Deep Convection and Column Water Vapor over Tropical Land versus Tropical Ocean: A Comparison between the Amazon and the Tropical Western Pacific

KATHLEEN A. SCHIRO AND J. DAVID NEELIN
Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, California

DAVID K. ADAMS
Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Mexico City, Mexico

BENJAMIN R. LINTNER
Department of Environmental Sciences, Rutgers, The State University of New Jersey, New Brunswick, New Jersey

(Manuscript received 18 April 2016, in final form 15 June 2016)

ABSTRACT
The relationships between the onset of tropical deep convection, column water vapor (CWV), and other measures of conditional instability are analyzed with 2 yr of data from the DOE Atmospheric Radiation Measurement (ARM) Mobile Facility in Manacapuru, Brazil, as part of the Green Ocean Amazon (GOAmazon) campaign, and with 3.5 yr of CWV derived from global positioning system meteorology at a nearby site in Manaus, Brazil. Important features seen previously in observations over tropical oceans—precipitation conditionally averaged by CWV exhibiting a sharp pickup at high CWV, and the overall shape of the CWV distribution for both precipitating and nonprecipitating points—are also found for this tropical continental region. The relationship between rainfall and CWV reflects the impact of lower-free-tropospheric moisture variability on convection. Specifically, CWV over land, as over ocean, is a proxy for the effect of free-tropospheric moisture on conditional instability as indicated by entraining plume calculations from GOAmazon data. Given sufficient mixing in the lower troposphere, higher CWV generally results in greater plume buoyancies through a deep convective layer. Although sensitivity of buoyancy to other controls in the Amazon is suggested, such as boundary layer and microphysical processes, the CWV dependence is consistent with the observed precipitation onset. Overall, leading aspects of the relationship between CWV and the transition to deep convection in the Amazon have close parallels over tropical oceans. The relationship is robust to averaging on time and space scales appropriate for convective physics but is strongly smoothed for averages greater than 3 h or 2.5°.

1. Introduction
Despite the complex relationships, interactions, and feedbacks that exist among the atmosphere, land, and ocean, a robust relationship exists between precipitation and column water vapor (CWV). Bretherton et al. (2004) identified a smooth relationship of CWV and precipitation in daily mean satellite observations. On shorter time scales, conditionally averaged precipitation rate increases sharply with increasing CWV (Peters and Neelin 2006; Holloway and Neelin 2009; Neelin et al. 2009). This sharp pickup is associated with the onset of conditional instability leading to deep convection. Furthermore, statistics of the transition to deep convection are analogous to properties of a continuous phase transition at a critical value of CWV (Peters and Neelin 2006; Neelin et al. 2009) and can be understood in terms of stochastic variations across the deep convective onset threshold (Stechmann and Neelin 2011). Evaluating this deep convective transition using radiosondes from the DOE Atmospheric Radiation Measurement (ARM) site at Nauru in the tropical western Pacific, Holloway and Neelin (2009) demonstrated that CWV represents a...
proxy for the impact of free-tropospheric humidity on the conditional instability of entraining plumes affecting the transition from shallow to deep convection and, thus, that the statistics quantifying this transition provide a substantial constraint on subgrid-scale processes that must be represented in climate models. It was previously unclear, however, the extent to which this simplifying CWV–precipitation relationship applies for convective transition statistics over tropical land, as fundamental differences exist in the convective environment over land compared with ocean—including a stronger diurnal cycle and greater variations in the boundary layer (Nesbit and Zipser 2003).

There is substantial evidence suggesting the importance of free-tropospheric humidity to the onset of deep convection (Austin 1948; Malkus 1954; Yoneyama and Fujitani 1995; Brown and Zhang 1997; Wei et al. 1998; Raymond and Torres 1998; Sherwood and Wahrlich 1999; Parsons et al. 2000; Raymond 2000; Raymond and Zeng 2000; Tompkins 2001a; Redelsperger et al. 2002; Ridout 2002; Grabowski 2003; Bretherton et al. 2004; Chaboureau et al. 2004; Derbysheir et al. 2004; Guichard et al. 2004; Sobel et al. 2004; Sherwood et al. 2004; Kuang and Bretherton 2006; Tian et al. 2006; Wu et al. 2009; Waite and Khoulider 2010; Zhang and Klein 2010; Kumar et al. 2013), yet many models are currently too insensitive to free-tropospheric humidity (Biasutti et al. 2006; Dai 2006; Oueslati and Bellon 2013). This insensitivity contributes to systematic errors and biases in simulated precipitation on a number of space and time scales: the erroneous appearance of a double intertropical convergence zone (Hirota and Takayabu 2013; Hirota et al. 2014), deficiencies in the simulation of the Madden–Julian oscillation (Grabowski and Moncrieff 2004; Hannah and Maloney 2011; Jiang et al. 2011; Del Genio et al. 2012; Kim et al. 2012; Holloway et al. 2013; Kim et al. 2014; Rowe and Houze 2015), and failure to represent the shallow-to-deep convective transition and diurnal cycle of deep convection (Randall et al. 1991; Yang and Slingo 2001; Betts and Jakob 2002; Dai and Trenberth 2004; Bechtold et al. 2004; Chaboureau et al. 2004; Guichard et al. 2004; Dai 2006; Del Genio and Wu 2010; Waite and Khoulider 2010). The effect of free-tropospheric humidity on the onset of deep convection can be explained through mixing between a convective plume and its surrounding environment, which greatly affects the plume’s buoyancy. Mixing assumptions must, therefore, be appropriately constrained in convective parameterizations. This has been a long-standing challenge, yet several studies have demonstrated significant model improvement with realistic representations of entrainment processes (Neale et al. 2008; Bechtold et al. 2008; Zhao et al. 2009; Neelin et al. 2010; Sahany et al. 2012). In this regard, the convective transition statistics developed over tropical oceans have proven useful as model diagnostics (Sahany et al. 2012, 2014) that help to constrain entrainment representations, along with other observational and modeling studies (Raymond and Blyth 1986; Brown and Zhang 1997; Jensen and Del Genio 2006; Kuang and Bretherton 2006; Li et al. 2008; Bacmeister and Stephens 2011; Luo et al. 2010; Romps and Kuang 2010). The transition to deep convection can also be examined in the temporal domain (Holloway and Neelin 2010; Adams et al. 2013) in which time scales, lead–lag relations, and the distinction between temporal onset and termination (Stechmann and Neelin 2014) can be important.

There are several additional variables and processes controlling the transition to deep convection that must also be understood and accurately represented in models: free-tropospheric moistening processes (Johnson et al. 1999; Benedict and Randell 2007; Kembal-Cook and Weare 2001; Mapes et al. 2006; Hohenegger and Stevens 2013; Kumar et al. 2013; Masunaga 2013; Hagos et al. 2014), the influence of the diurnal cycle (Betts and Jakob 2002; Bechtold et al. 2004; Chaboureau et al. 2004; Del Genio and Wu 2010; Zhang and Klein 2010; Bechtold et al. 2014), the larger-scale dynamics forcing vertical ascent (Kumar et al. 2013; Hohenegger and Stevens 2013), convective downdrafts and cold pool formation (Tompkins 2001b; Khairoutdinov and Randall 2006; Schlemmer and Hohenegger 2014), cloud size (Boing et al. 2012), moist static energy gradients (Neelin and Held 1987; Lintner and Neelin 2007, 2008; Raymond et al. 2009; Lintner and Neelin 2010; Ma et al. 2011), vertical wind shear (Rotunno et al. 1988; LeMone et al. 1998), and microphysical processes, including cloud–aerosol interactions (Andreae et al. 2004; Khain et al. 2005). Important differences likely exist in the way these processes and variables contribute to the conditional instability of the environment over tropical land versus tropical oceans.

Thus far, an insufficient observational record in the continental tropics has limited development of convective transition statistics, yet the Green Ocean Amazon (GOAmazon) campaign in Manacapuru, Brazil (2014/15; Martin et al. 2016) has provided a unique opportunity to evaluate the transition to deep convection over tropical land, to elucidate potential complexities compared to the ocean, and to develop simple, useful statistics as model diagnostics. Here, we derive the CWV–precipitation relationship and associated statistics with these data and with a complementary 3.5-yr dataset from the central Amazon using global positioning system (GPS) meteorology that provides continuous, all-weather observations of CWV at high temporal resolution over tropical land (Adams et al. 2013, 2015). Parallels are drawn between the
land and the ocean to assess whether free-tropospheric humidity is also of leading-order importance to the conditional instability of an entraining plume over land as it is over ocean. The robustness of the convective transition statistics is tested as a function of spatial and temporal scales to establish a benchmark for comparison between models and observations at various scales. Last, the CWV–precipitation relationship is examined physically by linking vertical profiles of key thermodynamic quantities and plume buoyancies computed using turbulent mixing to the observed onset of deep convection.

2. Data

A suite of observations is used to establish relationships between CWV and deep convection across various instruments, time periods, and tropical locations. The principal location examined is the DOE ARM Mobile Facility at Manacapuru, Brazil (3°12'S, 60°35'W; 50-m altitude), established as part of the GOAmazon field campaign (January 2014–December 2015). The GOAmazon data used in this study cover the period from 10 January 2014 to 20 October 2015. The results for the GOAmazon site are compared with those derived from two retired DOE ARM sites in the tropical western Pacific: Nauru (0°31’S, 166°54'’E; 7-m altitude) and Manus Island (2°3’S, 147°25’’E; 4-m altitude). The analysis period used in this study and in Holloway and Neelin (2009) for Nauru spans April 2001–August 2006, and the analysis period from Manus Island spans January 2008–December 2010. In terms of radiosonde launches, these periods yield roughly comparable numbers to the western Pacific sites (3320 for Nauru and 3309 for Manus), each somewhat larger than the 2379 for GOAmazon.

Additional observations from a GPS meteorological station in Manaus, Brazil, are included in this study; this station functioned from July 2008 to December 2011 as part of the National Oceanic and Atmospheric Administration/Earth System Research Laboratory (NOAA/ESRL) Real-Time Ground-Based GPS Meteorological Network and was located at the National Institute for Amazon Research/Large-Scale Biosphere-Atmosphere Experiment (INPA/LBA) in Manaus (2.61°S, 60.2°W) (Adams et al. 2013, 2015).

a. Column water vapor

Radiosonde measurements at all ARM sites were obtained from Vaisala Digi-Cora III sounding systems at 2-s resolution; the raw sounding data were interpolated to 5-mb intervals (1 mb = 1 hPa). Reported instrumental uncertainties are approximately 0.5°C for temperature and 5% for relative humidity below 500 mb. At the GOAmazon site, radiosonde launches occurred four times daily (6 hourly) at 0530, 1130, 1730, and 2330 UTC, with occasional launches at 1430 UTC during the wet season. At Nauru, launches took place at 0000 and 1200 UTC, with occasional launches at 0230 and 1430 UTC, while at Manus Island, most launches took place at either 1130 or 2330 UTC, with occasional launches at 0330 or 1530 UTC.

CWV data sampled by microwave radiometer (MWR) at the GOAmazon site are derived from measurements of absolute microwave radiances (expressed as brightness temperatures) obtained at two frequencies: 23.8 and 31.4 GHz. The retrieval uncertainty for brightness temperatures is 0.3 K and for CWV is typically ~0.5 mm. The data used here are derived from radiance measurements with a statistical retrieval algorithm that uses monthly derived and location-dependent linear regression coefficients (Liljegren 1999). Data from another CWV product derived from a more physically based retrieval (Turner et al. 2007) are compared to these data in the appendix. All data for which the brightness temperature exceeds 100 K are removed from this dataset (Morris 2006), as are data that are affected by direct sunlight near local noon (1500–1700 UTC) for roughly a 3-week period surrounding the equinoxes. To address the so-called wet-window problem, in which water collecting on the surface of the lens introduces measurement inaccuracy during rainy periods, we linearly interpolate CWV values across time periods of 6 h or less. While the existence of the wet-window problem introduces uncertainty, particularly within the highest CWV bins, we argue that the linear interpolation procedure likely underestimates peak CWV and is justified given that the data gaps are typically short and the temporal persistence of water vapor values for strong convective events is on the order of hours (Holloway and Neelin 2010). Additionally, the appendix (Fig. A1) illustrates that there is no obvious systematic bias at high CWV for the times sampled (15-min-average radiometer CWV surrounding radiosonde launch between 10 January and 31 July 2014), which suggests that this interpolation does not greatly affect the results presented in this study.

One way to overcome measurement inaccuracy during rainy times is through use of GPS technology, as its all-weather capability allows for CWV measurements during rainy times (Adams et al. 2011; Adams et al. 2015). The CWV from GPS is derived from water vapor–induced delays in the radio signals from the satellite to the ground-based receiver (Bevis et al. 1992), and its accuracy in the Amazon is on the order of 1–2 mm (Adams et al. 2011). The INPA site consisted of a dual-frequency, geodetic-grade GNSS receiver/antenna and meteorological station concurrently measuring pressure, temperature, relative humidity, winds, and precipitation at 1-min sampling.
frequency. NOAA/ESRL processed the GNSS data in near–real time (2-h latency), with 30-min-average CWV values used in this study.

b. Precipitation
The GOAmazon precipitation measurements analyzed in sections 3 and 4 are from the Meteorological Measurements associated with the Aerosol Observing System (AOSMET), measured by the acoustic gauge of a Vaisala WXT520. When related to radiosonde CWV, AOSMET precipitation is averaged at 1-h intervals surrounding the launch; for analyses with radiometer CWV, AOSMET precipitation is averaged at 15-min intervals. In section 3, the averaging intervals are varied to evaluate the robustness of the statistics. These data were chosen among many other datasets available because we deemed them the most reliable over the full 2014/15 period (a detailed comparison of the different precipitation observing systems available at the GOAmazon site is included in the appendix).

The precipitation measurements used in this study vary slightly across sites owing to differences in instrumentation availability and reliability. In the tropical western Pacific at the Nauru and Manus Island ARM sites, precipitation was measured with an Optical Scientific optical rain gauge (ORG815), and 1-h averages surrounding radiosonde launches are analyzed in section 2. Section 5 uses precipitation from a Vaisala WXT-520 at the INPA site in Manaus, Brazil (30-min averages) for the analysis with GPS-derived CWV.

Section 4 assesses the robustness of the statistics presented as the horizontal resolution of the precipitation measurements decreases. We thus average precipitation from the Tropical Rainfall Measuring Mission’s (TRMM) 3B42 version 7 product across various spatial scales. The 3B42 precipitation estimates (mm h$^{-1}$) have a 3-hourly temporal resolution on a 0.25° $\times$ 0.25° grid, covering 50°S–50°N from 1 January 1998 to the present. The TRMM 3B42 precipitation estimates are a combination of multiple independent precipitation estimates from various microwave retrievals and algorithms, while missing data in individual 3-hourly merged-microwave retrievals are filled with microwave-adjusted merged geo-infrared (IR) estimates. The Precipitation Radar (PR) and TRMM Microwave Imager (TMI) are used to calibrate all input microwave data, while the IR estimates are computed using monthly matched microwave–IR histogram matching (Huffman et al. 2007). Estimates of precipitation from the microwave instruments are derived from several versions of the Goddard profiling algorithm (GPROF), a multichannel physical approach used to retrieve rainfall and vertical structure information (Kummerow et al. 2001). Over the oceans, GPROF uses signals from emission at low frequencies and scattering at higher frequencies. Over land, the algorithm reduces to a scattering-type procedure using only the higher-frequency channels. All of these estimates are adjusted to a best estimate using probability matching of precipitation rate histograms assembled from coincident data. Note that both the microwave and IR data are snapshots, except for small regions in which two or more overlapping microwave scenes are averaged. Generally, however, each satellite provides a sparse sampling of precipitation. As a result there can be significant gaps in the 3-hourly coverage by passive microwave estimates. Because of this, precipitation estimates can be thought of as instantaneous values, representative of the 3-h period in which they fall.

3. The relationship between deep convection and CWV over tropical land versus tropical oceans

a. The GOAmazon site—Manacapuru, Brazil
To illustrate the relationship between CWV and deep convection at the GOAmazon site, we conditionally average precipitation rate by CWV in Fig. 1. Figure 1a is the 1-h-average precipitation rate conditioned on radiosonde CWV, with the average centered at the time of radiosonde launch. Measurements for all available times (0530, 1130, 1730, and 2330 UTC, and occasionally 1430 UTC) were included in the averages. Note that for the statistics presented throughout, CWV bins are typically of equal 1.5-mm widths and range from 28 to 70 mm; exceptions to this will be noted where appropriate, such as here, where the highest CWV bin spans 6 mm from 64 to 70 mm, in order to include sufficient counts.

Beyond a threshold CWV value, a sharp increase in rain rate is evident. This confirms that the CWV–precipitation relationship and associated behavior exists over tropical land as it does over tropical oceans (Peters and Neelin 2006; Neelin et al. 2009; Holloway and Neelin 2009). The limited sampling of high CWV in the GOAmazon radiosonde observations, reflected in the large error bars (plus or minus one standard error), limits the precision with which the behavior above the pickup can be estimated; nevertheless, the data are sufficient to establish the occurrence of the pickup, and the radiosonde observations are key to analyzing the vertical structure, which will be discussed in section 6 below.

The larger sample size of radiometer CWV affords better quantification of the behavior at high CWV (Fig. 1d). For this purpose, Figs. 1d–f include four additional 1.5-mm bins at high CWV, in comparison to Figs. 1a–c. A sharp pickup is clearly evident in this dataset. Additionally, the conditionally averaged rain rates in the 61–64-mm range in Fig. 1a and the magnitudes observed in the 61–64-mm range of Fig. 1d mimic
each other, demonstrating the robustness of the results across various instruments. A strong correlation \((R = 0.91)\) between the 15-min-average radiometer CWV and radiosonde CWV (see Fig. A1) further highlights this consistency. Note that standard error bars in Figs. 1a and 1d reflect only the precipitation variance and number of counts in each bin as an estimate of sampling error and do not account for potential errors in CWV estimation.

The value of CWV at which the rapid pickup in precipitation begins, referred to as the critical value, is a useful measure in characterizing this onset. For the short, in situ datasets used here, empirical fits involve relatively few points with large error bars, so we simply use the point at the beginning of the rapidly increasing range as a rule of thumb. Estimating the critical value by a linear fit through the range over which precipitation is rapidly increasing, as in Sahany et al. (2014), and choosing a range of above 1 mm h\(^{-1}\) (appropriate for these data) yields a CWV value of \(\sim 60\) mm where the interpolation crosses 1 mm h\(^{-1}\) (Fig. 1d). This range is, however, instrument dependent.

Compared with the results from Neelin et al. (2009), the mean-tropospheric temperature at the GOAmazon site is 271.4 K, so the location of the pickup for GOAmazon occurs at lower CWV (\(\sim 61\) mm) than for comparable temperatures in the tropical eastern Pacific (\(\sim 65\) mm, interpolated between 271 and 272 K). This is consistent with the expectation that the mean-tropospheric temperature is only one of several controls on the onset of conditional instability and thus the location of the pickup and indicates that other key factors differing between tropical land and ocean are reflected in the onset. Specifically, boundary layer dynamics introduce additional complexity to the transition to deep convection over land, as the diurnal cycle is stronger over land and the partitioning of surface net radiation between latent and sensible heat fluxes depends on the interactions between several surface attributes (e.g., vegetation growth and soil moisture) and the atmosphere.

The curvature above the critical CWV in the radiometer analysis qualitatively resembles the behavior observed over the tropical oceans (Peters and Neelin 2006; Neelin et al. 2009), but we are cautious in drawing conclusions about this given the scatter at high values and limitations of the radiometer. The quantitative
values of the conditionally averaged precipitation in the pickup region are slightly smaller than those in microwave retrievals in Neelin et al. (2009) and Sahany et al. (2014), presumably in part a result of inherent uncertainties at high rain rates, particularly in the satellite observations where precipitation is inferred from cloud liquid water. Comparing the 15-min averages from the GOAmazon site to microwave retrievals over the tropical oceans (effectively snapshots) may also play a role.

Figures 1b and 1e illustrate an equally sharp increase in probability of precipitation as a function of CWV comparable to that shown for conditionally averaged rain rate in Figs. 1a and 1d, respectively. The fraction of precipitating points per CWV bin is defined as the number of CWV observations with rain rates greater than a small threshold (here 0.5 mm h⁻¹), divided by the total number of CWV samples in each bin. The probability increases dramatically above the critical value, sharply increasing to values greater than 50% in the highest CWV bins.

Figures 1c and 1f show the frequency of occurrence of different CWV values for all times and for precipitating times (where precipitation rates exceed 0.5 mm h⁻¹) at the GOAmazon site for radiosonde and radiometer CWV, respectively. Curves are scaled with respect to CWV bin sizes, similar to a probability density function (PDF) but in counts per millimeter—referred to here as frequency density. We chose not to normalize to instead yield PDFs to make the counts for each bin visible, as the lengths of the available datasets vary by instrument and location. The peak in the distribution of CWV, for both the radiometer and radiosonde analysis, occurs between 55 and 60 mm. The occurrence of the peak in the distribution occurs just below the critical point, consistent with the findings of Peters and Neelin (2006) and Neelin et al. (2009). The highest probability state of the system is near the beginning of the intense precipitation regime, as is shown by the distribution of precipitating points (the peak occurs in the 61–62.5-mm bin in the radiometer analysis and is slightly more spread out in the radiosonde data). Below 45 mm, no events exceeding the 0.5 mm h⁻¹ threshold are observed.

The longer-than-Gaussian tails of this distribution are also consistent with those seen in previous studies (Neelin et al. 2009, 2010), seen here with different instrumentation. Because of the lower number of radiosonde observations, we focus on radiometer observations (Fig. 1f). First, there is a long tail extending toward lower CWV in the distribution for precipitating points. The peak occurs just below or near the critical point, with a sharp decrease in frequency toward higher CWV in the region of rapid pickup of precipitation, consistent with the dissipative effects of precipitation on CWV (and of convection on buoyancy). Beyond the critical value, there is evidence of a long tail with roughly exponential decay as CWV increases, suggesting that the system is characterized by a higher frequency of extremes than would be expected from Gaussian statistics. This behavior is particularly evident in the radiometer analysis shown in Fig. 1f, but low counts in the high CWV bins and measurement uncertainty limit confidence in this feature.

Many of the onset statistics can be qualitatively and quantitatively captured by a simple stochastic model (Stechmann and Neelin 2011). This model suggests that the long tail for precipitating points in the low CWV regime is associated with a transition probability in which it typically takes some time to transition to a nonraining state when CWV decreases from the raining regime. The same hysteresis affects the position and value of the peak in the distribution for precipitating points, consistent with results here, suggesting it may be interesting in further work to distinguish temporal aspects of the transition, including distinguishing precipitating shallow and congestus convection or the formation of stratiform rain from deep precipitating convection (Ahmed and Schumacher 2015), as this is also a possible explanation of such characteristics of the distribution. The behavior of the distribution for all points at low CWV is expected to be rather dependent on the dynamics of the dry regime and has been noted to have various forms over ocean basins, including a second maximum. This may occur near the balance between evaporation and moisture divergence (Lintner and Neelin 2009).

b. The tropical western Pacific

Figure 2 illustrates the CWV–precipitation relationship for two sites in the tropical western Pacific—Nauru (Figs. 2a–c) and Manus Island (Figs. 2d–f). Compared to the Amazon in Fig. 1, the tropical western Pacific sites show very similar behavior. Radiosonde estimates are shown in Fig. 2, which can be directly compared to Figs. 1a–c. Both pickups of precipitation, for Nauru (Fig. 1a) and Manus Island (Fig. 1d), occur at higher values of CWV (~67 mm) than in the Amazon. As was discussed in section 1a, this is due in part to small differences in the mean-tropospheric temperatures (272.0 K at Nauru and 271.9 K at Manus Island) but is also likely due to fundamental differences in the convective environments of a tropical land site versus a tropical oceanic site. For reference, the values of column-integrated saturation specific humidity q_sat for the three sites are 76.0, 75.2, and 73.0 mm at Nauru, Manus Island, and the GOAmazon site, respectively, although it is known for tropical ocean basins spanning a wider range of tropospheric temperatures that q_sat poorly captures the temperature dependence (Neelin et al. 2009; Sahany et al. 2012) because the
relevant physical control is conditional instability rather than large-scale saturation.

As in the GOAmazon case, the fractions of precipitating points (Figs. 2b and 2f, for Manus Island and Nauru, respectively) sharply increase to 50% or greater beyond a critical CWV. This, again, illustrates that a sharp transition occurs not only in rain rate but also in the probability of precipitation beyond a threshold CWV. In Fig. 2c, Manus Island exhibits distinct peaks in its distributions: the peak of the CWV distribution occurs between 58 and 60 mm, whereas the peak in the distribution of precipitating points occurs between 60 and 63 mm. This is consistent with the findings from previous studies, where the peak in the precipitating points occurs at slightly lower CWV than the critical value. Also consistent is the sharp decrease in the frequency of CWV between the distribution peak and the CWV values where precipitation picks up rapidly. These characteristics are also observed for Nauru, but the peaks in the distributions of CWV and the precipitating points are broader in this sample from radiosondes; that is, the CWV distribution peak spans roughly 8 mm (≈50–58 mm), whereas the peak in the distribution of precipitating points spans roughly 10 mm (≈57–67 mm).

Even though marginal differences can be observed across locations, the main features of these statistics are consistent and robust across all three tropical locations. Though these statistics were not examined near regions of strong topographic forcing, land–sea contrasts, or other inhomogeneity, it has been suggested that these complexities can impact the statistics (Bergemann and Jakob 2016), as additional processes driving deep convection in these regions are not adequately captured with proxies such as CWV or column saturation. Nevertheless, the results presented here suggest that CWV is a good proxy for conditional instability and has a clear relationship to the onset of deep convection throughout much of the tropics.

4. The robustness of the observed statistics at various scales

a. The effects of temporal averaging

To explore how averaging over differing temporal scales can affect the statistics describing the transition to deep convection, we compute the transition statistics at various averaging intervals with in situ precipitation and radiometer CWV from the GOAmazon site. Four averaging intervals were chosen for this analysis: 15-min, 1-h, 3-h, and 1-day averages. These intervals were chosen to be most comparable to the current output available from models and observations.

In Fig. 3a, the magnitude of the conditionally averaged precipitation in the highest four CWV bins diminishes considerably as the averaging interval increases. Up to
and including 3-hourly averages, the magnitude and sharpness of the pickups are largely preserved. Beyond 3 h, the pickup is noticeably degraded. However, adapting the scale of the precipitation to the decrease in resolution (not shown) illustrates that despite the decreased sharpness of the pickup, the precipitation still increases with increasing CWV. Despite some variability in the shapes of the curves, the overall locations of the pickups are robust for temporal resolutions of 3 h or less. The location of the probability curve pickup in Fig. 3b, however, varies substantially as the size of the averaging interval increases: larger averages pick up sooner and have a higher probability of precipitating at high CWV. This can be explained by the fact that the 3-hourly and daily averages are more likely to span times when it is raining than the shorter averages are. The overall shape of the distribution is preserved with averaging (Fig. 3c), but the mean shifts to lower CWV and the tails of the distribution are shorter. Overall, these results illustrate how the statistics vary with temporal resolution, which should be considered when applying them as model diagnostics.

b. The effects of spatial averaging

The relationship between spatially averaged TRMM 3B42 3-hourly instantaneous precipitation (see section 2) and radiometer CWV (15-min averages) over the GOAmazon site is shown in Fig. 4 for 0.25° × 0.25° (Figs. 4a–c), 1.25° × 1.25° (Figs. 4d–f), and 2.5° × 2.5° (Figs. 4g–i). At either 0.25° or 1.25°, the relationship is comparable to the results in Fig. 1 and thus robust. At 2.5°, however, it starts to deteriorate, as the pickup of precipitation and the percentage of precipitating points occur too soon in comparison to Figs. 1d and 1e. These results are encouraging, as they suggest that resolutions up to about 1.25° × 1.25° are still of sufficient spatial resolution to reproduce robust statistics that explain the CWV–precipitation relationship, given that the temporal resolution is also adequate. This implies that these statistics are reproducible using the horizontal resolutions available with many current generation GCMs. In such comparisons, it should be borne in mind that a GCM with, for example, 2° resolution may respond at the finest scale available to it (i.e., the grid scale) in a manner similar to the convective response occurring at finer scales in observations.

5. Use of GNSS meteorological networks in the tropics

For two decades, GNSS/GPS meteorology has offered relatively inexpensive, high-frequency (~5 min), all-weather retrievals of CWV and is thus ideal for analyses requiring long, continuous records of observed CWV over land. This is particularly useful for studies in the tropics, where collecting in situ measurements is a challenge. We thus evaluate the convective transition statistics here for GPS data from a site in Manaus. In Fig. 5, the statistics are reproduced for GPS CWV and coincident measurements of precipitation (30-min averages) as in Fig. 1. Note that the precipitation measurements from the INPA site are biased low (see appendix). Therefore, for better comparison to the statistics in Fig. 1, the range shown on the precipitation axis (0–1.28 mm h⁻¹) is reduced relative to the range on the other pickup plots (0–6 mm) by a factor of 4.7—the ratio of means of precipitating points between the 30-min-averaged precipitation from both sites. Additionally, the threshold for identifying precipitating points is lowered to 0.1 mm h⁻¹ to more appropriately complement the 0.5 mm h⁻¹ threshold used in Fig. 1. When measurement differences are properly accounted for, Figs. 1 and 5 compare well: the location and shape of the pickup of precipitation is consistent, the probability of precipitation is just below 50% in the highest bin, and the distribution of CWV and precipitating points resides
near to the transition, with a sharp drop in frequency between the peak and the transition and a long tail extending out to high CWV. This suggests GPS technology will be valuable in observing characteristics of convection at high temporal resolution throughout tropical land regions.

6. Characterizing the variability of column moisture

a. Vertical thermodynamic profiles

Vertical profiles of thermodynamic quantities—specific humidity $q$, relative humidity (RH), and equivalent potential temperature $\theta_{\text{e}}$—are conditionally averaged on CWV in Figs. 6a–c, respectively. In Fig. 6a, it is evident that profiles of $q$ are most variable in the layers above 800 mb at the GOAmazon site. This differs slightly from the western Pacific case, as the variability in free-tropospheric $q$ (850–500 mb) with respect to CWV is slightly less over the Amazon than it is for Nauru [see Fig. 3a of Holloway and Neelin (2009)], presumably owing to stronger horizontal moisture gradients near Nauru (Lintner et al. 2011). Additionally, the contribution from the boundary layer is greater at the GOAmazon site with an approximately 4 g kg$^{-1}$ difference in humidity between the lowest and highest bins at 975 mb (Fig. 6a), whereas the difference is less than half this value at Nauru.

Profiles of RH (RH with respect to water) belonging to the highest CWV bins at the GOAmazon site (61 mm) are at least 90% saturated throughout the lower troposphere. At Nauru, this is the case for CWV greater than 66 mm [see Fig. 4a of Holloway and Neelin (2009)].
suggesting that the column is saturated for lower CWV in the Amazon than it is over the tropical western Pacific. Additionally, the boundary layer RH is roughly twice as variable over the Amazon as it is over Nauru, ranging from an approximately 35% difference at 975 mb between the highest and lowest bins versus an approximately 15% difference at Nauru. The variability observed in RH is highly consistent with variability in column moisture, since free-tropospheric temperature variations tend to be modest.

Equivalent potential temperature $\theta_e$, calculated reversibly following Emanuel (1994) in Fig. 6c, is an approximate measure of nonentraining parcel buoyancy, as convective available potential energy (CAPE) can be approximated by drawing a vertical line upward from the initial $\theta_e$. Where this line crosses the $\theta_e$ curve is roughly the level of free convection (LFC) of the unmixed parcel; the area to the left of the vertical line and to the right of the $\theta_e$ curve is roughly proportional to CAPE. The pattern of $\theta_e$ at the GOAmazon site shows similar overall variability in the vertical as it does at Nauru. In the absence of entrainment, many of the profiles belonging to the highest CWV bins have sufficient $\theta_e$ to support deep convection, providing that the convective inhibition (CIN) residual from the nighttime hours (seen in the $\theta_e$ profile) could potentially be overcome. This will be discussed further in section 7.

b. Moisture anomalies

Figure 7 illustrates the differences in $q$ at 1.5–3 h leading (red) and 1.5–3 h lagging (blue) precipitation, between profiles corresponding to precipitation events (rain rate $> 0.5$ mm h$^{-1}$) and those that do not correspond...
to a precipitation event (rain rate < 0.01 mm h\(^{-1}\)) for January–April soundings only. Leading an event, moisture anomalies exceeding 0.5 g kg\(^{-1}\) and as large as 1 g kg\(^{-1}\) are seen clearly throughout the lower troposphere, with a particularly large increase between 750 and 950 mb. This is consistent with evidence that increased low-tropospheric humidity supports deep convective initiation. These moisture anomalies are also seen in the tropical western Pacific at Nauru [Fig. 5 of Holloway and Neelin (2009)], where anomalies as large as 3 g kg\(^{-1}\) occur in the lower to midtroposphere within 3 h of a precipitation event. It is worth noting that the use of a higher precipitation threshold as in Holloway and Neelin (2.5 mm h\(^{-1}\)) does not change the results presented here with a 0.5 mm h\(^{-1}\) threshold. The anomaly in the lowermost layer at the GOAmazon site is more distinct from and almost as large as the anomaly within the hour of rainfall compared to the anomalies at Nauru, which show a consistent vertical structure at all times leading and lagging rainfall. This illustrates the greater moisture variability in the boundary layer over land preceding convection. It can be seen at both sites, however, that a distinct increase in moisture is present in the lower troposphere. There are likely multiple causes of this moistening in the lower troposphere preceding convection at both locations, including mixing and detrainment of water from shallow cumulus convection, in concert with convergence or lifting. Regardless of the cause, it will become evident in the next section as we examine the role of entrainment that sufficient lower-free-tropospheric moistening is essential to the onset of deep convection.

During a precipitation event, the anomaly in the 750–950-mb layer decreases, suggesting that moist air is lofted by updrafts, with drier air from downdrafts diluting the layer’s moisture content. The increased anomaly of mid- to upper-tropospheric humidity between 200 and 700 mb is likely due to the effects of convection via some combination of lofting and detrainment of moist air and large-scale ascent induced by convective heating, potentially including stratiform effects. As precipitation dissipates, this mid- to upper-tropospheric anomaly persists for hours afterward, which may aid in supporting subsequent convective events. These anomalies are present in the tropical western Pacific case as well, but the vertical structure is more consistent throughout the 6-h period than it is for the GOAmazon case; that is, the maximum \(q\) anomaly at all times is around 800 mb. In the GOAmazon case, on the other hand, the maximum 1.5 h before precipitation is found around 900 mb, during precipitation it is around 700 mb, and after precipitation it is found at about 500 mb. Additionally, separating the analysis out by time of day (not shown) indicates that these moisture anomalies are consistent for events occurring at all times of day.

Overall, in both the Amazon and the tropical western Pacific, humidity is enhanced throughout most of the troposphere for several hours leading and lagging the original precipitation event. Free-tropospheric humidity appears to behave similarly in land and ocean cases (although with larger amplitude variation in the western Pacific), whereas boundary layer moisture is more variable on short time scales in the land case. The Amazon also more clearly exhibits upper-tropospheric moisture anomalies in the hours following precipitation, while both the Amazon and western Pacific exhibit reduced boundary layer moisture.

c. Dependence on time of day

Considering the strength of the diurnal cycle over land, it is natural to wonder whether CWV is a good proxy for conditional instability at all times of day, given how conditions contributing to instability can vary diurnally. Figure 8 suggests that the relationship between CWV and precipitation is robust at all times of day. Figure 8a shows the relationship between 15-min-average radiometer CWV and precipitation for nighttime hours (2300–1200 UTC). The time intervals were chosen to complement the radiosonde launch times and the analysis presented in section 7. Figure 8d shows this relationship for the midday hours, which are the most convective hours of the day (1400–2000 UTC). The pickups of both
conditionally averaged precipitation (Figs. 8a and 8d) and the probability of precipitation (Figs. 8b and 8e) affirm that the relationship is robust throughout all times of day. The frequencies of occurrence of precipitation (Figs. 8c and 8f) are also consistent with the results in Fig. 1. Despite the fact that more convection occurs in the midday hours over the Amazon, the relationship holds true for all times of day.

7. The sensitivity of plume buoyancy to entrainment under simple freezing assumptions

In this section, we focus on connecting the observed pickup of precipitation to observed increases in buoyancy and the sensitivity to entrainment. We calculate the buoyancy perturbation profiles, the virtual temperature difference between the environment and the plume ($\Delta T_v = T_v,\text{plume} - T_v,\text{env}$) for plumes rising from the subcloud layer (1000 mb), with mixing occurring at each pressure level as described by

$$r_k = (1 - X_{k-1})r_{k-1} + X_{k-1} \bar{r}_{k-1},$$

(1)

where $X$ is the mixing coefficient, $r$ is a conserved variable (with $\bar{r}$ its environmental value), and $k$ denotes pressure level if $X$ varies. Here we calculate the mixing coefficient proportional to $z^{-1}$, where $z$ is height, in the layer in which plume mass flux is growing. This mixing assumption was referred to in Holloway and Neelin (2009) as deep inflow A (DIA) and corresponds to the Siebesma et al. (2007) LES-based dependence. DIA is chosen here because of its realistic representation of buoyancy perturbation profiles and overall consistency with the pickup of precipitation observed in Fig. 1 and is described as

$$X_k = c_i z_k^{-1} \Delta z,$$

(2)

where $X_k$ is the coefficient in (1), $\Delta z$ is a positive finite-difference-layer depth, and $c_i = 0.4$. Following Holloway and Neelin (2009), a simplified limiting case of freezing microphysics is also used: all condensate is conserved and freezing is assumed to take place very rapidly when the parcel reaches the freezing level.

The individual perturbation profiles are shown in Fig. 9 and have been conditionally averaged by CWV, with bin spacing as in Fig. 1a. Figure 9a illustrates the profiles of the radiosondes from all times of day. It is evident that only the highest CWV bins could be convective through a deep layer, as both the CIN (between the surface and ~850 mb) and reduced buoyancies in the mid- to upper troposphere are more likely to inhibit deep convection at lower CWV. Since the afternoon is the most convective time of day in the Amazon, we also separate the profiles by time of day to examine key thermodynamic differences and how stability in the nighttime hours could be contributing to the CIN observed in Fig. 9a.

Figure 9b, which includes nighttime soundings only (2330, 0530, and 1130 UTC), shows the larger layer of CIN seen in Fig. 9a, which suggests that this CIN is most...
characteristic of the nighttime soundings. At these times of day, it is unlikely for convection to fire as a result of local instability, but the profiles at high CWV appear supportive of deep convection if forced (Fig. 9b). In Fig. 9c, only profiles from late morning (1430 UTC) and early afternoon (1730 UTC) soundings were conditionally averaged by CWV. In contrast to the evening and morning soundings, there is very little CIN. The variability observed in the upper CWV bins is due to the low counts of profiles contributing to the average. Overall, it appears that a variety of CWV values would be conducive to convective activity in the afternoon hours, with CWV bins less than 60 mm acting to support shallow convection, whereas only the highest CWV bins act to support deep convection.

The entrainment calculations suggest that CWV has a causal effect on deep convection, but they do not indicate the extent to which this effect dominates the observed relationship. Evidence suggests that CWV can also covary with convective precipitation (Raymond et al. 2015). A concurrent study evaluating causality with modeling experiments (Y.-H. Kuo et al. 2016, personal communication), shows that without sufficient entrainment the pickup behavior observed in the relationship between precipitation and CWV (Figs. 1 and 2) does not occur.

Some caveats on this analysis should be noted:

1) The plume buoyancies sorted by CWV are considerably smaller in the lower troposphere compared to the tropical western Pacific case for the same computation (Holloway and Neelin 2009, their Fig. 8c). The onset of deep convection is thus likely dependent on other factors unique to tropical land cases—in particular, the greater variability of the boundary layer, as both sensible and latent heat fluxes are more variable. Additionally, there are key thermodynamic differences between the convective environments in the wet and dry seasons in the Amazon and thus likely differing thermodynamic controls on deep convection; that is, during the wet season, there is less CIN, less CAPE, and more moisture available throughout the column, whereas in the dry season there is more CIN, more CAPE, and less moisture available in the column (Collow et al. 2016).

2) Entrainment assumptions can affect the details of the buoyancy profiles seen in Fig. 9c. In particular, smaller (larger) values of the mixing coefficient in the lower troposphere yield larger (smaller) buoyancy values. Here it is useful to use a scheme that has already been applied to corresponding soundings at the western Pacific ARM site, but one could evaluate in further work whether there is evidence for differences in entrainment characteristics for deep convection over land in comparison to the ocean. More complex entrainment assumptions would obviously also have an impact—for example, rescaling the entrainment from the cloud base (de Rooy et al. 2013; Bechtold et al. 2014), the entrainment rate weakening as convection over land deepens (Del Genio and Wu 2010; Stirling and Stratton 2012), having a parameterized dependence on environmental humidity (Zhang and Klein 2010; Stirling and Stratton 2012), or a dependence on cloud size (Simpson 1971; Grabowski 2006; Khairoutdinov and Randall 2006; Stirling and Stratton 2012). However, the computations here indicate a strong dependence on free-tropospheric humidity can be found even with fixed entrainment.

3) Associated with the smaller buoyancy in the lower free troposphere compared to the oceanic case, the role of freezing is more important to occurrences of
positive buoyancy in the upper troposphere. If freezing is completely omitted, the jump in buoyancy seen near 550 mb in Fig. 9 does not occur, and profiles in the upper troposphere decrease slightly faster with height, yielding little buoyancy even in the high CWV cases. Microphysical differences are known to exist in continental versus maritime deep convective clouds: robust mixed-phase processes dominate continental convection, whereas maritime deep convection rarely involves vigorous microphysical processes in mixed-phase regions, although ice processes are still important (Rosenfeld and Lensky 1998; Xu and Zipser 2012). Proper representation of microphysical processes in models can affect simulation of clouds and precipitation (Khain et al. 2015; Zhang and Song 2016), and these processes can be significantly altered by different concentrations of aerosols (Andreae et al. 2004; Rosenfeld et al. 2008; Grabowski and Morrison 2016). Empirically examining the effects of ice microphysics on buoyancy in various convective environments of the tropics is thus of interest for future work.

The discussion of caveats above points to some interesting aspects in which representation of deep convection over tropical land can be expected to be more sensitive than over the ocean. The additional involvement of the boundary layer is no surprise. However, the dependence of the deep convective instability through the upper troposphere on contributions to buoyancy from the freezing process over land, even under highly favorable conditions (sufficient free-tropospheric water vapor and a favorable time of day), points to a potentially greater sensitivity to freezing microphysics than over ocean. These sensitivities will be addressed in future work. Nonetheless, the overall results for the leading order effects of lower-free-tropospheric water vapor on convection in the Amazon have striking parallels to the oceanic case.

8. Conclusions

This study compares and contrasts the relationship between CWV and deep convection in the Amazon to that in the tropical western Pacific using measurements from two neighboring sites at each location: specifically, results from the GOAmazon site in Manacapuru, Brazil, and the GNSS site at INPA in Manaus, Brazil, are compared to results from the DOE ARM sites at Nauru and Manus Island. The relationships evident at all locations are robust, with an increase in conditionally averaged rain rate as a function of CWV. The probability of precipitation often increases beyond 50% in the highest CWV bins. The distribution of CWV is consistent with the distributions observed in microwave retrievals over ocean (Neelin et al. 2009) for both precipitating points and all points, with the distribution for precipitating points peaking just below the critical value at which precipitation increases sharply and decreasing rapidly over the pickup region. All cases with sufficient data counts are consistent with a longer-than-Gaussian tail extending out to high CWV. Much of the variability in column moisture is due to variability in free-tropospheric humidity, suggesting that the onset of deep convection is strongly dependent on free-tropospheric humidity at tropical land sites, much like over tropical ocean sites.

The relationship between CWV and precipitation is generally robust across time of day. While there is a smaller fraction of precipitating points of a given CWV in nighttime hours compared with those occurring near midday, the conditionally averaged precipitation exhibits a very comparable pickup that increases beyond a threshold value of CWV. Thus, while the probability of nighttime precipitation likely depends on boundary layer factors, CWV remains an important proxy for the effects of lower-free-tropospheric water vapor on deep convection.

Because convection occurs at small time and space scales, spatial and temporal averaging can degrade the statistics describing the transition to deep convection. In daily averages, a highly smoothed version of the behavior may still be seen, but much information about the underlying physics—particularly the sharp onset of conditional instability associated with deep convection—is largely lost. Daily averages are thus suboptimal for examining this behavior over both land and ocean, and their use for such an analysis is not recommended. Examining these statistics at various averaging intervals closer to the appropriate time scales for convection indicates that the pickup curves are robust over averages from 15 min to 3 h. One-hour averages yield results very similar to 15-min averages, while 3-h averages slightly reduce the sharpness of the pickup. Similarly, using satellite retrievals of precipitation for a region surrounding the GOAmazon site at different spatial resolution yields convective transition statistics that reasonably reflect the in situ observations at 0.25° resolution, but the sharpness of the pickup of precipitation becomes smoothed as the domain of the spatial averaging increases.

Examining the temporal and vertical structure, lower-tropospheric moisture increases prior to convection and precipitation at the GOAmazon site. This is consistent with findings for the tropical western Pacific ARM sites (Holloway and Neelin 2009). However, for the land case the lower-tropospheric moisture is more variable in time and there are clear indications that moisture has been lofted following the convection, whether by the detrainment of water at various levels during the convective event or by heating and stratiform processes.
After convection, the subcloud layer becomes cooler and slightly drier over both land and ocean. The before and after moisture profiles in this tropical land case thus illustrate the two-way interaction between convection and water vapor, with increases in lower-tropospheric water vapor prior to convection consistent with impacts on buoyancy in entraining convection.

The latter impacts are tested by computing buoyancy profiles with a previously used profile of turbulent entrainment, which are then conditionally averaged by CWV to assess whether buoyancy through a deep convective layer is comparable to the onset of precipitation as a function of CWV. This is examined for soundings from all times of day and for nighttime and midday ensembles of profiles separately. For nighttime conditions, CIN may be an additional control, although the CIN is considerably smaller for the highest CWV values. The nighttime results may point to the presence of preexisting disturbances or boundary layer conditions not captured by CWV. The midday soundings show buoyancies sufficient for shallow convection over a middle range of CWV. However, only the highest CWV bins would be convective through a deep layer for each case—nighttime, midday, and all times—consistent with the pickup of precipitation. Some differences relative to the ocean are worth noting: there are likely greater contributions from the boundary layer to the conditional instability of the environment that cannot be sufficiently explained by variability in CWV, and there is evidence that freezing microphysics exerts greater influence on the development of buoyancy above the freezing level. Nevertheless, the dependence of deep convective onset on free-tropospheric humidity is robust and of leading order over both tropical land and tropical ocean.

Acknowledgments. U.S. Department of Energy Atmospheric Radiation Measurement (ARM) Climate Research Facility GOAmazon and Tropical West Pacific field campaign data were essential to this work. This research was supported by the Office of Biological and Environmental Research of the U.S. Department of Energy Grant DE-SC0011074 and DE-SC0011069, National Science Foundation Grants AGS-1102838 and AGS-1505198, and National Oceanic and Atmospheric Administration Grants NA14OAR4310274 and NA15OAR4310097. We thank C. Holloway for providing analysis routines.

APPENDIX

Comparison of CWV and Precipitation Datasets

To illustrate the consistency between radiometer CWV and radiosonde CWV, Fig. A1 shows 15-min-average radiometer CWV scattered against radiosonde CWV. CWV is thus sampled every 6 h within the period 10 January–30 September 2014. It is evident that there are no systematic biases observed at high values of CWV, which could have resulted from interpolation or measurement inaccuracy. Overall, while our ability to confirm consistency between instruments is limited to the sampling of the radiosondes, it is evident from this sample that the CWV values agree well across instruments.

Figure A2 shows the probability density functions (PDFs) of the five precipitation datasets used throughout this study: AOSMET at the GOAmazon site, ORG at Nauru, ORG at Manus Island, TRMM, and the dataset from INPA in Manaus, Brazil, coincident with the GPS CWV measurements. It is evident that the PDFs of the precipitation data from the GOAmazon site, Nauru, and Manus Island are all consistent with one another, whereas the TRMM and INPA datasets are biased low, with much lower probability of high rain rates (Fig. A2a). This contributes to differences in the magnitudes of the pickup curves between those seen in Fig. 4 (TRMM) and Fig. 5 (INPA), in comparison with Figs. 1 and 2. The TRMM data in Fig. 4 require a unique precipitation axis to those of Figs. 1 and 2, since these data have a different spatial footprint than all others used in this study. Figure 5, on the other hand, adopts an axis that is scaled according to the ratio of 30-min-mean INPA data and 30-min-mean radiometer data. This value (4.68 mm h$^{-1}$) is divided by the range used in Fig. 1d (6 mm h$^{-1}$) to instead yield a range of 1.28 mm h$^{-1}$ for the axes in Fig. 5. It is also apparent that the PDFs of the precipitation datasets at Nauru, Manus Island, and the GOAmazon site all exhibit power-law behavior (Fig. A2b) out to high precipitation (∼40 mm h$^{-1}$), after
which point the distribution drops off exponentially. The exponent of the power-law range is approximately $-1.25$, while the decay scale of the exponential range is roughly $26 \text{ mm h}^{-1}$.

Figure A3 compares the available precipitation observing systems at the GOAmazon site by scattering the 15-min-average precipitation rates of each system against the chosen dataset, AOSMET. Between 1 January and 15 October 2014, four instruments recorded precipitation: an ORG, a present weather detector (PWD), a Vaisala WXT520 from the AOSMET, and a Vaisala WXT520 from a system including a three-channel microwave radiometer (MWR3C). Comparison to MWR3C precipitation is not included in this analysis, but the data compare well with the AOSMET precipitation chosen for use in this study (A. Theisen, ARM Climate Research Facility Data Quality Office, 2016, personal communication).
Figure A3a shows the PWD and ORG datasets scattered against the AOSMET dataset. Two main features are worth noting: 1) the plateau of rain rates in the PWD data (blue) and 2) the erroneous rainfall measured by the ORG (green). The plateau of PWD rain rates indicates that the instrument records a maximum value of \( \sim 8–10 \text{ mm h}^{-1} \); this leads to the systematic recording of erroneously low rain rates above an unknown threshold. These data could be used to confirm the incidence of rain, but analysis of the rain-rate magnitudes using these data is not recommended. The ORG had many operational problems throughout the specified time period and thus often recorded precipitation when it was not raining, as is evident from the scatter on the ordinate. Less evident are all of the erroneous values at low rain rates recorded as a result of instrument malfunction. Eliminating all points less than 0.5 mm h\(^{-1}\) in the ORG data would likely remedy some issues on the low end, but a threshold would not likely help to eliminate erroneous data on the high end. Therefore, these data must be extensively examined and errors must be corrected for before using these data prior to 15 October 2014 when the instrument was repaired (A. Theisen, ARM Climate Research Facility Data Quality Office, 2016, personal communication).

After 15 October 2014, five instruments measured precipitation at the GOAmazon site; all data besides that from the MWR3C system are included in Fig. A3b. It is evident that the ORG data are consistent with the AOSMET precipitation after 15 October 2014, as are the data from the Parisvel laser disdrometer (PARS) and the tipping-bucket rain gauge (RAIN). Overall, however, the AOSMET precipitation dataset is the most reliable for use throughout the entire GOAmazon campaign, and is thus chosen for use in this analysis. Prior to 15 October 2014, use of neither the PWD nor the ORG precipitation datasets is recommended.

As previously noted in section 2, a value-added CWV product from the MWR was archived by the DOE ARM Data Archive in March 2016 and made available in June 2016 (R. Jundt, ARM Climate Research Facility, Pacific Northwest National Laboratory, 2016, personal communication). This dataset, referred to as MWRRET, is retrieved using a physically based methodology that includes more information about the atmospheric state in the retrieval process, an optimal estimation in an iterative scheme to retrieve CWV and liquid water path, and a radiative transfer model (Turner et al. 2007).

Even after linearly interpolating over periods of 6 h or less, the MWRRET contains 60% less data in comparison to the MWR data, and the missing data are likely biased toward periods with substantial precipitation. The length of missing values over which the MWRRET data are linearly interpolated is roughly 10%–50% greater than for the MWR data. In the MWRRET CWV dataset, differences in the linear interpolation of CWV values across wet-window periods (rainy times) tend to shift many of the values previously belonging to high CWV bins to lower CWV bins. The above contributes significantly to the difference in the statistics.

Figure A4 shows a comparison of this dataset to the analysis in Fig. 1. The results using MWRRET yield identical locations of the precipitation and probability pickups. However, the magnitude of the precipitation and the probability of precipitating points in the last four CWV bins (Figs. A2a and A2b) are reduced compared with Figs. 1d and 1e. Additionally, the long tails of the CWV distributions for all points and precipitating points at high CWV, seen in Fig. 1f, are less pronounced (Fig. A2c) since the occurrence of high CWV is significantly reduced. Overall, this comparison of the less-filtered MWR retrieval in the body of the paper to the more conservative MWRRET retrieval in Fig. A4 helps to bound the sensitivity of the statistics to retrieval choices. Regardless of the CWV data retrieval algorithm and method, the key features of the statistics observed in Fig. 1 are robust.
In summary, although the qualitative convective transition statistics are robust across a broad set of instrumentation, careful consideration must be given to the observing systems and various data retrieval methods for quantitative aspects, as systematic biases and instrument error could affect comparisons to model output.

REFERENCES


