Distributions of Tropical Precipitation Cluster Power and Their Changes under Global Warming. Part II: Long-Term Time Dependence in Coupled Model Intercomparison Project Phase 5 Models

KEVIN M. QUINN AND J. DAVID NEELIN

Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, California

ABSTRACT

Distributions of precipitation cluster power (latent heat release rate integrated over contiguous precipitating pixels) are examined in 1°–2°-resolution members of phase 5 of the Coupled Model Intercomparison Project (CMIP5) climate model ensemble. These approximately reproduce the power-law range and large event cutoff seen in observations and the High Resolution Atmospheric Model (HiRAM) at 0.25°–0.5° in Part I. Under the representative concentration pathway 8.5 (RCP8.5) global warming scenario, the change in the probability of the most intense storm clusters appears in all models and is consistent with HiRAM output, increasing by up to an order of magnitude relative to historical climate. For the three models in the ensemble with continuous time series of high-resolution output, there is substantial variability on when these probability increases for the most powerful storm clusters become detectable, ranging from detectable within the observational period to statistically significant trends emerging only after 2050. A similar analysis of National Centers for Environmental Prediction (NCEP)–U.S. Department of Energy (DOE) AMIP-II reanalysis and Special Sensor Microwave Imager and Imager/Sounder (SSM/I and SSMIS) rain-rate retrievals in the recent observational record does not yield reliable evidence of trends in high power cluster probabilities at this time. However, the results suggest that maintaining a consistent set of overlapping satellite instrumentation with improvements to SSM/I–SSMIS rain-rate retrieval intercalibrations would be useful for detecting trends in this important tail behavior within the next couple of decades.

1. Introduction

Characterizing the current state of organized tropical convection and projected changes under global warming is important because of the potentially large socioeconomic impacts associated with such changes. A survey of studies examining phases 3 and 5 of the Coupled Model Intercomparison Project (CMIP3 and CMIP5) coupled climate models shows a projected increase in extreme precipitation event frequency and intensity by the end of the twenty-first century (e.g., Tebaldi et al. 2006;

Kharin et al. 2007, 2013; Sillmann et al. 2013), although uncertainties emerge in both observational (e.g., Easterling et al. 2000; Alexander et al. 2006; Kharin et al. 2007, 2013; Lenderink and Van Meijgaard 2008; Allan et al. 2010) and global-scale modeling research (e.g., Tebaldi et al. 2006; Kharin et al. 2007, 2013; Allan and Soden 2008; Allan et al. 2010; Sillmann et al. 2013) in recent and future climate. One potential source of uncertainty in extreme precipitation projections is an inadequate representation of important physical processes associated with convection in climate models (e.g., Kharin et al. 2007) including convective organization (e.g., Tan et al. 2015). With the advent of high spatial and temporal spaceborne passive microwave imagers, much has been learned about convective organization based on satellite retrieved precipitation products (e.g., Huffman et al. 2007; Allan and Soden 2008; Allan et al. 2010), although such platforms are not without limitations (e.g., McCollum and Ferraro 2003; Hilburn and Wentz 2008; Bowman et al. 2009; Allan et al. 2010; Chen et al. 2013). There is thus a need for studies that examine

---

Corresponding author: Kevin M. Quinn, kquinn@atmos.ucla.edu

measures of convective organization in observations, in comparison to models and in terms of their potential changes under global warming.

In Quinn and Neelin (2017, hereafter Part I), we build an observational baseline of tropical precipitation integrated over contiguous clusters using high-resolution satellite data with complete spatial coverage. This is motivated by previous observational studies of cluster behavior in various space or time measures (Mapes et al. 2009; Peters et al. 2009, 2010, 2012; Wood and Field 2011; Skok et al. 2013) and by theory and modeling, suggesting time-domain cluster measures can be useful for understanding behavior under global warming (Neelin et al. 2017). The integrated precipitation over a cluster can equivalently be expressed as an integrated latent heat release or cluster power. The probability distribution of cluster power is shown to follow a long, scale-free power law, with a distinct cutoff (i.e., a more rapid drop in probability at high power). We then show that the High Resolution Atmospheric Model (HiRAM) with prescribed sea surface temperatures (SST) at two resolutions accurately simulates these observed cluster power statistics. We also examine HiRAM output from mid- and end-of-century simulations, finding large increases in the frequency of the most powerful clusters under a "business as usual" global warming scenario. These increases exceed those expected under simple Clausius–Clapeyron scaling of precipitation.

Here, we compare Special Sensor Microwave Imager (SSM/I) and Special Sensor Microwave Imager/Sounder (SSMIS) cluster power distributions with Tropical Rainfall Measuring Mission (TRMM) 3B42 data, to ensure consistency across satellite rain-rate retrievals. We then examine cluster power distributions over the SSM/I–SSMIS observational record to check for changes in the frequency of the most powerful storm clusters in recent climate. Trends uncovered in the analysis of the satellite data record are next compared to an analysis of cluster power distributions from the National Centers for Environmental Prediction (NCEP)–U.S. Department of Energy (DOE) AMIP-II reanalysis (NCEP Reanalysis 2) observationally constrained modeled precipitation dataset. We then analyze cluster power distributions from an ensemble of high-resolution models from the CMIP5 under the representative concentration pathway 8.5 (RCP8.5) global warming scenario. We (i) check if the model ensemble adequately reproduces cluster power distributions in current climate in observations and in HiRAM; (ii) verify that coupled models yield end-of-century changes consistent with those simulated in HiRAM; and (iii) conduct a time series analysis of the probabilities of clusters from the

highest power intervals to inform expectations for detectability.

2. Data and methods

Observational rain-rate data are retrieved from the SSM/I and SSMIS platforms onboard polar-orbiting Defense Meteorological Satellite Program (DMSP) satellites. The SSM/I platform flew onboard the *F08*, *F10*, *F11*, *F13*, *F14*, and *F15* DMSP satellites from 1987 to 2009 (Wentz 2013). SSMIS data come from the *F16* and *F17* DMSP satellites, beginning in May 2004 (Wentz et al. 2012). Our goal is to use cluster power computed from the SSM/I–SSMIS time series as a reliable indicator of recent historical cluster power behavior because the period for which SSM/I–SSMIS data are available (1987–present) is longer than the period over which TRMM 3B42 data (1998–2015) are available, presenting a more complete observational record of tropical precipitation data.

SSM/I–SSMIS rain-rate data are retrieved using the version 7 (V7) algorithm (mm h^{-1}) and are available twice daily, over oceans, at $0.25° \times 0.25°$ latitude–longitude resolution (Wentz et al. 2012), and, to remain consistent with Part I, we confine our time domain to 1 May–30 September. Rain rates and other data retrieved from these platforms are highly intercalibrated, with further details noted in Wentz and Spencer (1998), Wentz et al. (2012), Wentz (2013), and Yan and Weng (2008). Since SSMIS swath width (NSIDC 2016b) is 300-km wider than SSM/I swath width (NSIDC 2016a), we also test SSM/I–SSMIS cluster power distributions for sensitivity to the swath width difference over the complete SSM/I–SSMIS platform overlap period (2004–09) by comparing SSM/I, SSMIS, and SSMIS cluster power distributions that have had their swath width narrowed to match SSM/I swath width. On average, each SSM/I swath is 16 pixels narrower than each SSMIS swath over the tropics between 30°S and 30°N, so to match SSM/I and SSMIS swath widths, we run a simple procedure that masks eight pixels on each side of every SSMIS swath prior to running our clustering and binning procedures.

We first compare SSMIS cluster power distributions for 2004–09, examining cluster power sensitivity to the same minimum rain-rate thresholds used in Part I. Then we compare the cluster power distribution at the highest minimum rain-rate threshold used in this study (0.7 mm h^{-1}) from the TRMM 3B42 retrieval to SSM/I–SSMIS distributions at the same threshold retrieved over 2004–09, assessing storm cluster behavior across multiple satellite rain-rate retrievals. As noted in Part I, TRMM 3B42 rain-rate data (here, 2004–09 only) are merged from sensors onboard the TRMM spacecraft and other satellites to provide 3-hourly rain-rate

retrievals (mm h^{-1}) at 0.25° × 0.25° latitude–longitude resolution, over land and ocean (Huffmann et al. 2007; Huffmann and Blovin 2014; TRMM 2015). TRMM 3B42 rain-rate retrievals from only 0000 and 1200 UTC are included in this study, in order to optimize the comparison with twice-daily SSM/I–SSMIS data.

Next, we evaluate cluster power behavior in recent climate using an observationally constrained model. Modeled precipitation flux data (kg m^{-2} s^{-1}) from the NCEP Reanalysis 2 dataset present a data record extending back to 1979 and so are compared against satellite retrievals to characterize trends in cluster power behavior in the historical period. NCEP Reanalysis 2 data are available every 6 h (and not every 3 h like SSM/I–SSMIS, TRMM 3B42, and CMIP5 datasets employed here) at a horizontal resolution of approximately 2° latitude × 2° longitude (Kanamitsu et al. 2002), although only data from 0000 and 1200 UTC are incorporated here.

Last, we examine if a suite of coupled CMIP5 models (Table S1 in the supplemental information), chosen for relatively high spatial (e.g., less than 2°) and temporal (e.g., 3 hourly) resolution, exhibits changes in frequency and intensity of the most powerful storm clusters in recent, midcentury, and end-of-century climate under RCP8.5. In Part I we find that the portion of the distribution showing the most change corresponds to the events likely to have the most impact (i.e., highest cluster power), so we (i) analyze if the probability of storm clusters from the highest power intervals increases in the future, (ii) quantify the magnitude of such probability changes, and (iii) determine when changes first become detectable (as a statistically significant linear trend) compared to climatology. As in Part I, we only include cluster power data from 0000 and 1200 UTC, follow the same clustering procedure and cluster power calculation, and confine our analysis to the tropics (globally from 30°N to 30°S) from 1 May to 30 September.

To calculate cluster power, we follow a similar contiguous precipitating pixel approach described in Peters et al. (2012) and Part I, first masking cells not meeting the 0.7 mm h^{-1} minimum rain-rate threshold. Sensitivity to this threshold was analyzed in Part I. It has little effect in the observational datasets, but model behavior is more robust when regions of low rain rate are excluded. Contiguous precipitating cells with rain rates meeting this threshold then have their rain rates converted to the instantaneous latent heat release per grid cell, which are in turn integrated over a cluster to obtain cluster power, in units of gigawatts. We convert rain rate into a measure of latent heat release as it quantifies the instantaneous amount of energy released over a storm and directly ties into Earth's energy budget.

The binning procedure in building probability density functions (PDFs) for these distributions follows that of Part I, with slight modifications, as noted below. As in Part I, bin width increases smoothly as probabilities drop, using a bin width that is approximately constant in log space of cluster power. Recognizing that increments of cluster size are quantized to multiples of the minimum cluster size, one must ensure that the bin spacing is consistent with this. Bin widths are therefore adjusted to the integer multiple of the minimum cluster size that is closest to the chosen asymptotic constant bin width. In practice, the variations in bin width are small; Table S3 of the supplemental information (SI) shows both bin width and histogram counts N_i prior to normalization by the width of bin i and the total counts for each analysis presented. Error bars are given by $\pm N_i^{1/2}$, with the same normalization as the PDF. The minimum cluster size is set by the grid size and the minimum precipitation threshold, so the same bin boundaries apply to historical and future climate runs of the same dataset. Robustness to the choice of asymptotic bin width has been examined, and, for the key case of CMIP5 data changing under global warming, two versions are presented. The basic version simply has the asymptotic bin width set and then the PDF is computed. In a second version that ensures similar error bars among models in the high power range of most interest, the number of counts in the highest power nonzero bin in the historical period is examined over a small range of asymptotic bin widths, choosing the value for which the number of counts in the highest power nonzero bin in the historical period is a specified target that keeps fractional error at a uniform modest value that is uniform among models (Table S2 of the SI). Because probabilities drop steeply near the highest bin, a very small shift of the bin boundaries is sufficient to increase cluster counts in the highest bin. We then apply the same bin boundaries to future climate runs of the same dataset. In practice, this simple adjustment procedure yields improved statistical significance for the highest bin for comparison between recent climate and future climate output. PDFs are shown both pre- and postadjustment. For each time series, the bin structure is invariant in time (for the models, including both historical and RCP8.5).

3. Analysis

a. Cluster power distributions and probability trends for high power clusters in observations

We evaluate the current state of tropical cluster power behavior across multiple satellite retrievals (SSM/I, SSMIS, and TRMM 3B42) and an observationally
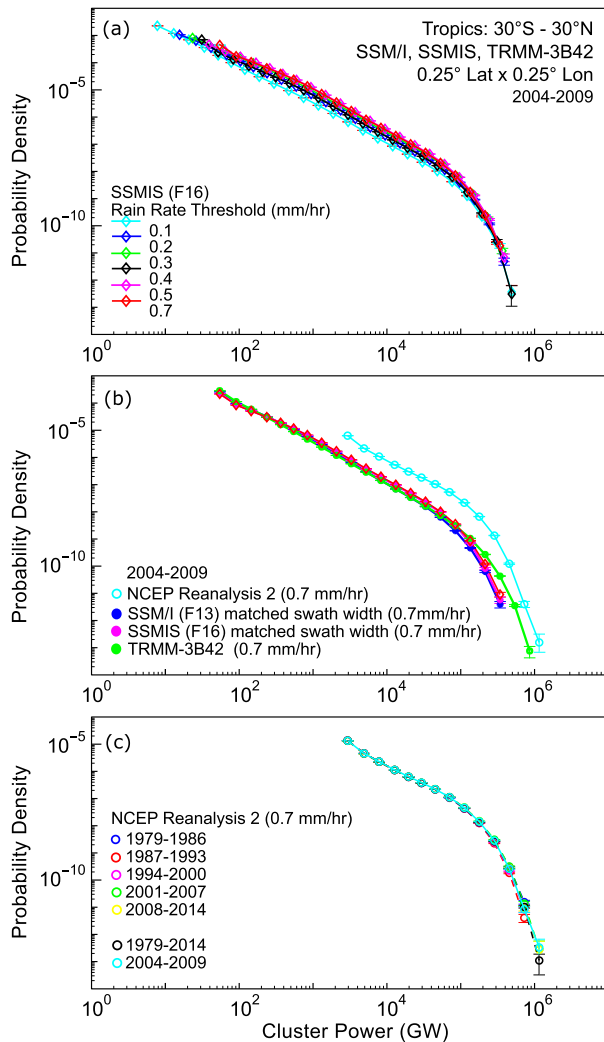
FIG. 1. (a) Probability distributions of cluster power (i.e., precipitation integrated over clusters of contiguous pixels exceeding the specified rain-rate threshold) from the SSMIS (*F16*) precipitation product over the tropics for 1 May–30 Sep 2004–09. As in Part I, cluster power can be equivalently expressed in terms of a mass budget as the integrated mass of water lost per hour (kg h^{-1}), with 1 GW equal to 1.4×10^6 kg h^{-1}. (b) Comparison of cluster power probability distributions for the 0.7 mm h^{-1} rain-rate threshold from TRMM 3B42, SSM/I (*F13*), matched-swath-width SSMIS (*F16*), and NCEP Reanalysis 2 precipitation products, 1 May–30 Sep 2004–09. (c) NCEP Reanalysis 2 precipitation product cluster power probability distributions for the 0.7 mm h^{-1} rain-rate threshold for 1979–2014, subdivided into periods shown.

FIG. 2. Scatterplot of rain-rate percentiles meeting the 0.7 mm h^{-1} rain-rate threshold from SSM/I (*F13*) and matched-swath-width SSMIS (*F16*) precipitation products, 1 May–30 Sep 2004–09. The correlation coefficient between precipitation products, least squares best-fit line, and one-to-one line are also plotted for reference (from 0 to 10 mm h^{-1}).

constrained model (NCEP Reanalysis 2) in Fig. 1, investigating if cluster power PDFs for recent climate mirror the general behavior described in Part I. In Figs. 2 and 3, we examine the SSMIS retrieval for changes in the probability of the most powerful storm clusters resulting from the swath width difference between the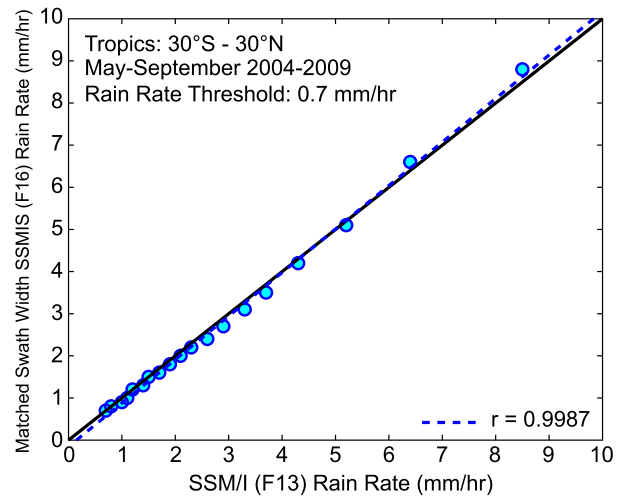 SSM/I and SSMIS platforms. We first intercompare rain-rate percentiles from the SSM/I and matched-swath-width SSMIS datasets, then analyze probability densities from the highest power bins for the SSM/I, SSMIS, and matched-swath-width SSMIS datasets. As shown in Part I, the largest changes in cluster power behavior occur in the high power bins beyond the cutoff. So, in Figs. 3 and 4, we show probability densities for the highest cluster power intervals beyond the cutoff using a time series analysis of SSM/I–SSMIS datasets and NCEP Reanalysis 2 output for which sufficient counts are available, describing any changes to the frequency of the most powerful storm clusters in recent climate.

Cluster power PDFs (Fig. 1a) from the SSMIS platform (*F16*) at multiple rain-rate thresholds display little sensitivity while also having similar scale-free power-law ranges and cutoffs as the TRMM 3B42 cluster power distributions in Part I. In Fig. 1b, the observed cluster power distributions for SSM/I (*F13*), SSMIS, and matched-swath-width SSMIS datasets at the 0.7 mm h^{-1} rain-rate threshold all have the same scale-free power-law range and cutoff, closely paralleling the TRMM 3B42 cluster power distribution with a sharp drop in the frequency of storm clusters in the highest power bins beyond 10^5 GW. Beyond the cutoff, the tail of the TRMM 3B42 distribution shifts toward higher power and includes two extra high power bins compared to the SSM/I–SSMIS distributions. The TRMM 3B42 dataset contains merged rain rates from multiple platforms, so unlike SSM/I–SSMIS platforms, there are no gaps in spatial coverage that limit cluster area. This implies that
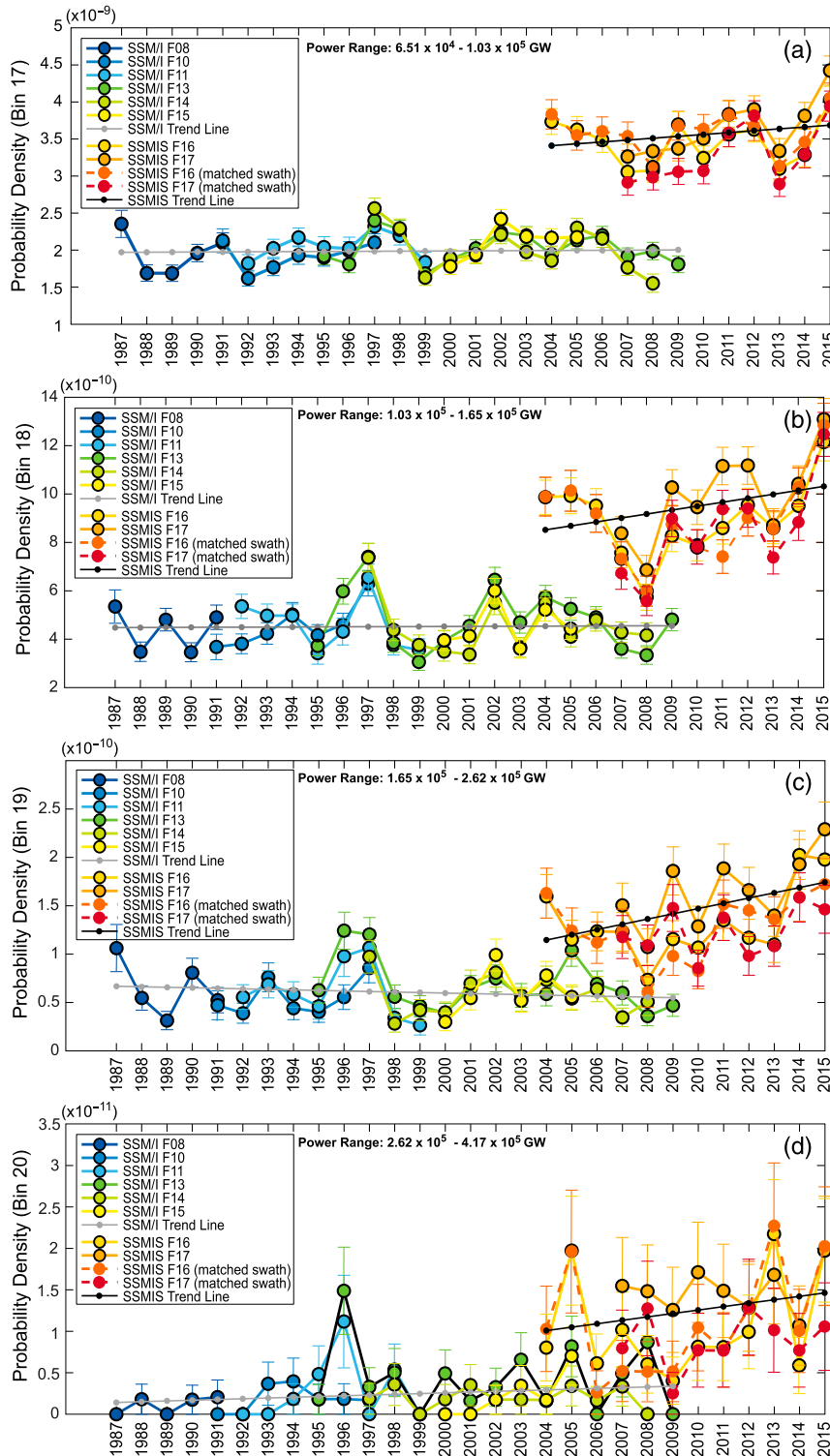
FIG. 3. Time series plots of probability densities from the four highest cluster power intervals (power intervals noted next to legends) for which statistics can be computed for each SSM/I, SSMIS, and matched-swath-width SSMIS precipitation product for the $0.7\,\mathrm{mm\,h^{-1}}$ rain-rate threshold, 1 May–30 Sep 1987–2015. Also plotted are least squares best-fit trend lines for the SSM/I and SSMIS platforms.
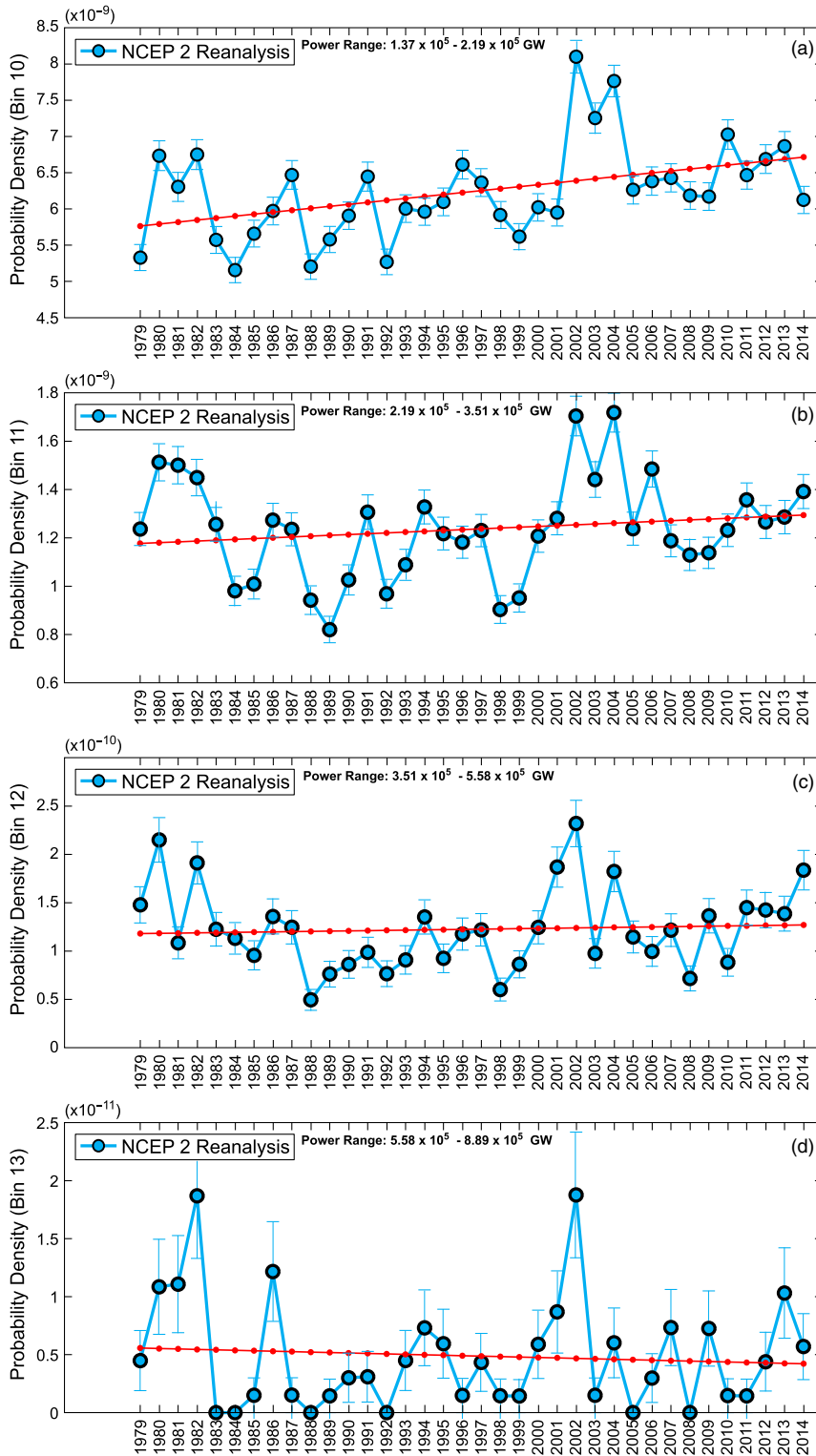
FIG. 4. As in Fig. 3, but for the NCEP Reanalysis 2 precipitation product.

TRMM 3B42 output includes larger clusters, which results in the TRMM 3B42 distribution containing two extra high power bins, and probability density drops slightly less steeply above the cutoff.

SSM/I–SSMIS and TRMM 3B42 data records only extend back to 1987 and 1998, respectively. For a more complete picture of cluster power behavior in recent climate, we analyze cluster power behavior in the NCEP Reanalysis 2 precipitation dataset, the data record of which begins in 1979. NCEP Reanalysis 2 cluster power distributions for 1979–2014 at the $0.7\,\mathrm{mm\,h^{-1}}$ rain-rate threshold (for entire timeframe and 7–8-yr subsets) are plotted in Fig. 1c. For reference, the NCEP Reanalysis 2 cluster power distribution from 2004–09 is replotted in Fig. 1b to compare with observed data. Because the NCEP Reanalysis 2 output has a comparatively coarse resolution, its scale-free, power-law range begins at higher power and contains fewer bins than the observed scale-free, power-law range. Even so, NCEP Reanalysis 2 cluster power distributions qualitatively agree with observed data, with a slight shift in the cutoff toward higher power. Furthermore, we find little variation in cluster power behavior between 1979 and 2014, the implications of which will be further examined in Fig. 4. The slightly different averaging interval compared to TRMM 3B42 (6 vs 3 h) and other caveats on retrieval comparisons (e.g., Adler et al. 2001) do not appear to have strong effects on the form of the distribution, and in examining changes below, each dataset is compared only to itself.

SSM/I (*F13*) and matched-swath-width SSMIS (*F16*) rain-rate retrievals for May–September 2004–09 display a high degree of agreement, suggesting consistency of intercalibration (Figs. 1 and 2), consistent with previous SSM/I–SSMIS calibration research noted in Yan and Weng (2008) and Sun and Weng (2008). SSM/I-matched-swath-width SSMIS cluster power distributions closely parallel each other (Fig. 1), rain-rate percentiles correlate highly ($r = 0.9987$; Fig. 2), and the least squares best-fit line through the rain-rate percentiles is nearly one-to-one (Fig. 2), suggesting the current intercalibration should suffice for most purposes. In principle, by matching SSMIS and SSM/I swath widths, one might expect to be able to construct a complete time series of observed cluster power behavior since 1987. However, as shown below, for trend detection in the probabilities of large precipitation clusters, the degree of accuracy required in intersatellite calibration is more demanding than the levels shown by the tests so far.

Figure 3 shows time series of probability densities for four cluster power intervals (6.51×10^4–1.03×10^5, 1.03–1.65×10^5, 1.65–2.62×10^5, and 2.62–4.17×10^5 GW), characterizing the high cluster power range for all

satellites carrying the SSM/I and SSMIS platforms for 1987–2015. For brevity, we will refer to these intervals as bins 17–20. As noted in Part I, the model-projected changes tend to occur disproportionately for the most powerful storm clusters ranging in the high power regime beyond the cutoff, with the fractional change in probability density increasing with power in this range. We thus examine whether a signature of increase can be seen for intervals of power across this range, conducting a time series analysis of the four highest power bins individually. Integrated probability above a given power value tends to be dominated by the interval near that power value (because of the steep drop in baseline probability density with increasing power), so showing probability density change as a function of power intervals provides a more complete view. The bin boundaries are the same for each platform: that is, bin 17 boundaries for SSM/I platforms *F08*, *F10*, *F11*, *F13*, *F14*, and *F15* are the same as for SSMIS platforms *F16* and *F17* bin 17 boundaries. The cluster power values for these bins remain constant between SSM/I and SSMIS platforms, while over time and platform the number of counts per bin varies.

A naïve analysis might infer from Figs. 1b and 2 that because PDFs and rain-rate percentiles are very close between SSM/I, SSMIS, and matched-swath-width data, one could analyze for a precipitation cluster trends through the entire time series without consideration of instrumentation differences. However, it turns out to be important to break out each satellite time series separately, using the overlap in time periods to explicitly test whether the calibration holds at this demanding level of accuracy. In these overlap periods (subsets of 2004–09 for the various satellites in Fig. 3), the SSM/I and SSMIS time series exhibit mismatches that indicate that the precision of intercalibration needed to construct a continuous SSM/I–SSMIS time series of probability in these upper bins does not appear to be met by the current intercalibration. It may also be noted that the swath-width-matching procedure appears to have only a modest impact in connecting the SSM/I–SSMIS time series. When trends (i.e., the slopes of the least squares best-fit lines) in the cluster probabilities in the four highest bins 17–20 are computed over the entire time series without consideration of instrumentation differences, each gives a statistically significant positive trend (by the Student's *t* test at the 95% level). Thus, a naïve blending of the satellite series might appear to give a positive trend. However, given the apparent calibration jump between the SSM/I and SSMIS time series for bins 17–19, this serves as a cautionary statement on the necessity of carefully examining the calibration. Probability trends for cluster probabilities for power within

the highest bins 17–19 are computed separately for each of the SSM/I and SSMIS platforms; for the short time series for the separate platforms, these trends fail to pass the Student's *t* test at the 95% level. For the highest power bin (bin 20), there is less calibration jump between the SSM/I and SSMIS time series. However, one should be duly cautious regarding the statistically significant increase in probability in this bin, given the calibration issues noted in the others. A more careful intercalibration of these satellites could potentially permit trend detection for cluster probabilities.

In Fig. 4 we also check for an upward trend (i.e., positive slope of the least squares best-fit line) in the probability of high power storm clusters in recent climate by analyzing time series plots of probability densities from four of the five highest power bins using NCEP Reanalysis 2 model output. These bins correspond to cluster power intervals of 1.37–2.19×10^5 (bin 10), 2.19–3.51×10^5 (bin 11), 3.51–5.58×10^5 (bin 12), and $5.58 \times 8.89 \times 10^5$ GW (bin 13). The time series from the highest power intervals (bin 14) is excluded because of insufficient storm counts over the observed period. The upward trend line in bin 10 (Fig. 4a), of approximately 15% increase over the 26-yr period passes the Student's *t* test at the 95% level. However, the trends in bins 11–13 do not pass this test at the 95% level, possibly associated with the fewer counts per bin in these higher bins. Additionally, the cluster power interval associated with bin 13 (5.58–8.89×10^5 GW) lies beyond the highest cluster power interval seen in SSM/I–SSMIS observations (2.62–4.17×10^5 GW). For the signature seen in Part I in end-of-century runs, the expectation would be for consistency of trend among the high power bins: that is, consistency of the upward trend displayed in bin 10 with a corresponding increase in the probability of storm clusters in bins 11–13. While results in Figs. 3 and 4 do not preclude trends in extreme precipitation being detectable in other measures, this analysis does indicate difficulty in detecting significant trends for changes in probability density of high power storm clusters in recent climate in these datasets.

b. Cluster power distributions and probability trends for high power clusters in CMIP5 models

In Part I, historical and future cluster power distributions from a high-resolution atmosphere model at two horizontal resolutions (HiRAM-C360 and HiRAM-C180) are analyzed. When compared to observed (e.g., TRMM 3B42) cluster power statistics, cluster power distributions at both resolutions compare favorably with each other and observations, with a similar scale-free, power-law range cutoff near 10^5 GW and sharp decrease in the probability of the most intense storm clusters
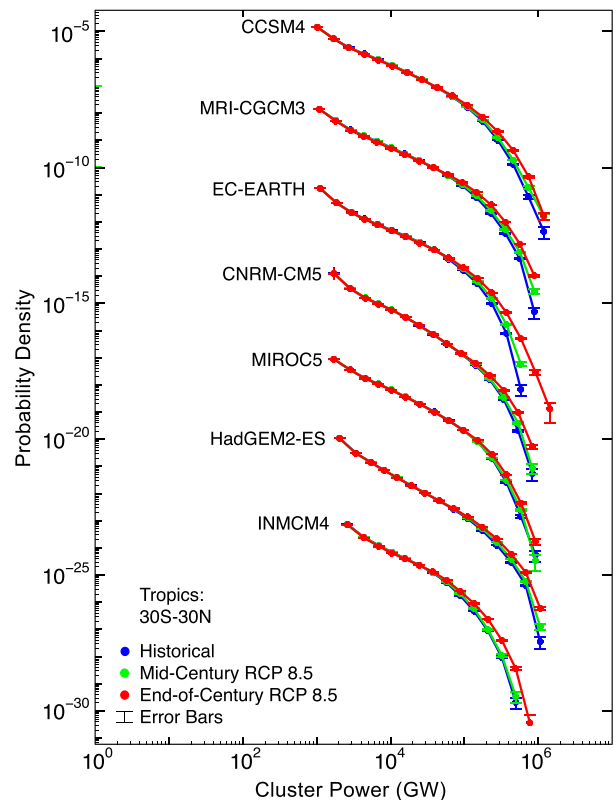

FIG. 5. Historical, midcentury, and end-of-century cluster power PDFs from seven high-resolution CMIP5 models for the $0.7 \, \text{mm h}^{-1}$ rain-rate threshold. For readability, all cluster power probability distributions except for CCSM4 have been shifted vertically down by 3 decades between each model. Relative probability values on the log-*y* axis are unchanged by this procedure; absolute values may be read on the *y* axis by increasing the exponent by 3 for each successive model (e.g., read 10^{-8} as 10^{-5} for MIROC5). Historical period: 1979–2005; midcentury RCP8.5 period: 2026–45 (CNRM-CM5, CCSM4, HadGEM2-ES, and INM-CM4.0) and 2025–50 (MRI-CGCM3, EC-EARTH, and MIROC5); and end-of-century RCP8.5 period: 2081–2100 (CNRM-CM5, CCSM4, HadGEM2-ES, and INM-CM4.0) and 2075–2100 (MRI-CGCM3, EC-EARTH, and MIROC5).

thereafter. The behavior of the HiRAM cluster power distributions shows little resolution dependence, but, in the range of the highest cluster power bins beyond the cutoff, the minimum rain-rate threshold has to be greater than about $0.3 \, \text{mm h}^{-1}$ to yield robust results consistent with observations. Without attempting to examine when a signal emerges, we also find that the probability of storm clusters from the high cluster power range increases, relative to recent climate, by a factor that increases with increasing cluster power. This increase can exceed a factor of 10 for the highest cluster power interval that can be adequately resolved in current climate. In Figs. 5–8, we investigate cluster power behavior from a suite of seven high-resolution CMIP5
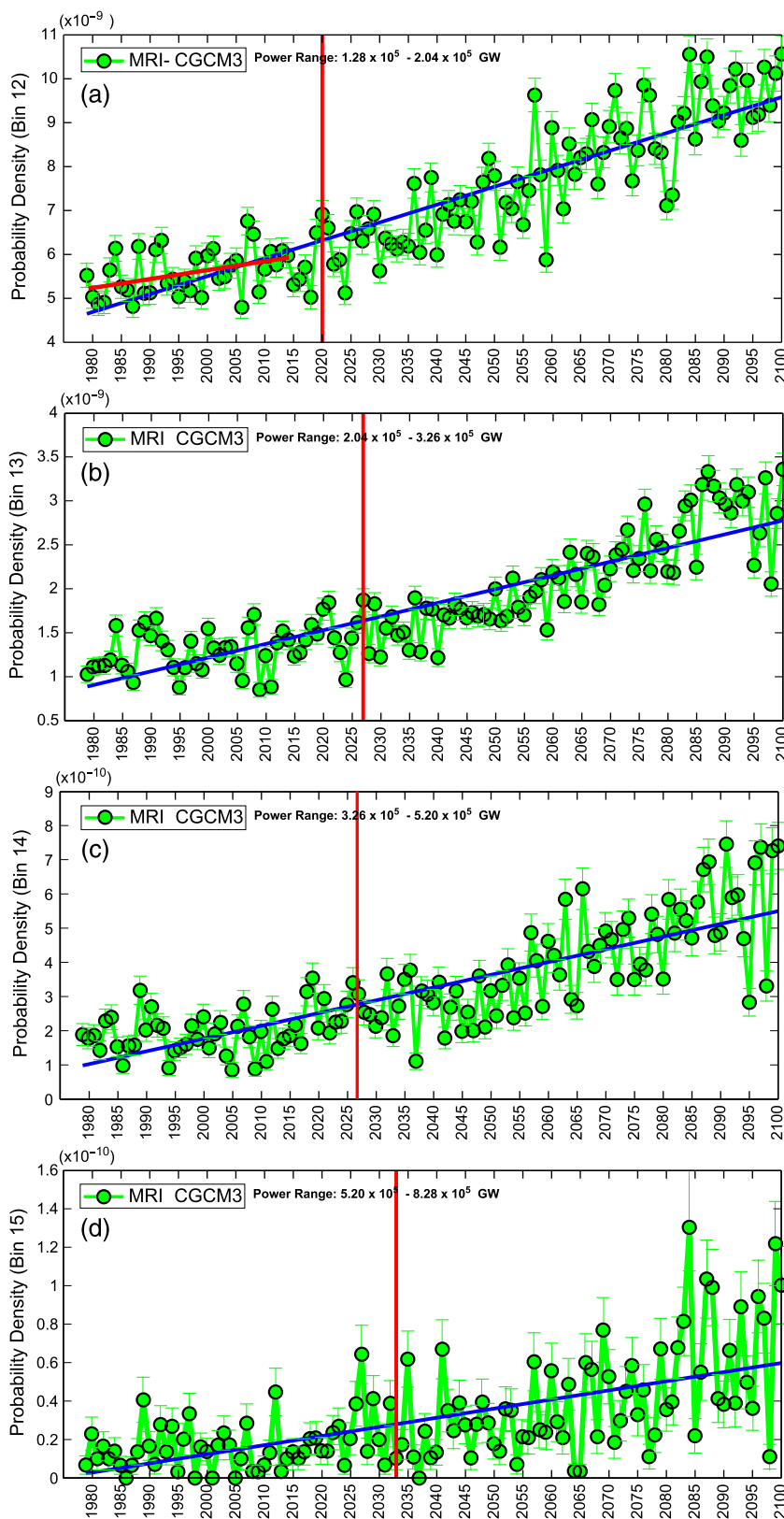
FIG. 6. As in Fig. 4, but for MRI-CGCM3 modeled precipitation clusters for 1979–2100. Recent historical trend (if significant) is shown in red, and year where trend becomes significant with all remaining years significant is depicted with a vertical red line. Trend from 1979–2100 is shown in blue.
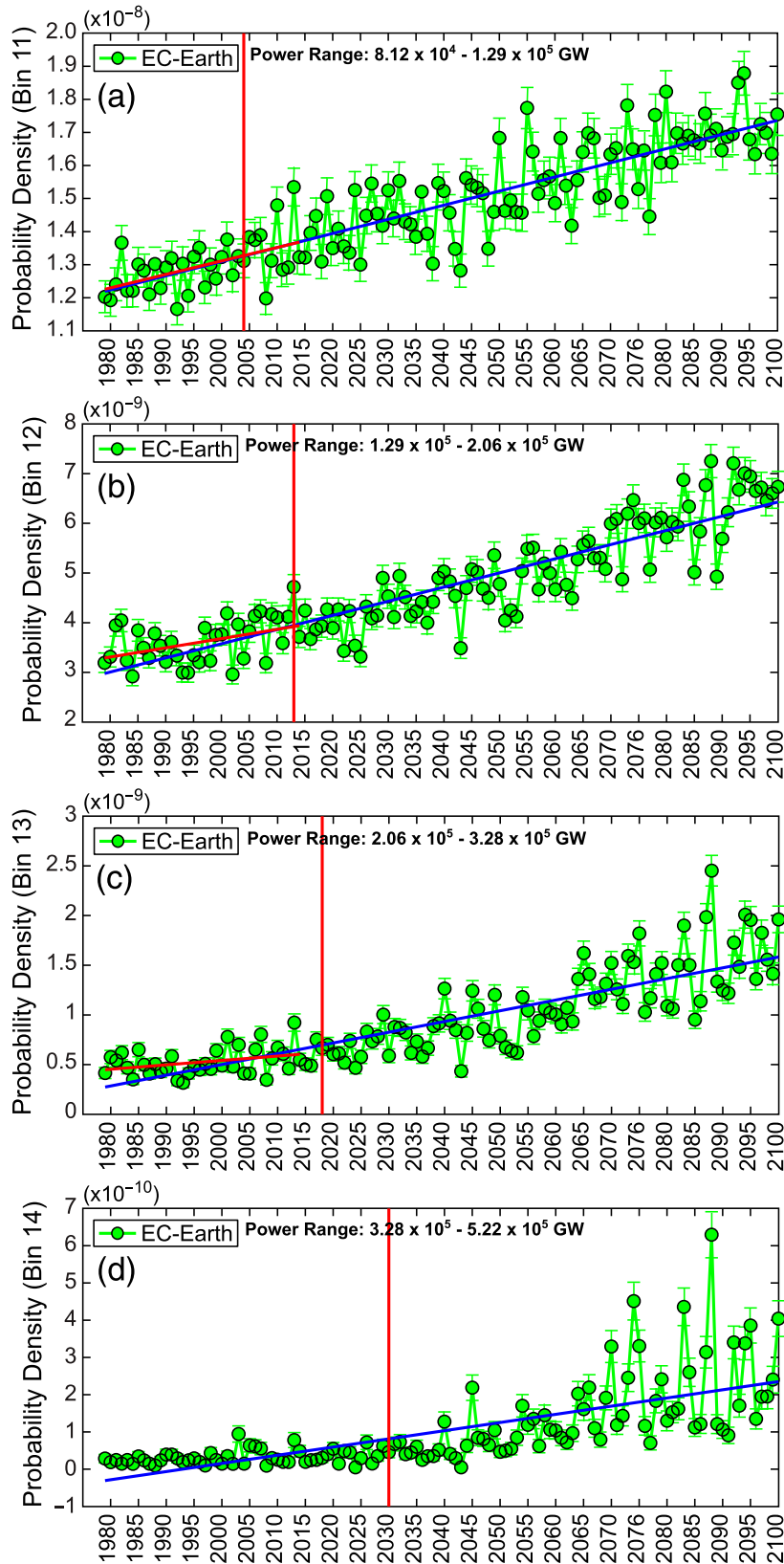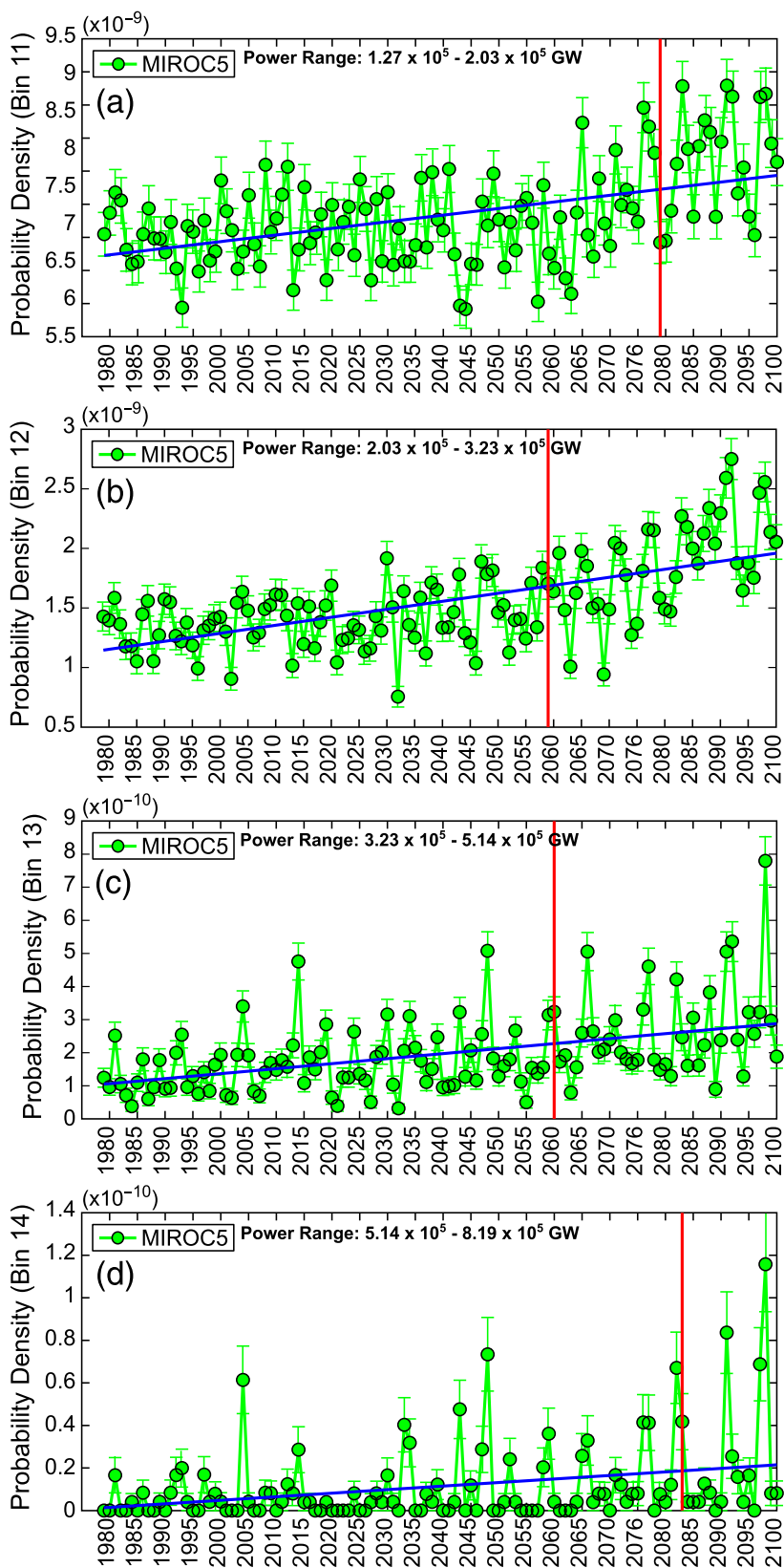
FIG. 7. As in Fig. 6, but for EC-EARTH model.

FIG. 8. As in Fig. 6, but for MIROC5 model.

models under the RCP8.5 scenario for recent, mid-century, and end-of-century climate. Figure 5 displays cluster power distributions from the model ensemble, from which corresponding changes to the frequency of the highest power storm clusters by mid- and end of century, relative to recent climate, can be inferred. Figure 5 is based on the binning procedure described in section 2 that ensures relative error bars for all models do not exceed a specified value during the historical period (see Table S2). To demonstrate robustness, Figure S1 and Table S3 in the SI show the corresponding results when this criterion is not enforced, exhibiting highly consistent results. Figures 6–8 examine when a statistically significant signal emerges for the three CMIP5 models in our study that have a continuous time series of data available for 1979–2100. We set 1979 as the start date for our model analysis for consistency with the NCEP Reanalysis 2 discussed in section 3. The cluster power intervals examined in these models are chosen to cover the power range above the cutoff.

Cluster power distributions in Fig. 5 indicate that our ensemble of CMIP5 models generally have similar scale-free, power-law ranges, cutoffs, and sharp decreases thereafter as observations and HiRAM model output. There is a substantial increase in the frequency of clusters (up to an order of magnitude) with power in the highest bins by end of century, also similar to HiRAM projections. Changes by midcentury are less clear. Three CMIP5 models (CNRM-CM5, INM-CM4.0, and MIROC5) exhibit little to no change in cluster power behavior by midcentury, while the other models (HadGEM2-ES, CCSM4, EC-EARTH, and MRI-CGCM3) are more consistent with midcentury projections displayed by the HiRAM simulations in Part I.

A number of methods have been used to assess the timing of detectability of statistically significant changes under global warming. Mahlstein et al. (2011) consider statistically significant differences between moving 30-yr windows, while Hawkins and Sutton (2012) estimate the time of emergence using signal-to-noise ratios and linear regression. Thompson et al. (2015) develop analytic expressions for time of emergence based on the Student's t test for significance of a linear regression trend (von Storch and Zwiers 1999). Here we use a similar linear regression test, applied numerically, to detect when a statistically significant trend in the probability increases for large cluster sizes emerges. This is done for the three models in this suite of CMIP5 models with a complete time series through the end of the century of the 3-hourly precipitation from which these clusters are computed. A least squares regression is carried out from 1979 to each year between 1984 and the end of the century. These are then evaluated for the first end year

for which (i) the regression fit passes the Student's t test at the 95% level, and (ii) regression lines (from 1979) extending to all subsequent end years pass the Student's t test at the same significance level. This year is used as a measure of when the trend emerges relative to interannual variability. This is done for a series of cluster power intervals corresponding to the uppermost bins in Table S3 that exceed an average of four counts per year in the historical period. These provide a convenient means of describing the evolution in the cluster power range above the cutoff where the most substantial changes are noted in Fig. 5. We describe the trends in detail for each cluster power interval in each model but draw conclusions based on the overall behavior.

The three models with continuous data through the end of the century (Figs. 6–8) show considerable spread as to when changes to the probability of the most intense storms become detectable. For the MRI-CGCM3 model (Fig. 6), we examine trends for cluster power intervals 1.28–2.04×10^5, 2.04–3.26×10^5, 3.26–5.20×10^5, and 5.20–8.28×10^5 GW. Similar to section 3a, we refer to these power intervals as bins 12–15, which correspond to those given in Table S3. A statistically significant upward trend (by the criteria above, which include statistical significance in all subsequent years) emerges by 2020 for bin 12 and 2033 for bin 13. The power interval for bin 12, 1.28–2.04×10^5 GW, is approximately equivalent to the interval 1.37–2.19×10^5 GW (bin 10) of the NCEP Reanalysis 2 dataset that shows a similar upward trend that is statistically significant by 2014. Storm clusters having power falling within bins 12–14 become from 2 (bin 12) to 4 times (bin 14) more frequent by end of century, relative to their historical mean values (1979–2014), and the frequency of storms from the highest power bin (bin 15) increases by a factor of 7.

For the EC-EARTH model (Fig. 7), probability densities for cluster power intervals 8.12×10^4–1.29×10^5 (bin 11), 1.29–2.06×10^5 (bin 12), 2.06–3.28×10^5 (bin 13), and 3.28–5.22×10^5 GW (bin 14) are shown. Statistically significant upward trends emerge by 2004 (bin 11), 2013 (bin 12), and 2018 (bin 13). The cluster power interval for bin 12 in the EC-EARTH model is approximately equivalent to cluster power intervals in bin 12 (MRI-CGCM3) and bin 10 (NCEP Reanalysis 2), and, notably, statistically significant frequency increases of storms within this power interval are shown in all three datasets. In the highest power bin shown (bin 14), an upward trend is not significant until 2030. Modeled EC-EARTH storms with their power falling within bins 11–13 become 1.4–4 times more frequent by end of century, respectively, relative to their historical mean values. The power intervals associated with bins 12 and 13 from EC-EARTH approximate power intervals associated

with bins 12 and 13 from MRI-CGCM3, and the magnitude of increase for both models within these power intervals is the same (roughly a factor of 2–4). Storm clusters from the highest power interval (bin 14) become 13 times more frequent in the EC-EARTH model, roughly the same power interval as MRI-CGCM3 bin 14, but almost twice as frequent. The end-of-century frequency increases for all MRI-CGCM3 and EC-EARTH bins shown in Figs. 6 and 7 fall above one standard deviation around their historical mean values, implying that the end-of-century increase in the frequency of high power storm clusters is unprecedented compared to recent climate.

The frequency of the most powerful storm clusters from the MIROC5 model (Fig. 8) show much greater interannual variability, compared to the EC-EARTH and MRI-CGCM3 output. The upward trends for MIROC5 for all of the power intervals shown do not become statistically significant until after 2060, possibly in part due to the large degree of interannual variability. The relative magnitude of the increased probability of powerful storms in the MIROC5 model only ranges from a factor of 1.16–1.51, respectively, less than the increases projected in the MRI-CGCM3 and EC-EARTH models in bins of comparable power. Additionally, only the trends from MIROC5 bins 11 and 12 (power interval equivalent to bins 12 and 13 in both MRI-CGCM3 and EC-EARTH models), not bins 13 and 14, fall above one standard deviation, relative to their historical mean.

4. Summary and discussion

Here we examine distributions of storm cluster power over the tropics using satellite rain-rate retrievals (SSM/I–SSMIS and TRMM 3B42) and NCEP Reanalysis 2 precipitation data, with SSM/I–SSMIS and TRMM 3B42 distributions, in general, closely paralleling each other. Specifically, SSM/I–SSMIS cluster power distributions show little variation by rain-rate threshold and have the same scale-free power-law region and cutoff at high cluster power as TRMM 3B42, indicating that swath width does not affect the cutoff. NCEP Reanalysis 2 cluster power distributions begin at higher power as a result of coarser resolution and cut off at slightly higher power than observations but, even so, approximate the same form as the observed cluster power distributions.

Since the SSM/I–SSMIS data record extends back to 1987, we also analyze SSM/I–SSMIS rain-rate data for changes to storm cluster power behavior in recent climate. SSMIS swaths are approximately 300-km wider than SSM/I swaths, so, to test cluster power statistics for sensitivity to swath width, we run a procedure where we

narrow SSMIS swaths to match SSM/I swath width at the $0.7\,\mathrm{mm\,h^{-1}}$ rain rate. The swath width calibration has only a modest impact, as SSM/I and matched-swath-width SSMIS rain rate percentiles are highly calibrated (e.g., the correlation coefficient between retrievals equals 0.9987) and have a least squares best-fit line that is nearly one to one. Despite the high degree of calibration, the probabilities of the most intense storm clusters from the four highest power bins broken out into time series by satellite display a discontinuity between SSM/I, SSMIS, and matched-swath-width SSMIS retrievals, such that, for the level of difference we are trying to detect here, the current SSM/I–SSMIS calibration is insufficient. Some of the highest power bins in both the SSM/I–SSMIS and NCEP Reanalysis 2 datasets do show an increase in the probability of the most powerful storm clusters in recent climate, but, given the instrumentation, these trends may not be reliable. Improvements to SSM/I–SSMIS rain-rate retrieval intercalibrations and overlapping satellite instrumentation would be useful for detecting trends in cluster power behavior within the next couple of decades. Any gap in satellite coverage, especially with consistent instrumentation, potentially hinders such efforts. Given the termination of the DMSP program, canceled launch of the F20 satellite, degradation of data from the *F17* satellite, and failure of the *F19* satellite, data record continuity may be an important consideration.

Cluster power distributions from our ensemble of CMIP5 models begin at higher power than our observed datasets and thus have shorter scale-free regions. Even so, all of the CMIP5 models reasonably reproduce the scale-free region with approximately the same exponent and cutoff as in observations here and in Part I. By end of century, we see a consistent trend toward more frequent high power storm clusters across all models. These increases occur above the cutoff and are consistent with an extension of the cutoff toward higher power, as occurs for probability distributions of precipitation accumulations (Neelin et al. 2017), which results in increases in probability density relative to the historical period that rise roughly exponentially across this range. For some models (MRI-CGCM3, EC-EARTH, HadGEM2-ES, and INM-CM4.0), the probability of storm clusters in the highest power bin for which statistics can be reliably computed in the historical period increase by up to an order of magnitude, relative to historical climate, matching HiRAM trends noted in Part I. For the other models (CCSM4, CNRM-CM5, and MIROC5), end-of-century changes are likewise substantial, with the probability of the most powerful storm clusters increasing by a factor of 3–7. These numbers typify the changes in the highest cluster power interval with

sufficient statistics in a 27-yr sample from historical climate: end-of-century simulations also exhibit events of magnitude unprecedented in the historical period, consistent with the extension of the cutoff.

Pendergrass et al. (2016) use an idealized suite of global radiative–convective equilibrium simulations to examine convective organization as a function of specified SST. They find a sharp shift to larger, more organized clusters under warmer conditions (global SST >303 K), with circulation changes as the primary driver. Although the idealized, fixed-SST configuration differs considerably from the realistic simulations considered here, one can ask whether the cluster power increases found here emerge gradually or exhibit threshold behavior, as well as when trends emerge from natural variability.

Given that the CMIP5 models analyzed agree on qualitative behavior of end-of-century changes, we conduct a time series analysis on the three models in our suite for which continuous data are available from the historical period through the end of the century, to investigate when statistically significant increases in the probability of the most intense storms appear. Trends in probability are examined for intervals of cluster power above the cutoff. In all three models, over these cluster power intervals, trends tend to emerge gradually without strong evidence of a rapid change that might be associated with a threshold, although in the uppermost interval examined it would be difficult to exclude nonlinearity because of lower sampling in the historical period. While relative increases are larger for higher cluster power, statistical significance of probability trends tends to emerge earlier for intervals that are not too far above the cutoff, because of the smaller sample size typically available at higher power. One of these three models (EC-EARTH) exhibits trends that are statistically significant prior to 2020 in three of four power intervals examined (with cluster power ranging from 0.5–5×10^5 GW); one model (MRI-CGCM3) shows similar emergence by about 2030, even for these metrics that break down the probability trends as a function of storm size. On the other hand, the third model (MIROC5) does not display a detectable signal in these metrics until after 2050, likely due to much larger interannual variability. Having more high-resolution CMIP models with continuous data records extending through the end of the century would help narrow the uncertainty surrounding the timing of detectability in these measures.

REFERENCES

Adler, R. F., C. Kidd, G. Petty, M. Morissey, and H. M. Goodman, 2001: Intercomparison of global precipitation products: The third Precipitation Intercomparison Project (PIP-3). *Bull. Amer. Meteor. Soc.*, **82**, 1377–1396, doi:10.1175/1520-0477(2001)082<1377:IOGPPT>2.3.CO;2.

Alexander, L. V., and Coauthors, 2006: Global observed changes in daily extremes of temperature and precipitation. *J. Geophys. Res.*, **111**, D05109, doi:10.1029/2005JD006290.

Allan, R. P., and B. J. Soden, 2008: Atmospheric warming and the amplification of precipitation extremes. *Science*, **321**, 1481–1484, doi:10.1126/science.1160787.

——, ——, V. O. John, W. Ingram, and P. Good, 2010: Current changes in tropical precipitation. *Environ. Res. Lett.*, **5**, 025205, doi:10.1088/1748-9326/5/2/025205.

Bowman, K. P., C. R. Homeyer, and D. G. Stone, 2009: A comparison of oceanic precipitation estimates in the tropics and subtropics. *J. Appl. Meteor. Climatol.*, **48**, 1335–1344, doi:10.1175/2009JAMC2149.1.

Chen, Y., E. E. Ebert, K. J. E. Walsh, and N. E. Davidson, 2013: Evaluation of TRMM 3B42 precipitation estimates of tropical cyclone rainfall using PACRAIN data. *J. Geophys. Res. Atmos.*, **118**, 2184–2196, doi:10.1002/jgrd.50250.

Easterling, D. R., G. A. Meehl, C. Parmesan, S. A. Chagnon, T. R. Karl, and L. O. Mearns, 2000: Climate extremes: Observations, modeling, and impacts. *Science*, **289**, 2068–2074, doi:10.1126/science.289.5487.2068.

Hawkins, E., and R. Sutton, 2012: Time of emergence of climate signals. *Geophys. Res. Lett.*, **39**, L01702, doi:10.1029/2011GL050087.

Hilburn, K. A., and F. J. Wentz, 2008: Mitigating the impact of RADCAL beacon contamination on F15 SSM/I ocean retrievals. *Geophys. Res. Lett.*, **35**, L18806, doi:10.1029/2008GL034914.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55, doi:10.1175/JHM560.1.

—— and D. T. Blovin, 2014: TRMM and other data precipitation data set documentation. NASA Tech. Rep., 44 pp. [Available online at ftp://meso-a.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf.].

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643, doi:10.1175/BAMS-83-11-1631.

Kharin, V. V., F. W. Zwiers, X. Zhang, and G. C. Hegerl, 2007: Changes in temperature and precipitation extremes in the IPCC Ensemble of Global Coupled Model Simulations. *J. Climate*, **20**, 1419–1444, doi:10.1175/JCLI4066.1.

——, ——, ——, and M. F. Wehner, 2013: Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, **119**, 345–357, doi:10.1007/s10584-013-0705-8.

Lenderink, G., and E. Van Meijgaard, 2008: Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nat. Geosci.*, **1**, 511–514, doi:10.1038/ngeo262.

Mahlstein, I., R. Knutti, S. Solomon, and R. W. Portmann, 2011: Early onset of significant local warming in low latitude countries. *Environ. Res. Lett.*, **6**, 034009, doi:10.1088/1748-9326/6/3/034009.

Mapes, B., R. Milliff, and J. Morzel, 2009: Composite life cycle of maritime tropical mesoscale convective systems in scatterometer and microwave satellite observations. *J. Atmos. Sci.*, **66**, 199–208, doi:10.1175/2008JAS2746.1.

McCollum, J. R., and R. R. Ferraro, 2003: Next generation of NOAA/NESDIS TMI, SSM/I, and AMSR-E microwave land rainfall retrievals. *J. Geophys. Res.*, **108**, 8382, doi:10.1029/2001JD001512.

NSIDC, 2016a: Special Sensor Microwave Imager (SSM/I). Accessed 1 February 2016. [Available online at https://nsidc.org/data/docs/daac/ssmi_instrument.gd.html.]

——, 2016b: Special Sensor Microwave Imager/Sounder (SSMIS). Accessed 1 February 2016. [Available online at https://nsidc.org/data/docs/daac/ssmis_instrument/.]

Neelin, J. D., S. Sahany, S. N. Stechmann, and D. Bernstein, 2017: Global warming precipitation accumulation increases above the current-climate cutoff scale. *Proc. Natl. Acad. Sci. USA*, **114**, 61 258–61 263, doi:10.1073/pnas.1615333114.

Pendergrass, A. G., K. A. Reed, and B. Medeiros, 2016: The link between extreme precipitation and convective organization in a warming climate: Global radiative-convective equilibrium simulations. *Geophys. Res. Lett.*, **43**, 11 445–11 452, doi:10.1002/2016GL071285.

Peters, O., J. D. Neelin, and S. W. Nesbitt, 2009: Mesoscale convective systems and critical clusters. *J. Atmos. Sci.*, **66**, 2913–2924, doi:10.1175/2008JAS2761.1.

——, A. DeLuca, A. Corral, J. D. Neelin, and C. E. Holloway, 2010: Universality of rain event size distributions. *J. Stat. Mech.*, **11**, P11030, doi:10.1088/1742-5468/2010/11/P11030.

——, K. Christensen, and J. D. Neelin, 2012: Rainfall and dragon-kings. *Eur. Phys. J. Spec. Top.*, **205**, 147–158, doi:10.1140/epjst/e2012-01567-5.

Quinn, K., and J. D. Neelin, 2017: Distributions of tropical precipitation cluster power and their changes under global warming. Part I: Observational baseline and comparison to a high-resolution atmospheric model. *J. Climate*, **30** (20), 8045–8059, doi:10.1175/JCLI-D-16-0683.1.

Sillmann, J., V. V. Kharin, F. W. Zwiers, X. Zhang, and D. Bronaugh, 2013: Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *J. Geophys. Res. Atmos.*, **118**, 2473–2493, doi:10.1002/jgrd.50188.

Skok, G., J. Bacmeister, and J. Tribbia, 2013: Analysis of tropical cyclone precipitation using an object-based algorithm. *J. Climate*, **26**, 2563–2579, doi:10.1175/JCLI-D-12-00135.1.

Sun, N., and F. Weng, 2008: Evaluation of Special Sensor Microwave Imager/Sounder (SSMIS) environmental data records. *IEEE Trans. Geosci. Remote Sens.*, **46**, 1006–1016, doi:10.1109/TGRS.2008.917368.

Tan, J., C. Jakob, W. B. Rossow, and G. Tselioudis, 2015: Increases in tropical rainfall driven by changes in frequency of organized deep convection. *Nature*, **519**, 451–460, doi:10.1038/nature14339.

Tebaldi, C., K. Hayhoe, J. M. Arblaster, and G. A. Meehl, 2006: Going to the extremes: An intercomparison of model-simulated historical and future changes in extreme events. *Climatic Change*, **79**, 185–211, doi:10.1007/s10584-006-9051-4.

Thompson, D. W. J., E. A. Barnes, C. Deser, W. E. Foust, and A. S. Phillips, 2015: Quantifying the role of internal climate variability in future climate trends. *J. Climate*, **28**, 6443–6456, doi:10.1175/JCLI-D-14-00830.1.

TRMM, 2015: TRMM_3B42: TRMM (TMPA) rainfall estimate L3 3 hour 0.25 degree × 0.25 degree, version 7. Goddard Earth Sciences Data and Information Services Center (GES DISC), accessed 1 July 2015, https://disc.gsfc.nasa.gov/datacollection/TRMM_3B42_7.html.

von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.

Wentz, F. J., 2013: SSM/I version-7 calibration report. Remote Sensing Systems Tech. Rep. 011012, 46 pp. [Available online at http://images.remss.com/papers/rsstech/2012_011012_Wentz_Version-7_SSMI_Calibration.pdf.]

——, and R. W. Spencer, 1998: SSM/I rain retrievals within a unified all-weather ocean algorithm. *J. Atmos. Sci.*, **55**, 1613–1627, doi:10.1175/1520-0469(1998)055<1613:SIRRWA>2.0.CO;2.

——, K. A. Hilburn, and D. K. Smith, 2012: Remote Sensing Systems DMSP SSM/I and SSMIS daily environmental suite on 0.25 deg grid, version 7. Remote Sensing Systems, accessed 1 February 2016. [Available online at www.remss.com/missions/ssmi.]

Wood, R., and P. R. Field, 2011: The distribution of cloud horizontal sizes. *J. Climate*, **24**, 4800–4816, doi:10.1175/2011JCLI4056.1.

Yan, B., and F. Weng, 2008: Intercalibration between special sensor microwave/imager sounder and special sensor microwave imager. *IEEE Trans. Geosci. Remote Sens.*, **46**, 984–995, doi:10.1109/TGRS.2008.915752.