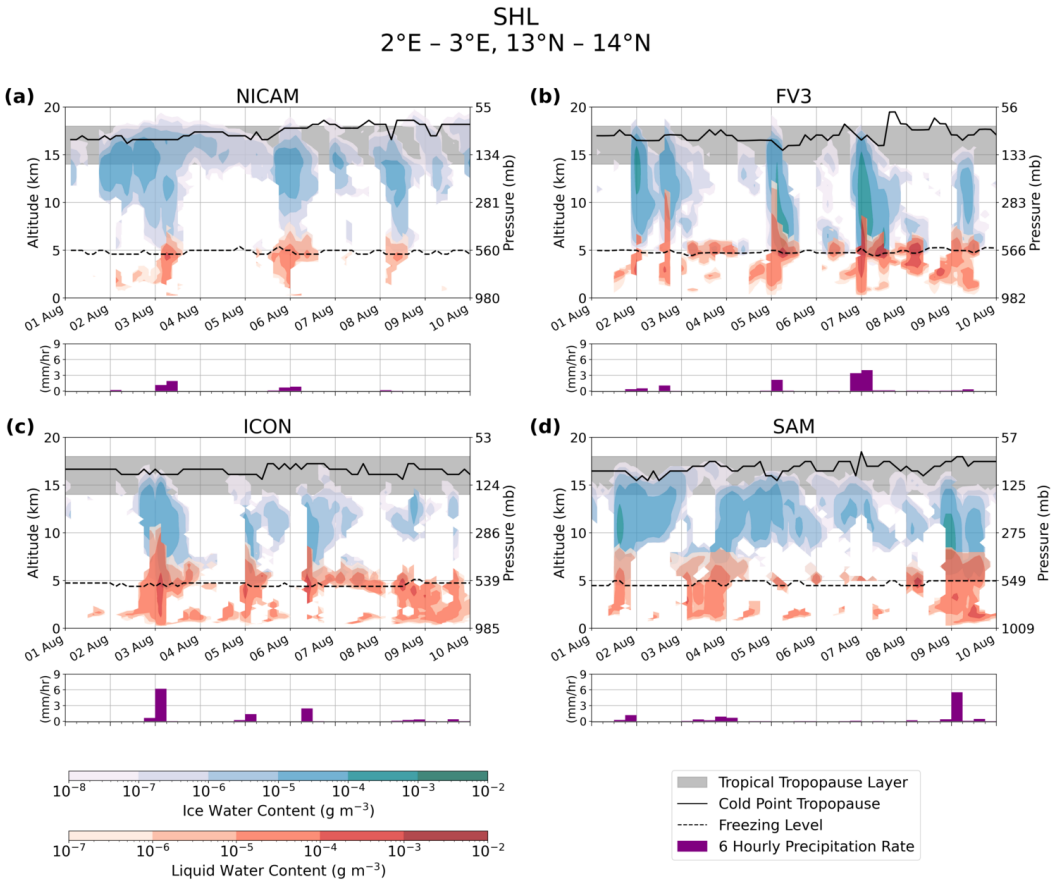

Figure 6. Time series of the convective cloud structure for the NFIS models in the Sahel. Ice water content is shaded in blue and liquid water content is shaded in red. Purple bars indicate the 6 hourly precipitation rate. The solid black line shows the cold point tropopause while the dashed black line shows the melting level. The TTL is shaded in gray. Model output is averaged over a $1° \times 1°$ box located at 2°E–3°E, 13°N–14°N, approximately centered over the AMMA site in Niamey, Niger. The first 48 hours (1–2 August) should be regarded as the model spin-up period.

As with precipitation, the vertical structure of convective clouds is plausibly simulated in the NFIS models, but with clear intermodel differences. Figures 6 (SHL) and 7 (TWP) show time series of the vertical cloud structure within a 1° × 1° box during the first 10 days of the model run. These figures highlight the qualitative differences in land and ocean convection (e.g., Liu et al., 2007), which the models capture nicely. Convection in the Sahel (Figure 6) is more sporadic and often deeper than in the TWP, where convection is nearly constant (Figure 7). The simulated deep convective clouds in the Sahel lack support from low clouds below the melting level in NICAM, FV3, and ICON.

For all models, the cloud structures and 6 hourly precipitation rates are consistent with the precipitation texture in Figures 4 and 5. The excessive TTL ice is consistent with the pervasive FWP in NICAM in both regions. For ICON in the TWP, the precipitation rates are substantially lower than in the other models, which agrees with the popcorn texture in the FWP and precipitation snapshots. The more intense convection in FV3 in both regions is evident in the depth of convection in Figures 6b and 7b as well as the higher FWPs within the deep convective cells in Figures 4d and 5d.

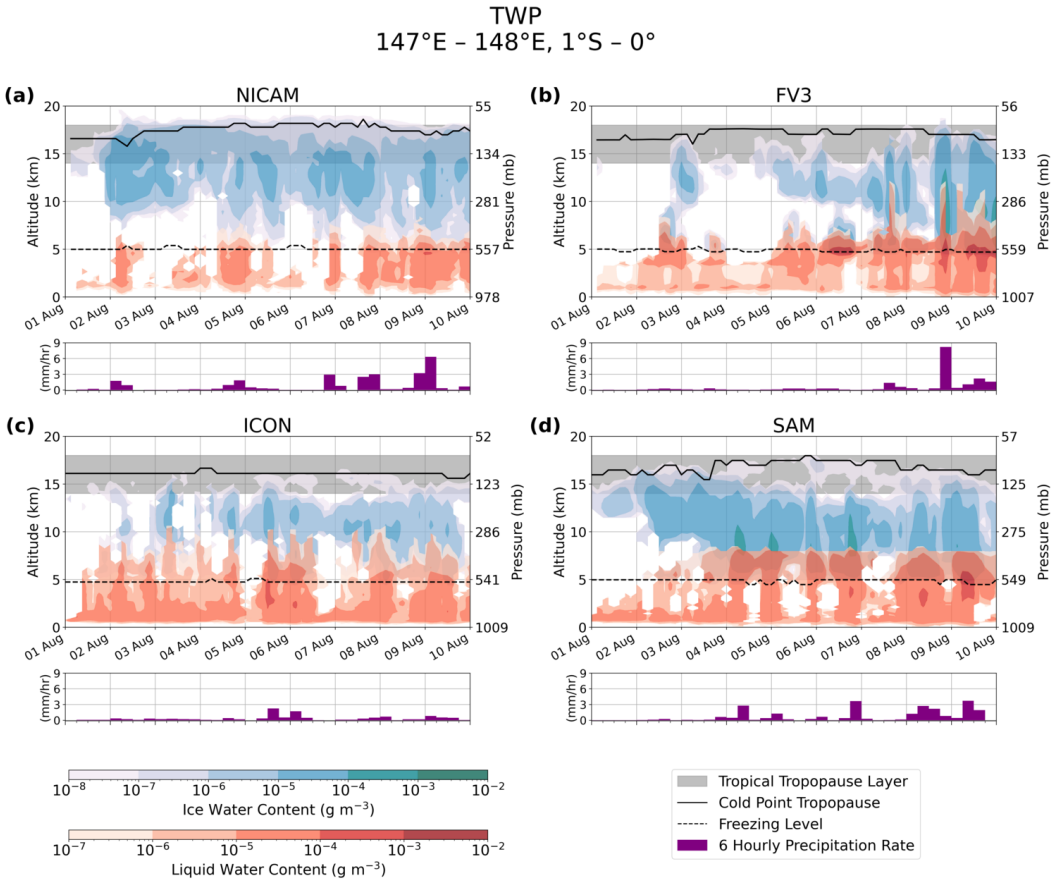

Figure 7. As in Figure 6, but for the TWP. Model output is averaged over a 1° × 1° box located at 147°E–148°E, 1°S–0°, approximately centered over the ARM site in Manus, Papua New Guinea.

All NFIS models have overshooting convection in the Sahel, which often penetrates past the cold point in NICAM and FV3 (Figure 6a–b). As in the average temperature profiles (Figure 2a–b), ICON places the cold point tropopause lower in the TTL than NICAM, FV3, or SAM. The cold points within the Sahel 1° × 1° box vary more over

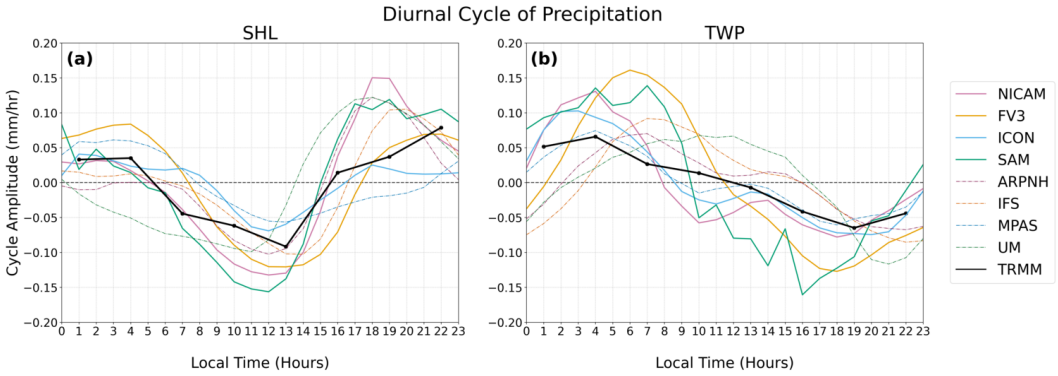

Figure 8. Amplitude of the regional-mean diurnal cycle of precipitation in (a) the Sahel and (b) the TWP for all DYAMOND models and for the 2006–2016 TRMM climatology. Precipitation rates are averaged at each hour for days 3–40 of the model run.

time than in the TWP. The frequency of convection reaching past the cold point in the models is surprising because very few precipitation features observed in the tropics extend above the lapse rate tropopause level (Liu & Zipser, 2005).

The typical vertical cloud structure differs between models. NICAM contains significantly more ice in the TTL than any of the other models (Figure 6a). Much of this additional ice comprises anvil or cirrus clouds with a low FWC. For FV3 in the Sahel (Figure 6b), convection reaches higher than other models, occasionally penetrating above 18 km. As shown in Figure 4c, the squall lines in ICON are followed by a region of lighter precipitation; this feature is apparent in the cloud structure where regions of high liquid water content indicative of low liquid clouds follow the convective cores near the melting level (Figure 6c). SAM is marked by persistent thick ice clouds with tops extending into the TTL (Figure 6d). Some of these clouds are thick anvils detraining off of the convective cores, but others appear detached with little liquid cloud below.

The vertical cloud structure in the TWP generally matches the qualitative texture expected of ocean convection, but with model-specific biases similar to the Sahel. NICAM has relatively little liquid cloud but persistent TTL ice up to an altitude of 18 km (Figure 7a). The ice within the TTL is almost entirely located above or near the tops of deep convective cores, whereas in the Sahel there are more frequent detached cirrus and anvils. ICON has a higher concentration of liquid water to support the ice cloud than in any other model, with supercooled liquid at altitudes as high as 10 km in updrafts (Figure 7c). SAM (Figure 7d) also has a lot of low liquid cloud beneath the deep convective clouds but has much more FWC than ICON in these columns, as in the Sahel (Figure 6d). These intermodel differences are also apparent in the joint albedo-OLR histograms analyzed in Part II.

3.4 Diurnal Cycle of Precipitation

The timing and amplitude of the 38-day diurnal cycle of precipitation for all models is similar to the TRMM 2006–2016 climatology, especially in the Sahel (Figure 8). Simulated rainfall peaks 2–3 hours too early in the Sahel (Figure 8a), but 2–3 hours too late in the TWP (Figure 8b). Nevertheless, the explicit deep convection in the DYAMOND GSRMs vastly reduces known biases in the diurnal timing of precipitation over land in GCMs (Bechtold et al., 2004).

4 Microphysics

Next, we examine the differences in the model output of microphysical variables, both within the TTL and integrated over the entire grid column.

4.1 Frozen Water Content and TTL Composition

NICAM is the only DYAMOND model for which we have the 3D graupel and snow output needed to plot their horizontally-averaged profiles. The vertical profiles of ice water content (IWC), IWC + snow water content (SWC), and IWC + SWC + graupel water content (GWC) for NICAM are plotted alongside DARDAR-estimated FWC in Figure 9a. Cloud ice dominates in the TTL, while snow and graupel peak at 6–7 km, just above the melting level. NICAM's total FWC peaks around the same level as the observations, but is overestimated. In the TWP, NICAM has much larger FWC on average at all levels than in the Sahel, but the relative levels and magnitudes of IWC, SWC, and GWC have the same trend across regions (see Figure S1).

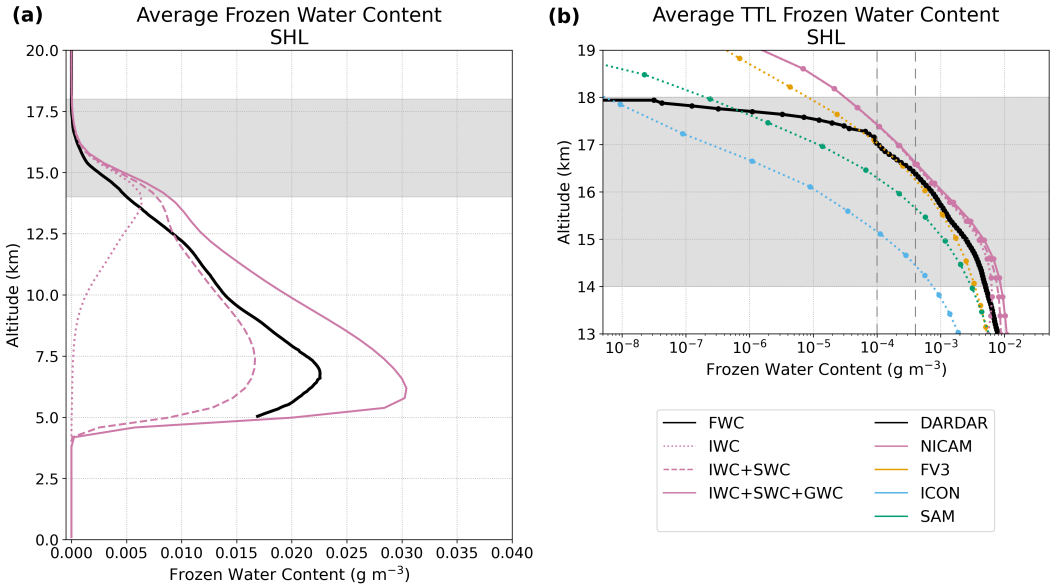

Figure 9. Area- and time-mean vertical profiles of frozen water contents (FWCs) in the Sahel. (a) Ice (dotted pink line), ice + snow (dashed pink line), and ice + snow + graupel (solid pink line) water contents for NICAM. The DARDAR FWC profile is shown in black. (b) Ice water content (dotted lines) for all NFIS models; ice + snow (dashed pink line) and ice + snow + graupel (solid pink line) water contents for NICAM; and DARDAR FWC (solid black line). Note that the x-axis of panel (b) is logarithmic. The dashed vertical lines indicate the approximate detection limits of the lidar at night (1×10^{-4} g m^{-3}) and during the day (4×10^{-4} g m^{-3}) (Avery et al., 2012).

Inside the TTL, the main frozen condensate type is cloud ice, which was archived for all four NFIS models. Thus, we compare the model IWC profiles in the TTL to DARDAR. NICAM, FV3, and SAM scatter within a plausible range around the observational estimate between 14–17 km. The FV3 IWC profile is closest to the data, while SAM has lower IWCs and NICAM has slightly higher IWCs. ICON greatly underestimates the observational profile of IWC throughout the entire TTL. For NICAM, the total FWC is slightly larger than the IWC at 14 km, but the contributions of snow and graupel be-

come negligible above ~17 km. It is unclear if adding snow and graupel would make up for the differences in IWC between models seen in Figure 9b.

4.2 Vertically Integrated Water Paths

In addition to IWP, GWP, and SWP, the DYAMOND intercomparison stored the column mass of the rain water path (RWP) and LWP. Figure 10 compares the time and regional averages of these quantities in the Sahel to DARDAR (see Figure S2 for the TWP equivalent). Since hydrometeors affect radiative fluxes, we also compare the time-mean simulated OLR to CERES CCCM data. The output for snow, graupel, and rain water paths was not saved for ARPNH or UM, so there is more frozen and liquid water in those models than shown here. These models are separated from the models with complete information by the vertical dashed lines in Figure 10.

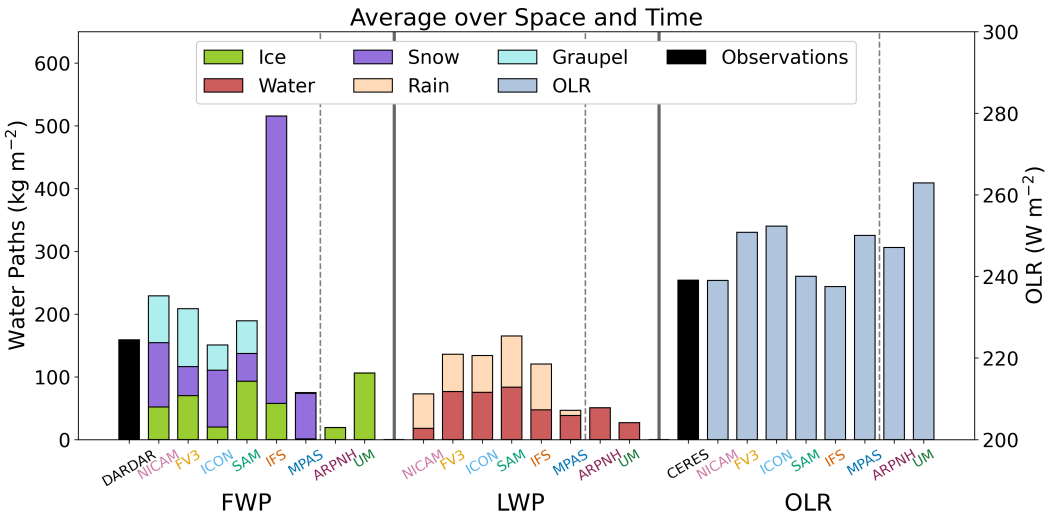

Figure 10. Bar chart of the time- and regional-mean frozen water paths (FWPs), total liquid water paths (LWPs), and outgoing longwave radiation (OLR) in the Sahel for all DYAMOND models. Black bars indicate observations of FWP from DARDAR and OLR from CERES. The vertical black dashed lines separate the models with complete information from those which did not save all column-integrated water paths. Note that the IFS microphysics scheme does not include graupel.

The left group of bars in Figure 10 shows a large intermodel variation in the FWP and its partitioning into frozen hydrometeor types. The FWP in IFS is almost twice that of DARDAR; however, the uncertainty in DARDAR FWP retrievals may be quite large, especially within thick clouds. In NICAM, FV3, ICON, SAM, and MPAS, the FWPs are much closer to the observed value, but the relative amounts of ice, snow, and graupel vary drastically between these models. For example, MPAS has almost no graupel or ice while ice makes up about half of the total FWP in SAM. These differences emphasize the point made in section 2.1 that the cloud ice alone does not comprise the total FWC in the models and that the hydrometeors (e.g., snow) may have a different definition in each model.

The center bars in Figure 10 compare the condensed water paths, which are more consistent across models. FV3, ICON, and IFS all have nearly identical amounts of cloud liquid and rain water, whereas NICAM and MPAS have much less.

The right bars in Figure 10 show a wide intermodel range in the average OLR. The models divide into two groups: those with average OLR close to the CERES observations (NICAM, SAM, and IFS) and those with more OLR than the observations (FV3, ICON, ARPNH, MPAS, and UM). The spread in OLR between these groups is over 25 W m^{-2}, whereas the uncertainty in the CERES OLR is only 3–4 W m^{-2} (Kato, 2021).

The differences in FWPs do not fully account for the OLR variation. It would be reasonable to assume that the models with more OLR have less frozen water than the others on average, but this is not always the case. For example, NICAM and IFS have a much higher FWP than DARDAR and the other models, but their average OLR is still very close to the observed mean. The mean OLR is also affected by model differences in the assumed radiative properties of each hydrometeor type in the microphysics scheme, such as the specification of their effective radii and which types are radiatively active. Additionally, mean OLR varies with the fractional coverage of clouds in each model.

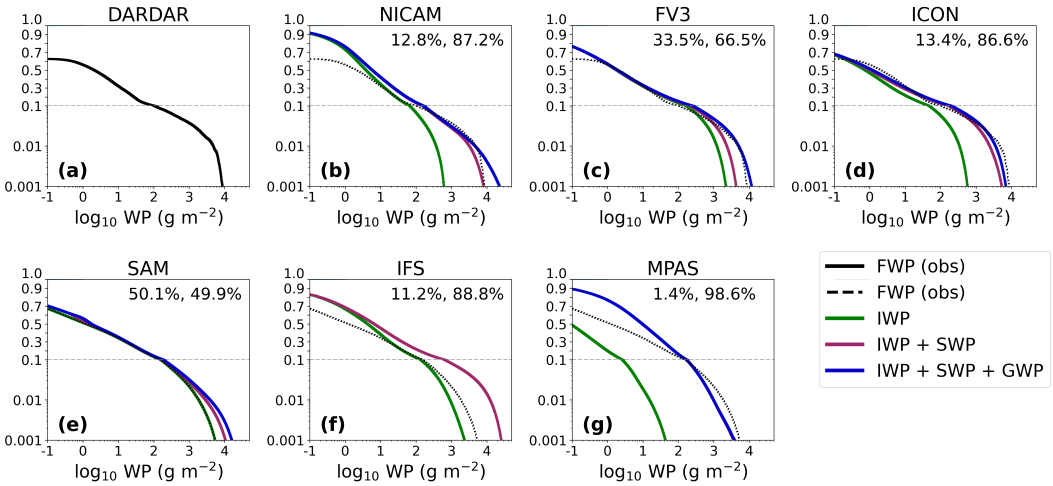

Figure 11. Cumulative distribution functions for ice (IWP; green line), ice + snow (IWP + SWP; pink line), and ice + snow + graupel (IWP + SWP + GWP; blue line) water paths for all models in the Sahel compared to DARDAR frozen water path observations (FWP; solid black line in a), dashed in b-g)). Values are accumulated from largest to smallest. ARPNH and UM are omitted because they do not save the 2D output of column-integrated snow or graupel water paths. The y-scale is logarithmic below 0.1 and linear above; the apparent discontinuity at 0.1 is an artifact of this split. The numbers in the upper-right corner of each panel show (left) the percentage of the total FWP from ice and (right) the percentage of total FWP from snow and graupel. These numbers cannot be deduced directly from this figure but are included to emphasize the large contributions of snow and graupel. Again, note that the IFS microphysics scheme does not include graupel.

Figure 11 compares the cumulative distribution functions (CDFs) of FWP between DARDAR and the DYAMOND models. These CDFs are computed from large FWPs to small ones so that, for example, the value at 10 g m^{-2} indicates the fraction of grid columns with an FWP exceeding 10 g m^{-2}. Since we define cirrus-free columns to be those with FWP $<$ 0.1 g m^{-2} (see section 2.5), the CDF at 0.1 g m^{-2} is equivalent to the ice cloud fraction. The dependence of simulated FWP distribution on the different hydrometeors is shown with additional lines for the CDFs of IWP and IWP+SWP. The steeper slope approaching 0.1 g m^{-2} indicates that the smaller FWPs are more frequent in the models than in the data. Still, the smallest values of FWP may not be radiatively

significant, and thus the thinnest simulated cirrus may not necessarily affect the heating profiles or top-of-atmosphere radiation balance.

The addition of SWP and GWP highlights the substantial contribution of snow and graupel to the overall frozen water mass in the DYAMOND models. The percentages in each panel of Figure 11b–h list the proportion of frozen mass from ice and snow + graupel, respectively. Ice is the most significant in SAM, constituting about half of the total frozen water mass. In all other models, snow and graupel make up the majority of the frozen mass, from 66.5% in FV3 to 98.6% in MPAS. These differences in partitioning mostly apply to the thickest clouds with high FWPs; the IWP makes up most of the FWP at values below $\sim$10 g m^{-2} for all except MPAS. Likewise, as shown in Figure 9, thin cirrus within the TTL are primarily composed of cloud ice in NICAM. Nevertheless, the large contributions of snow and graupel further stress that we cannot directly compare the simulated cloud ice to the observed frozen water in areas of deep convection.

As mentioned above, the value approached by the CDFs at 0.1 g m^{-2} represents an ice cloud fraction. All models overestimate the observed ice cloud fraction of $\sim$0.6, with most falling between $\sim$0.7 and $\sim$0.8. NICAM and MPAS have a much larger ice cloud fraction of 0.9. Qualitatively, the TWP results are similar to the Sahel (see Figure S3). All but IFS have a larger ice cloud fraction in the TWP than in the observations, but the proportion of frozen mass from snow + graupel falls within a few percent of the Sahel values. NICAM still has the largest ice cloud fraction, which is a surprising 0.999 in the TWP (Figure S3). Overall, the microphysics schemes in the models produce consistent cloud property biases specific to each model across the tropics.

5 TTL Cirrus

Finally, we use the NFIS models to try to understand how TTL cirrus are related to convection and what role convective mass fluxes play in setting the composition of the TTL.

5.1 TTL Cirrus and Convection

The diurnal cycle of convection and the collocation of convection with TTL cirrus are shown in the 40-day time series in Figure 12. Only FV3 is shown here for brevity; equivalent figures for NICAM, ICON, and SAM are included in the Supporting Information (Figures S4–S6). The diurnal cycle in the Sahel is much stronger and more distinct than in the TWP as expected (e.g., Yang & Slingo, 2001; Nesbitt & Zipser, 2003). In the Sahel, the deep convection is marked by periodic bouts of very low OLR, high precipitation rates, and high IWP indicative of deep cumulonimbus towers (Figure 12a, c, e). The regions of intense precipitation and high IWP occur more frequently and for longer periods of time in the TWP, characteristic of the more persistent and widespread convective storms observed over tropical oceans (Figure 12b, d, f). The convection and diurnal cycles shown in Figure 12 are qualitatively consistent with those for the other models (see Figures S4–S6) despite the intermodel differences in convection described in section 3.

Peaks in TTL IWP coincide with areas of large total-column IWP (Figure 12e–h). This suggests that the TTL ice is mostly located over deep convection, which qualitatively agrees with the collocation between total-column FWP and precipitation demonstrated in Figures 4 and 5. In both regions, the NFIS models also have frequent thinning anvils evident in the smaller IWPs extending away from the convective cores, but isolated TTL cirrus are not readily apparent. Therefore, TTL ice is mostly located in the vicinity of convection, with the majority of the ice mass concentrated directly over deep convection.

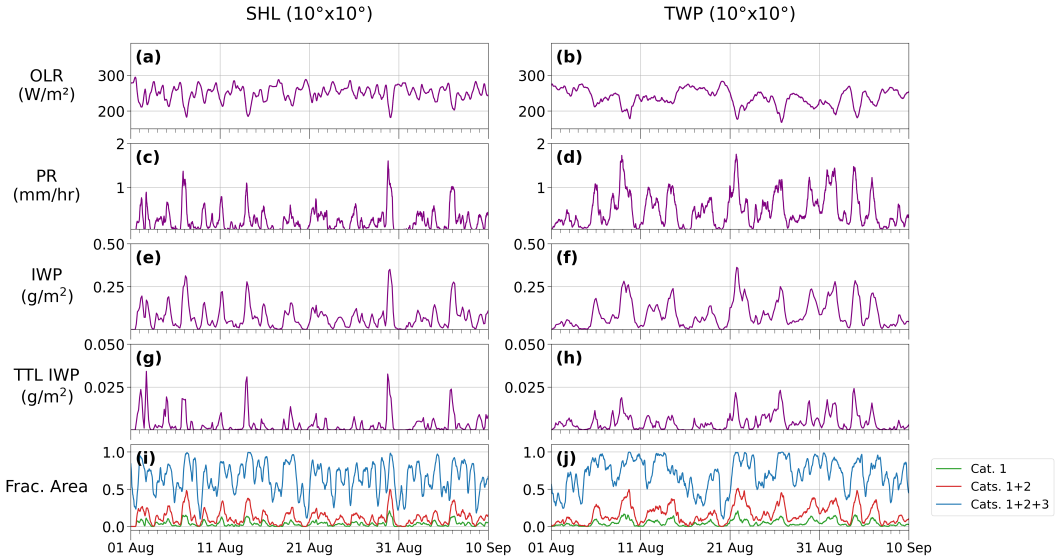

Figure 12. Time series for FV3 over days 1–40 of the model run: (a–b) area-mean outgoing longwave radiation (OLR); (c–d) area-mean precipitation rate (PR); (e–f) area-mean total-column ice water path (IWP); (g–h) area-mean TTL IWP; and (i–j) fractional areas of Category 1, Categories 1–2, and Categories 1–3 columns for FV3. Area-means and fractional areas are computed over the 10°×10° (left) Sahel and (right) TWP regions. For the TTL IWP, ice water content is integrated between the model levels closest to 14 and 18 km, which sometimes lie slightly outside of this range.

Similarly, the area spanned by anvil cirrus in the 10°×10° regions strongly correlates with that of deep convection (Figure 12i–j). Categories 1 and 2 are tightly coupled with each other and with convection in terms of low OLR, high precipitation rate, and large IWPs (Figure 12). For FV3, the fractional area of Category 2 lags behind Category 1 by about 3 hours in the Sahel and 5.5 hours in the TWP with a a maximum correlation coefficient of 0.9 in both regions. The coupling between Categories 1 and 2 is qualitatively similar in NICAM, ICON, and SAM (see Figures S4–S6). Category 3 thin cirrus track more closely with deep convection and anvils in ICON and SAM while the thin cirrus columns cover almost the entire 10°×10° area in NICAM. These relationships indicate that most anvils and thick cirrus are likely generated from deep convection, but other physical processes must contribute to generating and maintaining thinner cirrus.

5.2 Injection of Water Vapor and Frozen Water into the TTL

In Figure 13a and c, we bin cloud IWC and water vapor content (WV) in Sahel grid cells at an altitude of 14 km (the nominal base of the TTL) by the collocated vertical velocity for the NFIS models. The w bins are subjectively chosen to span the range of simulated updrafts and downdrafts; the probability distribution functions (PDFs) of vertical velocity are shown in Figure 13b. Figure 13a also shows the total FWC (i.e., including snow and graupel) for NICAM, the only model for which these variables were stored. The graupel and snow water contents may be substantial (up to 10 times larger than IWC) in areas of deep convection with strong updrafts and downdrafts. Thus, graupel and snow may dominate the vertical advective fluxes of frozen water into the TTL, so we only compare the w-binned vertical mass fluxes of frozen water and water vapor for NICAM (Figure 13d).

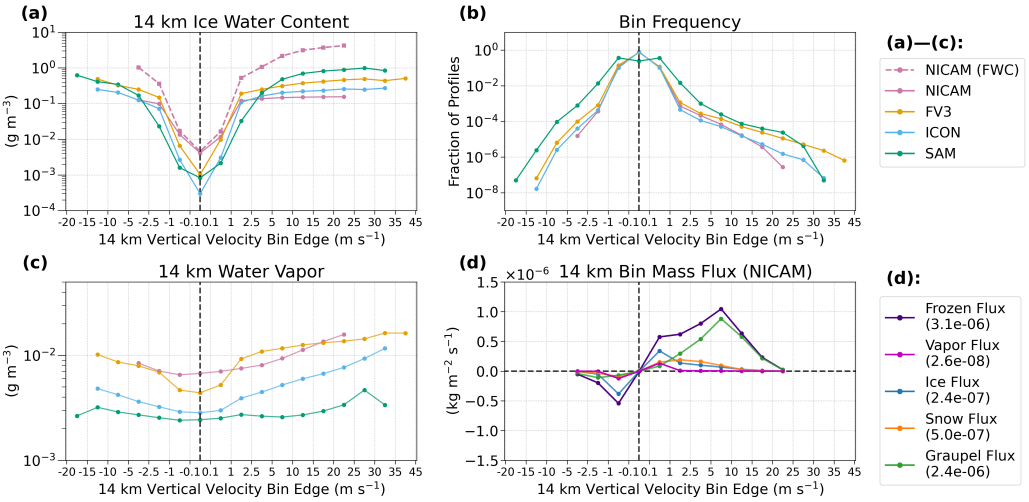

Figure 13. Selected variables sampled at all 14 km grid points for days 3–40 in the Sahel and averaged over vertical velocity bins: (a) cloud ice water content (pink dashed line shows NICAM frozen water content); (b) bin frequency; (c) water vapor content; and (d) frozen water, water vapor, ice, snow, and graupel mass fluxes for NICAM only.

The cloud IWC at 14 km varies greatly between models, but the relative differences across w bins are qualitatively consistent (Figure 13a). The IWC in FV3, ICON, and SAM increases by about three orders of magnitude from the weakest to strongest updrafts and downdrafts, which occur within deep convective clouds. The increase in IWC across bins is not as drastic in NICAM because its IWC is larger in the weakest w bins. For all models, the IWC is relatively symmetric between the largest positive and negative w bins, suggesting that cloud ice is typically circulated between nearby convective updrafts and downdrafts. Despite the differences in the model simulation of land and ocean convection discussed in section 3, the trends in IWC at different updraft and downdraft speeds are remarkably similar between regions (see Figure S7a).

Although the 14 km WV varies much less across w bins than the IWC, the trends differ somewhat between models (Figure 13c). For example, WV in SAM is nearly constant with w, but WV in FV3, ICON, and NICAM increases with updraft speeds. Nevertheless, the 14 km WV values are strikingly consistent; for each model and each w bin, all WV values lie within one order of magnitude from each other. As with IWC, the trends in WV are consistent across the tropics in the NFIS models (see Figure S7c).

The general shape of the vertical velocity PDF is consistent between the NFIS models, but there is intermodel spread in the infrequent extreme values (Figure 13b). Except for SAM, the models agree that in 99.9% of the grid cells, w lies within the -1 to 1 m s^{-1} bins; this is expected given the intermittency of deep convection that penetrates into the TTL. SAM simulates more 14 km updrafts and downdrafts of intermediate strength (1–10 m s^{-1}) than other models. SAM also produces downdrafts up to -20 m s^{-1} while NICAM does not simulate any downdrafts below -5 m^{-1} in either region. The strongest convective updrafts reach 45 m s^{-1} in FV3, which is 10–20 m s^{-1} larger than in the other models. The overall shape of the vertical velocity PDF is the same in the TWP as in the Sahel, but there is even more disagreement among the strongest updrafts and downdrafts (see Figure S7b). ICON and NICAM both have stronger maximum 14 km updraft speeds in the Sahel, but the updrafts in FV3 or SAM do not seem to be systematically stronger

in the Sahel. Overall, these differences in extreme values reflect diverse formulations of both model dynamics and microphysics that bear further scrutiny.

Figure 13d shows the mass fluxes at 14 km of water vapor and frozen water for NICAM only. The frozen mass flux is further broken down into the cloud ice, snow, and graupel fluxes. Because the WV is relatively constant across w bins (Figure 13c), the w-partitioning of 14 km water vapor flux is qualitatively similar to that of dry air and can be used as a proxy for the total air motion. The bin-averaged vapor flux in NICAM is largest in the ± 0.1–1 m s^{-1} bins and negligible everywhere else (Figure 13d, dark pink line). In contrast, the frozen mass flux distribution (Figure 13d, purple line) is weighted toward updrafts exceeding 1 m s^{-1} because they contain much more frozen water than nearly quiescent air. In all w bins, the frozen mass flux is much larger than the water vapor flux, consistent with the observationally-based estimates from Bolot and Fueglistaler (2021). Similarly, the integrated mass flux values given in the lower legends of Figures 13 and S7 show that in both regions, the upward frozen mass flux in NICAM at 14 km is much larger than the vapor mass flux and thus is the principal source of water in the TTL. Ice sedimentation is not considered in the distributions in Figures 13d and S7d. We presume that the ice sedimentation quickly offsets much of this upward flux, but we still expect that some injected frozen water remains and spreads into extensive TTL cirrus or ultimately sublimates and moistens the TTL (see Figure 2c–d).

Among the frozen hydrometeor fluxes in NICAM (Figure 13d), cloud ice dominates in the ± 0.1 m s^{-1} bins that represent the vast majority of the grid points over time (Figure 13b). In the 1 to 15 m s^{-1} bins, cloud ice and snow fluxes make up a small but nearly equal proportion of the total frozen mass flux. At updrafts stronger than 5 m s^{-1}, however, the graupel flux constitutes nearly all of the frozen mass flux. In the TWP, there is less upward frozen mass flux overall. The graupel flux there makes up a much smaller component of the total bin-integrated frozen water flux, but still exceeds the cloud ice and snow fluxes for updrafts stronger than 5 m s^{-1} (see Figure S7d).

5.3 Categorization of Water Injection by FWP

Partitioning the mass fluxes into the three FWP categories described in section 2.5 further reveals the importance of deep convection in transporting water into the TTL, even outside of columns with strong updrafts. Table 2 (see Table S1 for the TWP equivalent) contains the frequencies of occurrence of each category as well as the category-conditional frequencies of vertical velocities above ± 2.5 m s^{-1} and category-conditional average mass fluxes. Category 1 is the least frequent, containing 3.4% or less of all columns in the Sahel. Category 1 is slightly more common in the TWP for all models except ICON (see Table S1), which is consistent with the more frequent convection over the oceans. There is a large intermodel spread in the Category 2 and 3 values and consequently in the amount of clear sky. We expect most columns to fall into Category 3 because the TTL is dominated by very thin cirrus (e.g., Sassen et al., 2009); this is true in the Sahel, but Category 2 occurs most often in NICAM and SAM in the TWP. As shown in the vertical velocity PDFs (Figure 13b), the overwhelming majority of vertical motion in the TTL is very weak, even in areas of deep convection. Vertical velocities exceed ± 2.5 m s^{-1} in only 0.2% of Category 1 columns in SAM and less than 0.024% of Category 1 columns in NICAM, FV3, and ICON (Table 2). The strong upward mass fluxes in Category 1 are therefore not restricted to the Category 1 columns having strong updraft speeds.

The category-conditional average mass fluxes show the differences in roles played by each hydrometeor within the different physical processes in the TTL. Category 1 accounts for the majority of the total frozen mass flux in both regions even though it contains a small fraction of columns. As with the integrated values, most of the average Category 1 frozen mass flux comes from graupel flux in NICAM. Graupel flux is slightly less important in the TWP, but still accounts for about half of the average frozen flux (see

Table 2. Results from Frozen Water Path Categorization in the Sahel

Model	Cat.	Freq.	Freq. of $\lvert w\rvert \geq 2.5$ m s^{-1}	Avg. vapor flux (kg m^{-2} s^{-1})	Avg. frozen flux (kg m^{-2} s^{-1})	Avg. ice flux (kg m^{-2} s^{-1})	Avg. snow flux (kg m^{-2} s^{-1})	Avg. graupel flux (kg m^{-2} s^{-1})
NICAM	1	3.1 %	0.011 %	7.3×10^{-10}	3.1×10^{-6}	2.6×10^{-7}	4.9×10^{-7}	2.4×10^{-6}
	2	38.8 %	0.018 %	6.3×10^{-11}	-1.2×10^{-8}	-1.9×10^{-8}	6.8×10^{-9}	5.9×10^{-10}
	3	50.0 %	0.002 %	2.8×10^{-9}	1.9×10^{-10}	8.8×10^{-11}	1.0×10^{-10}	1.2×10^{-14}
FV3	1	3.4 %	0.024 %	5.0×10^{-8}		1.3×10^{-6}		
	2	21.4 %	0.033 %	-1.1×10^{-8}		1.9×10^{-8}		
	3	39.4 %	0.004 %	5.7×10^{-9}		1.8×10^{-9}		
ICON	1	3.2 %	0.010 %	3.7×10^{-9}		2.4×10^{-7}		
	2	20.3 %	0.011 %	-2.9×10^{-9}		6.3×10^{-9}		
	3	26.5 %	0.001 %	2.1×10^{-9}		4.6×10^{-10}		
SAM	1	3.4 %	0.200 %	1.1×10^{-8}		2.9×10^{-6}		
	2	32.1 %	0.025 %	4.1×10^{-9}		1.4×10^{-7}		
	3	35.2 %	0.003 %	2.1×10^{-8}		1.1×10^{-8}		

Note. The columns list frequency of each frozen water path category, percentage of columns in each category where $\lvert w\rvert \geq 2.5$ m s^{-1}, and regional- and time-mean mass fluxes for the NFIS models in the Sahel. The mass fluxes are weighted by category frequency.

Table S1). In Category 2, however, most of the frozen flux comes from cloud ice flux. Both the average frozen and cloud ice fluxes are negative, but this does not indicate a net downward flux; these fluxes integrate to nearly zero. The average vapor flux in Category 2 is also much larger than in Category 1, particularly in the TWP (see Table S1), but is still dwarfed by the frozen mass flux. Unsurprisingly, the average frozen mass flux is very small in Category 3 since the thin cirrus contain very little frozen water. Instead, the vapor flux is about 10 times larger than the frozen mass flux within Category 3 and nearly 4 times larger than the Category 1 vapor flux in both regions (Tables 2 and S1).

When the mass fluxes in Categories 2 and 3 are binned by w as with the total-column fluxes in Figure 13d, the distributions do not reach large w values and are approximately symmetric around zero (not shown). This symmetry suggests that in areas of both anvil and thin cirrus, the mass fluxes indicate a recycling of air through the 14 km level instead of a systematic upward flux. By extension, this indicates that there is likely much less sedimentation in these areas compared to deep convection. Altogether, the relationships between mass fluxes and FWP categories are surprisingly similar between the Sahel and TWP despite the differences in the characteristics of simulated land and ocean convection described in section 3.

Within Category 1, the $10° \times 10°$ area-integrated frozen mass flux in NICAM at each time step is strongly correlated with the instantaneous mass of frozen water within the TTL (Figure 14). The frozen mass flux is only weakly related to the total spatial area spanned by Category 1 columns at each time step, with a correlation coefficient half as large as that in Figure 14 (0.42 vs. 0.86). Therefore, a larger area of deep convection does not necessarily bring more frozen water into the TTL. The relationship in Figure 14 suggests that there is a regional-mean time scale of about 722 seconds (the reciprocal of the slope of the regression line) in the Sahel for the frozen water injected into the TTL to fall out or evolve into thinner cirrus (i.e., Categories 2 and 3). The same relationship is present in the TWP, but the time scale is longer at 1575 seconds (Figure S8). Compared to the Sahel, the weaker updrafts in the TWP tend to loft smaller ice particles that will take longer to sediment out of the TTL. Since NICAM has much more cloud ice in the TTL in both regions than FV3, ICON, or SAM, it is unclear if this ice removal timescale will quantitatively hold for the other models.

In NICAM, deep convection injects frozen water and vapor at 14 km and deposits it throughout the depth of the TTL. Figure 15 shows vertical profiles of the average frozen water and vapor mass fluxes (normalized by the 38-day average precipitation rate in the Sahel) as well as the largest 0.1% of vertical velocities within the updrafts or downdrafts (see Figure S9 for the TWP equivalent, normalized by the TWP average precipitation

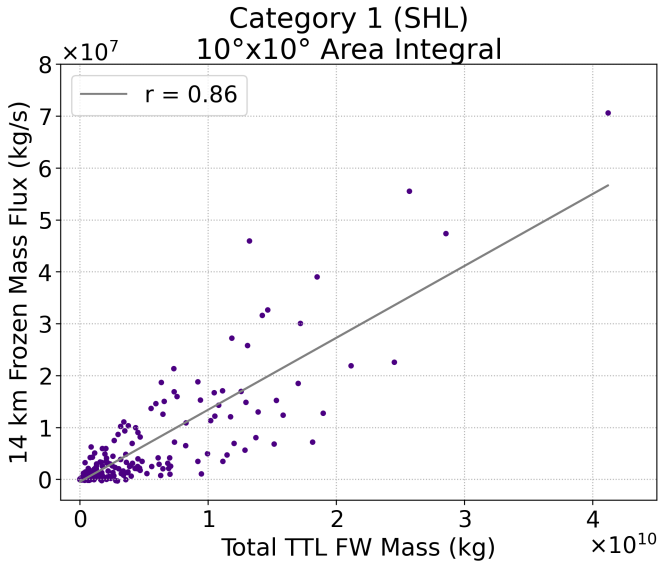

Figure 14. Correlations between the area-integrated Category 1 frozen mass flux at 14 km and the total mass of frozen water (FW) in the TTL in Category 1 for NICAM in the Sahel. Each point represents one instantaneous time step over days 3–40 of the model run with variables integrated over the 10°×10° Sahel region. The least-squares regression line is shown in gray with the correlation coefficient (r) in the legend.

rate). These profiles are for all columns, but as previously discussed, the Category 1 frozen mass flux makes up nearly all of the total frozen mass flux. In both regions, the frozen mass flux dominates throughout the lower TTL up to about 16.5 km, at which point it is overtaken by the vapor flux. This patterns agrees qualitatively with Bolot and Fueglistaler (2021), who found that deep convection is more important for moisture transport than large-scale advection (i.e., our Category 3 mass fluxes) below the cold point (∼16–17 km). However, the frozen mass flux values in NICAM are larger by a factor of ∼100 in both regions than the observationally-based, tropics-wide estimates from Bolot and Fueglistaler (2021). We normalize the estimates of ice flux from Figure 3 in Bolot and Fueglistaler (2021) by the area of their 30°S–30°N domain as well as the 38-day average Sahel or TWP precipitation rate. These estimates are then plotted alongside the NICAM values in Figures 15 and S9.

5.4 Frozen Water Flux into the TTL in Other Models

We cannot generalize the mass flux results in NICAM to the other models. The regional- and time-mean cloud ice mass fluxes at 14 km vary by a factor of 10 in the NFIS models within both regions (Table 2). We also expect the specific values of 14 km frozen and vapor mass fluxes to differ significantly between the NFIS models based on the vertical velocities, FWC, and WV. Although the 14 km IWC is very similar between models in the bins above ±5 m s^{-1} (Figure 13a), there are substantial differences in the relative proportions of cloud ice, snow, and graupel (Figure 11). Thus, the 14 km FWC will likely vary more than the 14 km IWC. Likewise, FV3 and ICON both have very different frequencies of updrafts exceeding 10 m s^{-1} than NICAM (Figure 13), which determine the most intense mass flux values in those models.

Regardless of the anticipated differences in the amounts of mass injected into the TTL, we still expect deep convection to be the most important process in supplying wa-

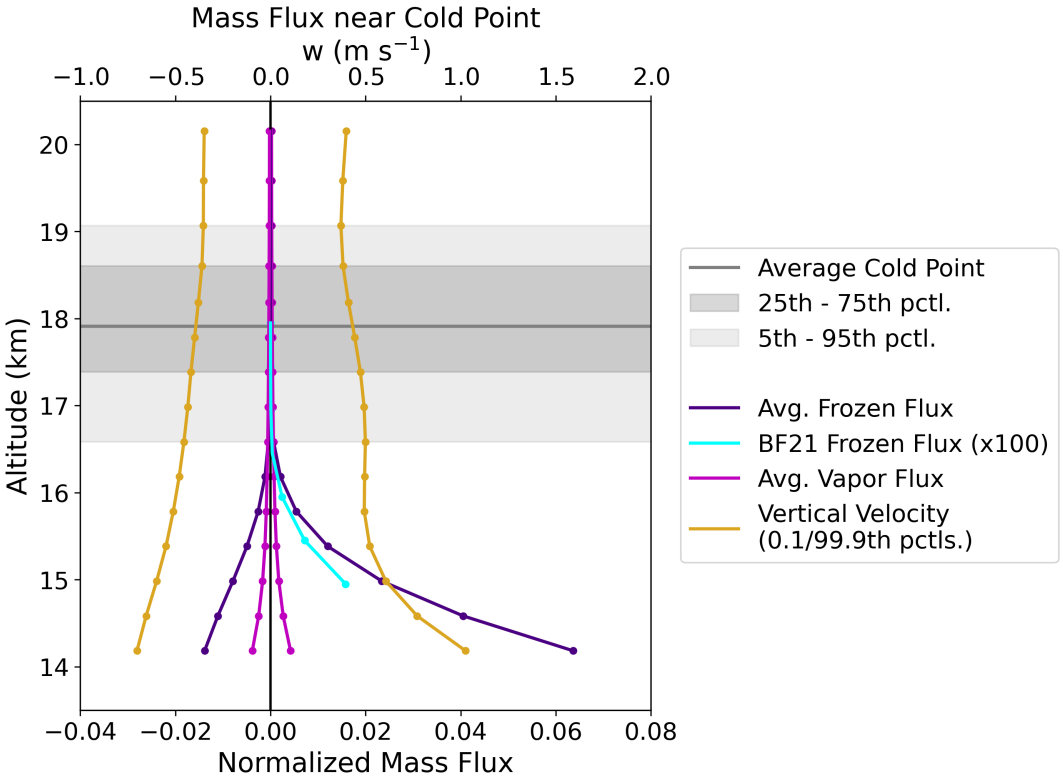

Figure 15. Vertical profiles of normalized frozen water (purple lines) and vapor mass fluxes (dark pink lines) for NICAM in the Sahel as well as the normalized estimates of frozen mass flux from Bolot and Fueglistaler (2021) times 100 (BF21; cyan line). The mass fluxes for NICAM are split between updrafts and downdrafts, averaged over the 10°×10° region and time, and weighted by the relative frequency of the updrafts or downdrafts. The gold lines show the 0.1th and 99.9th percentiles of vertical velocity. The average cold point height is shown in the solid gray line while the 5th–95th (25th–75th) percentiles are shaded in light (dark) gray. All percentiles are taken over the 10°×10° Sahel region and days 3–40 of the model run.

ter into the TTL. In all of the NFIS models, the concentration of cloud ice over deep convection is very high (see section 5.1). This proportion will only increase once graupel and snow are considered because these hydrometeors make up 50–87% of the total integrated FWP in FV3, ICON, and SAM (Figure 11). Since the 14 km WV in FV3, ICON, and SAM is similar to that in NICAM (Figure 13c), we can assume that the frozen mass flux will be much larger than the vapor flux in the other models. Nevertheless, we cannot draw any definitive conclusions about the frozen mass flux in other models without access to the 3D profiles of snow and graupel in FV3, ICON, or SAM.

6 Summary and Discussion

In this paper, we have used satellite and reanalysis data to evaluate the simulation of convectively generated cirrus in GSRMs from the DYAMOND project. These GSRMs all have horizontal grid spacing below 5 km, which permits explicit convection (Stevens et al., 2019). We focused on four models which archived cloud ice at the full native grid resolution: NICAM, FV3, ICON, and SAM. We analyzed two representative 10°×10°

latitude-longitude regions, one over the Sahel in western Africa and one over the tropical West Pacific.

Before analyzing how convection generates tropical cirrus in the DYAMOND GSRMs, we first assessed how well the GSRMs reproduce deep convection. As a group, the DYAMOND GSRMs simulate deep convective precipitation, its organization, and its diurnal cycle fairly well over both the land and ocean regions, though intermodel scatter is large. In contrast, GCMs struggle with these tasks. The GSRM-simulated convective updraft and downdraft strengths have qualitatively similar frequency distributions. All the GSRMs produce frequent overshooting convection that reaches the TTL; in most models, convective updrafts occasionally penetrate the cold point tropopause. In all GSRMs and both regions, the time-mean relative humidity with respect to ice is ∼70–90% in the TTL.

For the convective cloud properties (water content vertical profiles and the cloud microphysics, including partitioning between cloud liquid, ice, snow, and graupel) and the convective injection of frozen water into the TTL, differences between GSRMs exceed differences between land and ocean regions, consistent with the findings of Turbeville et al. (2021) in Part II of this study. In NICAM in both regions, the most intense convective updrafts inject nearly all of the incoming frozen water into the TTL. We presume that most of this injected ice sediments out of the TTL near the deep convection, although we did not have the necessary information to be sure. We found that across the models, the mass of frozen water within the TTL correlates with the deep convective frozen mass flux into the TTL; their ratio suggests an ice residence time of about ten to thirty minutes.

Our analysis was limited by the output variables available for most of the DYAMOND GSRMs. Having the 3D profiles of snow and graupel water contents would allow for evaluation of intermodel differences in the frozen mass fluxes and TTL FWC. Because cloud microphysics parameterizations are a major source of intermodel differences, saving the full hydrometeor profiles should be prioritized in future GSRM simulations.

We have focused on how water vapor and frozen water reach the TTL. What happens to the injected mass afterward is also critical to understanding the relationship between TTL cirrus and deep convection. In Part II of this paper, Turbeville et al. (2021) explore the full distribution of tropical cirrus in the DYAMOND GSRMs, including transitions between the three FWP categories introduced here. The models simulate the observed transition between deep convection and thin cirrus, though with clear model-specific biases (Turbeville et al., 2021). That analysis could not address two scientifically important questions: how quickly air moves from Category 1 (deep convection) to Category 2 (thick cirrus) and how much frozen water is transported between categories vs. how much sediments out.

Despite their present limitations, GSRMs are a promising complement to climate models for the simulation of cirrus initiated by deep convection, including within the TTL. This study shows that subgrid processes drive the substantial variation between models. Specifically, the model microphysics and, to a lesser extent, the dynamical flow solver may both be important. We suspect that improved microphysics would more readily advance the simulation of tropical cirrus than higher spatial resolutions. The sub-5 km horizontal and 200–500 m vertical grid spacings seem sufficient to simulate cirrus and may capture gravity wave activity around deep convection that is thought to be important for TTL cirrus, though this has not been clearly shown. Radiatively-driven circulations within TTL cirrus (Durran et al., 2009) would require finer grids to explicitly resolve; whether subgrid turbulence parameterizations can represent this process adequately remains to be seen. Overall, this work has demonstrated that GSRMs model deep convection realistically enough to form a good foundation for simulating convectively generated cirrus and their impacts on the TTL, climate more broadly, and climate change.

Careful observationally-based improvements of GSRMs are needed to fully realize this potential.

Acknowledgments

We thank Adam Sokol and Blaž Gasparini for their help with obtaining and analyzing the DARDAR and CCCM data. This material is based upon work supported by the National Science Foundation through the Partnerships in International Research and Education program under grant number OISE-1743753. DYAMOND data management was provided by the German Climate Computing Center (DKRZ) and supported through the projects ESiWACE and ESiWACE2, which have received funding from the European Union's Horizon 2020 research and innovation program under grant agreements No. 675191 and 823988. DARDAR data were obtained from the ICARE Data and Services Center (`http://www.icare.univ-lille1.fr/projects/dardar`). The TRMM 3B42 data were obtained from `https://doi.org/10.5067/TRMM/TMPA/3H/7`. The satellite data from the A-Train Integrated CALIPSO, CloudSat, CERES, and MODIS Merged Product Release B1 (CCCM) were obtained from `https://doi.org/10.5067/AQUA/CERES/NEWS_CCCM-FM3-MODIS-CAL-CS_L2.RELB1`. The ERA5 hourly data on pressure levels were obtained from the Copernicus Climate Data Store (`https://doi.org/10.24381/cds.bd0915c6`).

References

Arteta, J., Marecal, V., & Riviere, E. D. (2009). Regional modelling of tracer transport by tropical convection – Part 1: Sensitivity to convection parameterization. *Atmos. Chem. Phys.*, 20. doi: 10.5194/acp-9-7081-2009

Avery, M., Winker, D., Heymsfield, A., Vaughan, M., Young, S., Hu, Y., & Trepte, C. (2012). Cloud ice water content retrieved from the CALIOP space-based lidar: CALIOP ICE WATER CONTENT. *Geophysical Research Letters*, *39*(5). doi: 10.1029/2011GL050545

Bechtold, P., Chaboureau, J.-P., Beljaars, A., Betts, A., Köhler, M., Miller, M., & Redelsperger, J.-L. (2004). The simulation of the diurnal cycle of convective precipitation over land in a global model. *Quarterly Journal of the Royal Meteorological Society*, *130*(604), 3119–3137. doi: 10.1256/qj.03.103

Berthou, S., Rowell, D. P., Kendon, E. J., Roberts, M. J., Stratton, R. A., Crook, J. A., & Wilcox, C. (2019). Improved climatological precipitation characteristics over West Africa at convection-permitting scales. *Climate Dynamics*, *53*(3-4), 1991–2011. doi: 10.1007/s00382-019-04759-4

Bolot, M., & Fueglistaler, S. (2021). Tropical Water Fluxes Dominated by Deep Convection Up to Near Tropopause Levels. *Geophysical Research Letters*, *48*(4). doi: 10.1029/2020GL091471

Bousquet, O., & Chong, M. (2000). The oceanic mesoscale convective system and associated mesovortex observed 12 December 1992 during TOGA-COARE. *Quarterly Journal of the Royal Meteorological Society*, *126*(562), 189–211. doi: 10.1002/qj.49712656210

Chao, W. C. (2013). Catastrophe-Concept-Based Cumulus Parameterization: Correction of Systematic Errors in the Precipitation Diurnal Cycle over Land in a GCM. *Journal of the Atmospheric Sciences*, *70*(11), 3599–3614. doi: 10.1175/JAS-D-13-022.1

Delanoë, J., & Hogan, R. J. (2008). A variational scheme for retrieving ice cloud properties from combined radar, lidar, and infrared radiometer. *Journal of Geophysical Research*, *113*(D7), D07204. doi: 10.1029/2007JD009000

Delanoë, J., & Hogan, R. J. (2010). Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *Journal of Geophysical Research*, *115*, D00H29. doi: 10.1029/2009JD012346

Deng, M., Mace, G. G., Wang, Z., & Lawson, R. P. (2013). Evaluation of Several

A-Train Ice Cloud Retrieval Products with In Situ Measurements Collected during the SPARTICUS Campaign. *Journal of Applied Meteorology and Climatology*, *52*(4), 1014–1030. doi: 10.1175/JAMC-D-12-054.1

Dessler, A. E., Palm, S. P., Hart, W. D., & Spinhirne, J. D. (2006). Tropopause-level thin cirrus coverage revealed by ICESat/Geoscience Laser Altimeter System. *Journal of Geophysical Research*, *111*(D8), D08203. doi: 10.1029/2005JD006586

Duda, J. D., & Gallus, W. A. (2013). The Impact of Large-Scale Forcing on Skill of Simulated Convective Initiation and Upscale Evolution with Convection-Allowing Grid Spacings in the WRF*. *Weather and Forecasting*, *28*(4), 994–1018. doi: 10.1175/WAF-D-13-00005.1

Durran, D. R., Dinh, T., Ammerman, M., & Ackerman, T. (2009). The Mesoscale Dynamics of Thin Tropical Tropopause Cirrus. *Journal of the Atmospheric Sciences*, *66*(9), 2859–2873. doi: 10.1175/2009JAS3046.1

Fierli, F., Orlandi, E., Law, K. S., Cagnazzo, C., Cairo, F., Schiller, C., ... Volk, C. M. (2011). Impact of deep convection in the tropical tropopause layer in West Africa: in-situ observations and mesoscale modelling. *Atmospheric Chemistry and Physics*, *11*(1), 201–214. doi: 10.5194/acp-11-201-2011

Fueglistaler, S., Dessler, A. E., Dunkerton, T. J., Folkins, I., Fu, Q., & Mote, P. W. (2009). Tropical tropopause layer. *Reviews of Geophysics*, *47*(1). (Publisher: John Wiley & Sons, Ltd) doi: 10.1029/2008RG000267

Futyan, J. M., Russell, J. E., & Harries, J. E. (2004). Cloud Radiative Forcing in Pacific, African, and Atlantic Tropical Convective Regions. *JOURNAL OF CLIMATE*, *17*, 11. doi: 10.1175/1520-0442(2004)017⟨3192:CRFIPA⟩2.0.CO;2

Haladay, T., & Stephens, G. (2009). Characteristics of tropical thin cirrus clouds deduced from joint CloudSat and CALIPSO observations. *Journal of Geophysical Research*, *114*, D00A25. doi: 10.1029/2008JD010675

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. doi: 10.1002/qj.3803

Huffman, G. J., Adler, R. F., Bolvin, D. T., & Nelkin, E. J. (2010). The TRMM Multi-Satellite Precipitation Analysis (TMPA). In M. Gebremichael & F. Hossain (Eds.), *Satellite Rainfall Applications for Surface Hydrology* (pp. 3–22). Dordrecht: Springer Netherlands. doi: 10.1007/978-90-481-2915-7_1

Huffman, G. J., & Bolvin, D. T. (2018). *TRMM and Other Data Precipitation Set Documentation.* NASA/GSFC. Retrieved 2021-02-19, from `https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/3B42_3B43_doc_V7.pdf`

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., ... Stocker, E. F. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, *8*(1), 38–55. doi: 10.1175/JHM560.1

Inoue, T., Satoh, M., Miura, H., & Mapes, B. (2008). Characteristics of Cloud Size of Deep Convection Simulated by a Global Cloud Resolving Model over the Western Tropical Pacific. *Journal of the Meteorological Society of Japan*, *86A*, 1–15. doi: 10.2151/jmsj.86A.1

Jensen, E. J., Toon, O. B., Pfister, L., & Selkirk, H. B. (1996a). Dehydration of the upper troposphere and lower stratosphere by subvisible cirrus clouds near the tropical tropopause. *Geophysical Research Letters*, *23*(8), 825–828. (Publisher: John Wiley & Sons, Ltd) doi: 10.1029/96GL00722

Jensen, E. J., Toon, O. B., Selkirk, H. B., Spinhirne, J. D., & Schoeberl, M. R. (1996b). On the formation and persistence of subvisible cirrus clouds near the tropical tropopause. *Journal of Geophysical Research: Atmospheres*, *101*(D16), 21361–21375. doi: 10.1029/95JD03575

Jensen, E. J., Ueyama, R., Pfister, L., Bui, T. V., Alexander, M. J., Podglajen, A., ... Schoeberl, M. R. (2016). High-frequency gravity waves and homogeneous ice nucleation in tropical tropopause layer cirrus: WAVES AND ICE NUCLEATION. *Geophysical Research Letters*, *43*(12), 6629–6635. doi: 10.1002/2016GL069426

Kato, S. (2021, August). [Email]. Private communication.

Kato, S., Rose, F. G., Sun-Mack, S., Miller, W. F., Chen, Y., Rutan, D. A., ... Collins, W. D. (2011). Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties. *Journal of Geophysical Research*, *116*(D19), D19209. doi: 10.1029/2011JD016050

Kato, S., Sun-Mack, S., Miller, W. F., Rose, F. G., Chen, Y., Minnis, P., & Wielicki, B. A. (2010). Relationships among cloud occurrence frequency, overlap, and effective thickness derived from CALIPSO and CloudSat merged cloud vertical profiles. *Journal of Geophysical Research*, *115*, D00H28. doi: 10.1029/2009JD012277

Khairoutdinov, M., Randall, D., & DeMott, C. (2005). Simulations of the Atmospheric General Circulation Using a Cloud-Resolving Model as a Superparameterization of Physical Processes. *Journal of the Atmospheric Sciences*, *62*(7), 2136–2154. doi: 10.1175/JAS3453.1

Krämer, M., Rolf, C., Luebke, A., Afchine, A., Spelten, N., Costa, A., ... Avallone, L. (2016). A microphysics guide to cirrus clouds – Part 1: Cirrus types. *Atmospheric Chemistry and Physics*, *16*(5), 3463–3483. doi: 10.5194/acp-16-3463-2016

Lee, J., Yang, P., Dessler, A. E., Gao, B.-C., & Platnick, S. (2009). Distribution and Radiative Forcing of Tropical Thin Cirrus Clouds. *Journal of the Atmospheric Sciences*, *66*(12), 3721–3731. doi: 10.1175/2009JAS3183.1

Liu, C., & Zipser, E. J. (2005). Global distribution of convection penetrating the tropical tropopause. *Journal of Geophysical Research*, *110*(D23), D23104. doi: 10.1029/2005JD006063

Liu, C., Zipser, E. J., & Nesbitt, S. W. (2007). Global Distribution of Tropical Deep Convection: Different Perspectives from TRMM Infrared and Radar Data. *Journal of Climate*, *20*(3), 489–503. doi: 10.1175/JCLI4023.1

Long, C. N., Mather, J. H., & Ackerman, T. P. (2016). The ARM Tropical Western Pacific (TWP) Sites. *Meteorological Monographs*, *57*, 7.1–7.14. doi: 10.1175/AMSMONOGRAPHS-D-15-0024.1

Luo, Z., & Rossow, W. B. (2004). Characterizing Tropical Cirrus Life Cycle, Evolution, and Interaction with Upper-Tropospheric Water Vapor Using Lagrangian Trajectory Analysis of Satellite Observations. *Journal of Climate*, *17*(23), 4541–4563. doi: 10.1175/3222.1

Mace, G. G., Deng, M., Soden, B., & Zipser, E. (2006). Association of Tropical Cirrus in the 10–15-km Layer with Deep Convective Sources: An Observational Study Combining Millimeter Radar Data and Satellite-Derived Trajectories. *Journal of the Atmospheric Sciences*, *63*(2), 480–503. doi: 10.1175/JAS3627.1

Massie, S., Gettelman, A., Randel, W., & Baumgardner, D. (2002). Distribution of tropical cirrus in relation to convection: DISTRIBUTION OF TROPICAL CIRRUS. *Journal of Geophysical Research: Atmospheres*, *107*(D21), AAC 19–1–AAC 19–16. doi: 10.1029/2001JD001293

Murphy, D. M., & Koop, T. (2005). Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Quarterly Journal of the Royal Meteorological Society*, *131*(608), 1539–1565. doi: 10.1256/qj.04.94

Nesbitt, S. W., & Zipser, E. J. (2003). The Diurnal Cycle of Rainfall and Convective Intensity according to Three Years of TRMM Measurements. *JOURNAL OF CLIMATE*, *16*, 20. doi: 10.1175/1520-0442(2003)016⟨1456:TDCORA⟩2.0.CO;

2

Putman, W. (2020, August). [Email]. Private communication.

Redelsperger, J.-L., Diongue, A., Diedhiou, A., Ceron, J.-P., Diop, M., Gueremy, J.-F., & Lafore, J.-P. (2002). Multi-scale description of a Sahelian synoptic weather system representative of the West African monsoon. *Quarterly Journal of the Royal Meteorological Society*, *128*(582), 1229–1257. doi: 10.1256/003590002320373274

Redelsperger, J.-L., & Lafore, J.-P. (1988). A Three-Dimensional Simulation of a Tropical Squall Line: Convective Organization and Thermodynamic Vertical Transport. *Journal of Atmospheric Sciences*, *45*(8), 1334 – 1356. (Place: Boston MA, USA Publisher: American Meteorological Society) doi: 10.1175/1520-0469(1988)045⟨1334:ATDSOA⟩2.0.CO;2

Sassen, K., Wang, Z., & Liu, D. (2009). Cirrus clouds and deep convection in the tropics: Insights from CALIPSO and CloudSat. *Journal of Geophysical Research*, *114*, D00H06. doi: 10.1029/2009JD011916

Schoeberl, M. R., Jensen, E. J., Pfister, L., Ueyama, R., Avery, M., & Dessler, A. E. (2018). Convective Hydration of the Upper Troposphere and Lower Stratosphere. *Journal of Geophysical Research: Atmospheres*, *123*(9), 4583–4593. doi: 10.1029/2018JD028286

Slingo, A., Bharmal, N. A., Robinson, G. J., Settle, J. J., Allan, R. P., White, H. E., ... Miller, M. (2008). Overview of observations from the RADAGAST experiment in Niamey, Niger: Meteorology and thermodynamic variables. *Journal of Geophysical Research*, *113*, D00E01. doi: 10.1029/2008JD009909

Sokol, A. B., & Hartmann, D. L. (2020). Tropical Anvil Clouds: Radiative Driving Toward a Preferred State. *Journal of Geophysical Research: Atmospheres*, *125*(21). doi: 10.1029/2020JD033107

Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J., & Plattner, G.-K. (2010). Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming. *Science*, *327*(5970), 1219–1223. doi: 10.1126/science.1182488

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., ... the CloudSat Science Team (2002). THE CLOUDSAT MISSION AND THE A-TRAIN: A New Dimension of Space-Based Observations of Clouds and Precipitation. *Bulletin of the American Meteorological Society*, *83*(12), 1771–1790. doi: 10.1175/BAMS-83-12-1771

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., ... Zhou, L. (2019). DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Progress in Earth and Planetary Science*, *6*(1), 61. doi: 10.1186/s40645-019-0304-z

Turbeville, S. M., Nugent, J. M., Ackerman, T. P., Bretherton, C. S., & Blossey, P. N. (2021). *Tropical cirrus in global storm-resolving models. Part II: Cirrus Life Cycle and Top-of-Atmosphere Radiative Fluxes.* (Unpublished manuscript)

Ueyama, R., Jensen, E. J., & Pfister, L. (2018). Convective Influence on the Humidity and Clouds in the Tropical Tropopause Layer During Boreal Summer. *Journal of Geophysical Research: Atmospheres*. doi: 10.1029/2018JD028674

Virts, K. S., & Houze, R. A. (2015). Clouds and Water Vapor in the Tropical Tropopause Transition Layer over Mesoscale Convective Systems. *Journal of the Atmospheric Sciences*, *72*(12), 4739–4753. doi: 10.1175/JAS-D-15-0122.1

Yang, G.-Y., & Slingo, J. (2001). The Diurnal Cycle in the Tropics. *MONTHLY WEATHER REVIEW*, *129*, 18. doi: 10.1175/1520-0493(2001)129⟨0784:TDCITT⟩2.0.CO;2

Zipser, E. J., Cecil, D. J., Liu, C., Nesbitt, S. W., & Yorty, D. P. (2006). WHERE ARE THE MOST INTENSE THUNDERSTORMS ON EARTH? *Bulletin of the American Meteorological Society*, *87*(8), 1057–1072. doi: 10.1175/BAMS-87

931 -8-1057