

Evaluating Tropical Precipitation Relations in CMIP6 Models with ARM Data

TODD EMMENEGGER,^a YI-HUNG KUO,^a SHAOCHENG XIE,^b CHENGZHU ZHANG,^b
CHENG TAO,^b AND J. DAVID NEELIN^a

^a *Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, California*

^b *Lawrence Livermore National Laboratory, Livermore, California*

(Manuscript received 21 May 2021, in final form 16 May 2022)

ABSTRACT: A set of diagnostics based on simple, statistical relationships between precipitation and the thermodynamic environment in observations is implemented to assess phase 6 of the Coupled Model Intercomparison Project (CMIP6) model behavior with respect to precipitation. Observational data from the Atmospheric Radiation Measurement (ARM) permanent field observational sites are augmented with satellite observations of precipitation and temperature as an observational baseline. A robust relationship across observational datasets between column water vapor (CWV) and precipitation, in which conditionally averaged precipitation exhibits a sharp pickup at some critical CWV value, provides a useful convective onset diagnostic for climate model comparison. While a few models reproduce an appropriate precipitation pickup, most models begin their pickup at too low CWV and the increase in precipitation with increasing CWV is too weak. Convective transition statistics compiled in column relative humidity (CRH) partially compensate for model temperature biases—although imperfectly since the temperature dependence is more complex than that of column saturation. Significant errors remain in individual models and weak pickups are generally not improved. The conditional-average precipitation as a function of CRH can be decomposed into the product of the probability of raining and mean precipitation during raining times (conditional intensity). The pickup behavior is primarily dependent on the probability of raining near the transition and on the conditional intensity at higher CRH. Most models roughly capture the CRH dependence of these two factors. However, compensating biases often occur: model conditional intensity that is too low at a given CRH is compensated in part by excessive probability of precipitation.

KEYWORDS: Atmosphere; Convection; Convective clouds; Precipitation; Climate models; Clouds

1. Introduction

The U.S. Department of Energy (DOE)'s Atmospheric Radiation Measurement (ARM) program provides unique field observations at its tropical and subtropical sites for exploring the complex interactions and relationships between clouds, aerosols, and radiation and their associated large-scale environments (Mather and Voyles 2013; Xie et al. 2010; Zhang et al. 2020). Bretherton et al. (2004) identified a robust nonlinear relationship between column water vapor (CWV) and precipitation, $\langle P \rangle$, conditionally averaged by CWV bin. At sufficiently high temporal and spatial resolution, conditionally averaged precipitation ramps up once a critical CWV has been reached. That is, for CWV values past a certain threshold, the expected rate of precipitation begins to increase rapidly or “pickup.” The CWV–precipitation relationship applies to both mesoscale and smaller convective scales and is observed over tropical land and ocean in observational records (Neelin et al. 2009; Schiro 2017) and some models (Kuo et al. 2020).

This sharp increase in precipitation represents the occurrence of conditional instability in the transition to deep convection as a function of thermodynamic conditions (Schiro et al. 2016). Holloway and Neelin (2009) assessed this transition with radiosondes from the tropical western Pacific ARM site Nauru, concluding that CWV serves as a proxy for the role of free tropospheric humidity on the buoyancy of entraining plumes. Plume buoyancy is highly sensitive to entrainment and mixing of environmental air; the importance of free-tropospheric humidity to the onset of deep convection is well known and is reflected in the pickup (Kuo et al. 2018; Bretherton et al. 2004; Ahmed and Schumacher 2015; Neelin et al. 2009). Recent work has focused on the conditional average temporal evolution associated with this transition (Wolding et al. 2020).

The onset of deep convection is a complex process, and realistic representations of it in climate models are necessary for accurate simulations of convective-precipitation statistics (Sahany et al. 2012). In past generations of global climate models (GCMs), model precipitation has been shown to be too insensitive to free-tropospheric humidity, leading to errors on a range of spatial and temporal scales. A number of studies (Neale et al. 2008; Bechtold et al. 2008; Zhao et al. 2009; Neelin et al. 2010; Sahany et al. 2012) have identified significant model improvement with more realistic representations of entrainment and mixing (Schiro et al. 2016). Kuo et al. (2020) compared pairs of models which differ primarily in their moist convective parameterization schemes [e.g., the GFDL AM4 variants AM4G9 (double-plume convective scheme) and the AM4B6 (Donner convective scheme)] and

 Denotes content that is immediately available upon publication as open access.

 Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-21-0386.s1>.

Corresponding author: Todd Emmenegger, temmen@atmos.ucla.edu

DOI: 10.1175/JCLI-D-21-0386.1

© 2022 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

concluded that the basic convection onset statistics can distinguish convective parameterizations. A large intermodel spread in tropical precipitation presents a challenge toward future climate projections (Biasutti and Sobel 2009; Biasutti 2013; Voigt et al. 2016; Su et al. 2017; Maloney et al. 2019). Tropical precipitation contributes significantly to biases in climatologies and large-scale modes of tropical variability such as the MJO (Del Genio 2012; Zhu and Hendon 2015; Jiang et al. 2021), the diurnal cycle of precipitation (Del Genio and Wu 2010; Rio et al. 2009; Hourdin et al. 2013; Covey et al. 2016), and the double ITCZ (Mapes and Neale 2011; Hwang and Frierson 2013; Oueslati and Bellon 2013; Hirota et al. 2014; Tian and Dong 2020).

Model diagnostics of tropical convection play an integral part in understanding the limitations of model behavior in future warming scenarios; this study explores biases in tropical precipitation related processes across a model cohort from the recent phase 6 of the Coupled Model Intercomparison Project (CMIP6). The analysis tools presented in this work form part of the ARM data-oriented metrics and diagnostics package (ARM-DIAGS; Zhang et al. 2020), to facilitate process-level evaluation of climate models using ARM datasets. After overviewing data and models (section 2), in section 3 we present the convective transition statistics at two tropical western Pacific sites, Manus and Nauru, highlighting the important features of the observed datasets and discuss and compare CMIP6 model behavior. In section 4 we discuss the influence of temperature in the surrounding large scale environment and identify potential biases of models. In section 5 we present the statistics compiled in column-relative humidity as a way of reducing model temperature bias in the diagnostics.

2. Data and models

The ARM program datasets of precipitation and CWV are used as an observational baseline in the model comparison in evaluating the relationship of CWV and precipitation. Here we present the results from two tropical western Pacific ARM sites: Nauru (0°31'S, 166°54'E) for the period 1999–2009 and Manus (2°3'S, 147°25'E) for 1998–2010. ARM sites use a microwave radiometer for CWV measurements and an Optical Scientific optical rain gauge (ORG815) for precipitation (hereinafter referred to as ORG). CWV measurements are recorded every 20 s but exhibit gaps due to the “wet-window” problem in which water collects on the surface of the lens and introduces erroneous measurements during intense precipitation. As in Schiro et al. (2016), the wet-window problem is addressed by a linear interpolation of CWV values across time periods of 6 h or less. The wet-window problem primarily affects high CWV bins, and the interpolation is often across short time scales and underestimates peak CWV (Schiro et al. 2016). CWV during strong precipitation events does not diminish rapidly—typically on the scale of hours (Schiro et al. 2016).

In addition to the in situ precipitation data, satellite retrieval Tropical Rainfall Measuring Mission (TRMM) 3B42 version 7 precipitation measurements (TRMM 2011) are augmented with the in situ CWV. In model evaluation, TRMM 3B42 data are

averaged spatially over 1° from its $0.25^\circ \times 0.25^\circ$ resolution, centered on the latitude–longitude point nearest to the sites.

Here we analyze 11 CMIP6 models and include the 1-hourly European Centre for Medium-Range Weather Forecasts Reanalysis Fifth Generation (ERA5) atmospheric reanalysis product (Hersbach et al. 2020); model and the reanalysis data used are summarized in Table 1. The 6-hourly snapshots of three-dimensional moisture and temperature, and 3-hourly averaged precipitation rates are extracted from the model cohort for the years 1995–2015 of their historical experiment. The Department of Energy (DOE) E3SM project provides data from two simulations: E3SM-1-0 $0.25^\circ \times 0.25^\circ$, a 21-yr, high-resolution coupled simulation forced with repeating annual cycle forcing approximating observed conditions in 1950 (Caldwell et al. 2019) and E3SM-1-0-EAM $1.00^\circ \times 1.00^\circ$, a 6-yr atmosphere-only Atmospheric Model Intercomparison Project (AMIP) climatology run using present-day climate forcing for the year 2000, with climatological sea surface temperature and sea ice prescribed from observations (Rasch et al. 2019; Xie et al. 2018; Zhang et al. 2019). These models were chosen based on availability of 6-hourly three-dimensional snapshots of temperature and moisture and 3-hourly average precipitation data in the Earth System Grid Federation archive at the time of analysis.

The data for each model is extracted from the nearest point to each ARM site. The column for CWV is defined as from the surface to 200 mb (1 mb = 1 hPa). Precipitation data for NASA-GISS, MPI-ESM1-2-LR, MIROC-E2SL, MIROC6, MRI-ESM2-0, NESM3, E3SM-1-0-EAM, and E3SM-1-0 is 3-hourly averaged output, while MPI-ESM1-2-LR, NorESM2-LM, and NorESM2-MM output is 6-hourly averaged. The effects of using different temporal averaging windows in the statistics are discussed in the beginning of section 3. With the exception of the E3SM-1-0-EAM, all model outputs are from coupled ocean-atmosphere GCMs. Model CWV data for all models are 6-hourly instantaneous snapshots. CWV is analyzed as causal to precipitation; the start of the precipitation averaging window coincides with the time of the CWV snapshot, or precipitation lags CWV. The statistics presented in the following sections were also evaluated with precipitation leading CWV and the statistics remained robust (not shown). We note the caveat that the ARM radiometer CWV is measured locally, as opposed to the model values from a grid cell on the order of (100^2) km² and is available at higher time resolution. The comparison remains reasonable given that CWV tends to have longer spatial and temporal autocorrelation than precipitation (Holloway and Neelin 2010; Abbott et al. 2016; Kuo et al. 2018) and that coarse-graining of data on scales from 25 to 200 km has modest impacts (Kuo et al. 2018). The spatial dependence of the convective onset statistics is further explored in section 3.

In section 4, we address the temperature dependence of the moisture–precipitation relationship which requires a calculation of the column-integrated (from the surface to 200 mb) saturation specific humidity $\overline{q_{\text{sat}}}$, $\overline{q_{\text{sat}}} \equiv g^{-1} \int q_{\text{sat}}[T(p), p] dp$, where $q_{\text{sat}}[T(p), p]$ is the saturation humidity with respect to liquid water, is used as a bulk measure of tropospheric temperature. The column relative humidity for observations is calculated from 1-hourly snapshots of CWV and temperature

TABLE 1. List of models analyzed in this study. Table is similar to the table in Pathak et al. (2019) for CMIP5 models. The references column gives the data source and the source for the convective trigger/closure.

Model name	Institute	Resolution (lon $\times$ lat), vertical levels	Convective trigger/closure	References
NASA-GISS	Goddard Institute for Space Studies, NASA	$2.00^\circ \times 2.00^\circ$, 40	Higher moist static energy of an adiabat to environment at level above/cloud-base buoyancy	(NASA/GISS) (2018); DelGenio and Yao (1993); Schmidt et al. (2014)
MPI-ESM1-2-HAM	Max Planck Institute for Meteorology	$1.85^\circ \times 1.88^\circ$, 47	Net positive moisture convergence/CAPE	Wieners et al. (2019); Möbis and Stevens (2012)
MPI-ESM1-2-LR	Max Planck Institute for Meteorology	$1.85^\circ \times 1.88^\circ$, 47	Net positive moisture convergence/CAPE	Wieners et al. (2019); Möbis and Stevens (2012)
MIROC-E2SL	Japan Agency for Marine-Earth Science and Technology	$2.77^\circ \times 2.81^\circ$, 40	CAPE/prognostic convective kinetic energy	Hajima et al. (2019); Ando et al. (2021)
MIROC6	Japan Agency for Marine-Earth Science and Technology	$1.39^\circ \times 1.41^\circ$, 81	CAPE/prognostic convective kinetic energy	Tatebe and Watanabe (2018); Ando et al. (2021)
MRI-ESM2-0	Meteorological Research Institute Japan Meteorological Agency	$1.11^\circ \times 1.13^\circ$, 80	CAPE/CAPE	Yukimoto et al. (2019, 2011)
NorESM2-LM	Norwegian Meteorological Institute	$1.89^\circ \times 2.50^\circ$, 32	CAPE/CAPE	Seland et al. (2019, 2020)
NorESM2-MM	Norwegian Meteorological Institute	$0.94^\circ \times 1.25^\circ$, 32	CAPE/CAPE	Bentsen et al. (2019); Seland et al. (2020)
NESM3	Nanjing University of Information Science and Technology	$1.85^\circ \times 1.88^\circ$, 47	Net positive moisture convergence/CAPE	Cao and Wang (2019); Cao et al. (2018)
E3SM-1-0	DOE E3SM Project	$0.25^\circ \times 0.25^\circ$, 72	CAPE/CAPE	Caldwell et al. (2019)
E3SM-1-0-EAM	DOE E3SM Project	$1.00^\circ \times 1.00^\circ$, 72	CAPE/CAPE	Rasch et al. (2019)

of the ARM Best Estimate data (ARMBE; Xie et al. 2010), an hourly integrated product assembled from various ARM measurements for use in climate model evaluation. Stringent quality controls were applied to raw ARM data used in producing ARMBE. The 1-hourly snapshots at the beginning of the precipitation averaging window are chosen for the ARMBE statistics, remaining consistent with the causal relationship of CWV to precipitation. Model $\widehat{q_{\text{sat}}}$ is calculated from 6-hourly snapshots and treated in the same fashion as CWV in that it leads the precipitation averaging window. In the following discussion, we use “ARM” to refer to the data directly from ARM individual instruments, which are often at much higher temporal frequencies and applied less data quality controls than the ARMBE product.

3. Convective transition statistics in CWV

Figures 1 and 2 show the basic set of convective onset statistics for ensemble of CMIP6 models at the Nauru and Manus site. Schiro et al. (2016) and Kuo et al. (2018) examined the effects of temporal averaging on the convection onset statistics, concluding that up to and including a 3-h averaging window (the temporal resolution of the majority of the cohort of CMIP6 models examined here) the prominent characteristics of the pickups (Figs. 1a and 2a) are largely preserved; beyond 3 h, the pickups begin to be slightly smoothed by the averaging.

Observational data for Figs. 1 and 2 use 3-h averages, the effects of temporal averaging are shown in Fig. S1 of the supplemental material.

The convective transition statistics of three observational products: TRMM Microwave Imager processed by Remote Sensing Systems algorithm v7.1 (TMI; Wentz et al. 2015), $0.25^\circ \times 0.25^\circ$ resolution, ARMBE, and ARM are compared in Fig. 3. CWV measurements between observational products are compared in Fig. S2 of the supplemental material, and overall, observational CWV measurements are consistent, with a slight low bias in TMI measurements. In ground-based observational products, ARM and ARMBE, the drop of precipitation at the highest CWV can be attributed to the wet-window problem (Kuo et al. 2018). That is, the CWV values during high precipitation events are likely missing from the record, and the gap-filling can only partially restore the missing information. Heavily raining times are thus preferentially missing, or are associated with gap-filled portions of the CWV that tend to miss the very highest values, resulting in a low bias in precipitation at the highest CWV. The $\langle P \rangle$ of TMI reaches much higher precipitation rates ($\sim 6 \text{ mm h}^{-1}$) at the highest CWV bin (Fig. 3). ARMBE data exhibits a $\langle P \rangle$ with nonzero values at low CWV due to its hourly average.

The probability of precipitation curves (Figs. 1b and 2b) display the most sensitivity to temporal averaging; as the temporal averaging increases, the pickup occurs at lower CWV

Nauru annual Averaged over 3 hours

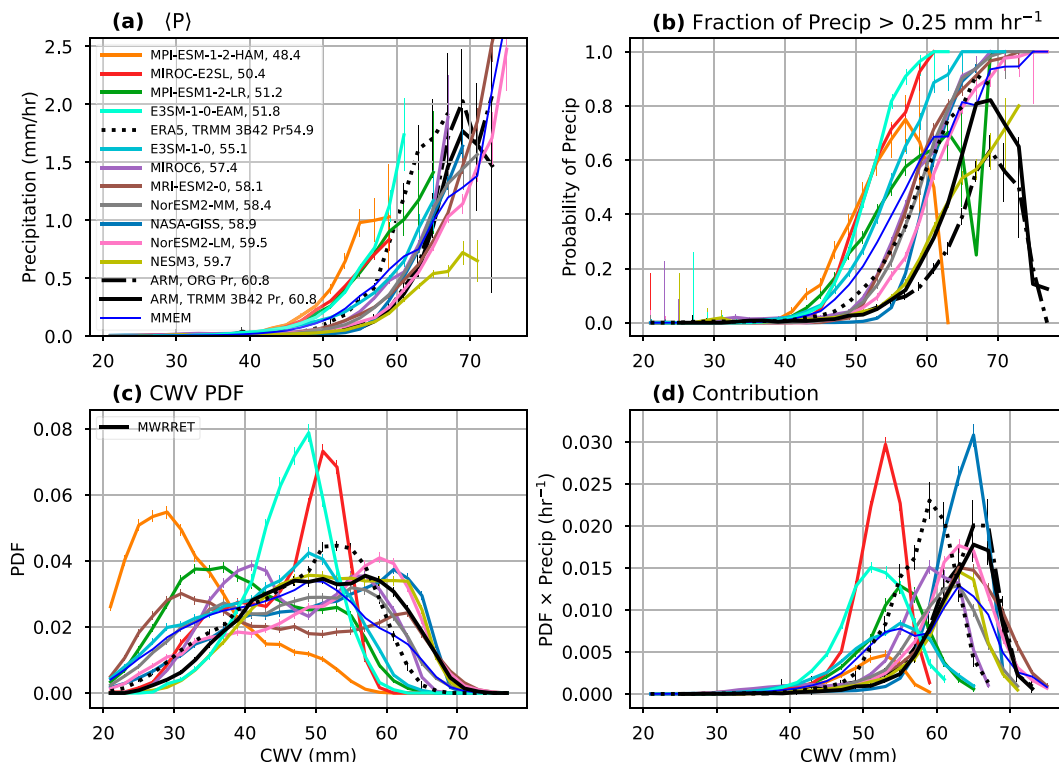


FIG. 1. (a) CWV conditionally averaged precipitation rate ($\langle P \rangle$), the critical water vapor value of the pickup (w_c) is quantified by a linear fit to $\langle P \rangle$ and is listed in the legend; models are listed in ascending order according to their estimate of w_c , error bars represent the standard error probability of precipitation, (b) fraction of precipitation $> 0.25 \text{ mm h}^{-1}$ in respective CWV bin, error bars are the Wilson score interval for 90% confidence, (c) the CWV PDF, with 90% Wald interval error bars, and (d) the precipitation contribution, or the $\langle P \rangle$ -weighted PDF ($\langle P \rangle \times \text{CWV PDF}$), with error bars that represent the error propagation of the weighting [the sum of fractional error in (a) and (c)] for ARM in situ ORG (dashed black), TRMM-3B42 1° spatial average precipitation (solid black), CMIP6 models (colors), and the multimodel mean (blue) for the Nauru ARM site. The observational and all model precipitation data with the exception of MPI-ESM1-2-HAM, NorESM2-LM, and NorESM2-MM are averaged over 3-h windows. MPI-ESM1-2-HAM, NorESM2-LM, and NorESM2-MM precipitation data are 6-h averages. Both observational precipitation products are augmented with ARM in situ microwave radiometer CWV. The CWV bin width is 2 mm. Statistics are not calculated for bins with less than 10 counts.

and higher CWV values display a higher probability of precipitation. Temporal averages over 3 h or greater intervals are more likely to span raining times. The wet-window problem similarly affects the observed probability at the highest CWV.

The important feature of the conditionally averaged precipitation is the sharp, sudden increase of $\langle P \rangle$ or “pickup” above a critical CWV value. We refer to this critical value of CWV as w_c . Previous quantitative approximations of w_c have been calculated through the CWV-axis intercept of a linear fit in a representative range of $\langle P \rangle$ (Sahany et al. 2012; Kuo et al. 2018). The range for the fit used here corresponds to consecutive monotonically increasing $\langle P \rangle$ above 0.25 mm h^{-1} and w_c is defined as the point in which the linear fit crosses a precipitation value of 0.25 mm h^{-1} . Applying these conditions to approximating w_c lessens low w_c bias caused by slower increasing $\langle P \rangle$ (see NESM3 $\langle P \rangle$ in Figs. 1a and 2a for example of weakly increasing $\langle P \rangle$) and/or nonzero $\langle P \rangle$ values at low CWV caused

by potential drizzle biases. Approximated w_c values are displayed in the legend of Figs. 1 and 2.

An important takeaway when comparing observational products in Fig. 3, is that the critical value of the pickup remains robust across the satellite retrieval and ARM datasets at $\sim 60 \text{ mm}$. Although the form of the pickup remains the same, the w_c value can depend on factors that vary between land and ocean such as the mean tropospheric temperature, boundary layer dynamics and a stronger diurnal cycle over land (Schiro et al. 2016). In evaluating model performance with the convective transition statistics, whether the model point nearest to the ARM site used in analysis is one over land or ocean could have an impact on the location of its w_c . In Fig. 3, w_c remains robust at the Nauru sites across the ARM ground-based data and the TMI satellite retrieval. Note that only Nauru is shown here as TMI retrieval data are limited to points over oceans; the spatial extent of Nauru [$4 \text{ km} \times 6 \text{ km}$ (Long and McFarlane 2012)] is

Manus annual Averaged over 3 hours

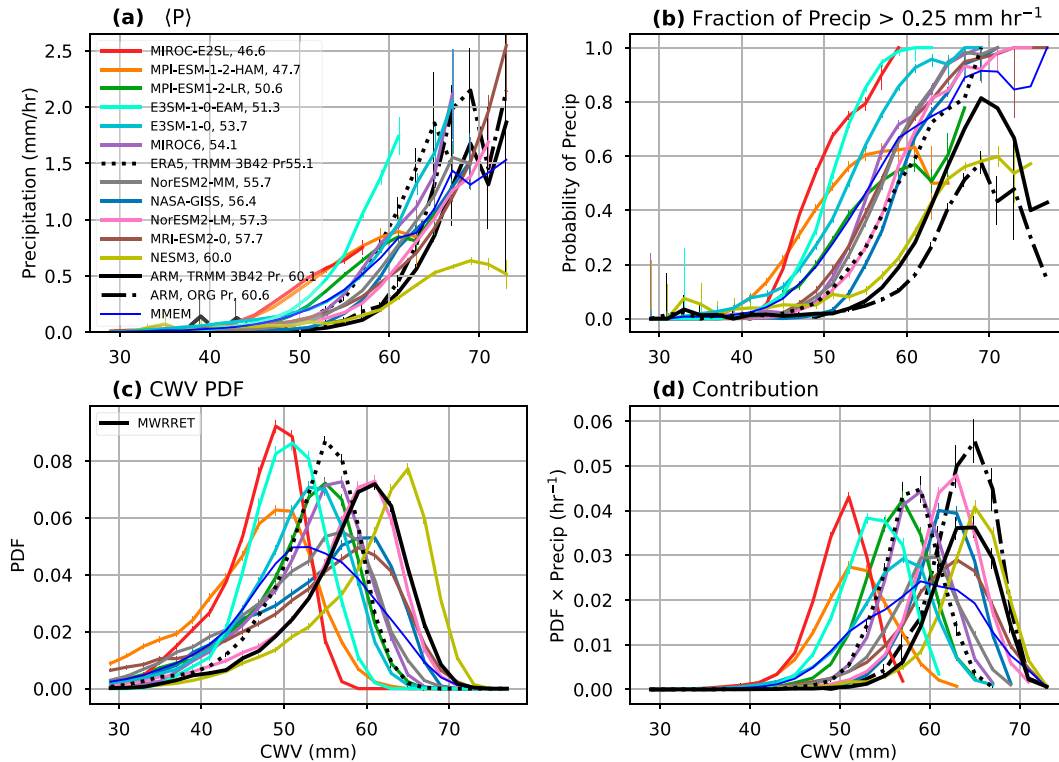


FIG. 2. As in Fig. 1, but for the Manus Island ARM site.

small enough such that TMI records data for the site, while the $2^\circ \times 2^\circ$ data centered on Manus contains missing points. Nauru may exhibit more maritime characteristics due to its small spatial extent, but still has some island heating characteristics (Long and McFarlane 2012), impacting boundary layer structure and generating downwind cloud streets (Matthews et al. 2007).

The spatial dependence of the convective transition statistics was discussed in Kuo et al. (2018); the authors find $\langle P \rangle$ and the CWV PDF insensitive to the spatial resolution due to the large autocorrelation on spatial scales of moisture and temperature, while the probability of precipitation shifts slightly toward lower CWV values. Figure 3 shows a comparison of observational products and the effects of spatial averaging in the convective onset statistics. The spatial averages taken from boxes of 0.5° to 2.0° centered on Nauru do not substantially differ from each other. The w_c of the spatial averages occur slightly earlier than that of the non-averaged points. The probability of precipitation $\geq 0.25 \text{ mm h}^{-1}$ occurs earlier for spatial averages, and increases at a slightly faster rate, particularly at lower CWV values where the probability transitions from nonzero values.

The robustness of $\langle P \rangle$ over different spatial scales up to a 2° box is due to the large spatial autocorrelation of CWV. The TMI curve (green) in Fig. 3 includes all points within a 1° box centered on Nauru; compared to the 1° average in the plot (purple), the pickup is slightly later and CWV PDF extends to higher values, suggesting that the spatial average reduces the probability of high CWV values.

The characteristic shape of the CWV PDFs of the observations features a gradual increase at a low CWV toward a broad peak which drops off rapidly around w_c and is determined by the large-scale flow interacting with convective physics (Kuo et al. 2020). The CWV PDF is limited on the left by the large-scale flow—regions dominated by ascent (descent) spend more time at higher (lower) CWV—and the right by precipitation. The sharp decrease in the CWV PDF at higher $\langle P \rangle$ is consistent with precipitation becoming an effective moisture sink as the CWV reaches w_c (Schiro et al. 2016). The aforementioned wet window could also affect the highest portion of the CWV PDF, reducing the frequency at which high CWV values are recorded by the radiometer during strong precipitation. In comparing observational products, (satellite to ground-based measurements; Fig. 3) the most notable differences in the statistics are the characteristics of the CWV PDFs. At both sites, the CWV PDFs of TMI data rise and peak at lower CWV values and drop off before the in situ radiometer data. It is likely this low bias in the TMI data stems from the retrieval algorithms of the instruments, as the bias is noticeable in the scatterplots of the CWV from the ARM in situ and TMI (Fig. S2). However, the start of the CWV PDF drop-off is robust between the two datasets, differing in the rate at which the PDF begins drop.

In general, models tend to have an early pickup. Around half of the models: NASA-GISS, NorESM2-LM, NorESM2-MM, MIROC6, and E3SM-1-0-EAM reach similar $\langle P \rangle$ at their highest CWV bin to observations at both sites. MRI-ESM2-0

Nauru annual Averaged over 3 hours

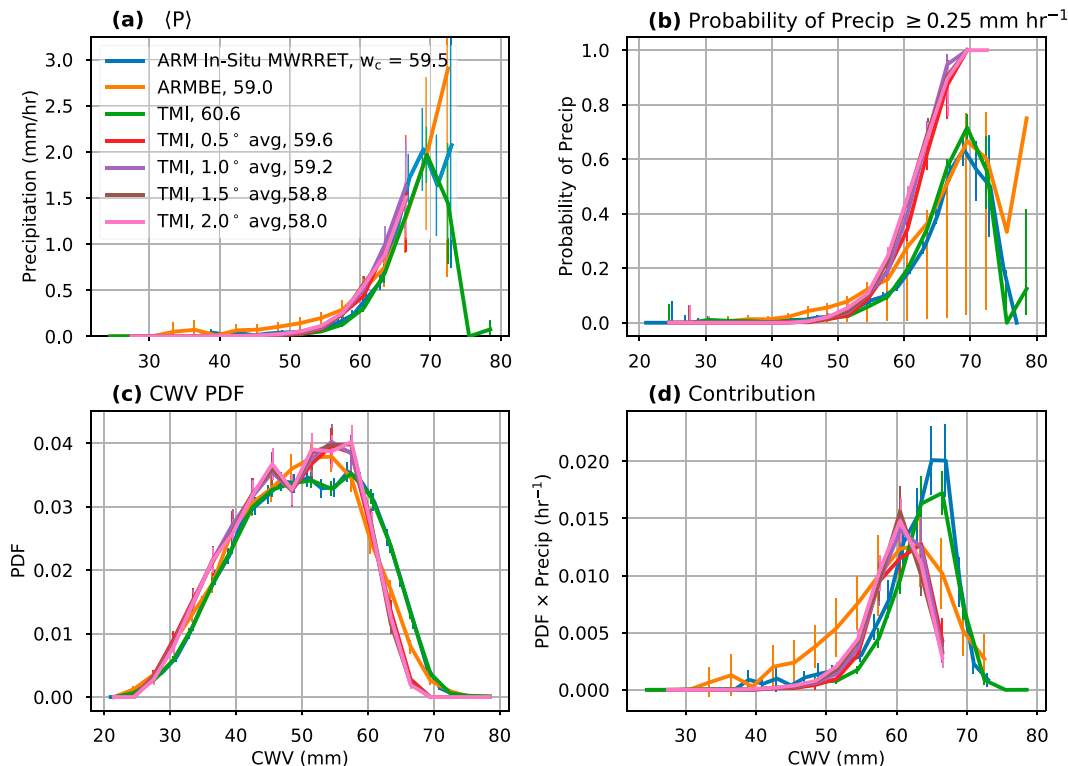


FIG. 3. As in Fig. 1, but for observational products and different spatial averages.

precipitates at a higher rate ($\sim 3 \text{ mm h}^{-1}$) at both sites, picking up earlier yet increasing its $\langle P \rangle$ at a slower rate. MPI-ESM-1-2-HAM, MIROC-ES2L, and NESM3 show low maximum rates of $\langle P \rangle$ at both sites. The approximated w_c for some models (MIROC6, MRI-ESM2-0, NorESM models, and NESM3) models are within a range of $\sim 3 \text{ mm}$ to observations at both sites; the difference in the approximated w_c in these models stems from the rate at which their $\langle P \rangle$ increases. MPI-ESM-1-2-HAM, NorESM2-LM, and NorESM2-MM precipitation data are 6-h time averages. As discussed in the beginning of this section, a 6-h temporal averaging window slightly smooths the pickup in $\langle P \rangle$ and reduces the precipitation rate only slightly, although NorESM2-LM and NorESM2-MM $\langle P \rangle$ curves agree fairly well with observations.

Respective model performance of $\langle P \rangle$ remains relatively consistent in magnitude, slope, and w_c of $\langle P \rangle$ across the two sites. Model pickup tends to be too weak with increasing CWV at both sites. NASA-GISS, MIROC6, and E3SM-1-0-EAM pickup at a rate similar to that of observations at Nauru, while only E3SM-1-0-EAM replicates the pickup slope at Manus. One known factor that can contribute to an early pickup (i.e., at low CWV) is insufficient effect of lower free tropospheric moisture on buoyancy through entrainment (Holloway and Neelin 2009; Kuo et al. 2017; Schiro and Neelin 2019). It is also plausible that the early pickup in models could be associated with the pervasive model drizzle problem (Rushley et al. 2018; Stephens et al. 2010; Jing et al. 2017; Chen and Dai 2019),

suggesting that the lack of sufficiently strong dependence on the moisture environment yields models that precipitate too frequently at low intensities. In addition, the unrealistically strong coupling of convection with surface heating, as assumed in many convection parameterizations, may also lead model convection to be triggered too easily (Xie et al. 2019).

Figures 1b and 2b show the fraction of precipitation over a threshold, 0.25 mm h^{-1} , in each CWV bin. The error bars represent the Wilson score 95% confidence interval for each bin. In comparison to observations, the fraction above the threshold increases at a similar rate to the ORG precipitation set of observations at Nauru, and slightly faster than the TRMM-3B42 precipitation set with increasing CWV. At Manus, the majority of models increase slightly faster than both observational precipitation datasets. We note that the probability of precipitation (Figs. 1b and 2b) is more sensitive to averaging than $\langle P \rangle$, both to temporal averaging as previously discussed, and to spatial averaging, as observed when comparing the higher precipitation fraction of the TRMM 3B42 curve (averaged over 1°) in comparison to the in situ ORG curve. Nonetheless there is high consistency between the probability measure of the pickup and the conditional average measure—models that pickup early in one also do so in the other.

Differences in the observational CWV PDFs between the Nauru and Manus site are easily observed in terms of their shapes. Nauru exhibits a broad peak from ~ 40 to 60 mm , while Manus shows a more narrow peak from ~ 55 to 65 mm .

This reflects the large-scale flow around each site. The Manus site is located at the heart of the western Pacific warm pool and exhibits a more pronounced ascent climatology compared to the Nauru site, which is located on the edge of the western Pacific warm pool (Xie et al. 2010). Nauru experiences both the upwelling and downwelling portions of the Walker circulation (Long and McFarlane 2012).

At the Nauru site, most models exhibit more complicated CWV PDFs than the observations. MPI-ESM-1-2-HAM has a strong peak toward lower CWV, with a long trailing tail to higher CWV at the right. Generally, models exhibit either a broad peak or bimodal distribution at Nauru (Fig. 1c). The more complicated PDF shapes observed in models could possibly be attributed to climatological biases. The Nauru site's location on the edge of the western Pacific warm pool is associated with a weaker ascent regime compared to the Manus site, leading to more time spent at lower CWV. The location of the sites in the tropical western Pacific would suggest likely effects of ENSO related forcing on vertical motion. The peak at lower CWV can be attributed to the balance between surface evaporation and moisture divergence (Kuo et al. 2018). E3SM-1-0-EAM and MIROC-E2SL show a prominent peak and narrow range of CWV, with a steep and early drop-off around their w_c for both Nauru and Manus (Figs. 1c and 2c). In the above discussion of the relation of w_c to the CWV PDF and the peaks of the contribution, models that pickup early (late) in general have CWV PDFs and contribution peaks shifted toward lower (higher) CWV (Figs. 1c and 2c).

Overall, the models replicate the characteristic shape of the CWV PDF at the Manus site, implying that the models are able to capture the ascent regime of the large scale flow at Manus. NorESM2-LM almost replicates the observed PDF exactly. The widths of the CWV PDF peaks are appropriate for the majority of models at Manus. Peaks of the CWV PDF for MRI-ESM2-0, Nor-ESM2-MM, NASA-GISS rise too slow, and these models have fatter tails to the left of their peaks. Models peak at lower CWV than in observations, consistent with the location of their w_c . Although NESM3's w_c is the closest to observations, its PDF peaks at a higher CWV. It is possible that the weaker pickup of this model allows the column to stay at higher CWV as precipitation is not as effective of a moisture sink as in other models. This calls for an additional consideration in the relationship between the pickup and CWV PDF. Typically, the CWV PDF is characterized by an abrupt drop on the right at higher CWV around values in the range of the pickup; the sudden drop is consistent with precipitation's dissipative effects of water vapor and buoyancy (Schiro et al. 2016). The slope of $\langle P \rangle$ reflects the behavior of a convective event in which the high precipitation rate cannot be sustained as CWV and buoyancy are consumed, and the precipitation rate starts to diminish. All models capture the prominent peak in the CWV PDF at Manus (Fig. 2c) and for the most part have CWV PDF drop-off rates similar to observations, yet have more shallow $\langle P \rangle$ rates of increase.

Distributions of precipitation contributions give how much of the total amount of precipitation comes from a given range of water vapor values (units are h^{-1} from the precipitation units mm h^{-1} and the PDF units of mm^{-1} giving the occurrences

per interval of water vapor). These are useful because they capture the combination of two effects: the sharp increase of conditionally averaged precipitation at w_c and the rapid drop in the frequency of occurrence above w_c . These provide similar information as the PDFs of precipitating points (the PDF of time spent at a CWV and precipitating greater than 0.25 mm h^{-1} ; not shown here) but without the need of setting a threshold.

The contribution peaks are generally wider than those of the observations. Contribution peaks for the models with the appropriate w_c to observations at Nauru, NESM3 and NorESM2-LM, show early peaks in their contribution. All models behave appropriately in their characteristics of CWV PDFs and precipitation contributions respective of their w_c : rapid drop-off of the CWV PDF in the range of high $\langle P \rangle$ and precipitation contribution peaks in a narrow range after w_c .

We note significant difference between the ERA5 reanalysis and the ARM CWV at both sites. The ERA5 w_c occurs at a lower value than the ARM data at both sites, resulting in a low-CWV environment bias (Figs. 1 and 2). A significant contributor to this is likely temperature bias, as discussed in section 5. Figure 3 illustrates that significant differences in $\langle P \rangle$ and the CWV PDFs are not readily attributable to those which arise from varying spatial-averaging domains and in situ measurements.

4. Temperature dependence

In evaluating the models, biases of the large-scale environment should be taken into account before definitive conclusions of the efficacy of model's convective-precipitation-related processes are reached. The temperature dependence of the convection onset statistics is inferred from column saturation: higher average column temperature requires higher CWV to reach saturation which shifts w_c toward higher CWV. The threshold w_c is dependent on the temperature through the convecting layer, sometimes parameterized as a fixed value of its saturation in models [e.g., Betts and Miller (1986) scheme]. Models that exhibit a cold bias in daily temperatures should therefore reach column saturation at lower CWV and pick up earlier (lower w_c). A traditional way to account for a temperature bias would be to examine convection onset statistics in column relative humidity [$\text{CRH}(\%) \equiv (\text{CWV}/\widehat{q_{\text{sat}}}) \times 100$], where CRH is calculated with its instantaneous CWV and $\widehat{q_{\text{sat}}}$ values (Bretherton et al. 2004; Wolding et al. 2020).

The convection onset statistics for observations (1° -averaged TRMM-3B42 precipitation) and three representative models—the rest of the models are shown in the supplemental material (Figs. S5 and S6)—conditioned on $\widehat{q_{\text{sat}}}$ are shown in Nauru (Fig. 4) and Manus (Fig. 5). The same data in the CWV domain is shown in the supplemental material Figs. S3 and S4. The models shown here were chosen to give a survey of various model biases with respect to observations: MRI-ESM2-0 picks up at a slow rate with a w_c within 3 mm of observations and reaches a higher rate of precipitation in its highest CWV bin, NESM3 exhibits the slowest pickup and its CWV PDF drops off at higher CWV than observations, E3SM-1-0 picks up at a rate similar to that of observations, yet much earlier than observations. For observational statistics conditioned on

Nauru Averaged over 3 hours

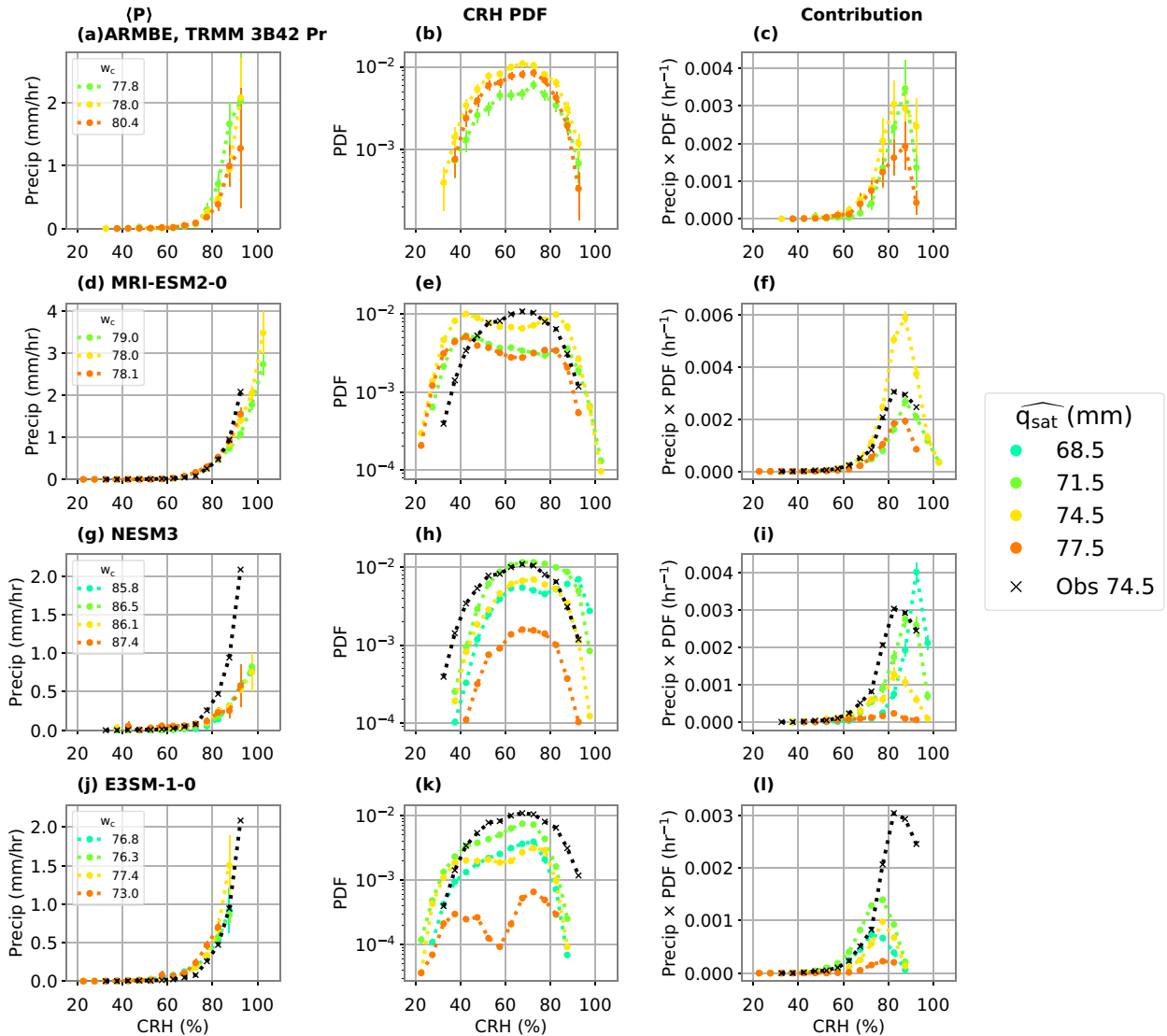


FIG. 4. (a) $\langle P \rangle$ conditioned on $\widehat{q}_{\text{sat}}$, error bars represent the standard error with the w_c for each respective $\langle P \rangle$ listed in the legend, (b) the CRH PDF conditioned on $\widehat{q}_{\text{sat}}$ with 90% Wald interval error bars, and (c) the $\langle P \rangle$ -weighted PDF with error bars that represent the error propagation of the weighting [the fractional sum of error of (a) and (b)] for each $\widehat{q}_{\text{sat}}$ bin for Nauru. The ARMBE CRH with TRMM 3B42 precipitation averaged over 1° centered on the nearest spatial point is shown in (a)–(c). (d)–(l) Three representative models are shown, all with w_c within a close range to that of the observations. The most populated $\widehat{q}_{\text{sat}}$ bin for observations (74.5 mm) statistics are overlaid on the model (black x). Statistics are not calculated for bins with less than 10 counts.

ARMBE derived $\widehat{q}_{\text{sat}}$, the temperature dependence discussed above is observed as $\widehat{q}_{\text{sat}}$ -binned pickup curves and PDFs collapse to a nearly single CRH value. An additional measure of $\widehat{q}_{\text{sat}}$ for observations is calculated from 1-hourly 0.25° European Centre for Medium-Range Weather Forecasts ERA5 atmospheric temperature reanalysis (Hersbach et al. 2020) for the years 2002–14. Results using ERA5 $\widehat{q}_{\text{sat}}$ are displayed in Figs. S5d, S5e, S5f, S6d, S6e, and S6f.

In general, models exhibit some degree of temperature dependence, the two exceptions being the MPI models (shown in supplemental material Figs. S5m–r and S6m–r). Models are

consistent in showing temperature dependence across the Nauru and Manus sites. The $\widehat{q}_{\text{sat}}$ value with the largest contribution (Figs. 4 and 5) depends on large-scale conditions in the frequency of temperature, as the magnitude of $\widehat{q}_{\text{sat}}$ -conditioned $\langle P \rangle$ at saturation does not display a discernible dependence on $\widehat{q}_{\text{sat}}$.

Basinwide convective transition statistics are consistent when conditioned on ERA5 and NCEP2 Reanalyses $\widehat{q}_{\text{sat}}$ (not shown). The $\widehat{q}_{\text{sat}}$ PDF for the ARMBE, NCEP Reanalysis 2, and the ERA5 reanalysis products and models is shown in Fig. 6. A cold bias is seen in the NCEP Reanalysis 2 $\widehat{q}_{\text{sat}}$

Manus Averaged over 3 hours

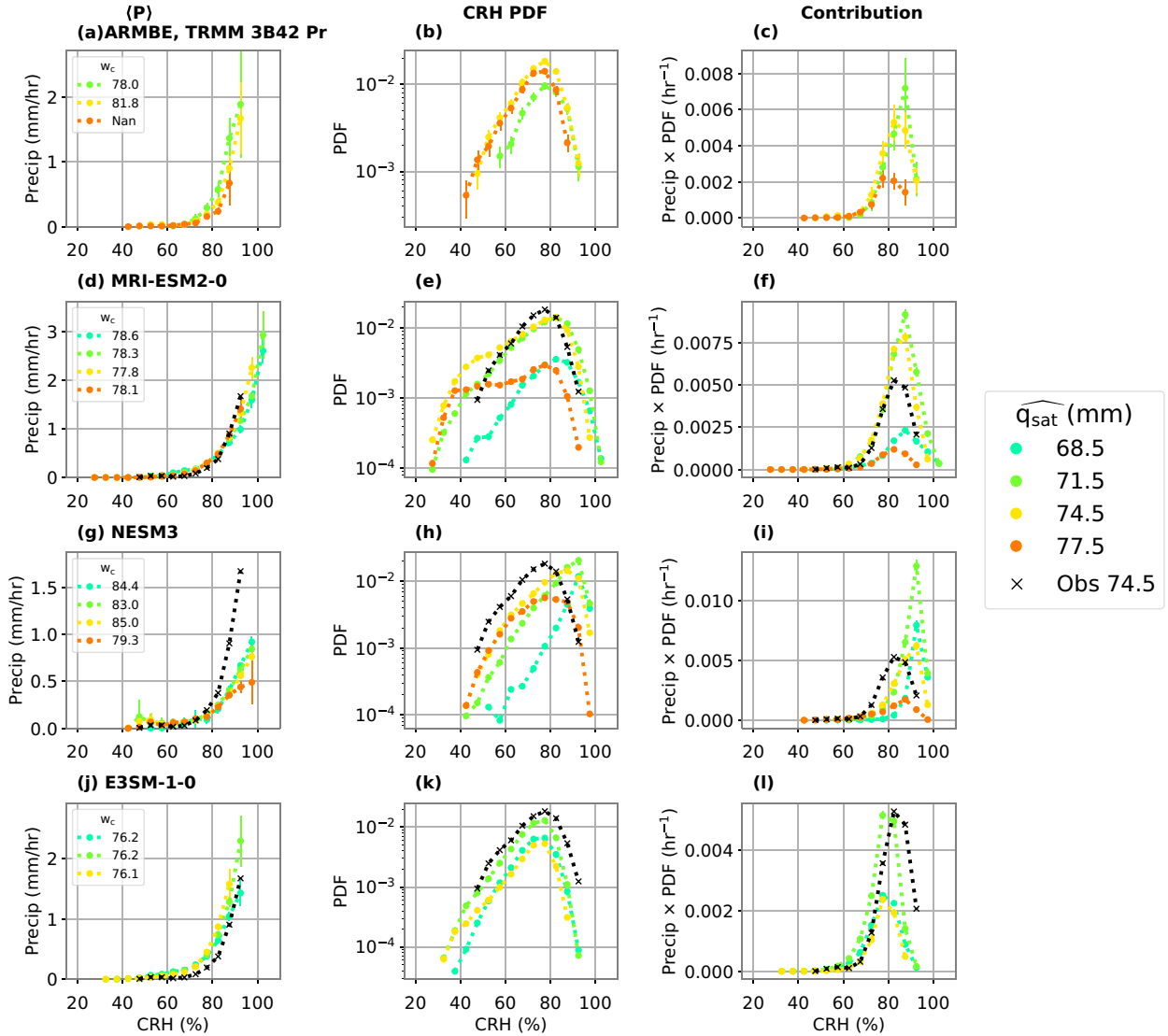


FIG. 5. As in Fig. 4, but for the Manus site. A w_c value of “Nan” in the legend indicates that $\langle P \rangle$ does not contain an adequate number of points in the precipitation rate range for estimation.

compared to ERA5 especially for Manus. ARMBE $\widehat{q_{\text{sat}}}$ is significantly warmer than both reanalysis products and the majority of models. The statistics for column variables of the ARMBE were compared to radiosonde measurements (Stokes and Schwartz 1994) and were found to be nearly identical (not shown).

A noticeable feature of the model comparisons in Fig. 6 is the $\widehat{q_{\text{sat}}}$ bias apparent in each model. In comparing the thermodynamic environment of models, NASA-GISS, MIROC-E2SL and MPI-ESM1-2-HAM have a strong cold bias (peaks ~ 5 mm low of ERA5), MRI-ESM2-0 is the only model with a large warm bias (peaks ~ 3 mm high) at both sites while the other models seem to crowd around the ERA5 $\widehat{q_{\text{sat}}}$ peak, mostly at a slightly higher (within ~ 1 mm) $\widehat{q_{\text{sat}}}$.

The observations and reanalysis products conditioned using $\widehat{q_{\text{sat}}}$ collapse to a nearly single critical CRH at the sites. In ARMBE and models, the collapse is not perfect, owing to the more nuanced relationship of convective onset to plume buoyancy and its associated factors: boundary layer moisture and temperature relative to the lower free troposphere (LFT), and sub-saturation of the LFT (Ahmed and Neelin 2018; Adames et al. 2021). Models are consistent in their temperature dependence in the pickup conditioned on $\widehat{q_{\text{sat}}}$ across sites (Figs. 4 and 5).

The characteristics shapes of the $\widehat{q_{\text{sat}}}$ -binned CRH PDFs in Figs. 4b,e,h,k and 5b,e,h,k are similar to the CWV PDFs (Figs. 1c and 2c) discussed earlier—a peak at low CRH with a sharp decrease toward lower CRH and a slow increase to the

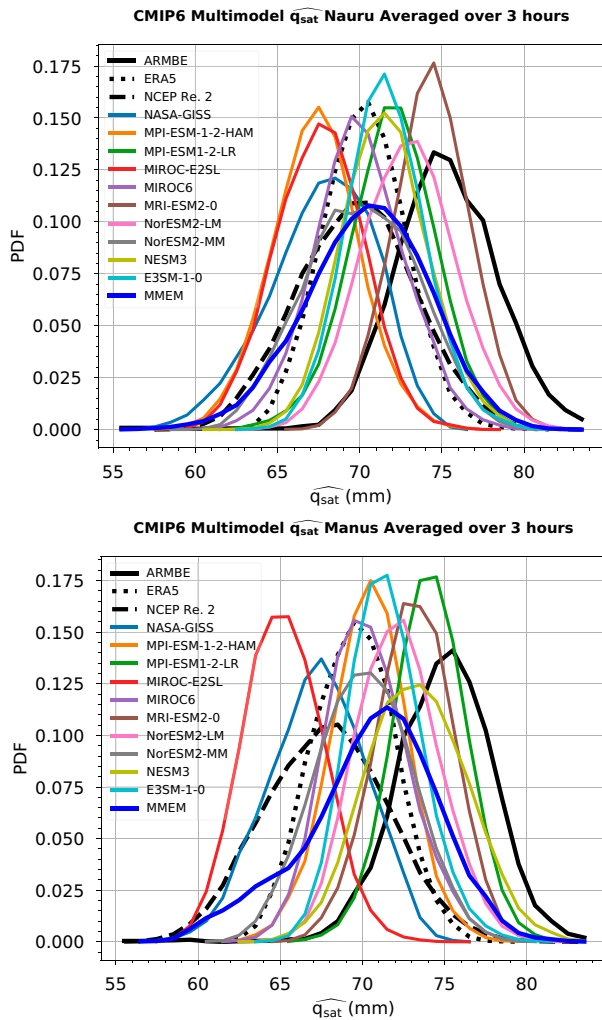


FIG. 6. The $\widehat{q}_{\text{sat}}$ PDF for (top) Nauru and (bottom) Manus of observations and reanalysis products and models.

right, peaking again around w_c , and dropping off sharply. The point at which the PDF (Figs. 4b,e,h,k and 5b,e,h,k) begins to decrease rapidly is referred to as the drop-off. It occurs above w_c in the range where convection and the moisture sink associated with precipitation tend to oppose increases of buoyancy and CWV, respectively. The highest $\widehat{q}_{\text{sat}}$ -bin PDF seems to spend more time at lower CRH and drops off before the other bins. The probability shifting to lower CRH values as $\widehat{q}_{\text{sat}}$ increases is expected, as reaching higher values of CRH at higher $\widehat{q}_{\text{sat}}$ requires more moisture. However, the inconsistency of the CRH value of the drop-offs across the $\widehat{q}_{\text{sat}}$ bins presents a caveat in using CRH in the statistics. CRH tends to overcompensate for temperature dependence as the location of the drop-off of the PDFs in CWV does not increase as fast with temperature as $\widehat{q}_{\text{sat}}$ (Neelin et al. 2009; Kuo et al. 2020). This results in the failure of normalization by $\widehat{q}_{\text{sat}}$ to completely collapse the statistics in Figs. 4b,e,h,k and 5b,e,h,k. In other words, CRH puts too much emphasis on saturation; other considerations with respect to temperature affect the drop-offs of

the $\widehat{q}_{\text{sat}}$ -binned PDFs. For convective precipitation, criteria for conditional instability have a more complex dependence on temperature than simply saturation, so it is unsurprising that temperature biases are not completely resolved when using CRH.

5. Convective transition statistics in CRH

a. Does CRH collapse temperature dependence?

In section 4, we discuss how compiling the convection onset statistics in CRH helps account for potential model temperature biases through $\widehat{q}_{\text{sat}}$ binning. In general, model convection onset statistics display temperature dependence at both sites, $\widehat{q}_{\text{sat}}$ -binned conditionally averaged precipitation in CRH collapse to a single critical CRH, CRH PDFs drop off at a similar CRH—also replicate the “overcompensation” behavior of the higher $\widehat{q}_{\text{sat}}$ bins noted in the observations—and contributions peak around a similar CRH. The early pickups of the models in CWV are suggestive of a cold bias. In Fig. 6 we note a large spread of temperature biases at Nauru compared to the ERA5 dataset: three models which show a relatively strong cold bias (NASA-GISS, MPI-ESM1-2-HAM, MIROC-E2SL), two models with a slight cold bias (MIROC6 and NorESM2-MM), three with a slight warm bias (MPI-ESM1-2-LR, NorESM2-LM, and NESM3), and one model with a strong warm bias (MRI-ESM2-0).

Figures 7 and 8 are similar to Figs. 1 and 2, but the statistics are now compiled in CRH. The statistics for two additional ARM sites are included in the supplemental material to expand on site comparisons included in the ARM-DIAGS package: Darwin, another ARM tropical western Pacific site (2°3' 36"S, 147°25'30"E) and Southern Great Plains (SGP; 36°36'18"N, 97°29'6"W) in Figs. S7 and S8, respectively (Gaus-tad and Riihimäki 1998; Holdridge and Kyrouac 1998; Turner et al. 2007).

b. Basic convective onset statistics in CRH

Despite the limitations of CRH in accounting for temperature dependence (because the CRH PDFs drop off earlier for higher $\widehat{q}_{\text{sat}}$, as discussed above), it does prove useful for accounting for systematic model bias. Specifically, while the majority of models tend to pick up early relative to observations and reanalysis products in CWV, in Figs. 7a and 8a, we see that accounting for temperature biases with CRH shifts model pickups so that they exhibit a spread above and below observations. The ARMBE $\widehat{q}_{\text{sat}}$ for both locations shown here (Fig. 6) is significantly warmer (~5 mm) than ERA5 and the multimodel ensemble mean. Accounting for the large difference in $\widehat{q}_{\text{sat}}$ by compiling the statistics in CRH shifts the pickup of the observations near the center of the model cohort at both sites. Note that despite differences in $\widehat{q}_{\text{sat}}$ between ARMBE and ERA5, the pickups in CRH agree fairly well, suggesting consistency between water vapor and $\widehat{q}_{\text{sat}}$ values within each product.

The CRH statistics are not limited to only adjusting the pickup. Similar to the CWV PDFs shown in section 3, the CRH PDFs of the models maintain their ability to capture

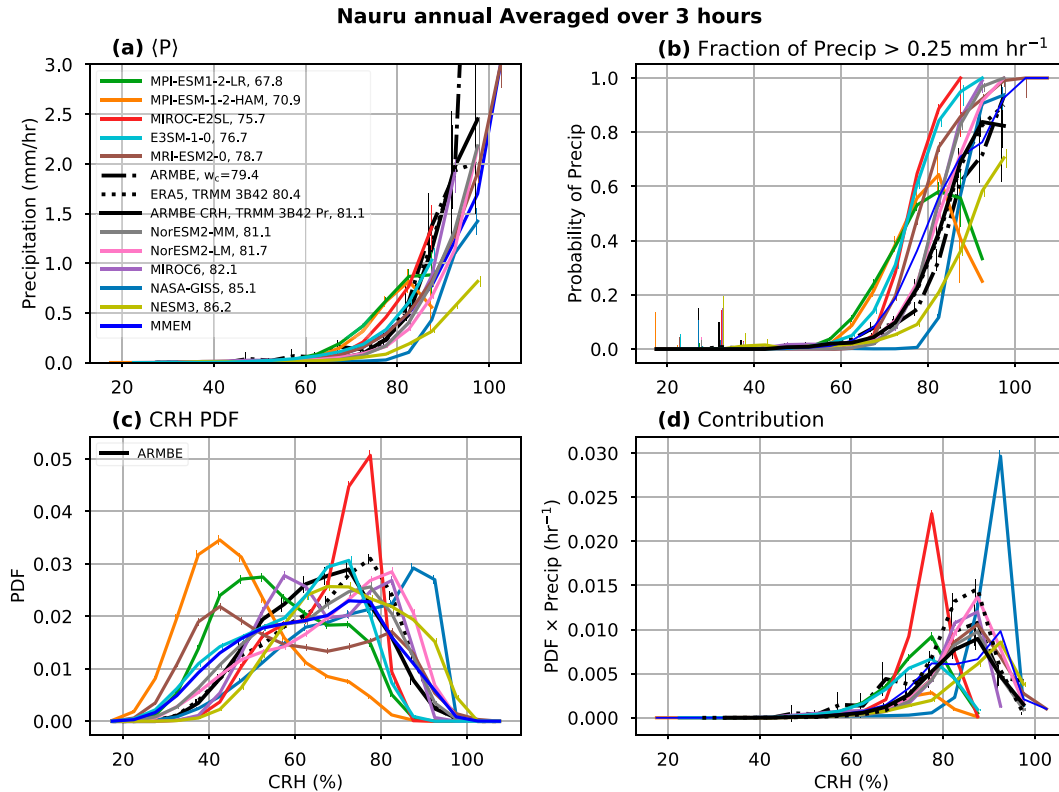


FIG. 7. As in Fig. 1, but the statistics are compiled in CRH for the Nauru site. The CRH bin width is 5%.

the characteristic shape seen in observations at Manus, and show similar discrepancies to their CWV in their CRH distribution at the Nauru site. The smaller peak in the observational CRH PDF at Nauru is an artifact of the large scale environment effect on the CWV PDF and determined by the dry regime dynamics.

Considering the temperature dependence of CWV and the $\widehat{q}_{\text{sat}}$ -binned CRH PDFs, the CRH PDFs collapse as (lower) higher CWV bins coincide with (lower) higher $\widehat{q}_{\text{sat}}$ and the drop-off of the $\widehat{q}_{\text{sat}}$ -binned CRH PDFs seen in Figs. 4b,e,h,k and 5b,e,h,k have a drop-off located around a similar CRH in Figs. 7c and 8c (e.g., just below 80% CRH in these cases). The collapsing effect of CRH is more notable at the Nauru site, the broad peak of the CWV PDF at Nauru in observations narrows into a more prominent CRH peak around its critical CRH. As discussed in section 2, the models behave appropriately in their CWV with respect to their pickup so models with temperature dependence should be able to replicate this behavior. In lower $\widehat{q}_{\text{sat}}$ bins, where overcompensation is not a problem (i.e., critical CWV increases relatively one-to-one with $\widehat{q}_{\text{sat}}$) the CRH PDF will tighten around its critical CRH. The effect to which the overcompensation has on the CRH PDFs is dependent on the value of the higher $\widehat{q}_{\text{sat}}$ bins and their frequency. With respect to both these factors, higher $\widehat{q}_{\text{sat}}$ bins only have a minor shift to the left (around a 1–2 CRH bins) and do not constitute a significant portion of the PDFs in both observations and models and may only shift some mass to the left, making the CRH PDF a little wider. A notable

feature of the three models with $\widehat{q}_{\text{sat}}$ dependence is that their CRH PDFs drop off faster than the observations, while they exhibited more appropriate drop-offs in the CWV PDF. An interesting feature of the CRH PDFs at Manus is that model PDFs which drop-off at a similar rate to observations in CWV maintain to do so in CRH, although the precipitation does not pick up in the models as fast as observations. A noticeable feature in the NASA-GISS model is the drastic drop-off close to saturation at both sites. The contribution peak of models follows the aforementioned discussion related to the CRH PDF and conditional precipitation, in that models that pickup early tend to peak their contributions too early at all sites.

c. Decomposition into probability of precipitation and conditional intensity, $\text{Pr}(P^+)$ and $\langle P^+ \rangle$

In our discussion of the slope of $\langle P \rangle$ in section 3, we draw comparisons between the drop-offs of the CWV PDFs in models and observations. The pickup depends strongly not only on the probability of raining times (e.g., Peters and Neelin 2006; Holloway and Neelin 2010; Brown et al. 2010) but also has a dependence on the sensitivity of precipitation intensity to environmental humidity during raining times (e.g., Kuo et al. 2018). A probability decomposition of the pickup such as the one performed in Igel and Biello (2019) for global precipitation, can help look at the relative contributions of these two factors. Consider a simple formulation of $\langle P \rangle$ based on the law of total expectation where the sample space is partitioned in two

Manus annual Averaged over 3 hours

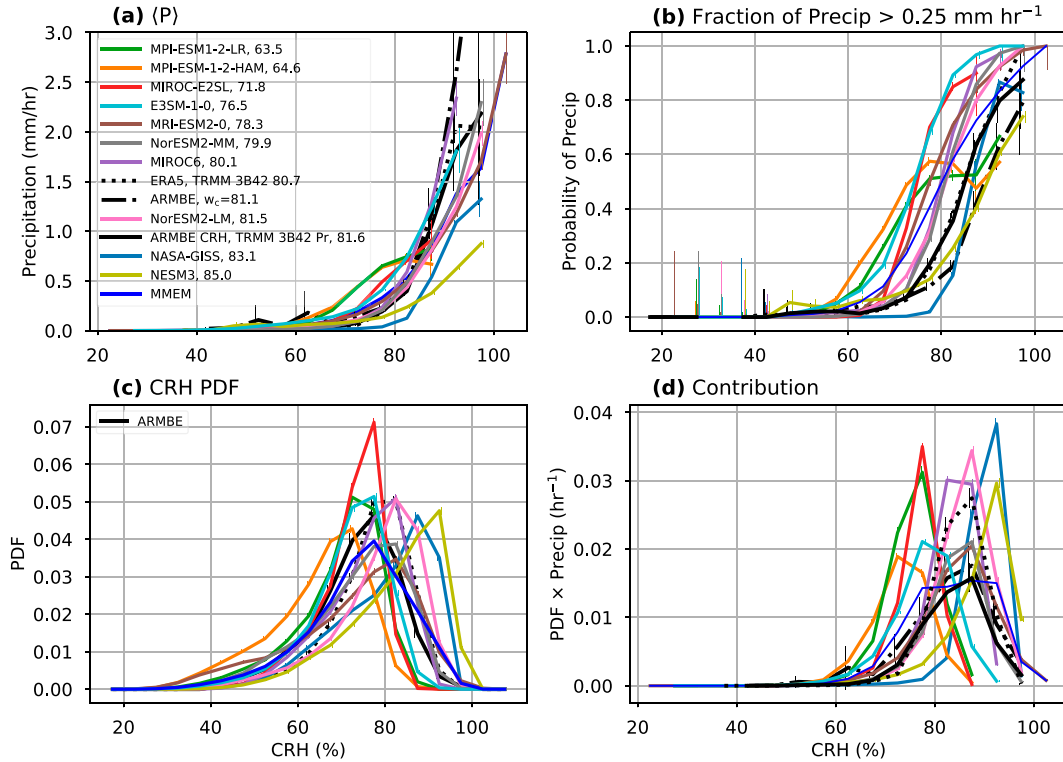


FIG. 8. As in Fig. 7, but for the Manus ARM site.

regimes, raining (precipitation above some threshold, P^+) and nonraining (precipitation measurement below some threshold, P^-):

$$\langle P \rangle_i = \mathbb{E}(P_i | P_i^+) \Pr(P_i^+) + \mathbb{E}(P_i | P_i^-) \Pr(P_i^-)$$

$$\langle P \rangle_i \approx \mathbb{E}(P_i | P_i^+) \Pr(P_i^+), \quad (1)$$

where the subscript i is used to denote the CRH bins, $\Pr$ is the probability operator, and $\mathbb{E}$ is the expectation operator. The threshold is set to 0.25 mm h⁻¹ here, consistent with the probability of precipitation metric (Figs. 1b, 2b, 7b, and 8b). Assuming that the average of precipitation values under the threshold is negligible compared to the average above, the conditional precipitation can be thought of as the product of probability of raining and the conditional mean precipitation during raining times, $\langle P^+ \rangle$, here termed conditional intensity for brevity. Figure 9 shows $\langle P^+ \rangle$ for the cohort.

Observations and model $\langle P^+ \rangle$, with the exception of NESM3, show sensitivity to CRH at both sites. The Probability of Precipitation (Figs. 7b and 8b) and $\langle P^+ \rangle$ exhibits a sharp increase at some point in the CRH domain, suggesting that $\langle P \rangle$ is the product of two pickups: $\langle P^+ \rangle$ and $\Pr(P^+)$. In ARMBE CRH and ERA5 CRH with TRMM-3b42 precipitation, $\langle P^+ \rangle$ increases sharply in the upper CRH range (black curves in Fig. 9). In models, with the exception of the MPI models, this occurs at even higher CRH.

This behavior may be summarized as: (i) the beginning of the $\langle P \rangle$ pickup is more dependent on increases in the Probability of Precipitation: (ii) where the probability caps out at high CRH, $\langle P^+ \rangle$ begins to govern the $\langle P \rangle$. This can be seen in Fig. 10, which shows the decomposition of the two factors as in Eq. (1). The logarithmic y axis of Fig. 10 allows the factors in Eq. (1) to be seen additively as $\ln \langle P^+ \rangle + \ln \Pr(P^+)$ (with the caveat that a constant could be added to one and subtracted from the other). Thus the change across the CRH range of the orange and green curves in Fig. 10 gives the relative contribution of $\langle P^+ \rangle$ and $\Pr(P^+)$, respectively. Note differences between $\ln \langle P \rangle$ and its estimation, $\ln \langle P^+ \rangle + \ln \Pr(P^+)$ are small except for at lower $\langle P \rangle$ values. In observations (Figs. 10a–c) the probability $\Pr(P^+)$ governs $\langle P \rangle$ through the early part of the increase, until $\Pr(P^+)$ flattens out, whereupon $\langle P^+ \rangle$ increases become important.

Most of the models capture this behavior at least qualitatively, with the following exceptions. The MPI models have a $\langle P^+ \rangle$ increase that is reasonable compared to observations (Figs. 9 and 10e,f) but the probability $\Pr(P^+)$ picks up substantially earlier (Figs. 7b and 8b) resulting in an early pickup in $\langle P \rangle$ (Figs. 7a and 8a). NESM3's $\langle P^+ \rangle$ lies relatively flat or shows very little sensitivity to CRH (Figs. 9 and 10l).

In neither the observational products nor the models is the $\langle P \rangle$ curve well described by an exponential, perhaps not surprisingly given the competing ingredients contributing to it. The observational products show a faster-than-exponential

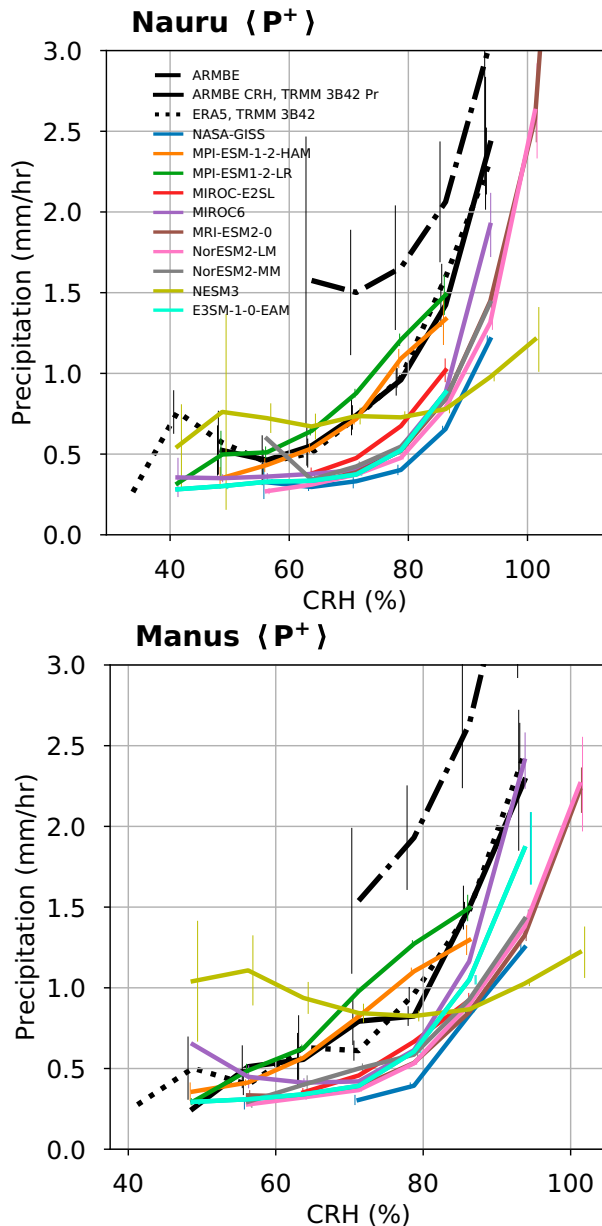


FIG. 9. As in Figs. 7a and 8a, but the conditional average, $\langle P^+ \rangle$ is calculated over raining times only (precipitation greater than or equal to 0.25 mm h^{-1}). Error bars represent the standard error of the precipitating points.

increase (upward curvature in Fig. 10) in the vicinity of w_c , and a slower-than-exponential increase at very high CRH. Most of the models share this feature, with the exception of the MPI models and MRI-ESM2-0, and with NASA-GSS exhibiting an exaggerated version.

The competing contributions of $\langle P^+ \rangle$ and $\text{Pr}(P^+)$ may be expected to depend on spatial resolution for very fine scales. However, Fig. S9 shows that for the range of spatial averaging relevant to model comparison, the observations are not sensitive to spatial scale.

With respect to averaging over raining times, the sensitivity of the precipitation threshold needs to be considered. The larger the threshold, the larger the non-raining times average precipitation is compared to raining average. The choice of 0.25 mm h^{-1} captures the sharp increase in probability of raining in Figs. 1b, 2b, 7b, and 8b, while low enough such that its increase in probability remains robust when averaging over spatial scales used in models (Fig. 3b). The average over non-raining times also remains negligible compared to the average over raining times as seen in Fig. 10.

An analysis of precipitation biases in CMIP5 models by Pathak et al. (2019) found that biases in models were more closely related to convective trigger and closure assumptions, as opposed to the cloud model employed in the GCM. A popular trigger based on convective potential energy (CAPE) where convection is activated once CAPE is above a certain threshold [often with modifications such as CAPE generation by large scale forcing (Xie and Zhang 2000), additional relative humidity thresholds (Xie et al. 2002), diluting CAPE with entrainment assumptions (Zhang 2009), or some combination (Suhas and Zhang 2014)] is used in the majority of the models examined here with the exception of the NASA-GISS, the MPI models, and NESM3. NASA-GISS uses a convective trigger based on buoyancy considerations (a virtual temperature test) where convection is initiated if a parcel, lifted adiabatically, has a higher virtual temperature than the level above (DelGenio and Yao 1993). Suhas and Zhang (2014) evaluated convective triggers and found CAPE-based triggers to be among the best performing in activating convection at appropriate times. Convective triggers for each model are listed in Table 1.

Contrasting to the above, MPI models and NESM3 use the Tiedtke (1989) scheme with modifications for deep convection by Nordeng (1994). The convective triggers for these models have buoyancy requirements and also require net positive moisture convergence (Möbis and Stevens 2012) since this is used in the convective closure. The buoyancy component is simply buoyant surface-level air at the lifting condensation level, which may be too easily met.

We underline that biases in w_c are not only due to biases in a model's convective trigger, but are also related to the life cycle of a convective cloud (Wolding et al. 2020). During a precipitating event, a convective column evolves through different convective-cloud/precipitation-type regimes reflected in changes to the column's moisture and/or temperature content. The pickup is therefore not so readily attributable to the onset of convection, but also may reflect a later stage in the convective life cycle from which moisture is consumed by precipitation, and the column shifts to a lower CRH or CWV content. The diagnostic separation into $\langle P^+ \rangle$ and $\text{Pr}(P^+)$ here will depend on this full life cycle and thus on both convective triggers and closures.

The models with moisture-convergent-based triggers/closures, MPI models and NESM3, show a shallower rate of $\langle P \rangle$ increase along with more nonzero values (lower probability of precipitation past pickup bin) compared to the rest of the cohort (Figs. 7a,b and 8a,b, respectively). While their behavior differs in terms of $\langle P^+ \rangle$ (MPI performing well, while NESM3 is essentially flat), in their overall simulation of $\langle P \rangle$, $\langle P^+ \rangle$, and $\text{Pr}(P^+)$ (Figs. 7–10), they tend to be outliers in terms of poor

Nauru

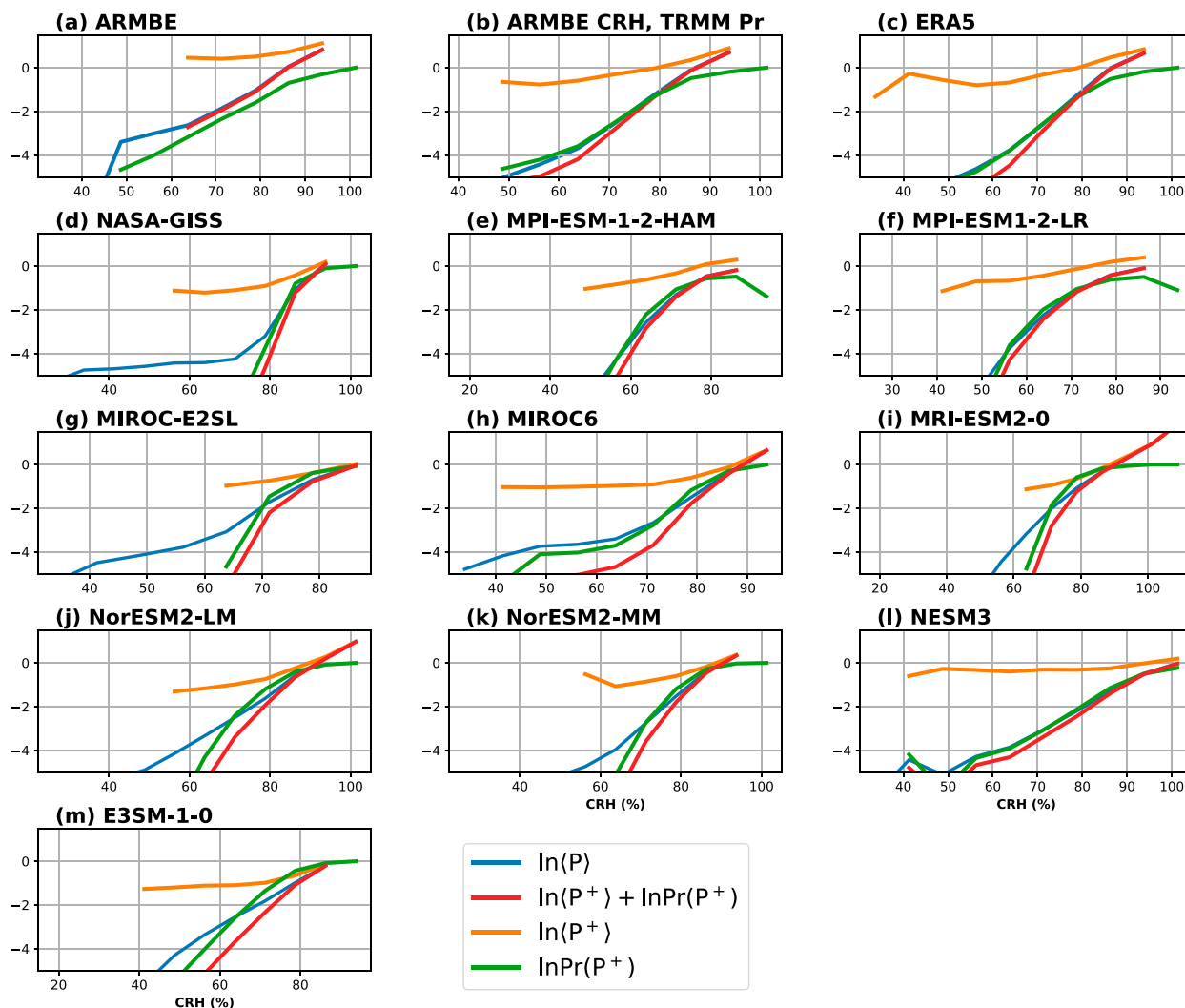


FIG. 10. The $\langle P \rangle$ (blue), $\ln\langle P^+ \rangle$ (orange), $\ln\Pr(P \geq 0.25)$ (green), and $\ln\langle P^+ \rangle + \ln\Pr(P \geq 0.25)$ (red) for the cohort examined at Nauru. Note that the y axis is in ln units. The orange curve here is the same as the $\langle P^+ \rangle$ curves shown in Fig. 9.

performance relative to models with buoyancy-based triggers/closures.

The w_c of buoyancy-based trigger (CAPE triggers and NASA-GISS) models lie in a close vicinity to that of observations and the probability of raining is too high at lower and midrange CRH values (Figs. 7a,b and 8a,b, respectively), implying that these models seem to compensate for the nearly constant values of precipitation in mid-values of CRH by raining more frequently. This is similar to the drizzle problem mentioned earlier—a consequence of models reacting too fast to surface heating and the diurnal cycle, as CAPE does not accumulate to allow for strong precipitation events (Xie et al. 2019).

Here we have focused on tying model biases in $\langle P \rangle$ to the frequency and magnitude of precipitation—for the purpose of attributing bias to behavior during the raining ($\langle P^+ \rangle$) or nonraining

(probability of precipitation) regime. Column integrated measures do not fully elucidate the biases in the generation of CAPE and buoyancy that contribute to these biases. Event-by-event statistics, their temporal evolution, and the associated vertical thermodynamic structures are explored in a companion paper, in which a similar set of statistics in the more physically insightful buoyancy domain is used, similar to the analysis done in Ahmed and Neelin (2021).

6. Conclusions and discussion

The relationship between CWV and precipitation is observed across ARM sites and in all CMIP6 models examined here. All models exhibit some form of the observed CWV-related pickup in precipitation in which the conditional-average precipitation rate increases rapidly past some critical value of

CWV, w_c . The physical relationship among statistics seen in observations is reflected in the behavior of the individual models in which w_c determines the CWV PDF drop-off. The majority of the models' convection onset statistics display some degree of temperature dependence in the CWV value of the pickup and collapse approximately to a common critical CRH value across $\widehat{q_{\text{sat}}}$ bins. However, prior results suggest that the onset of convective instability has a more complex dependence on temperature. The vertical structure of temperature and moisture, as well as the entrainment of free tropospheric air, affect the buoyancy of a rising convective plume, yielding an onset moisture-temperature dependence slightly different than that of bulk saturation (Holloway and Neelin 2009; Kuo et al. 2017; Schiro et al. 2018; Wolding and Maloney 2015; Wolding et al. 2020). Some aspect of these factors is apparently systematically different in the climate model representations than in the ARMBE data and ERA5 reanalysis. However, models are able to reproduce similar column saturation humidity PDFs to those of observations.

In section 4, we examine the temperature dependence of the convection transition statistics. We note the problems associated with compiling the statistics in CRH, despite the hope that this common practice might reduce the effect of the temperature biases of models. Compiling the statistics in CRH helps to account for model temperature biases in bringing the critical point of model pickup closer to that of observations, shifting the cohort of model pickups so that not all models pick up earlier than observations and reanalysis. This is especially important for the case of the ERA5 reanalysis product whose cold temperature bias at the ARM sites relative to ARMBE data (shown in Fig. 6) accounts for its earlier pickup in CWV (Figs. 1a and 2a). CRH PDF drop-offs remain consistent with the location of w_c . The rate of drop-off of model CRH PDFs remains similar to that of the observations for models with no temperature dependence, while models that display temperature dependence exhibited a faster rate in their CRH PDF drop-off. Further separating the conditional average precipitation into probability of precipitation, $\text{Pr}(P^+)$ and conditional intensity (conditional precipitation in the raining regime, $\langle P^+ \rangle$) reveals that models which produce $\langle P \rangle$ similar to observations do so through compensating biases. Through this framework, the $\langle P \rangle$ behavior is characterized as being primarily dependent on the probability of precipitation near the transition and the conditional intensity at higher CRH. Models which do perform well in reproducing $\langle P \rangle$ similar to that of observations do so by initiating convection too often at lower CRH, while their $\langle P^+ \rangle$ picks up at a value later than observations. Most models exhibit low conditional intensity $\langle P^+ \rangle$ relative to observations. A few models that pick up at too low CRH values, the MPI models, are able to replicate a similar $\langle P^+ \rangle$ to that of observations, but are limited in their CRH range, and rain too frequently at low CRH. Models using descendants of the Tiedtke scheme appear to fare worse overall by these measures than models with buoyancy-based schemes. Except for NASA-GISS, models rain too frequently at low CRH regardless of the convective trigger.

Models are able to capture the characteristic shape of the CWV and CRH PDFs fairly well at the Manus site, a site

dominated by the ascent regime. Models do have trouble with their PDFs at sites with seasonal variations in large-scale flow regime. At the Nauru site, the observational PDFs of CWV have a single broad peak whereas in some models, the CWV PDF is bimodal. In general, the CMIP6 model cohort examined here: 1) pickup too early in CWV and do not increase their conditionally averaged precipitation fast enough in both CWV and CRH; 2) associated with this, the precipitation contribution as a function of CWV or CRH peaks at too low a value; 3) the models capture the characteristic shape of CWV and CRH PDFs for sites that do not have strong variations in large-scale flow regime; and 4) exhibit behavior consistent in their CWV and CRH PDFs and peaks of contributions in relation to their critical CWV and CRH.

The ARM datasets allow for higher temporal resolution analysis. As discussed in section 3a, time-averaging up to 3 h has little effect on the important characteristics of the statistics; we expect the higher resolution Cloud Feedback Model Intercomparison Project data for the models discussed here to remain consistent with our results once released. Higher resolution data would benefit diagnosing the large intermodel spread in regard to the various convective parameterizations on faster time scales. In a companion paper, we explore event-by-event behavior of models in the context of their associated vertical thermodynamic structures. The analysis tools and datasets presented here will be continued to be added in the ARM-DIAGS package to facilitate use of these analyses by the climate community (Zhang et al. 2020).

Acknowledgments. This work was supported under U.S. Department of Energy Grant DE-SC0011074 and Subcontract B634021 and National Science Foundation AGS-1936810 (TE and JDN). We acknowledge the U.S. Department of Energy's Atmospheric Radiation Measurement (ARM) program and the Program for Climate Model Diagnosis and Intercomparison to Earth System Science Portals for making CMIP data available. E3SM data were obtained from the Energy Exascale Earth System Model project, sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. Work at Lawrence Livermore National Laboratory was supported by the DOE ARM program and performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The authors thank Dr. Kathleen Schiro for discussion and for providing an initial set of scripts that the current diagnostics built upon, as well as Dr. Wuyin Lin for providing the high-frequency E3SM output, and Drs. Brandon Wolding and M. Igel for helpful suggestions during the review process.

Data availability statement. ARM data are openly available through the U.S. Department of Energy's Atmospheric Radiation Measurement Portal for ARM, DOI:10.5439/1027369, cited in Gaustad and Riihimäki (1998) and ARMBE, DOI:10.5439/1333228, cite in Xiao and Shaocheng (1996). ERA5 data are available through the Climate Data Store portal, DOI:10.24381/cds.bd0915c6, as cited in (Hersbach et al. 2020).

TRMM-3b42 data are available through NASA's Earth Data portal, DOI:10.5067/TRMM/TMPA/3H/7, as cited in [TRMM \(2011\)](#). All model data are available through the Climate Model Diagnosis and Intercomparison to Earth Systems Science Portal at esgf-node.llnl.gov.

REFERENCES

- Abbott, T. H., S. N. Stechmann, and J. D. Neelin, 2016: Long temporal autocorrelations in tropical precipitation data and spike train prototypes. *Geophys. Res. Lett.*, **43**, 11 472–11 480, <https://doi.org/10.1002/2016GL071282>.
- Adames, Á. F., S. W. Powell, F. Ahmed, V. C. Mayta, and J. D. Neelin, 2021: Tropical precipitation evolution in a buoyancy-budget framework. *J. Atmos. Sci.*, **78**, 509–528, <https://doi.org/10.1175/JAS-D-20-0074.1>.
- Ahmed, F., and C. Schumacher, 2015: Convective and stratiform components of the precipitation–moisture relationship. *Geophys. Res. Lett.*, **42**, 10–453, <https://doi.org/10.1002/2015GL066957>.
- , and J. D. Neelin, 2018: Reverse engineering the tropical precipitation–buoyancy relationship. *J. Atmos. Sci.*, **75**, 1587–1608, <https://doi.org/10.1175/JAS-D-17-0333.1>.
- , and —, 2021: A process-oriented diagnostic to assess precipitation–thermodynamic relations and application to CMIP6 models. *Geophys. Res. Lett.*, **48**, e2021GL094108, <https://doi.org/10.1002/2016GL071282>.
- Ando, T., and Coauthors, 2021: Description of MIROC6 AGCM. CCSR Rep. 65, 48 pp., <https://doi.org/10.15083/0002000180>.
- Bechtold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo, 2008: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quart. J. Roy. Meteor. Soc.*, **134**, 1337–1351, <https://doi.org/10.1002/qj.289>.
- Bentsen, M., and Coauthors, 2019: NCC NorESM2-MM model output prepared for CMIP6 CMIP historical. Earth System Grid Federation, accessed 24 April 2020, <https://doi.org/10.22033/ESGF/CMIP6.8040>.
- Betts, A. K., and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.*, **112**, 693–709, <https://doi.org/10.1002/qj.49711247308>.
- Biasutti, M., 2013: Forced Sahel rainfall trends in the CMIP5 archive. *J. Geophys. Res. Atmos.*, **118**, 1613–1623, <https://doi.org/10.1002/jgrd.50206>.
- , and A. H. Sobel, 2009: Delayed Sahel rainfall and global seasonal cycle in a warmer climate. *Geophys. Res. Lett.*, **36**, L23707, <https://doi.org/10.1029/2009GL041303>.
- Bretherton, C. S., M. E. Peters, and L. Back, 2004: Relationships between water vapor path and precipitation over the tropical oceans. *J. Climate*, **17**, 1517–1528, [https://doi.org/10.1175/1520-0442\(2004\)017<1517:RBWVPA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<1517:RBWVPA>2.0.CO;2).
- Brown, J. R., C. Jakob, and J. M. Haynes, 2010: An evaluation of rainfall frequency and intensity over the Australian region in a global climate model. *J. Climate*, **23**, 6504–6525, <https://doi.org/10.1175/2010JCLI3571.1>.
- Caldwell, P. M., and Coauthors, 2019: The DOE E3SM coupled model version 1: Description and results at high resolution. *J. Adv. Model. Earth Syst.*, **11**, 4095–4146, <https://doi.org/10.1029/2019MS001870>.
- Cao, J., and B. Wang, 2019: NUIST NESMv3 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation, accessed 24 April 2020, <https://doi.org/10.22033/ESGF/CMIP6.8769>.
- , and Coauthors, 2018: The NUIST Earth System Model (NESM) version 3: Description and preliminary evaluation. *Geosci. Model Dev.*, **11**, 2975–2993, <https://doi.org/10.5194/gmd-11-2975-2018>.
- Chen, D., and A. Dai, 2019: Precipitation characteristics in the community atmosphere model and their dependence on model physics and resolution. *J. Adv. Model. Earth Syst.*, **11**, 2352–2374, <https://doi.org/10.1029/2018MS001536>.
- Covey, C., P. J. Gleckler, C. Doutriaux, D. N. Williams, A. Dai, J. Fasullo, K. Trenberth, and A. Berg, 2016: Metrics for the diurnal cycle of precipitation: Toward routine benchmarks for climate models. *J. Climate*, **29**, 4461–4471, <https://doi.org/10.1175/JCLI-D-15-0664.1>.
- Del Genio, A. D., 2012: Representing the sensitivity of convective cloud systems to tropospheric humidity in general circulation models. *Surv. Geophys.*, **33**, 637–656, <https://doi.org/10.1007/s10712-011-9148-9>.
- , and M.-S. Yao, 1993: Efficient cumulus parameterization for long-term climate studies: The GISS scheme. *The Representation of Cumulus Convection in Numerical Models*, K. A. Emanuel and D. J. Raymond, Eds., Springer, 181–184.
- , and J. Wu, 2010: The role of entrainment in the diurnal cycle of continental convection. *J. Climate*, **23**, 2722–2738, <https://doi.org/10.1175/2009JCLI3340.1>.
- Gaustad, K., and L. Riihimäki, 1998: MWR Retrievals (MWRRET1LILJCLOU). Atmospheric Radiation Measurement (ARM) user facility, accessed 23 March 2019, <https://doi.org/10.5439/1027369>.
- Hajima, T., and Coauthors, 2019: MIROC MIROC-ES2L model output prepared for CMIP6 CMIP historical. Earth System Grid Federation, accessed 25 April 2020, <https://doi.org/10.22033/ESGF/CMIP6.5602>.
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Hirota, N., Y. N. Takayabu, M. Watanabe, M. Kimoto, and M. Chikira, 2014: Role of convective entrainment in spatial distributions of and temporal variations in precipitation over tropical oceans. *J. Climate*, **27**, 8707–8723, <https://doi.org/10.1175/JCLI-D-13-00701.1>.
- Holdridge, D., and J. Kyrkouac, 1998: Surface meteorological instrumentation (met). ARM, accessed 23 March 2019, <https://doi.org/10.5439/1786358>.
- Holloway, C. E., and J. D. Neelin, 2009: Moisture vertical structure, column water vapor, and tropical deep convection. *J. Atmos. Sci.*, **66**, 1665–1683, <https://doi.org/10.1175/2008JAS2806.1>.
- , and —, 2010: Temporal relations of column water vapor and tropical precipitation. *J. Atmos. Sci.*, **67**, 1091–1105, <https://doi.org/10.1175/2009JAS3284.1>.
- Hourdin, F., and Coauthors, 2013: LMDZ5B: The atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dyn.*, **40**, 2193–2222, <https://doi.org/10.1007/s00382-012-1343-y>.
- Hwang, Y.-T., and D. M. W. Frierson, 2013: Link between the double-interropical convergence zone problem and cloud biases over the Southern Ocean. *Proc. Natl. Acad. Sci. USA*, **110**, 4935–4940, <https://doi.org/10.1073/pnas.1213302110>.

- Igel, M. R., and J. A. Biello, 2019: A reconstructed total precipitation framework. *npj Climate Atmos. Sci.*, **2**, 32, <https://doi.org/10.1038/s41612-019-0090-8>.
- Jiang, X., D. Kim, and E. D. Maloney, 2021: Progress and status of MJO simulation in climate models and process-oriented diagnostics. *The Multiscale Global Monsoon System*, C.-P. Chang et al., Eds., World Scientific, 315–326, <https://doi.org/10.1142/11723>.
- Jing, X., K. Suzuki, H. Guo, D. Goto, T. Ogura, T. Koshiro, and J. Mülmenstädt, 2017: A multimodel study on warm precipitation biases in global models compared to satellite observations. *J. Geophys. Res. Atmos.*, **122**, 11 806–11 824, <https://doi.org/10.1002/2017JD027310>.
- Kuo, Y.-H., C. R. Mechoso, and J. D. Neelin, 2017: Tropical convective onset statistics and causality in the water vapor–precipitation relation. *J. Atmos. Sci.*, **74**, 915–931, <https://doi.org/10.1175/JAS-D-16-0182.1>.
- , J. D. Neelin, and K. A. Schiro, 2018: Convective transition statistics over tropical oceans for climate model diagnostics: Observational baseline. *J. Atmos. Sci.*, **75**, 1553–1570, <https://doi.org/10.1175/JAS-D-17-0287.1>.
- , and Coauthors, 2020: Convective transition statistics over tropical oceans for climate model diagnostics: GCM evaluation. *J. Atmos. Sci.*, **77**, 379–403, <https://doi.org/10.1175/JAS-D-19-0132.1>.
- Long, C. N., and S. A. McFarlane, 2012: Quantification of the impact of Nauru Island on ARM measurements. *J. Appl. Meteor. Climatol.*, **51**, 628–636, <https://doi.org/10.1175/JAMC-D-11-0174.1>.
- Maloney, E. D., Á. F. Adames, and H. X. Bui, 2019: Madden–Julian Oscillation changes under anthropogenic warming. *Nat. Climate Change*, **9**, 26–33, <https://doi.org/10.1038/s41558-018-0331-6>.
- Mapes, B., and R. Neale, 2011: Parameterizing convective organization to escape the entrainment dilemma. *J. Adv. Model. Earth Syst.*, **3**, M06004, <https://doi.org/10.1029/2011MS000042>.
- Mather, J. H., and J. W. Voyles, 2013: The ARM Climate Research Facility: A review of structure and capabilities. *Bull. Amer. Meteor. Soc.*, **94**, 377–392, <https://doi.org/10.1175/BAMS-D-11-00218.1>.
- Matthews, S., J. M. Hacker, J. Cole, J. Hare, C. N. Long, and R. M. Reynolds, 2007: Modification of the atmospheric boundary layer by a small island: Observations from Nauru. *Mon. Wea. Rev.*, **135**, 891–905, <https://doi.org/10.1175/MWR3319.1>.
- Möbis, B., and B. Stevens, 2012: Factors controlling the position of the Intertropical Convergence Zone on an aquaplanet. *J. Adv. Model. Earth Syst.*, **4**, M00A04, <https://doi.org/10.1029/2012MS000199>.
- NASA/GISS, 2018: NASA-GISS GISS-E2.1G model output prepared for CMIP6 CMIP historical. Earth System Grid Federation, accessed 12 April 2020, <https://doi.org/10.22033/ESGF/CMIP6.7127>.
- Neale, R. B., J. H. Richter, and M. Jochum, 2008: The impact of convection on ENSO: From a delayed oscillator to a series of events. *J. Climate*, **21**, 5904–5924, <https://doi.org/10.1175/2008JCLI2244.1>.
- Neelin, J. D., O. Peters, and K. Hales, 2009: The transition to strong convection. *J. Atmos. Sci.*, **66**, 2367–2384, <https://doi.org/10.1175/2009JAS2962.1>.
- , B. R. Lintner, B. Tian, Q. Li, L. Zhang, P. K. Patra, M. T. Chahine, and S. N. Stechmann, 2010: Long tails in deep columns of natural and anthropogenic tropospheric tracers. *Geophys. Res. Lett.*, **37**, L05804, <https://doi.org/10.1029/2009GL041726>.
- Nordeng, T. E., 1994: Extended versions of the convective parametrization scheme at ECMWF and their impact on the mean and transient activity of the model in the tropics. ECMWF Tech. Memo. 206, 41 pp., <https://www.ecmwf.int/node/11393>.
- Oueslati, B., and G. Bellon, 2013: Convective entrainment and large-scale organization of tropical precipitation: Sensitivity of the CNRM-CM5 hierarchy of models. *J. Climate*, **26**, 2931–2946, <https://doi.org/10.1175/JCLI-D-12-00314.1>.
- Pathak, R., S. Sahany, S. K. Mishra, and S. Dash, 2019: Precipitation biases in CMIP5 models over the South Asian region. *Sci. Rep.*, **9**, 9589, <https://doi.org/10.1038/s41598-019-45907-4>.
- Peters, O., and J. D. Neelin, 2006: Critical phenomena in atmospheric precipitation. *Nat. Phys.*, **2**, 393–396, <https://doi.org/10.1038/nphys314>.
- Rasch, P., and Coauthors, 2019: An overview of the atmospheric component of the energy exascale Earth system model. *J. Adv. Model. Earth Syst.*, **11**, 2377–2411, <https://doi.org/10.1029/2019MS001629>.
- Rio, C., F. Hourdin, J.-Y. Grandpeix, and J.-P. Lafore, 2009: Shifting the diurnal cycle of parameterized deep convection over land. *Geophys. Res. Lett.*, **36**, L07809, <https://doi.org/10.1029/2008GL036779>.
- Rushley, S., D. Kim, C. Bretherton, and M.-S. Ahn, 2018: Reexamining the nonlinear moisture–precipitation relationship over the tropical oceans. *Geophys. Res. Lett.*, **45**, 1133–1140, <https://doi.org/10.1002/2017GL076296>.
- Sahany, S., J. D. Neelin, K. Hales, and R. B. Neale, 2012: Temperature-moisture dependence of the deep convective transition as a constraint on entrainment in climate models. *J. Atmos. Sci.*, **69**, 1340–1358, <https://doi.org/10.1175/JAS-D-11-0164.1>.
- Schiro, K. A., 2017: Thermodynamic controls on deep convection in the tropics: Observations and applications to modeling. Ph.D. thesis, University of California, 132 pp.
- , and J. D. Neelin, 2019: Deep convective organization, moisture vertical structure, and convective transition using deep-inflow mixing. *J. Atmos. Sci.*, **76**, 965–987, <https://doi.org/10.1175/JAS-D-18-0122.1>.
- , —, D. K. Adams, and B. R. Lintner, 2016: Deep convection and column water vapor over tropical land versus tropical ocean: A comparison between the Amazon and the tropical western Pacific. *J. Atmos. Sci.*, **73**, 4043–4063, <https://doi.org/10.1175/JAS-D-16-0119.1>.
- , F. Ahmed, S. E. Giangrande, and J. D. Neelin, 2018: GoAmazon2014/5 campaign points to deep-inflow approach to deep convection across scales. *Proc. Natl. Acad. Sci. USA*, **115**, 4577–4582, <https://doi.org/10.1073/pnas.1719842115>.
- Schmidt, G. A., and Coauthors, 2014: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *J. Adv. Model. Earth Syst.*, **6**, 141–184, <https://doi.org/10.1002/2013MS000265>.
- Seland, Ø., and Coauthors, 2020: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geosci. Model Dev.*, **13**, 6165–6200, <https://doi.org/10.5194/gmd-13-6165-2020>.
- Seland, Y., and Coauthors, 2019: NCC NorESM2-LM model output prepared for CMIP6 CMIP historical. Earth System Grid Federation, accessed 24 April 2020, <https://doi.org/10.22033/ESGF/CMIP6.8036>.

- Stephens, G. L., and Coauthors, 2010: Dreary state of precipitation in global models. *J. Geophys. Res.*, **115**, D24211, <https://doi.org/10.1029/2010JD014532>.
- Stokes, G. M., and S. E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) program: Programmatic background and design of the cloud and radiation testbed. *Bull. Amer. Meteor. Soc.*, **75**, 1201–1221, [https://doi.org/10.1175/1520-0477\(1994\)075<1201:TARMPP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1994)075<1201:TARMPP>2.0.CO;2).
- Su, H., and Coauthors, 2017: Tightening of tropical ascent and high clouds key to precipitation change in a warmer climate. *Nat. Commun.*, **8**, 15771, <https://doi.org/10.1038/ncomms15771>.
- Suhas, E., and G. J. Zhang, 2014: Evaluation of trigger functions for convective parameterization schemes using observations. *J. Climate*, **27**, 7647–7666, <https://doi.org/10.1175/JCLI-D-13-00718.1>.
- Tatebe, H., and M. Watanabe, 2018: MIROC MIROC6 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation, accessed 24 April 2020, <https://doi.org/10.22033/ESGF/CMIP6.5603>.
- Tian, B., and X. Dong, 2020: The double-ITCZ bias in CMIP3, CMIP5, and CMIP6 models based on annual mean precipitation. *Geophys. Res. Lett.*, **47**, e2020GL087232, <https://doi.org/10.1029/2020GL087232>.
- Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.*, **117**, 1779–1800, [https://doi.org/10.1175/1520-0493\(1989\)117<1779:ACMFSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2).
- TRMM, 2011: TRMM (TMPA) rainfall estimate L3 3 hour 0.25 degree $\times$ 0.25 degree V7 (TRMM_3B42). Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed 25 June 2020, <https://doi.org/10.5067/TRMM/TMPA/3H/7>.
- Turner, D. D., S. A. Clough, J. C. Liljegren, E. E. Clothiaux, K. E. Cady-Pereira, and K. L. Gaustad, 2007: Retrieving liquid water path and precipitable water vapor from the Atmospheric Radiation Measurement (ARM) microwave radiometers. *IEEE Trans. Geosci. Remote Sens.*, **45**, 3680–3690, <https://doi.org/10.1109/TGRS.2007.903703>.
- Voigt, A., and Coauthors, 2016: The tropical rain belts with an annual cycle and a continent model intercomparison project: TRACMIP. *J. Adv. Model. Earth Syst.*, **8**, 1868–1891, <https://doi.org/10.1002/2016MS000748>.
- Wentz, F., C. Gentemann, and K. Hilburn, 2015: Remote Sensing Systems TRMM TMI daily environmental suite on 0.25 deg grid, version 7.1. Remote Sensing Systems, accessed 25 June 2020, <http://www.remss.com/missions/tmi>.
- Wieners, K.-H., and Coauthors, 2019: MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP historical. Earth System Grid Federation, accessed 24 April 2020, <https://doi.org/10.22033/ESGF/CMIP6.6595>.
- Wolding, B. O., and E. D. Maloney, 2015: Objective diagnostics and the Madden-Julian Oscillation. Part II: Application to moist static energy and moisture budgets. *J. Climate*, **28**, 7786–7808, <https://doi.org/10.1175/JCLI-D-14-00689.1>.
- , J. Dias, G. Kiladis, F. Ahmed, S. W. Powell, E. Maloney, and M. Branson, 2020: Interactions between moisture and tropical convection. Part I: The coevolution of moisture and convection. *J. Atmos. Sci.*, **77**, 1783–1799, <https://doi.org/10.1175/JAS-D-19-0225.1>.
- Xiao, C., and X. Shaocheng, 1996: ARMBECLDRAD—ARMBE: Cloud Radiation measurements. ARM, accessed 23 March 2019, <https://doi.org/10.5439/1333228>.
- Xie, S., and M. Zhang, 2000: Impact of the convection triggering function on single-column model simulations. *J. Geophys. Res.*, **105**, 14 983–14 996, <https://doi.org/10.1029/2000JD900170>.
- , and Coauthors, 2002: Intercomparison and evaluation of cumulus parametrizations under summertime midlatitude continental conditions. *Quart. J. Roy. Meteor. Soc.*, **128**, 1095–1135, <https://doi.org/10.1256/003590002320373229>.
- , and Coauthors, 2010: Clouds and more: ARM climate modeling best estimate data. *Bull. Amer. Meteor. Soc.*, **91**, 13–20, <https://doi.org/10.1175/2009BAMS2891.1>.
- , and Coauthors, 2018: Understanding cloud and convective characteristics in version 1 of the E3SM atmosphere model. *J. Adv. Model. Earth Syst.*, **10**, 2618–2644, <https://doi.org/10.1029/2018MS001350>.
- , and Coauthors, 2019: Improved diurnal cycle of precipitation in E3SM with a revised convective triggering function. *J. Adv. Model. Earth Syst.*, **11**, 2290–2310, <https://doi.org/10.1029/2019MS001702>.
- Yukimoto, S., and Coauthors, 2011: Meteorological Research Institute-Earth System Model version 1 (MRI-ESM1)-model description. Meteorological Research Institute Tech. Rep. 64, 92 pp., http://www.mri-jma.go.jp/Publish/Technical/DATA/VOL_64/index_en.html.
- , and Coauthors, 2019: MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP historical. Earth System Grid Federation, accessed 24 April 2020, <https://doi.org/10.22033/ESGF/CMIP6.6842>.
- Zhang, C., and Coauthors, 2020: The ARM data-oriented metrics and diagnostics package for climate models: A new tool for evaluating climate models with field data. *Bull. Amer. Meteor. Soc.*, **101**, E1619–E1627, <https://doi.org/10.1175/BAMS-D-19-0282.1>.
- Zhang, G. J., 2009: Effects of entrainment on convective available potential energy and closure assumptions in convection parameterization. *J. Geophys. Res.*, **114**, D07109, <https://doi.org/10.1029/2008JD010976>.
- Zhang, Y., and Coauthors, 2019: Evaluation of clouds in version 1 of the E3SM atmosphere model with satellite simulators. *J. Adv. Model. Earth Syst.*, **11**, 1253–1268, <https://doi.org/10.1029/2018MS001562>.
- Zhao, M., I. M. Held, S. J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50-km resolution GCM. *J. Climate*, **22**, 6653–6678, <https://doi.org/10.1175/2009JCLI3049.1>.
- Zhu, H., and H. H. Hendon, 2015: Role of large-scale moisture advection for simulation of the MJO with increased entrainment. *Quart. J. Roy. Meteor. Soc.*, **141**, 2127–2136, <https://doi.org/10.1002/qj.2510>.