Statistical Modeling of Extreme Precipitation with TRMM Data

LEVON DEMIRDJIAN

Department of Statistics, University of California, Los Angeles, Los Angeles, California

YAPING ZHOU

Goddard Earth Sciences Technology and Research, Morgan State University, Baltimore, and Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, Maryland

GEORGE J. HUFFMAN

Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 27 January 2017, in final form 19 September 2017)

ABSTRACT

This paper improves upon an existing extreme precipitation monitoring system that is based on the Tropical Rainfall Measuring Mission (TRMM) daily product (3B42) using new statistical models. The proposed system utilizes a regional modeling approach in which data from similar locations are pooled to increase the quality of the resulting model parameter estimates to compensate for the short data record. The regional analysis is divided into two stages. First, the region defined by the TRMM measurements is partitioned into approximately 28 000 nonoverlapping clusters using a recursive *k*-means clustering scheme. Next, a statistical model is used characterize the extreme precipitation events occurring in each cluster. Instead of applying the block maxima approach used in the existing system, in which the generalized extreme value probability distribution is fit to the annual precipitation maxima at each site separately, the present work adopts the peak-over-threshold method of classifying points as extreme if they exceed a prespecified threshold. Theoretical considerations motivate using the point process framework for modeling extremes. The fitted parameters are used to estimate trends and to construct simple and intuitive average recurrence interval (ARI) maps that reveal how rare a particular precipitation event is. This information could be used by policy makers for disaster monitoring and prevention. The new method eliminates much of the noise that was produced by the existing models because of a short data record, producing more reasonable ARI maps when compared with NOAA's long-term Climate Prediction Center ground-based observations. Furthermore, the proposed method can be applied to other extreme climate records.

1. Introduction

The effective monitoring and measurement of extreme precipitation events form an integral component for understanding the underlying nature of extreme climate phenomena and are crucial for evaluating future changes and impacts of precipitation extremes. Many recent studies have found a marked increase in the frequency and intensity of extreme precipitation events occurring in the last few decades (Donat et al. 2016; Min et al. 2011; Alexander et al. 2006). Changes in the behavior of extreme precipitation phenomena are among the most important aspects of global climate change, with significant implications for human society and the environment. For example, a study of the spatial heterogeneity of such changes found that regions where high-intensity precipitation is less common are especially prone to increases in precipitation totals and extremes (Donat et al. 2016); unfortunately, the infrastructure in these regions is particularly ill adapted to deal with extreme precipitation. A rise in the frequency and severity of extreme climate events also exacts a large human and economic toll. For example, in October 2013, Typhoon Fitow led to record winds and flooding throughout eastern China, shutting down roadways, schools, and hospitals, and resulting in an estimated $10 billion USD in total damages (Typhoon Committee 2013). In mid-August 2016, a storm system in southern Louisiana resulted in unprecedented precipitation and flooding, with some areas receiving in excess of 280 mm of rain in a single day.

Corresponding author: Levon Demirdjian, levondem@ucla.edu

DOI: 10.1175/JAMC-D-17-0023.1

The storm, which brought roughly 3 times as much rain over Louisiana as Hurricane Katrina did in 2005, was later described as being an event occurring with 0.2% probability in any given year (Di Liberto 2016). More recently, in October 2016 Hurricane Matthew ravaged the western Atlantic Ocean, causing widespread power outages and flooding and over $8 billion in total damage. Hurricane Matthew led to the deaths of over 500 people in Haiti alone, and was the strongest storm to hit the country in over 50 yr.

Satellite-based retrieval algorithms based on the measurements made by the Tropical Rainfall Measuring Mission (TRMM) and the more recent Global Precipitation Measurement (GPM) satellites have provided a rich source of precipitation data at the global scale. The TRMM Multisatellite Precipitation Analysis (TMPA; Huffman et al. 2007) combines precipitation estimates from a variety of satellite systems to provide estimates at fine scales (3 hourly, 0.25° × 0.25°) with quasi-global coverage (50°S–50°N); moreover, TMPA estimates are available in both real-time (3B42-RT) and post-real-time (3B42) data products.

One of the most common approaches for modeling extreme values of hydrological variables is to adopt the framework of statistical extreme value theory, where precipitation intensities are assumed to be random draws from an underlying probability distribution, and characterizing extreme value behavior is equivalent to characterizing the upper tail of this distribution (Leadbetter et al. 1983; Katz et al. 2002; Shane and Lynn 1964; Chan et al. 2014). Although physical models can quite accurately describe the processes generating precipitation, from a probabilistic point of view, the true data-generating process producing precipitation intensities is almost never known in practice. Thus, one typically uses a set of data to select a distribution from a prespecified family of distributions that describe the tail behavior. To translate the estimates of the fitted model parameters to terms easily understood by policy makers and the general public, one can construct average recurrence intervals (ARIs) that describe the rarity of precipitation events. For example, a precipitation event with an ARI of 10 yr means that it occurs on average once every 10 yr. The amount of precipitation corresponding to the 10-yr ARI is referred to as the 10-yr return level. Note that a 10-yr ARI does not mean that the event will occur once every 10 yr; it simply means that in any given year, there is a 10% probability of such an event occurring, and that the occurrence of the event in one year does not preclude it from occurring in another year.

Extreme value distributions (EVDs) like the generalized extreme value (GEV) and generalized Pareto (GP) distributions have commonly been used for the modeling of precipitation and temperature extremes. EVDs have been used to analyze trends and changes in daily temperature (Brown et al. 2008), to project changes in seasonal precipitation extremes using ensembles of climate models (Kharin et al. 2007; Fowler and Ekström 2009), and to study the spatial and spatiotemporal behavior of extreme precipitation (Wang et al. 2017; Schindler et al. 2012). Serinaldi and Kilsby (2014) used the GP distribution to model precipitation extremes, focusing specifically on the impact of threshold selection on the tail behavior of the fitted GP distributions. Using a point process model, Heaton et al. (2011) discovered significant increases in the intensity of extreme weather in parts of the contiguous United States (CONUS). Schindler et al. (2012) modeled extreme precipitation across the United Kingdom using an inhomogeneous Poisson point process, accounting for annual cycles using a sinusoidal model for the location and scale parameters of the corresponding GEV distribution. The point process approach to extreme value analysis has also been used to detect trends in ozone levels (Smith 1989), as well as to generate stochastic climate scenarios to facilitate the modeling of precipitation extremes (Furrer and Katz 2008).

The extreme precipitation monitoring system proposed in Zhou et al. (2015) uses measurements taken from the TMPA data series to construct ARI maps for the purpose of disaster preparation and monitoring. While the TRMM extreme precipitation monitoring system is a highly effective framework in general, the statistical modeling of the system Zhou et al. (2015) used suffers from several limitations. First, data from each of the grid points in the TMPA domain are considered to be independent, an assumption that is questionable in practice. Second, only the annual maxima values for each grid location are considered to be extreme, meaning that only 16 data points are available for model fitting at each location. As a result, there is a high degree of uncertainty in the parameter estimates and resulting ARI maps. Furthermore, the annual maxima approach cannot accommodate multiple extreme events occurring during the same year, for example, during different seasons.

In this paper, we propose an alternative methodology for the statistical modeling of the TRMM extreme precipitation monitoring system that overcomes the above limitations. In section 2, we outline the two stages of our proposed algorithm, which first partitions the map into disjoint clusters of similar sites and then fits an appropriate statistical model to the pooled data in each cluster. In section 3, we present the results of our methodology when estimating return levels and trends in extreme precipitation and compare the return-level estimates with those in Zhou et al. (2015). Section 4 demonstrates that our procedure is general enough to be used to analyze

extreme climate events other than precipitation; in this case, we analyze surface air temperature data. We conclude with a discussion covering several possible extensions of our work.

2. Methodology

To overcome the above-mentioned shortcomings of the existing TRMM extreme precipitation monitoring system, we implement a two-stage methodology that 1) partitions the map into relatively homogeneous nonoverlapping regions, and 2) fits an appropriate statistical distribution to the data from each of the regions from the first stage. We are not proposing a completely novel methodology for extreme value analysis but rather an alternative framework for modeling the TRMM data that improves upon the methodology of Zhou et al. (2015). All of the results in this paper are based on the TRMM 3B42 daily precipitation record (NASA GES DISC 2016).

a. Regional clustering

The idea of pooling similar sites into one common region has a rich history in the hydrological literature (Cunnane 1989; Hosking et al. 1985; Hosking and Wallis 1988), has also been utilized in precipitation analysis (Buishand 1991), and fits into the broader framework of regional frequency analysis (Hosking and Wallis 1993, 1997).

There are two general approaches for clustering sites in a regional analysis of extreme climate events. In the first approach, regions are clustered based on their site characteristics (e.g., locational and topographic information) and not at-site statistics such as the time series of annual maxima or threshold exceedances (i.e., Smithers and Schulze 2001; Satyanarayana and Srinivas 2008; Wang et al. 2017; Hosking and Wallis 1997). An alternative framework for regional partitioning is to use the data themselves as input into the clustering algorithm. For example, the location similarity measures in Bernard et al. (2013) and Bador et al. (2015) use the time series of annual maxima themselves as variables in the clustering algorithm, the goal being to achieve maximum stability within each cluster. Despite the merits of these clustering methods, there are two potential drawbacks with this approach. First, using the same data both to form the regional clusters and to test for homogeneity within those clusters will almost certainly lead to a biased assessment of homogeneity (Hosking and Wallis 1997). Furthermore, the clustering results will change every time data are added to the model, for example, if data from the GPM "IMERG" data product were to be added to the statistical model.

With these considerations, here we adopt a clustering scheme based on site characteristics using a recursive k-means clustering algorithm with spatial location (longitude, latitude), topography [derived from 5′ National Geophysical Data Center (NGDC) TerrainBase Global Digital Terrain Model (DTM), version 1.0 (Row and Hastings 1994), and binned into 0.25° resolution], and the 90th percentile of precipitation values (all variables standardized) as input to the algorithm. The k-means algorithm seeks to partition the data (here, the map) into k nonoverlapping groups (where the number of clusters k is prespecified) so as to minimize the sum of squared distances from each data point to its assigned cluster's center in feature space. See Hastie et al. (2009) for more details about k-means clustering and its implementation.

Our recursive k-means algorithm first partitions the map into approximately 30 large clusters; each cluster is further partitioned into another set of 30 clusters, resulting in about 900 clusters in total. This process is repeated a final time for each of the resulting regions; if there are fewer than 30 grid points in a particular region, we skip this final step for that region. This entire process yields 28 221 nonoverlapping regions, for an average of about 20 grid points per cluster, which follows the guidelines set forth in Hosking and Wallis (1997). Figure 1 illustrates the idea behind the recursive clustering scheme. Note that the region a given cluster covers need not be contiguous, and one can weight the inputs of the algorithm to adjust their relative importance. The results of our algorithm are displayed in Fig. 2 for the first two clustering operations.

Next, we implement the homogeneity test given in Viglione et al. (2007) that combines the "Hosking and Wallis heterogeneity statistic" (Hosking and Wallis 1997) with the bootstrap Anderson–Darling statistic (Scholz and Stephens 1987) to decide if the distributions of extreme precipitation intensity for different sites within each cluster are the same. Of 28 221 regions, 21 112 were identified as being acceptably homogeneous. We did not correct for multiple testing since the Hosking and Wallis statistic is not a formal test statistic, and therefore the number of heterogeneous regions is almost certainly overestimated. Since regional analysis will produce more accurate statistical estimates than a single-site analysis even with slight or moderate degrees of homogeneity (Hosking and Wallis 1997), we do not expect our results to be greatly affected by the heterogeneity in some clusters.

b. Statistical modeling

The next stage is to fit an appropriate probability distribution to the pooled extreme precipitation data in each resulting cluster. The estimated parameters of the fitted distributions will then characterize the underlying behavior of extreme precipitation events in that region.

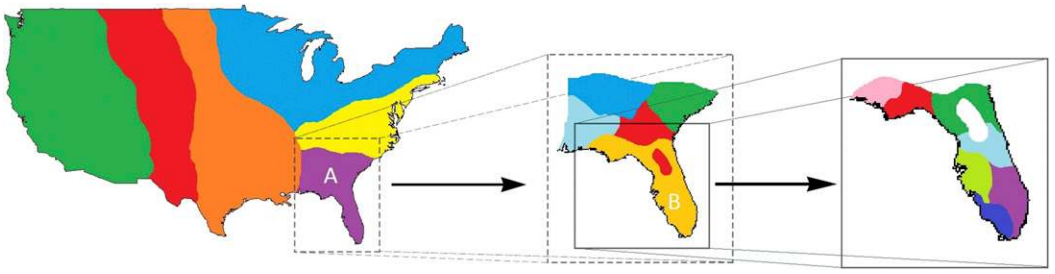

FIG. 1. Illustration of recursive clustering algorithm. In this example, the CONUS is initially clustered into six distinct regions (marked by different colors). Each region is further clustered (e.g., region A is itself partitioned into five clusters); this process is repeated for each resulting cluster (e.g., region B is further partitioned into seven clusters).

1) CHOICE OF AN APPROPRIATE DISTRIBUTION

We begin by reviewing some of the common approaches to extreme value modeling, motivating our choice to adopt the point process (PP) framework to model precipitation extremes.

To model extreme values, Zhou et al. (2015) utilize the block maxima approach where only the largest annual precipitation values are considered to be extreme, and where the GEV distribution is used to model the resulting extreme values. See Leadbetter et al. (1983) for the theoretical justification for using the GEV distribution to model sample maxima. The GEV cumulative distribution function is given by

$$F_{\mathrm{GEV}}(x;\mu,\sigma,\xi)=\begin{cases}\exp\left\{-\left[1+\dfrac{\xi(x-\mu)}{\sigma}\right]^{-\frac{1}{\xi}}\right\} & \text{if} \quad \xi\neq 0\\ \exp\left[-\exp\left(-\dfrac{x-\mu}{\sigma}\right)\right] & \text{if} \quad \xi=0,\end{cases} \tag{1}$$

where μ is the location parameter, $\sigma > 0$ is the scale parameter, and ξ is the shape parameter. Extreme value modeling using block maxima to fit the GEV distribution has widely been used for modeling hydrological extreme data [see, e.g., Katz et al. (2002) and the references therein], but has the obvious limitation that a large number of observations are discarded, resulting in a short data record. One approach for dealing with this limitation of the block maxima approach is to adopt the peak-over-threshold (POT) method, where observations are considered extreme if they exceed a prespecified threshold (Todorovic and Zelenhasic 1970; Davison and Smith 1990). For large enough thresholds, the distribution of threshold exceedances will approximately follow the GP distribution (Leadbetter et al. 1983).

The framework of PP unifies the two approaches discussed above [see Cox and Isham (1980) for the general theory of point processes; some applications to environmental modeling via the PP approach can be found in Smith (1989) and Smith and Shively (1995)]. According to PP theory, the occurrence time and

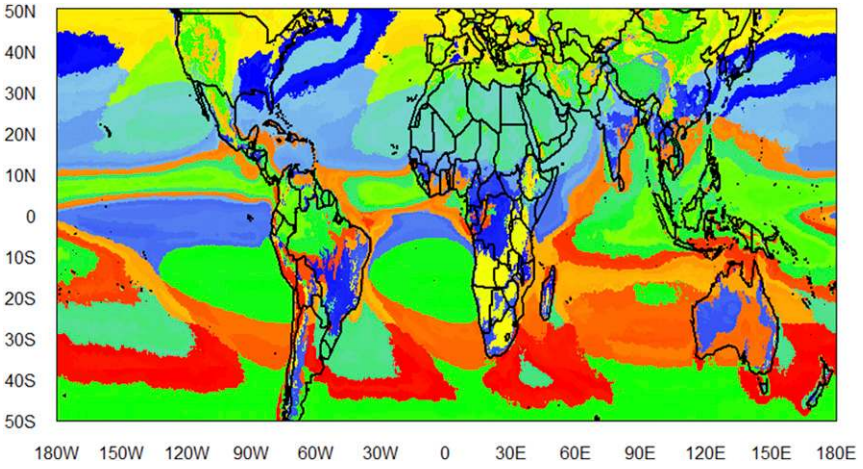

FIG. 2. Results of the clustering algorithm. Each color corresponds to a different cluster. While there are over 28 000 distinct clusters, only those created during the first two stages are depicted.

intensity of an event that exceeds a prespecified threshold will approximately follow a Poisson PP (assuming the threshold is sufficiently large). Moreover, the intensity function of the PP is parameterized by a GEV (μ, σ, ξ) distribution corresponding to the annual maximum distribution of the observed process (Leadbetter et al. 1983; Coles 2001). Using the PP framework offers the advantage that its likelihood is parameterized in terms of the GEV parameters in (1), since these parameters are invariant to the choice of threshold. Furthermore, this parameterization allows nonstationarity to easily be incorporated into the model by modeling the GEV parameters as functions of time or other covariates. These parameters are often easier to interpret than those of the corresponding GP models. See Coles (2001) for more details regarding the equivalence of the GP and PP approaches to extreme value modeling.

With the above considerations in mind, we proceed using the PP framework. Several practical considerations must be addressed before proceeding to fit a model to the data.

2) THRESHOLD SELECTION

The problem of selecting the threshold ω in both the GP and PP approaches is an instance of the bias–variance trade-off commonly encountered in statistics; a threshold that is too low may lead to model bias, while a threshold that is too large may yield larger variability in the resulting parameter estimates. See Serinaldi and Kilsby (2014) for more on the issue of threshold selection in POT models and methods to correct for model bias due to short data records.

There are many reasonable, data-driven methods for selecting the threshold ω. For instance, one can set ω equal to some large percentile of the data, for example, the 95th or 99th percentile of daily precipitation values. Another approach is to model the threshold as a time-varying function (Coles 2001), for example, as a step function:

$$\omega(t) = \omega_i \quad \text{if} \quad t \in T_i, \tag{2}$$

where the T_i are disjoint sets indexing time, and where the ω_i are predetermined constants. In our analysis, we adopt the threshold function in (2) where we let T_i, $i = 1, \ldots, 12$ correspond to the different months and where the ω_i in each region correspond to the 99th percentile of precipitation values for the pooled data in that region and month. Since the function in (2) has abrupt jumps at the end of each component, we smooth the threshold function in (2) via cubic splines.

3) SPATIAL AND TEMPORAL DEPENDENCE

Since extreme precipitation events tend to occur in temporal clusters (e.g., spans of 2–3 days at a time), in practice, the assumption of independent observations underlying the PP framework will be violated. To deal with this problem, we adopt a commonly used declustering procedure that first partitions the threshold exceedances at each site into separate temporal clusters, then only retains the cluster maxima for subsequent model fitting. Here, we add data points (precipitation values) to each temporal cluster until 5 consecutive points fall below the (99th percentile) threshold. For more details on this particular declustering scheme, see, for example, section 5.3.2 in Coles (2001).

There is also the problem of likely spatial dependence arising from the regional clustering procedure. It is not always clear how to effectively incorporate spatial dependence into an extreme value–based statistical model. Even recent attempts at incorporating spatial dependence into a regional analysis (see, e.g., Wang et al. 2014) require a subjective specification of a dependence structure. Misspecification of this dependence structure can introduce significant bias into the model, defeating the purpose of modeling such dependence in the first place. As pointed out in Katz et al. (2002) and Hosking and Wallis (1988), intersite correlation introduces little bias (if any) into point estimates of quantiles but results in underestimation of the standard errors of model parameters. For these reasons, we do not attempt to model the spatial dependence in this work.

4) MODEL FITTING

Several methods, such as maximum likelihood estimation (MLE) (Ferguson 1996), L moments (Hosking 1990, 2006), and Bayesian estimation can be used for model fitting and parameter estimation, though we found the Bayesian framework to be too computationally intensive for our analysis. When experimenting with these different model fitting techniques, we found there to be a minimal difference overall in the parameter estimates because of the relatively large sample sizes obtained as a result of the clustering step. Furthermore, the only way to obtain confidence intervals for parameter estimates in the L-moment framework is to apply the parametric bootstrap, making this approach relatively computationally expensive. Because of these considerations, we decided to proceed using the MLE approach. All model fitting was carried out using the "extRemes" package available in the R computing environment (Gilleland and Katz 2016).

5) NONSTATIONARITY

Under the assumption of stationarity in the time series, finding the return levels and recurrence intervals is straightforward. In the case of nonstationarity, however, the situation is more complicated since the properties of

the underlying distribution vary with time (we take the term "nonstationary" to refer to any statistical model whose parameters are expressed as a function of time). Risk forecasts based on stationary models will ignore time-dependent changes in the distribution of extreme precipitation intensity, leading to potentially unrealistic estimates of risk. Several measures have been recently proposed to address this difficulty: these include the effective return level (Katz et al. 2002; Cooley 2013), the design life level (Rootzén and Katz 2013), and the nonstationary extreme value analysis (NEVA) framework of Cheng et al. (2014). Here we have chosen to use the effective return level, though the other two methods can also be used depending on one's goals.

As a first approximation appropriate to many locations, we model the location and scale parameters of the PP model with the first-order sinusoidal functions:

$$\mu(t) = \alpha_0 + \alpha_1 \sin\left(\frac{2\pi t}{365.25}\right) + \alpha_2 \cos\left(\frac{2\pi t}{365.25}\right) \quad \text{and} \tag{3}$$

$$\log\sigma(t) = \beta_0 + \beta_1 \sin\left(\frac{2\pi t}{365.25}\right) + \beta_2 \cos\left(\frac{2\pi t}{365.25}\right); \tag{4}$$

the annual periodicity of these functions ensures that the effective return levels need only be computed for each day of the year (e.g., for $t = 1, \ldots, 365$ as opposed to each day in the entire time series), yielding one return-level map for each day of the year for any specified ARI.

For thoroughness, we compared the model defined by (3) alone, that is, assuming time-dependent location parameter and constant scale and shape parameters, with the model defined by both (3) and (4). The latter model better explains the data in 74% of the regions according to both the Akaike information criterion (AIC) and Bayesian information criterion (BIC) (it is worth noting that the AIC can result in model overfitting, whereas the BIC, which penalizes additional parameters, can lead to underfitting). Furthermore, both the AIC and BIC indicate the nonstationary model defined by (3) and (4) is superior to the stationary model in 94% of the regional clusters. Therefore, we adopt the nonstationary model defined by (3) and (4) throughout the rest of the paper unless stated otherwise.

3. Results

In this section, we discuss the return-level and trend estimates of the nonstationary PP model.

a. Return-level estimates

After fitting a distribution to the data in each region, the resulting parameter estimates are used to construct return-level maps that convey the rarity of precipitation events. It is important to note that because of the short data record, estimates of lower probability are subject to high uncertainty. As remarked in Parzybok et al. (2011), ARI results obtained from extreme value analysis are expected to be reliable for twice the data length. Since we are using 16 yr of TRMM data in our analysis, the model will be able to identify a 32-yr ARI event relatively accurately.

Some examples of the return-level maps for CONUS are given in Fig. 3. Comparing the maps for 1 January and 1 July reveals that there can be significant variability in the severity of extreme events throughout the year. For example, much of the West Coast has substantially higher return levels in January than in July, whereas the return levels are relatively stable among these two dates for much of the East Coast. Our findings are consistent with the results of Agel et al. (2015), who found that the intensity on extreme days in the Northeast is relatively invariant to the season.

Figures 4a and 4b show model diagnostic plots for the data from the cluster containing Los Angeles. To produce the Q–Q plot in Fig. 4a, the parameters of the fitted PP model are converted to the equivalent GP distribution (the quantiles are from threshold excesses of the data). The Q–Q plot in Fig. 4a indicates a reasonable model fit, with the empirical data distribution having a thicker upper tail than the fitted distribution. The Z plot in Fig. 4b is yet another gauge of model fit tailored specifically for the PP model fit (Smith and Shively 1995). Under the PP model, the waiting times between events should follow a mean-one exponential distribution. Therefore, the Z plot is a Q–Q plot that compares the quantiles of empirical waiting times with the quantiles of a mean-one exponential distribution. Figure 4b does not indicate any obvious departures from model assumptions.

Figure 4c shows some of the fitted return-level curves for several extreme precipitation events that occurred in Los Angeles from late 2004 to early 2005. The threshold for extreme events varies from around 1 mm in the summer (not surprising if one has ever spent a summer in Los Angeles) to about 38 mm in February. According to our model, one event crosses the 100-yr return-level curve, corresponding to an event that occurs in any given year with about 1% probability (as always, one should interpret such estimates after considering sampling variability, for example, via confidence bands for the return-level curves). In fact, the 2004/05 winter season proved to be one of the wettest seasons on record for Los Angeles county.

Finally, to capture the uncertainty in the parameter estimates used to make the return-level maps, we calculate 95% normalized confidence ranges (NCR) following the procedure in Zhou et al. (2015). For each

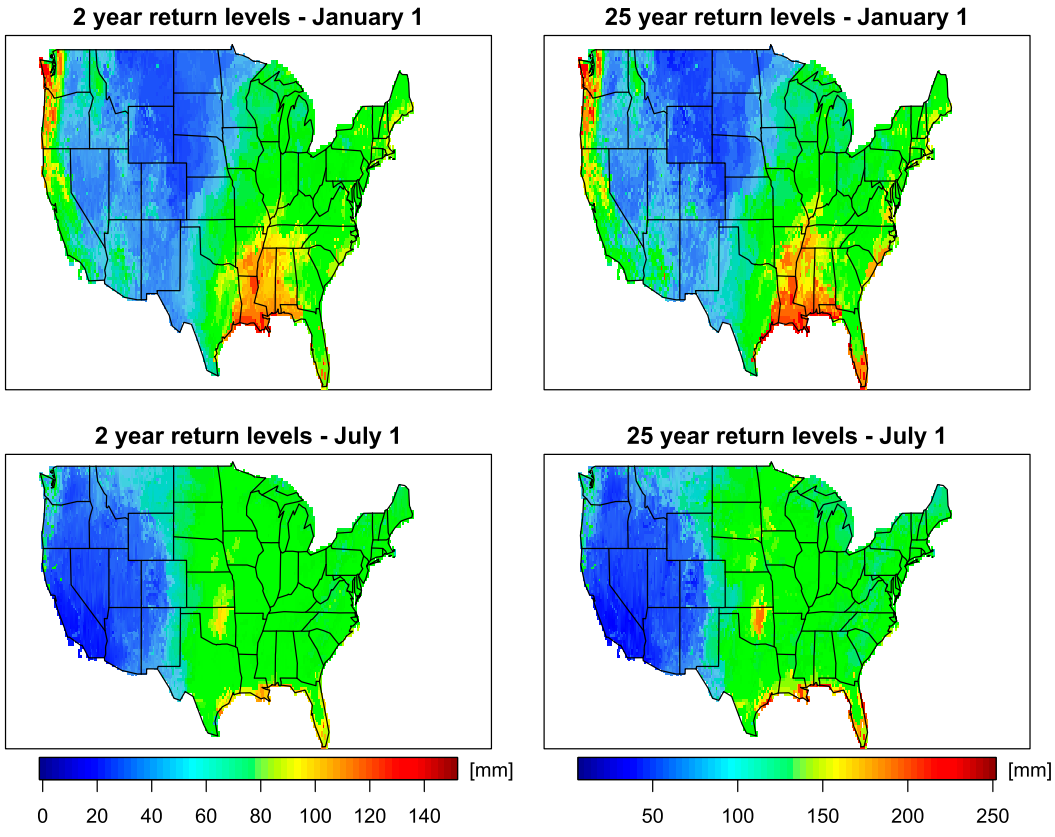

FIG. 3. (left) Two-year and (right) 25-yr return-level maps for the CONUS for (top) 1 Jan and (bottom) 1 Jul resulting from the nonstationary PP model using the TRMM 3B42 daily product.

region and for a given ARI (in years), we compute the difference between the upper and lower limits of the 95% confidence interval for the return levels, then divide this difference by the point estimate of the return levels. The NCR offers the advantage that it is independent of units of measurement, and can thus be used to compare regions with very different mean precipitation. Smaller values of the NCR imply a more confident estimate of the ARI; for example, an NCR value of 1 corresponds to an ARI estimate that lies within 100% of its magnitude with 95% confidence. Since the return-level estimates vary according to the time of year, we take a conservative approach and compute the maximum value the NCR obtains during the year. Figure 5 reveals that the majority of the regions on the map correspond to high confidence estimates (e.g., NCR < 1), both for 5- and 20-yr ARIs. The general pattern in the NCR maps is very similar to the results of Zhou et al. (2015), with low confidence regions primarily located in exceedingly dry areas such as northern Africa, the Arabian Peninsula, and the southeast Pacific Ocean, though the values in the 5- and 20-yr NCR maps based on our methodology are generally much smaller than those in Zhou et al. (2015). As pointed out in Zhou et al. (2015), as the data length of the TRMM-GPM precipitation records increases, the degree of confidence in the ARI estimates will increase even further.

b. *Comparison with previous models*

To put our results into perspective, we compare the return-level maps resulting from our proposed methodology with those based on the annual maxima–GEV framework as in Zhou et al. (2015). Both methods are applied to the same 3B42 daily precipitation data, but because Zhou et al. (2015) did not use the data from 2013 in their analysis, we restrict the data for our model to the 1998–2012 span to facilitate model comparisons. As a benchmark for comparison, we also show the return-level maps generated using NOAA's Climate Prediction Center (CPC) daily unified precipitation dataset, which is a gauge-based, gridded, and quality-controlled product derived from daily and hourly precipitation measurements from 1948 to 2012 in which

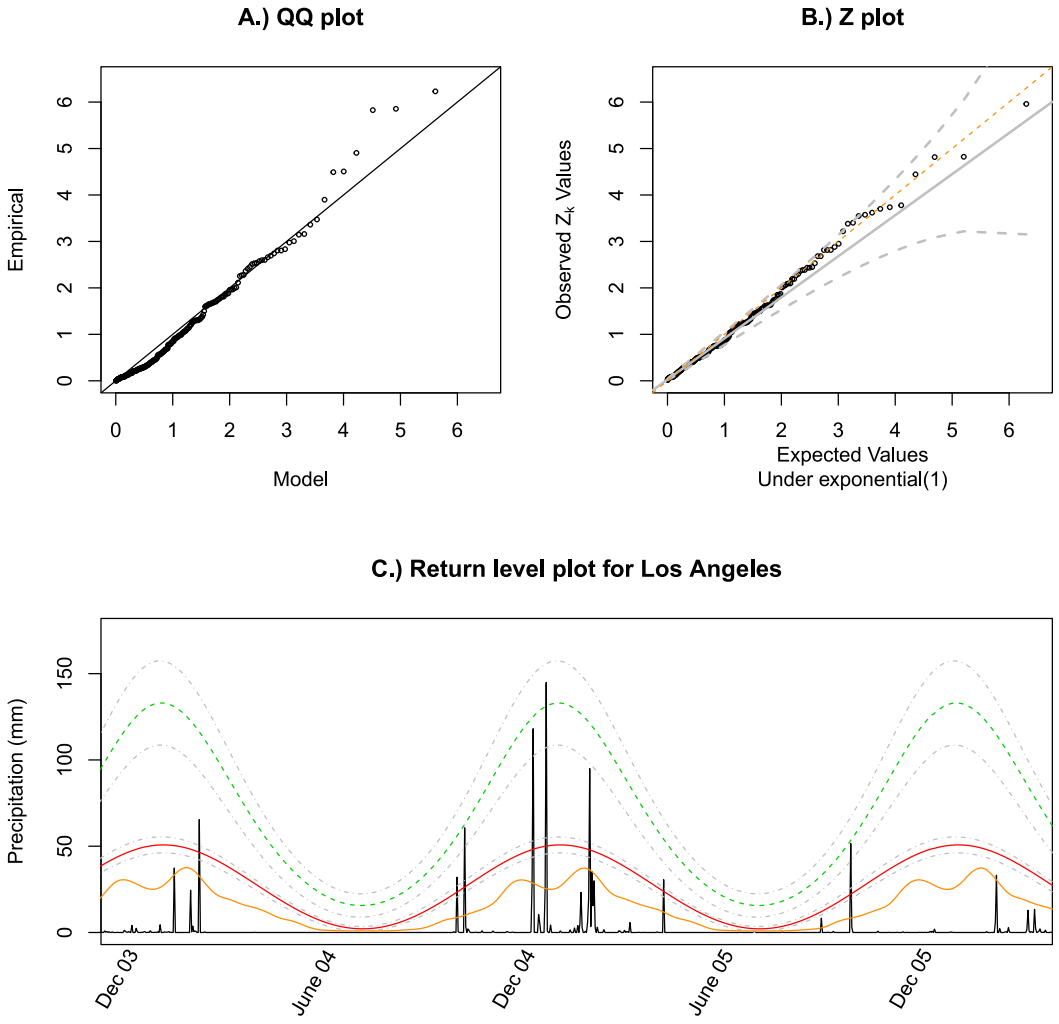

FIG. 4. Nonstationary PP model diagnostic plots and return-level plot for the cluster containing Los Angeles. (a) The Q–Q plot. (b) The Z plot: The solid gray line is the regression fit of Z_k on the expected values of the observed order statistics under the model. The dashed orange line is a 45° reference line, and the dashed gray lines are 95% confidence bounds. (c) Return-level plot: Fitted precipitation return levels in Los Angeles for December 2003–May 2006. The orange solid, red solid, and green dashed curves correspond to the seasonal threshold, the 2-yr return level, and the 100-yr return level, respectively. The 95% confidence bounds are indicated by gray dot–dashed curves.

measurements were taken from over 13 000 stations (8000 before 2012) over the CONUS. The CPC data also have the same 0.25° spatial resolution as the 3B42 data. The CPC data were modeled using the single-site, annual maxima–GEV framework in Zhou et al. (2015). To make our results directly comparable to both of these sets of return-level maps that were constructed under the assumption of stationarity (implying a single return-level map for the entire year), we also assume stationarity in our PP approach and thus do not allow for seasonality in the rest of this section. That is, we take a single threshold for the entire time series (the 99th percentile of precipitation values) of a given region, and assume that the location, scale, and shape parameters do not vary with time or other covariates.

In Fig. 6, we compare the return-level maps corresponding to ARIs of 2 and 25 yr produced using the three different approaches stated above. The most striking feature of these diagrams is the reduction in noise when using the regional analysis over the existing single-site methodology. In the return-level maps corresponding to an ARI of 25 yr, for example, the return-level map based on the TRMM data using the single-site block maxima approach is quite coarse, with many isolated grid points exhibiting return levels that are in sharp contrast to their surrounding

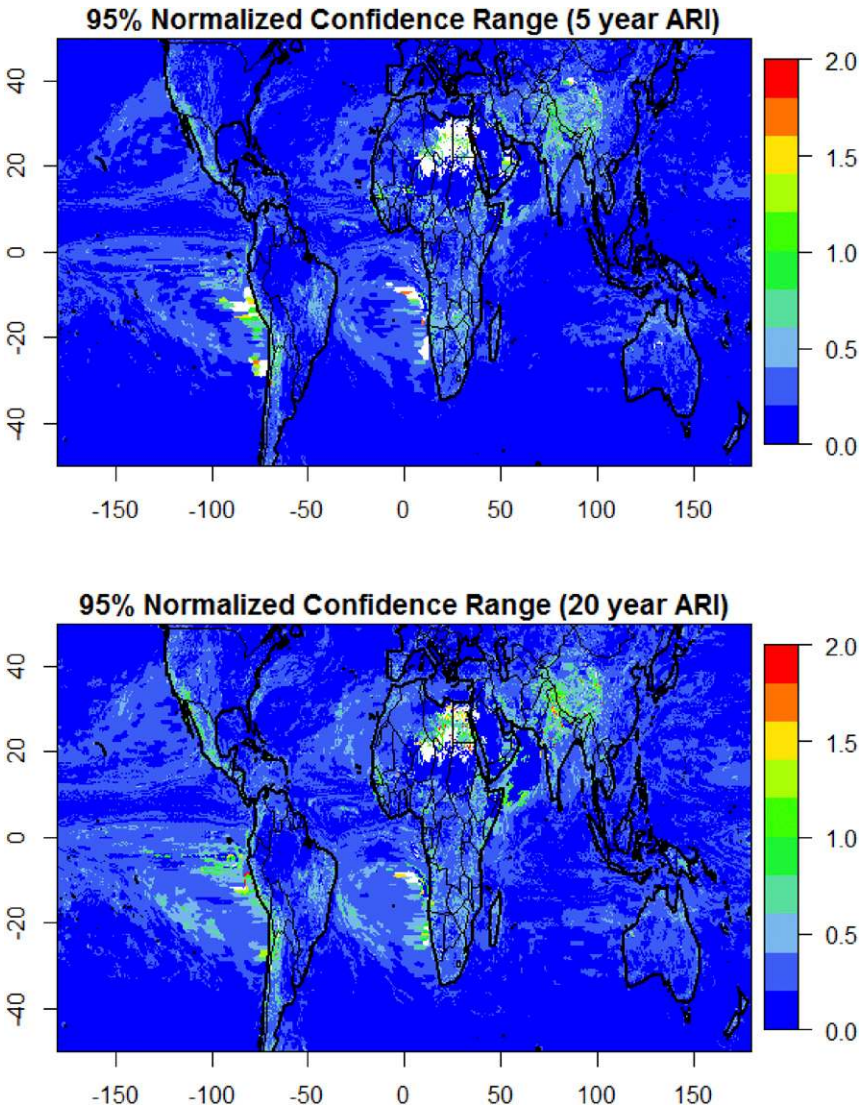

FIG. 5. Maximum of daily 95% NCRs of estimated (top) 5-yr and (bottom) 20-yr return levels from the nonstationary PP model. White values correspond to NCR values above 2.

neighbors. The short data record for this approach (15 data points per site) means that the GEV model fitting procedure could not effectively separate the signal from the statistical noise. Of course, it is possible that some of the isolated ''spikes'' in the return-level maps reflect actual contrasts in precipitation extremes. However, since the same GEV method was used on both the 65-yr CPC data and the 15-yr TRMM data, and since using a longer data record smoothed away most of the spikes, it is reasonable to conclude that most of the contrasts were indeed a result of the short data record. From the maps, it is apparent that our methodology results in a smoother return-level map when compared with the single-site, annual maxima framework, capturing the general pattern in the CPC results using fewer data.

c. *Model fit*

To assess how well the stationary PP approach models the observed data, we constructed several diagnostic plots including kernel density plots as well as Q–Q plots. The results for one randomly selected region, corresponding to four grid points in western Colombia, are displayed in Fig. 7. The density and Q–Q plots indicate that both the PP and single-site GEV models fit to the TRMM 3B42 series explain the data

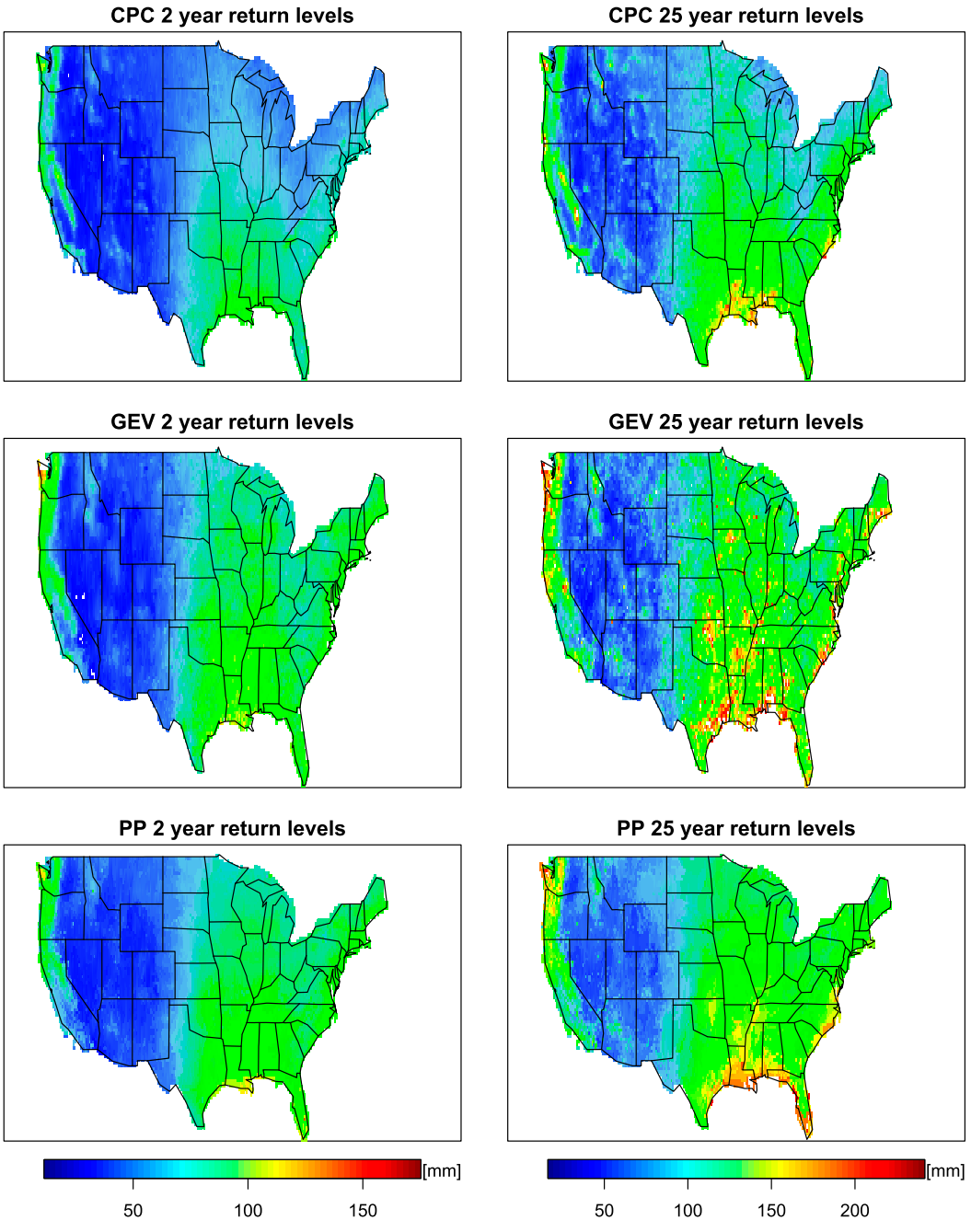

FIG. 6. Comparison of the return-level maps produced by the CPC measurements (GEV model; CPC daily unified product), GEV-based TRMM model (TRMM 3B42 product), and the stationary PP TRMM model (TRMM 3B42 product).

reasonably well (note the bimodality in the empirical distribution of the block maxima model—this issue is discussed further in the discussion section). Figure 7 also includes return-level plots for both methods, which plot the return levels (in millimeters) expected to occur on average once during the corresponding recurrence interval (given in years). The return-level plots suggest that the two models differ in their characterizations of the tail behavior of extreme events. Indeed, at the 5% level of significance, the PP model fit implies a finite upper bound for extreme precipitation intensity, while the GEV model fit indicates unbounded tail behavior. The 95% confidence limits (dashed gray lines) indicate a higher level of confidence

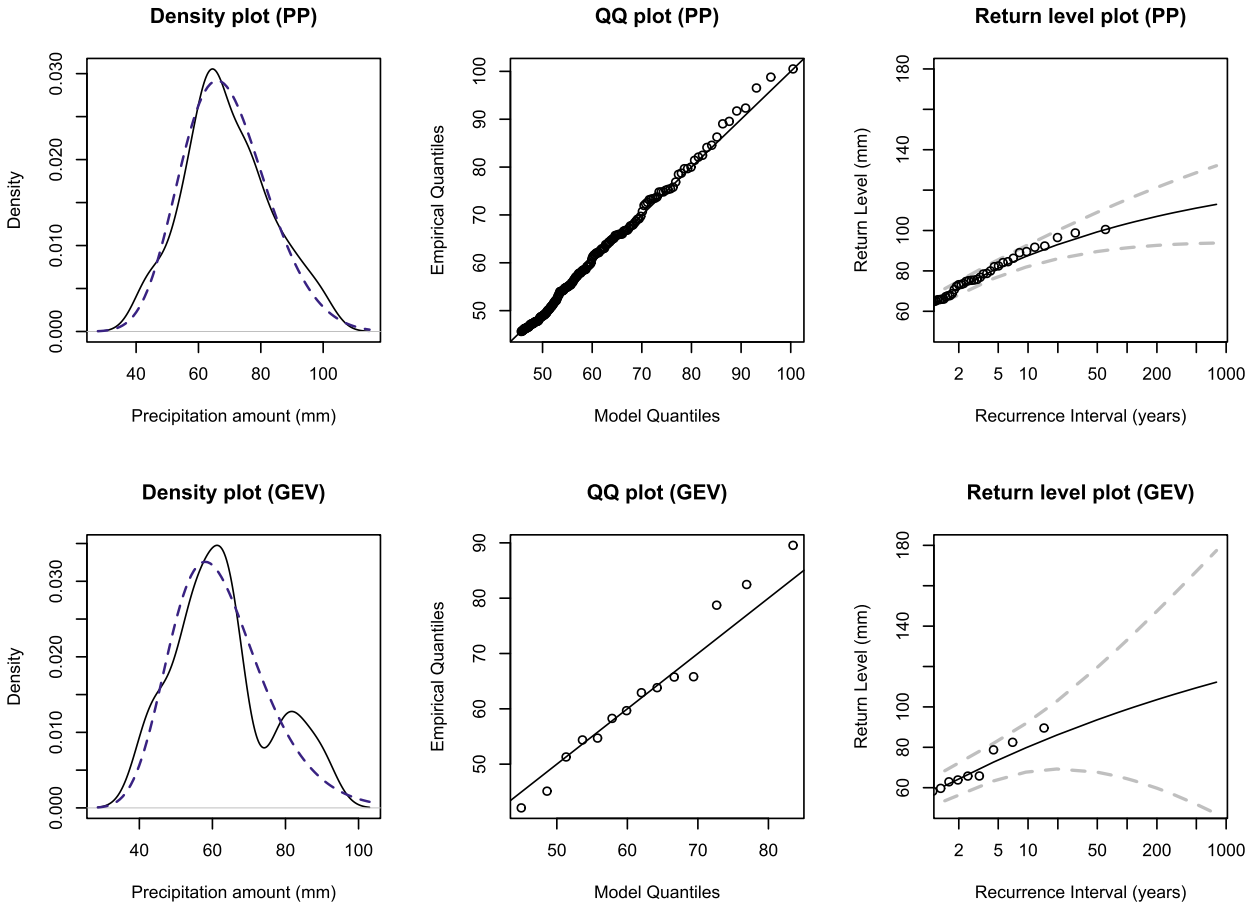

FIG. 7. Comparison of model fit using (top) the PP approach and (bottom) the block maxima/GEV approach for a randomly selected regional cluster corresponding to four grid points in western Colombia (for the block maxima approach, we randomly selected one of these four grid points). (left) Kernel density plots. Black (solid) curves are empirical data; blue (dashed) curves are model fit. To create the PP density plot, the empirical density of the annual maxima of the data is calculated (black solid line) and compared with the GEV distribution implied by the fitted PP (blue dashed curve). (center) Q–Q plots. (right) Return-level plots; the dashed curves are 95% confidence bounds.

in the results produced by the PP method than the single-site GEV approach. We also note that the 95% NCR maps corresponding to our method in the stationary setting (not included here for brevity) are very similar to those in Fig. 5, indicating an overall increase in statistical confidence.

d. Case study

We applied our methodology to evaluate the severity of a particular climate event, Typhoon Fitow, the strongest typhoon to hit mainland China in more than 60 years. Specifically, we estimated the annual probabilities of the precipitation event that occurred on 6 October 2013 for the nonstationary PP model with regional clustering, as well as for the stationary GEV model of annual maxima without regional clustering used in Zhou et al. (2015).

Figure 8 shows the 1-day precipitation total on 6 October 2013 over China's Zhejiang province, as well as the predicted annual probabilities of the corresponding precipitation intensities of both models. The estimated probabilities for the precipitation totals recorded during this event are generally higher under the PP model than those of the GEV model, implying that such extreme events are more common than the existing method in Zhou et al. (2015) would have predicted. Most of the probabilities under the GEV model are less than 0.01, and given the short length of the data record, the validity of such estimates is questionable. Although there are also low-probability events (<0.01) predicted by the PP model, more than 80% of the predicted probabilities are larger than 3%; thus the reliability of the PP estimates is less affected by the short data record. The PP model predictions in Fig. 8 reveal that there were three distinct

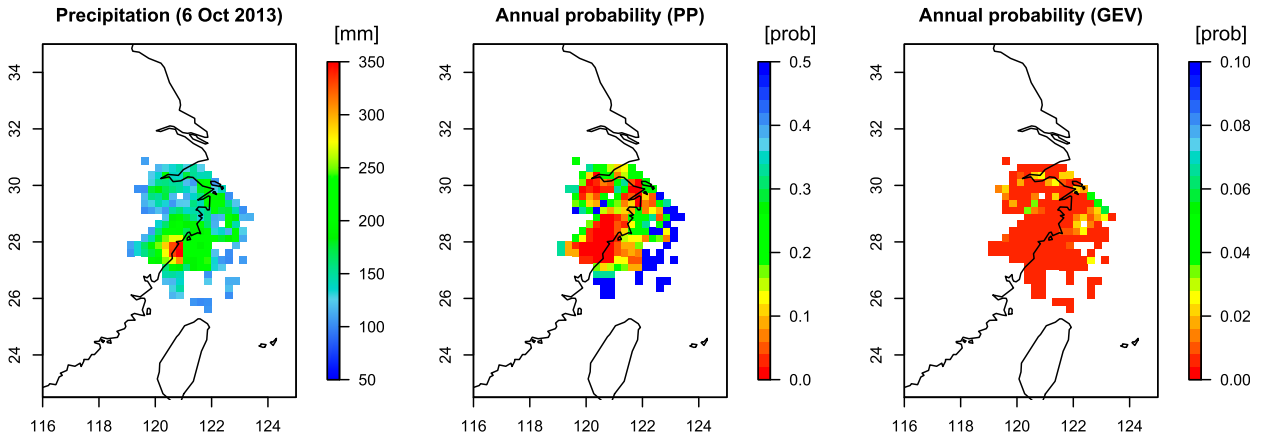

FIG. 8. Typhoon Fitow (6 Oct 2013) (left) precipitation (mm) and predicted annual probabilities for the (center) nonstationary PP model and (right) stationary GEV model. Only precipitation levels > 50 mm and their corresponding probabilities are shown for clarity. Note the different probability scales.

regions of particularly rare precipitation intensity, with the largest region overlapping with the area of heaviest precipitation. The GEV approach failed to make the distinction between these three regions.

e. *Trends in extreme precipitation intensity*

A straightforward modification of the nonstationary PP model allows an analysis of long-term trends. A simple starting point is to model the GEV location parameter as a linear function of time, that is,

$$\mu(t) = \mu_0 + \mu_1 t, \tag{5}$$

and to assume constant scale and shape parameters. In this setup and for any fixed probability p, the coefficient μ_1 measures the change in the GEV quantile function over the data period (given t is scaled to lie in [0, 1]); positive values of μ_1 reflect more intense extreme precipitation events and negative values reflect less intense extreme events. To visualize the results, we adopt the approach used in Katz et al. (2002) and set $p = 0.5$ and compute the percentage change in the median of the fitted GEV distribution over the data period; intuitively, we are calculating how much the underlying distributions of extreme precipitation intensities shifted from 1998 to 2013. The percent changes in the medians of extreme precipitation intensities are shown in Fig. 9 (only trends significant at the 5% level are shown). We stress that these results should not be extrapolated to periods outside of the data record and are only used here to study the behavior of extreme events from 1998 to 2013.

Figure 9 shows generally increasing intensities of extreme precipitation in the tropical ITCZ, including the tropical Indian Ocean, "Maritime Continent," western Pacific Ocean warm pool, Caribbean Sea, and Gulf of Mexico regions. Decreases in extreme precipitation are observed in most of the tropical and subtropical land regions, that is, South America, tropical and southern Africa, and north and west Australia, consistent with the results of Wu and Lau (2016). Negative trends are also observed over most of the CONUS, especially in the southwest United States, contributing to the drying trend in the region (Prein et al. 2016). However, decreases in extreme precipitation in the midlatitude oceans in the Pacific and North Atlantic, together with increases in extreme precipitation in the southern (north) edge of the subtropical jet in the Northern and Southern Hemisphere could indicate an equatorward shift of heavy precipitation regions as opposed to a general expansion of the ITCZ (Zhou et al. 2011; Lucas et al. 2014).

We emphasize that only linear trends in time have been investigated here, and therefore our model can only detect static increases/decreases in precipitation extremes. One possible workaround to this problem would be to use the average temperature within each cluster as a covariate instead of time; the resulting model could then capture more complex behaviors in the global precipitation system. In addition, since the data record is relatively short, the estimated trends might be capturing part of a longer-period fluctuation. For example, even models that correctly identify a trend over a short time period may fail to identify a reversal of the trend if such a reversal occurred over a time span longer than the data record (Fu et al. 2010; Kunkel et al. 2013).

4. Application to surface air temperature data

The generality of the PP framework implies that our clustering and model fitting procedures can easily be

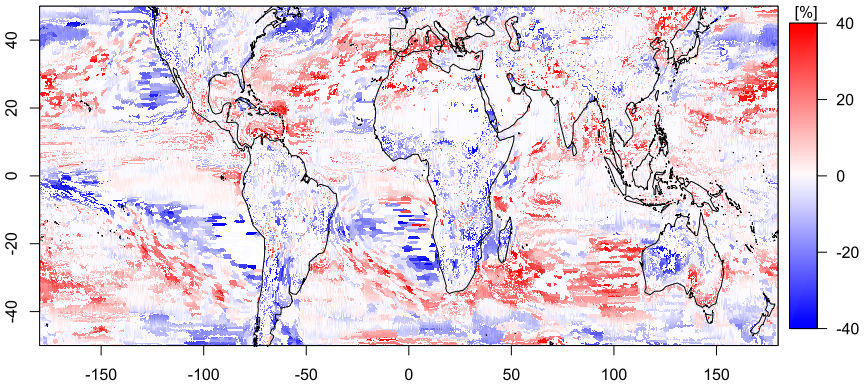

FIG. 9. Percent change in the median of the fitted GEV distribution of extreme precipitation intensities. Positive changes reflect more intense extreme precipitation events, and negative changes reflect less intense extreme events (only trends significant at the 5% level are shown).

applied to model various types of data other than precipitation data. As a proof of concept, in this section we apply our methodology to analyze trends in extreme temperature intensity. Specifically, we use surface air temperature data from NOAA's NCEP North American Regional Reanalysis (NARR) product (NOAA/NCEP 2004). The data are daily surface temperatures (°C) spanning from 1 January 1979 to 31 December 2013 over North America at a resolution of approximately 0.3° (32 km) at the lowest latitude, and the number of grid squares is 349 × 277. Here, we restrict our analysis to the CONUS. More information about the NARR product can be found online (http://www.esrl.noaa.gov/psd/data/gridded/data.narr.html; Mesinger et al. 2006). Again, we stress that the results from the short data record cannot be extrapolated into the future.

We used 50 clusters for the first round of k means and 30 clusters for the second round, resulting in a total of 1500 disjoint regions. We used location and the 90th percentile of temperature values as input for clustering, though more extensive analyses should consider a more comprehensive set of variables. The Viglione et al. (2007) homogeneity test identified 1431 of 1500 regions as being acceptably homogeneous. Next, we fit a nonstationary PP model to the data in each region following the procedure outlined in section 2. Since we will be examining long-term trends, for the threshold function in (2), we took one threshold per year, taken to be the 95th percentile of temperature intensities for that year (using the 95th percentile instead of the 99th percentile produced more stable parameter estimates). As before when examining trends in precipitation extremes, we assumed constant scale and shape parameters and a linear trend in the location parameter. A map showing the percent change of the median of the fitted extreme temperature distributions is shown in Fig. 10, along with a map of average temperatures for comparison. Only trends significant at the 5% level are shown.

According to our model, most of CONUS experienced an increase in the intensities of extreme temperature events during this time period. Figure 10 indicates that the largest increase in the medians of extreme temperatures was about 4% in southern Louisiana and eastern Texas. The East Coast also showed a consistent increase in extreme temperature intensities, with the largest increase of about 2% occurring in eastern Maryland and Delaware. The trends are reversed near parts of the Rocky Mountains, with decreases in the median of temperature intensities as large as 2% in western Colorado. Some smaller decreases are observed in the northern Great Plains and parts of central California. These results are generally in line with the analyses and projections of Schoof and Robeson (2016), who predict a consistent increase across the United States in the number of excessively warm days over the twenty-first century. Our findings are also consistent with the behavior of extreme heat waves over this time period, particularly with the increased number of extreme heat waves occurring from 2000 to 2010 (Kunkel et al. 2013). Notably, unlike the findings in Peterson et al. (2013), our results do not reflect any cooling trends over the "warming hole" (Meehl et al. 2012; Kunkel et al. 2006) in the southeastern United States. The phase reversal of the interdecadal Pacific oscillation in the tropical Pacific in the late 1990s may explain the disappearance of the warming hole after 2000 (Meehl et al. 2015), and therefore part of the difference in our findings may be due to differences in the data period [1950–2007 in Peterson et al. (2013) vs 1979–2013 here]. Once again, we emphasize that we have assumed a simple

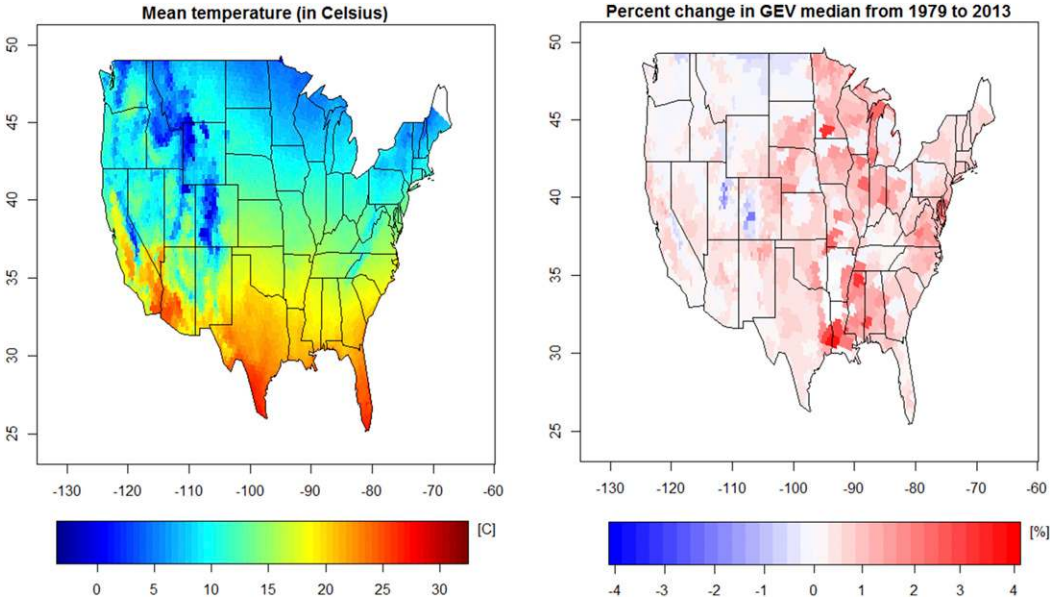

FIG. 10. (left) Mean surface air temperatures in the CONUS, 1979–2013 (NCEP NARR product). (right) Percent change in the median of the distribution of temperature extremes from the nonstationary PP model (only trends significant at the 5% level are shown).

linear trend in time, and that more complicated trend structures would be able to capture more sophisticated behavior in temperature extremes.

5. Discussion

In this paper, we propose an alternative methodology for the statistical modeling of the TRMM extreme precipitation monitoring system. Our regional clustering algorithm, in conjunction with the POT approach for modeling extremes, allows us to leverage more data than the single-site block maxima method, yielding more accurate estimates of the regional ARIs. The resulting return-level maps produced by our method (Fig. 6) reveal that our algorithm can more effectively separate out the statistical noise than the existing Zhou et al. (2015) approach. Our model provides a useful tool for studying the global and regional characteristics and trends of extreme variables, whether these are precipitation events or other climate events.

There are several possible extensions to our analysis. First, in this paper we only consider 1-day precipitation totals. More complete information about return levels and trends in extreme precipitation can be obtained by considering multiday cumulative precipitation totals, for example, 3- or 5-day precipitation totals reflecting the severity of multiple-day precipitation events. However, when modeling such accumulated precipitation events, we noticed significant multimodality in the intensity of the accumulated precipitation events. While multimodality in precipitation occurrences and intensity has been previously reported (Schindler et al. 2012; Tye et al. 2016), we are not aware of any statistical models that have specifically been developed to model multimodality in accumulated precipitation totals. We are currently developing a framework based on mixture modeling that would be able to deal with this realistic scenario.

Second, we did not attempt to model the spatial dependence among grid locations in each regional cluster. Future studies should aim at developing models that are flexible enough to accommodate a wide range of dependence structures while being careful to avoid overfitting.

Finally, we chose to adopt first-order sinusoidal functions to represent the GEV location and scale parameters when estimating return levels. While this choice may be a reasonable first approximation for modeling seasonality at all locations, a more flexible seasonal cycle would be more appropriate. Effectively modeling the seasonal cycle can be beneficial for assessing the variability in extreme events throughout the year at any location; the resulting effective return levels can be crucial for public policy and disaster relief planning, especially during months where extreme precipitation events are particularly intense. A more realistic and flexible seasonal cycle warrants further study.

Acknowledgments. Levon Demirdjian was supported by a Burroughs Wellcome Fund Population and Laboratory Based Sciences Award at UCLA, and he thanks the NASA Goddard Space Flight Center internship program. Yaping Zhou was supported by NASA Precipitation Measurement Mission (NNH12ZDA001N-PMM) and the Science of Terra and Aqua Program (NNH13ZDA001N-TERAQ). George J. Huffman was supported by NASA Precipitation Measurement Mission (Award 573945.04.18.02.78).

REFERENCES

Agel, L., M. Barlow, J.-H. Qian, F. Colby, E. Douglas, and T. Eichler, 2015: Climatology of daily precipitation and extreme precipitation events in the northeast United States. *J. Hydrometeor.*, **16**, 2537–2557, https://doi.org/10.1175/JHM-D-14-0147.1.

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

Bador, M., P. Naveau, E. Gilleland, M. Castellà, and T. Arivelo, 2015: Spatial clustering of summer temperature maxima from the CNRM-CM5 climate model ensembles & E-OBS over Europe. *Wea. Climate Extremes*, **9**, 17–24, https://doi.org/10.1016/j.wace.2015.05.003.

Bernard, E., P. Naveau, M. Vrac, and O. Mestre, 2013: Clustering of maxima: Spatial dependencies among heavy rainfall in France. *J. Climate*, **26**, 7929–7937, https://doi.org/10.1175/JCLI-D-12-00836.1.

Brown, S. J., J. Caesar, and C. A. T. Ferro, 2008: Global changes in extreme daily temperature since 1950. *J. Geophys. Res.*, **113**, D05115, https://doi.org/10.1029/2006JD008091.

Buishand, T. A., 1991: Extreme rainfall estimation by combining data from several sites. *Hydrol. Sci. J.*, **36**, 345–365, https://doi.org/10.1080/02626669109492519.

Chan, S. C., E. J. Kendon, H. J. Fowler, S. Blenkinsop, N. M. Roberts, and C. A. T. Ferro, 2014: The value of high-resolution Met Office regional climate models in the simulation of multihourly precipitation extremes. *J. Climate*, **27**, 6155–6174, https://doi.org/10.1175/JCLI-D-13-00723.1.

Cheng, L., A. AghaKouchak, E. Gilleland, and R. W. Katz, 2014: Non-stationary extreme value analysis in a changing climate. *Climatic Change*, **127**, 353–369, https://doi.org/10.1007/s10584-014-1254-5.

Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, Vol. 97, Springer, 221 pp., https://doi.org/10.1007/978-1-4471-3675-0.

Cooley, D., 2013: Return periods and return levels under climate change. *Extremes in a Changing Climate: Detection, Analysis and Uncertainty*, A. AghaKouchak, et al., Eds., Water Science and Technology Library, Vol. 65, Springer, 97–114, https://doi.org/10.1007/978-94-007-4479-0_4.

Cox, D. R., and V. Isham, 1980: *Point Processes.* Chapman and Hall, 188 pp.

Cunnane, C., 1989: Statistical distributions for flood frequency analysis. WMO Tech Rep. 718/Operational Hydrology Rep. 33, 61 pp.

Davison, A. C., and R. L. Smith, 1990: Models for exceedances over high thresholds. *J. Roy. Stat. Soc.*, **52B**, 393–442.

Di Liberto, T., 2016: August 2016 extreme rain and floods along the Gulf Coast. NOAA, accessed 12 November 2016, https://www.climate.gov/news-features/event-tracker/august-2016-extreme-rain-and-floods-along-gulf-coast.

Donat, M. G., A. L. Lowry, L. V. Alexander, P. A. O'Gorman, and N. Maher, 2016: More extreme precipitation in the world's dry and wet regions. *Nat. Climate Change*, **6**, 508–513, https://doi.org/10.1038/nclimate2941.

Ferguson, T. S., 1996: *A Course in Large Sample Theory*. Texts in Statistical Science, Vol. 38, Chapman and Hall/CRC, 256 pp.

Fowler, H. J., and M. Ekström, 2009: Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes. *Int. J. Climatol.*, **29**, 385–416, https://doi.org/10.1002/joc.1827.

Fu, G., N. R. Viney, S. P. Charles, and J. Liu, 2010: Long-term temporal variation of extreme rainfall events in Australia: 1910–2006. *J. Hydrometeor.*, **11**, 950–965, https://doi.org/10.1175/2010JHM1204.1.

Furrer, E. M., and R. W. Katz, 2008: Improving the simulation of extreme precipitation events by stochastic weather generators. *Water Resour. Res.*, **44**, W12439, https://doi.org/10.1029/2008WR007316.

Gilleland, E., and R. W. Katz, 2016: extRemes 2.0: An extreme value analysis package in R. *J. Stat. Software*, **72**, 1–39, https://doi.org/10.18637/jss.v072.i08.

Hastie, T., R. Tibshirani, and J. Friedman, 2009: *The Elements of Statistical Learning*. 2nd ed. Springer Series in Statistics, Springer, 745 pp.

Heaton, M. J., M. Katzfuss, S. Ramachandar, K. Pedings, E. Gilleland, E. Mannshardt-Shamseldin, and R. L. Smith, 2011: Spatio-temporal models for large-scale indicators of extreme weather. *Environmetrics*, **22**, 294–303, https://doi.org/10.1002/env.1050.

Hosking, J. R. M., 1990: L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. Roy. Stat. Soc.*, **52B**, 105–124.

——, 2006: On the characterization of distributions by their L-moments. *J. Stat. Plann. Inference*, **136**, 193–198, https://doi.org/10.1016/j.jspi.2004.06.004.

——, and J. R. Wallis, 1988: The effect of intersite dependence on regional flood frequency analysis. *Water Resour. Res.*, **24**, 588–600, https://doi.org/10.1029/WR024i004p00588.

——, and ——, 1993: Some statistics useful in regional frequency analysis. *Water Resour. Res.*, **29**, 271–281, https://doi.org/10.1029/92WR01980.

——, and ——, 1997: *Regional Frequency Analysis: An Approach Based on L Moments*. Cambridge University Press, 238 pp., https://doi.org/10.1017/CBO9780511529443.

——, ——, and E. F. Wood, 1985: Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, **27**, 251–261, https://doi.org/10.1080/00401706.1985.10488049.

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, https://doi.org/10.1175/JHM560.1.

Katz, R. W., M. B. Parlange, and P. Naveau, 2002: Statistics of extremes in hydrology. *Adv. Water Resour.*, **25**, 1287–1304, https://doi.org/10.1016/S0309-1708(02)00056-8.

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, https://doi.org/10.1175/JCLI4066.1.

Kunkel, K. E., X.-Z. Liang, J. Zhu, and Y. Lin, 2006: Can CGCMs simulate the twentieth-century "warming hole" in the central United States? *J. Climate*, **19**, 4137–4153, https://doi.org/10.1175/JCLI3848.1.

——, L. E. Stevens, S. E. Stevens, L. Sun, E. Janssen, D. Wuebbles, and J. G. Dobson, 2013: Climate of the contiguous United States. Part 9, Regional Climate Trends and Scenarios for the U.S. National Climate Assessment, NOAA Tech. Rep. NESDIS 142-9, 85 pp., https://scenarios.globalchange.gov/sites/default/files/NOAA_NESDIS_Tech_Report_142-9-Climate_of_the_Contiguous_United_States_3.pdf.

Leadbetter, M. R., G. Lindgren, and H. Rootzén, 1983: *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics, Springer, 336 pp., https://doi.org/10.1007/978-1-4612-5449-2.

Lucas, C., B. Timbal, and H. Nguyen, 2014: The expanding tropics: A critical assessment of the observational and modeling studies. *Wiley Interdiscip. Rev.: Climate Change*, **5**, 89–112, https://doi.org/10.1002/wcc.251.

Meehl, G. A., J. M. Arblaster, and G. Branstator, 2012: Mechanisms contributing to the warming hole and the consequent U.S. east–west differential of heat extremes. *J. Climate*, **25**, 6394–6408, https://doi.org/10.1175/JCLI-D-11-00655.1.

——, ——, and C. T. Y. Chung, 2015: Disappearance of the southeast U.S. "warming hole" with the late 1990s transition of the interdecadal Pacific oscillation. *Geophys. Res. Lett.*, **42**, 5564–5570, https://doi.org/10.1002/2015GL064586.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, https://doi.org/10.1175/BAMS-87-3-343.

Min, S.-K., X. Zhang, F. W. Zwiers, and G. C. Hegerl, 2011: Human contribution to more-intense precipitation extremes. *Nature*, **470**, 378–381, https://doi.org/10.1038/nature09763.

NASA GES DISC, 2016: TRMM Multi-Satellite Precipitation Analysis (TMPA) product. Accessed 20 June 2016, https://pmm.nasa.gov/data-access/downloads/trmm.

NOAA/NCEP, 2004: NCEP Reanalysis Data. NOAA/OAR/ESRL, accessed 26 July 2016, http://www.esrl.noaa.gov/psd/data/gridded/data.narr.monolevel.html.

Parzybok, T. W., B. Clarke, and D. M. Hulstrand, 2011: Average recurrence interval of extreme rainfall in real-time. *Earthzine*, http://earthzine.org/2011/04/19/average-recurrence-interval-of-extreme-rainfall-in-real-time/.

Peterson, T. C., and Coauthors, 2013: Monitoring and understanding changes in heat waves, cold waves, floods, and droughts in the United States: State of knowledge. *Bull. Amer. Meteor. Soc.*, **94**, 821–834, https://doi.org/10.1175/BAMS-D-12-00066.1.

Prein, A. F., G. J. Holland, R. M. Rasmussen, M. P. Clark, and M. R. Tye, 2016: Running dry: The U.S. Southwest's drift into a drier climate state. *Geophys. Res. Lett.*, **43**, 1272–1279, https://doi.org/10.1002/2015GL066727.

Rootzén, H., and R. W. Katz, 2013: Design life level: Quantifying risk in a changing climate. *Water Resour. Res.*, **49**, 5964–5972, https://doi.org/10.1002/wrcr.20425.

Row, L. W., III, and D. A. Hastings, 1994: TerrainBase worldwide digital terrain data, release 1.0 NOAA/National Geophysical Data Center, accessed 22 June 2016, ftp://ftp.ngdc.noaa.gov/Solid_Earth/cdroms/TerrainBase_94/data/global/tbase/.

Satyanarayana, P., and V. V. Srinivas, 2008: Regional frequency analysis of precipitation using large-scale atmospheric variables. *J. Geophys. Res.*, **113**, D24110, https://doi.org/10.1029/2008JD010412.

Schindler, A., D. Maraun, and J. Luterbacher, 2012: Validation of the present day annual cycle in heavy precipitation over the British Islands simulated by 14 RCMs. *J. Geophys. Res.*, **117**, D18107, https://doi.org/10.1029/2012JD017828.

Scholz, F. W., and M. A. Stephens, 1987: K-sample Anderson–Darling tests. *J. Amer. Stat. Assoc.*, **82**, 918–924, https://doi.org/10.1080/01621459.1987.10478517.

Schoof, J. T., and S. M. Robeson, 2016: Projecting changes in regional temperature and precipitation extremes in the United States. *Wea. Climate Extremes*, **11**, 28–40, https://doi.org/10.1016/j.wace.2015.09.004.

Serinaldi, F., and C. G. Kilsby, 2014: Rainfall extremes: Toward reconciliation after the battle of distributions. *Water Resour. Res.*, **50**, 336–352, https://doi.org/10.1002/2013WR014211.

Shane, R. M., and W. R. Lynn, 1964: Mathematical model for flood risk evaluation. *J. Hydraul. Div. Amer. Soc. Civ. Eng.*, **90**, 1–20.

Smith, R. L., 1989: Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Stat. Sci.*, **4**, 367–377, https://doi.org/10.1214/ss/1177012400.

——, and T. S. Shively, 1995: Point process approach to modeling trends in tropospheric ozone based on exceedances of a high threshold. *Atmos. Environ.*, **29**, 3489–3499, https://doi.org/10.1016/1352-2310(95)00030-3.

Smithers, J. C., and R. E. Schulze, 2001: A methodology for the estimation of short duration design storms in South Africa using a regional approach based on L-moments. *J. Hydrol.*, **241**, 42–52, https://doi.org/10.1016/S0022-1694(00)00374-7.

Todorovic, P., and E. Zelenhasic, 1970: A stochastic model for flood analysis. *Water Resour. Res.*, **6**, 1641–1648, https://doi.org/10.1029/WR006i006p01641.

Tye, M. R., S. Blenkinsop, H. J. Fowler, D. B. Stephenson, and C. G. Kilsby, 2016: Simulating multimodal seasonality in extreme daily precipitation occurrence. *J. Hydrol.*, **537**, 117–129, https://doi.org/10.1016/j.jhydrol.2016.03.038.

Typhoon Committee, 2013: Republic of Korea. ESCAP/WMO Typhoon Committee Member Rep., 42 pp., http://www.typhooncommittee.org/8IWS_2TRCG/docs/Members%20Report/MEMBER_REPORT(ROK).pdf.

Viglione, A., F. Laio, and P. Claps, 2007: A comparison of homogeneity tests for regional frequency analysis. *Water Resour. Res.*, **43**, W03428, https://doi.org/10.1029/2006WR005095.

Wang, Z., J. Yan, and X. Zhang, 2014: Incorporating spatial dependence in regional frequency analysis. *Water Resour. Res.*, **50**, 9570–9585, https://doi.org/10.1002/2013WR014849.

——, Z. Zeng, C. Lai, W. Lin, X. Wu, and X. Chen, 2017: A regional frequency analysis of precipitation extremes in mainland China with fuzzy c-means and L-moments approaches. *Int. J. Climatol.*, **37**, 429–444, https://doi.org/10.1002/joc.5013.

Wu, H.-T. J., and W. K.-M. Lau, 2016: Detecting climate signals in precipitation extremes from TRMM (1998–2013)—Increasing contrast between wet and dry extremes during the "global warming hiatus." *Geophys. Res. Lett.*, **43**, 1340–1348, https://doi.org/10.1002/2015GL067371.

Zhou, Y. P., K.-M. Xu, Y. C. Sud, and A. K. Betts, 2011: Recent trends of the tropical hydrological cycle inferred from Global Precipitation Climatology Project and International Satellite Cloud Climatology Project data. *J. Geophys. Res.*, **116**, D09101, https://doi.org/10.1029/2010JD015197.

——, W. K. M. Lau, and G. J. Huffman, 2015: Mapping TRMM TMPA into average recurrence interval for monitoring extreme precipitation events. *J. Appl. Meteor. Climatol.*, **54**, 979–995, https://doi.org/10.1175/JAMC-D-14-0269.1.