Representing Extremes in a Daily Gridded Precipitation Analysis over the United States: Impacts of Station Density, Resolution, and Gridding Methods

MELISSA GERVAIS, L. BRUNO TREMBLAY, JOHN R. GYAKUM, AND EYAD ATALLAH

Department of Atmospheric and Oceanic Science, McGill University, Montreal, Quebec, Canada

ABSTRACT

This study focuses on errors in extreme precipitation in gridded station products incurred during the upscaling of station measurements to a grid, referred to as representativeness errors. Gridded precipitation station analyses are valuable observational data sources with a wide variety of applications, including model validation. The representativeness errors associated with two gridding methods are presented, consistent with either a point or areal average interpretation of model output, and it is shown that they differ significantly (up to 30%). An experiment is conducted to determine the errors associated with station density, through repeated gridding of station data within the United States using subsequently fewer stations. Two distinct error responses to reduced station density are found, which are attributed to differences in the spatial homogeneity of precipitation distributions. The error responses characterize the eastern and western United States, which are respectively more and less homogeneous. As the station density decreases, the influence of stations farther from the analysis point increases, and therefore, if the distributions are inhomogeneous in space, the analysis point is influenced by stations with very different precipitation distributions. Finally, ranges of potential percent representativeness errors of the median and extreme precipitation across the United States are created for high-resolution ($0.25°$) and low-resolution areal averaged ($0.9°$ lat $\times$ $1.25°$ lon) precipitation fields. For example, the range of the representativeness errors is estimated, for annual extreme precipitation, to be from $+16\%$ to -12% in the low-resolution data, when station density is 5 stations per $0.9°$ lat $\times$ $1.25°$ lon grid box.

1. Introduction

The impact of climate change on precipitation is of great interest to society given the socioeconomic implications of changes in the distribution of precipitation. It has been theorized that future increases in Earth's temperature will result in a general change in the distribution of precipitation amounts toward fewer lighter precipitation events, more droughts, and more heavy precipitation events (Trenberth et al. 2003). This idea is supported by observations and reanalyses for the historical period (Groisman et al. 2005; Shiu et al. 2012), as well as in predictions of future climate using global climate models (GCMs) (Groisman et al. 2005; Sun et al. 2007). The ability of GCMs to accurately predict future changes in precipitation is crucial for the development of measures

associated with adaptation to climate change. Fundamental to the improvement of GCM prediction is the validation against observational data.

To conduct a fair comparison between simulated and observed precipitation, errors in the observed precipitation field must be quantified. When the observations consist of precipitation station data gridded to the model resolution, there are two main sources of error that must be considered: measurement error and representativeness error due to gridding (Tustison et al. 2001). Several studies have examined errors associated with precipitation gauge measurements and found systematic biases in precipitation on the order of 10% for liquid precipitation (Adam and Lettenmaier 2003) and one order of magnitude larger for solid precipitation (Goodison et al. 1998; Cherry et al. 2007). Representativeness error is defined by Tustison et al. (2001) as "the errors in representing data (i.e., either model output or observations) at a scale other than their own inherent scale." For gridded station data, the representativeness error can be impacted by

Corresponding author address: Melissa Gervais, Department of Atmospheric and Oceanic Sciences, McGill University, 805 Sherbrooke Street West, Montreal QC H3A 0B9, Canada.
E-mail: melissa.gervais@mail.mcgill.ca

the method of gridding employed and the density of stations.

The method used to grid precipitation station data depends on how GCM-simulated precipitation is interpreted, whether it is considered as an areal average over a grid box or as a point estimate (Osborn and Hulme 1997; Chen and Knutson 2008). If grid resolution is high relative to the scale of precipitation features, the two interpretations are identical. However, unlike many other climate variables, precipitation consists of small-scale structures that are discontinuous in nature (Hewitson and Crane 2005). This can result in significant differences between the areal average and point estimate interpretations, which leads to large differences in inferred precipitation statistics (e.g., median and extremes; Accadia et al. 2003; Chen and Knutson 2008). Specifically, Chen and Knutson (2008) found that interpreting precipitation as an area average resulted in generally lower extreme precipitation values and a higher number of wet days than the point value interpretation.

Furthermore, precipitation in GCMs is parameterized (i.e., not explicitly resolved). This parameterization represents smaller-scale structures, such as updrafts and downdrafts, by a single area averaged output (Osborn and Hulme 1997; Chen and Knutson 2008). We thus consider precipitation as an area averaged quantity over a model grid cell similar to previous modeling studies using regional climate models (RCMs) and GCMs (Osborn and Hulme 1997; Tustison et al. 2001; Hewitson and Crane 2005; Chen and Knutson 2008; Gober et al. 2008; Hofstra et al. 2010).

When using station observations for model validation, the method employed to upscale the station data should reflect the consideration of model precipitation as an area average within a grid box. To this end, Hewitson and Crane (2005) and Chen and Knutson (2008) recommend gridding station data to a higher resolution than the model grid using an objective analysis (OA), and subsequently using an area weighted averaging procedure to remap onto the model grid. The purpose of the OA is to create a set of regularly spaced point observations (Hewitson and Crane 2005). This reduces the impact of irregular spacing of station observations, and may be necessary to bridge gaps between locations where the station density is low. These OAs are typically conducted using distance weighted methods, which inherently include some smoothing, even in regions with a high density of observations (Ensor and Robeson 2008; Chen and Knutson 2008). The amount of smoothing that occurs during the OA stage generally affects extremes in precipitation more than the means (Ensor and Robeson 2008; Hofstra et al. 2010). The area weighted averaging procedure takes the area of overlap between the

high-resolution OA grid boxes and the model grid box into account, while conducting the remapping. This results in further smoothing of precipitation, especially extreme events (Chen and Knutson 2008).

In addition to the method of gridding utilized, the density of station measurements can have a large effect on precipitation analyses. In a discussion by Daly (2006), various factors such as elevation, terrain-induced climate transitions, and coastal zones are identified that may influence the ability of objective analysis schemes to accurately portray precipitation. They suggest that regions farther than 100 km from coastlines are easier to represent, whereas regions with significant coastal influence on precipitation or terrain features are more difficult for objective analyses. As such they would require lower or higher station densities, respectively, to accurately represent precipitation.

The impacts of changing station density on gridded precipitation have been studied for different resolution and using different gridding methods (Osborn and Hulme 1997; Kursinski and Zeng 2006; Hofstra et al. 2010; Chen et al. 2008). Hofstra et al. (2010) examined the impact of reducing station density in regions over western Europe for a sample of 10 grid points. They first grid the data onto a higher-resolution 0.1° lat–lon grid using an OA and then remap to a lower resolution, either a 0.22° or 0.44° grid. They conducted repeated gridding of station data for their selected grid points, with decreasing input stations and for many combinations of stations removed. They found that the variance and mean of precipitation typically decreased with reduced station density. Chen et al. (2008) conduct an intercomparison of objective analyses of station data within the United States onto a 0.5° lat–lon grid using several objective analysis methods. They also examine the impacts of reducing the percent of input stations employed in the analysis. They find an increase in errors in the aggregate statistics across the United States, with decreased percent station inclusion, that is higher in the summer than the winter.

Osborn and Hulme (1997) and Kursinski and Zeng (2006) conducted similar studies but at a lower resolutions [e.g., 2.5° lat–lon in Kursinski and Zeng (2006)] and using a simple average of stations within the grid box to compute the area average. Using this method, Osborn and Hulme (1997) found that the variance of daily precipitation in western Europe, China, and Zimbabwe increases with decreased input stations. Kursinski and Zeng (2006) used hourly station data in Ohio and found similar results to Osborn and Hulme (1997). They observed that the average precipitation amount per hour varied more widely and with generally higher precipitation rates as the number of input stations decreases. The
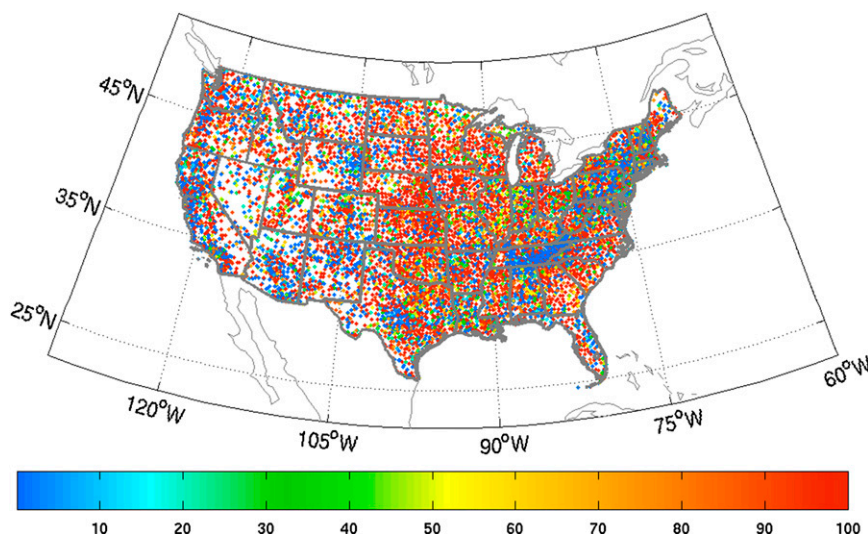
FIG. 1. Average reporting rate (%) for each GHCN station.

reduction of station density in these studies thus had an effect opposite to that found by Hofstra et al. (2010), who conducted their study at both a different resolution and using a different gridding method. These studies leave open questions of how gridding method, resolution, and region of study could impact the relationship between precipitation statistics and station density.

The central purpose of this study is to provide a quantitative assessment of the representativeness errors associated with gridded station data, used in particular for the validation of GCMs. The United States provides a useful test bed for such studies as it has a high density of stations and encompasses a wide range of climate regimes. Station data within the United States are used to examine precipitation statistics at several scales, from point measurements to OA high-resolution grids to area averaged low-resolution grids. This repeats the experiment conducted by Chen and Knutson (2008) but at a higher resolution for the area average grid, typical of the current generation of GCMs, and including the representativeness error during the transition from station to OA data. Furthermore, the impact of station density is assessed for both the high-resolution OA and low-resolution areal-average precipitation data. This is accomplished by conducting an experiment of successively gridding station data within the United States with a decreasing number of input stations. The purpose of this experiment is to provide a measure of the potential representativeness errors due to station density. We will investigate how the relationship between station density and precipitation errors depends on seasonality, characteristic length scales of precipitation, and geographic location. The methodology in this study is specifically geared toward examining the representativeness errors in gridded precipitation data

for the purpose of model validation; however, the results may be used to understand the impact of gridding methods and station density for any application of OA or areal-averaged precipitation. In another study, we follow up on this work with an application of these results in a study on errors in the distribution and extremes of precipitation in a GCM (Gervais et al. 2014).

2. Data

We use daily precipitation station data from the Global Historical Climatology Network–Daily version 1.0 (GHCN) dataset, from the National Climatic Data Center. This dataset consists of over 30 000 stations worldwide, recording temperature and precipitation. Extensive quality control procedures have been applied to the data to address issues such as formatting, duplicate stations, and outliers (Menne et al. 2012), and no further quality controls are applied in this study. For our analysis, stations from the contiguous United States are used (over 10 000 stations) over the time period of 1979–2003. The reporting times of each station vary depending on the data source (Menne et al. 2012). All stations are used regardless of their reporting rates in order to maximize the information ingested. On average, 41% of the total number of stations are reporting daily; the percentage of time for which each station is reporting over the period of study is shown in Fig. 1.

3. Methodology

a. Station decorrelation lengths

The decorrelation length is used in this study to provide a measure of the length scale of precipitation

processes. Specifically, it is defined as the distance at which the correlation between a given location and those at this distance away falls to $1/e$ (Osborn and Hulme 1997). It is produced for each station by computing the distance and correlation with all other stations using a Kendall's tau rank method, which is valid for non-normally distributed fields (Wilks 2006). Following Osborn and Hulme (1997), an exponential of the form

$$r = e^{-(x/x_0)}$$

is then fit to the correlation versus distance data for each station, where r is the interstation correlation, x is the distance between the stations, and x_0 is the decorrelation length. Maps of the decorrelation lengths of all stations are then created to demonstrate how the decorrelation length varies geographically.

b. Gridding methods

The GHCN station data are first gridded to a 0.25° lat–lon grid. This grid is chosen to be consistent with the Climate Prediction Center's Daily Unified Precipitation Data (UPD) (Chen et al. 2008), a widely used gridded station precipitation product for the United States created using a more sophisticated OA method. This high-resolution gridded GHCN precipitation field (HRES) is constructed by conducting an OA on the GHCN station data using a three-pass Cressman scheme (Cressman 1959), with smaller radii of influence and successive corrections at each pass. The three radii of influence in the Cressman scheme are 6, 3, and 1.5 times the average minimum station distance or on average 120, 60, and 30 km when all available stations are employed. This method was chosen over the optimal interpolation scheme used in the creation of the UPD in an effort to reduce computational costs. For reasons discussed in section 3c the interpolation is conducted numerous times, so a simpler method is preferred. The Cressman scheme, however, is less accurate than the optimal interpolation method (Chen et al. 2008) and does not include orographic adjustments. Precipitation stations tend to be located at lower elevations and precipitation tends to increase with elevation. This typically results in a bias toward lower precipitation amounts in mountainous regions, and so many gridded gauge analyses conduct an orographic adjustment to account for this effect (e.g., Xie et al. 2007; Hutchinson et al. 2009). The spatial patterns in precipitation statistics are similar between the HRES gridded precipitation and UPD datasets (not shown), implying that the Cressman OA scheme is adequate for our purposes.

The HRES data are used to create low-resolution precipitation fields on a 0.9° × 1.25° lat–lon grid,

a common resolution for current GCMs, consistent with the treatment of precipitation data for GCM validation. Two methods of transformation of the HRES data are utilized to be consistent with either the point or area average interpretation of precipitation. For the point interpretation, a simple bilinear interpolation is applied to the grid nodes of the HRES data to produce a low-resolution product (LRES-interp). For the area average interpretation, the HRES data is remapped using the Spherical Coordinate Remapping and Interpolation Package (SCRIP) from the Los Alamos National Laboratory (Jones 1999). SCRIP is a flux-conserving method that computes weights for each input grid based on the area overlap between the input grid and the output grid. As discussed in the introduction, the area average interpretation of GCM precipitation is considered to be the most appropriate. The low-resolution gridded GHCN precipitation created through remapping, which is hereafter called LRES, is thus considered the better product for GCM validation. A schematic diagram of the transformation of the data from station to HRES, to LRES and LRES-interp products is shown in Fig. 2.

Following Chen and Knutson (2008), precipitation statistics are computed after the OA and remapping procedures are applied. All statistics are calculated bimonthly and annually, and then averaged over all years. The bimonthly periods used are January–February (JF), March–April (MA), May–June (MJ), July–August (JA), September–October (SO), and November–December (ND). We define precipitating days as those with $>1\,\mathrm{mm\,day^{-1}}$ of precipitation, consistent with the World Climate Research Programme/Climate Variability and Predictability (WCRP/CLIVAR) Expert Team on Climate Change Detection, Monitoring, and Indices (ETCCDMI) group. This value is arbitrary and is larger than the minimum detectability threshold of a rain gauge; however, it is employed in many other studies (e.g., Dai 2006; Sun et al. 2006; Chen and Knutson 2008). Some results using $0.25\,\mathrm{mm\,day^{-1}}$ as the threshold to define precipitating days are discussed for comparison; however, results are shown using the $1\,\mathrm{mm\,day^{-1}}$ threshold unless otherwise specified. We compute the median and the 97th percentile (herein referred to as extreme) of precipitating days as metrics of the non-Gaussian distribution of precipitation.

c. Station density experiment

An experiment is conducted to determine the impact of reducing station density on the statistics of gridded precipitation products. The experiment consists of producing HRES and LRES fields for the entire time period (following the methodology in section 3b) and calculating the statistics of these fields (median and extreme),
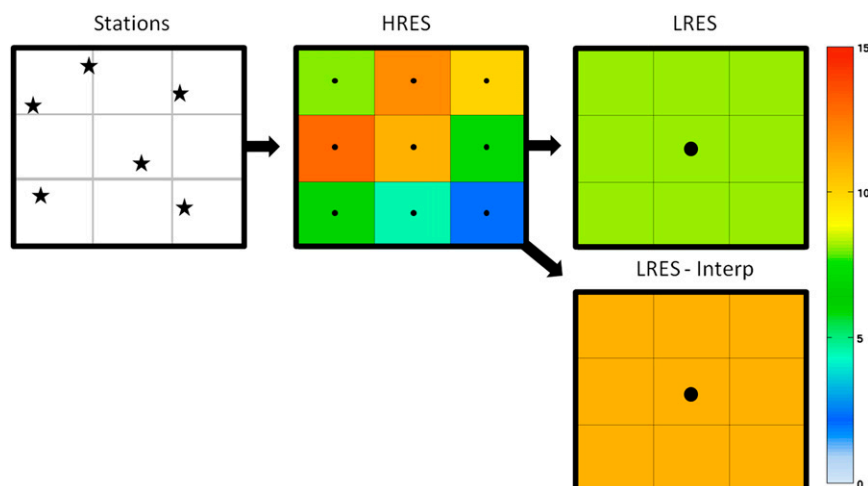
FIG. 2. Schematic diagram displaying a simplified view of the grid transformation between station, HRES, LRES, and LRES-interp grids. Stations are shown as stars, HRES grid boxes as gray lines, HRES grid points as small black circles, LRES/LRES-interp grid boxes as thick lined boxes, and LRES/LRES-interp grid points as larger black circles. Grid box precipitation amounts are shown in color and are taken from a sample of actual data during a precipitation event. Note that in this simplified example the HRES and LRES grid points are collocated and 9 HRES grid boxes fit inside 1 LRES grid box. The LRES procedure therefore consists of averaging the HRES grid values. The LRES-interp method therefore simplifies to assuming the value of the middle HRES grid node.

using subsequently fewer input stations. The initial number of input stations is shown in Fig. 3. This reduction process is repeated 20 times, each time successively removing a randomly chosen set of stations, amounting to 5% of the initial number of stations. Assuming the distribution of precipitation is well represented during the first step of the experiment, utilizing 100% of the total stations, then any deviation from the initial value of a precipitation metric (median or extreme) during subsequent steps represents a climatological error in the precipitation metric resulting from a change in station density.

We are interested in characterizing the representativeness errors associated with a given station density. To this end, we represent the station density for the LRES data as simply the number of stations within the LRES grid box. For the HRES data, however, the radii of influence of the OA are larger than HRES grid boxes, implying that stations outside of a grid box influence the analysis. Furthermore, a large portion of HRES grid boxes contain few or no input stations. Consequently, a larger area is chosen for the calculation of the HRES station density. We defined the HRES data density as the number of stations within a box of the same dimension as an LRES grid box ($0.9°$ lat $\times$ $1.25°$ lon), but centered on the HRES grid points. Choosing the same area as the LRES density calculations allows for direct comparison between impacts of station density on the HRES and LRES fields. To take

station reporting rates into consideration, the densities of stations reporting each day are calculated, averaged over the time period of interest (bimonthly or annually), and then averaged climatologically.

Since station density is highly variable throughout the domain, the change in station density for each percent removal step varies considerably across the United States. Therefore, results are presented as a function of the station density at each grid box, as opposed to the percentage of stations removed. This allows for the association of climatological errors at a given removal step, for all grid points, with their station density.

In studies on the impact of reduced station density on gridded precipitation, Osborn and Hulme (1997), Kursinski and Zeng (2006), and Hofstra and New (2009) found that errors are dependent on the combination of stations removed; in other words, there is a spread in the distribution of errors when various combinations of station removals are conducted. Unlike these studies in which station density experiments were conducted within a small sample of grid boxes, our study is conducted using an OA technique over a large domain. This method is more computationally intensive, and therefore the experiment here is not repeated for the many possible combinations of station removals, as done in Osborn and Hulme (1997), Kursinski and Zeng (2006), and Hofstra and New (2009). Instead, we define the percent climatological errors of each grid point as the
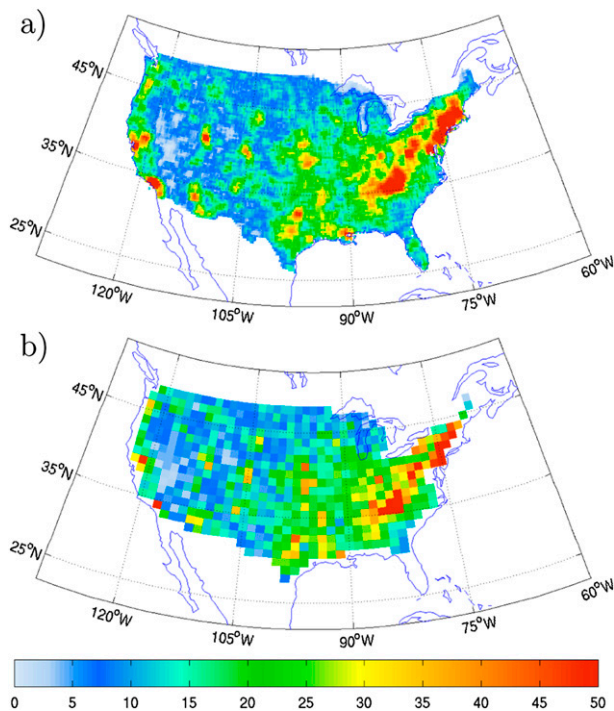
a)

b)

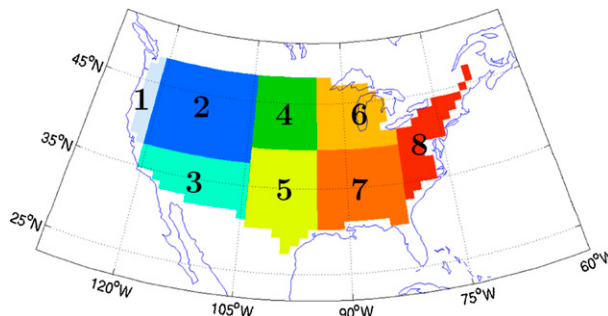FIG. 3. Initial number of stations per grid box for the (a) HRES and (b) LRES data.

FIG. 4. Map of regions in the United States: 1) West Coast, 2) Rockies, 3) North American monsoon, 4) northern Great Plains, 5) southern Great Plains, 6) Great Lakes, 7) Gulf, and 8) East Coast.

percent difference between the initial value of a precipitation metric and the value at subsequent steps. This normalized measure of climatological error allows for the intercomparison of grid boxes across a region, which is used to create a distribution of errors analogous to one produced when varying combinations of stations removed are compared. This method implicitly assumes that error structures are similar between grid boxes, which is not necessarily true. However, if we are concerned with a general definition of climatological errors that can be applied to the entire domain, then it is advantageous to take into account all of the potential error responses within a given region.

Distributions of errors with respect to station density are produced by employing all grid boxes within the United States, or different subregions of the United States, to create scatterplots of percent climatological error. This is repeated for the median and extreme precipitation, for the LRES and the HRES datasets, and for different periods. The domain was separated into regions with similar precipitation climatologies (Fig. 4), assuming that the dependence of the errors on station removal is similar within a climatological region. Climatological errors are included for all grid points regardless of the initial station density to maximize the information that is being ingested and to avoid biasing the results to more populated regions. The concentration of

computed climatological percent error points is represented by coloring each error point to correspond to the number of points occurring within 1% error bins. There is a high density of points at 0% climatological error, even at low station density; this is due to grid boxes with lower station density being normalized, and then only accruing errors slowly with further removal of stations. This could also imply that in regions with low station density, such as the Rockies, the climatological percent errors may be underestimated. Upper and lower bounds on the climatological percent error for the United States are defined by exponential curves of the form

$$y = ae^{bx} + ce^{dx}$$

fitted to the 99th and 1st percentiles of the error distribution. These percentiles and fits were chosen to best represent the outer limits of the climatological error distribution, while also excluding outliers. The coefficients of determination (R^2) of these fits are very high (typically $R^2 > 0.9$), supporting the use of the exponential fit.

4. Results

a. Impacts of gridding on precipitation statistics

In this section, we investigate how the gridding of station data onto a high-resolution grid and remapping onto a lower-resolution grid (typical of a GCM) alters the statistics of precipitation. In addition, we quantify the impact of interpreting model precipitation as an area average versus a point estimate. This is accomplished through the intercomparison of the median and extreme precipitation of GHCN data in various forms, from original station data through to various gridded products (HRES, LRES-interp, and LRES).

Annual station precipitation climatologies show a wide range of median (4–15 mm day^{-1}) and extreme (25–80 mm day^{-1}) values across the United States
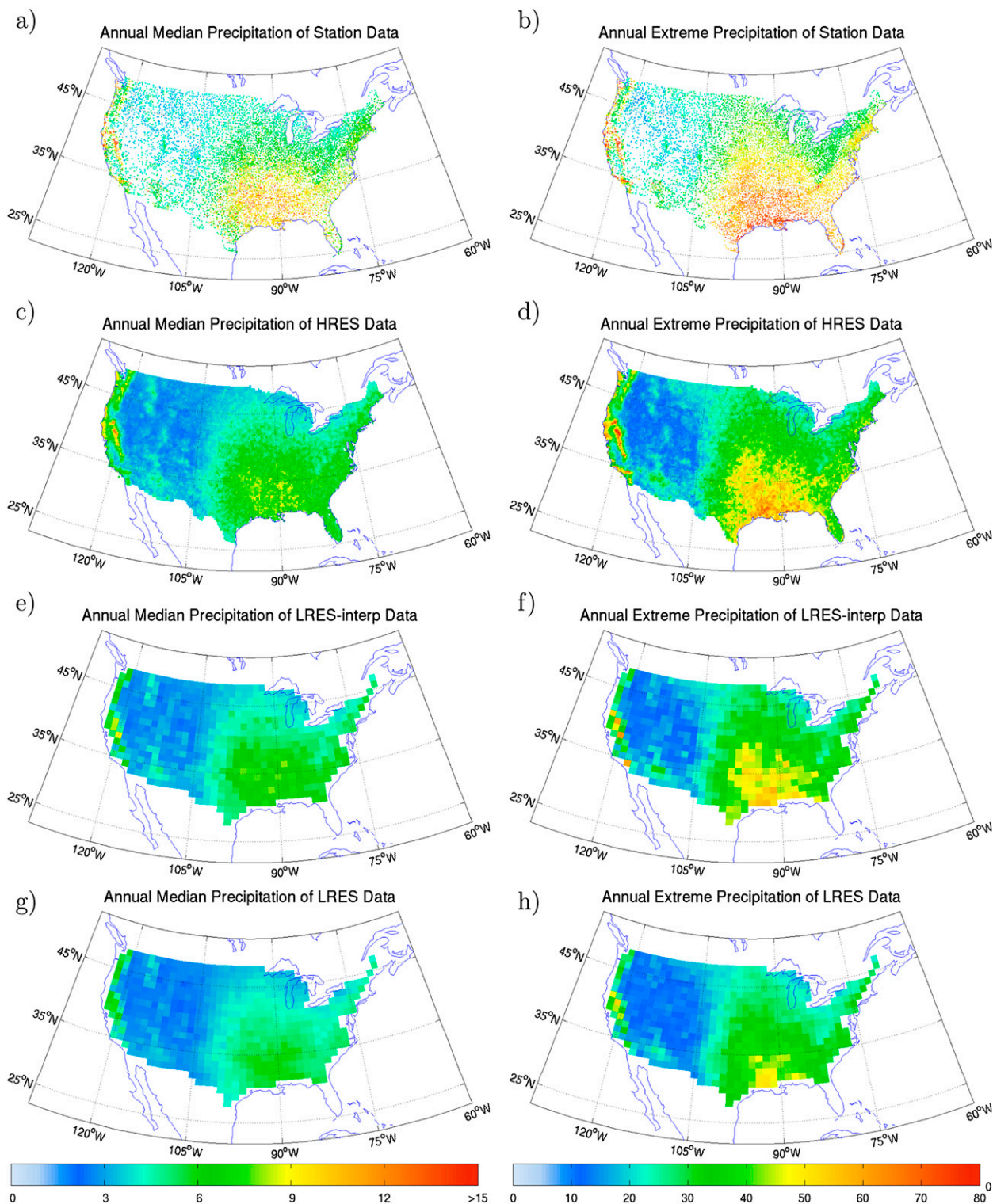
FIG. 5. Average annual (left) median and (right) extreme precipitation (mm day^{-1}) calculated at each station or grid box for (a),(b) the GHCN station, (c),(d) HRES, (e),(f) LRES-interp, and (g),(h) HRES data.

(Figs. 5a,b). Regions with heavy precipitation are present along the West Coast as well as in the Cascade and Sierra Nevada mountain ranges. In the Rocky Mountains lower median and extreme precipitation are recorded. A large area of high precipitation in the southeast United States and the Eastern Seaboard is seen in both the median and extreme precipitation. There is a defined region of high precipitation east of the Appalachians, especially in the extreme precipitation. When the threshold used to define a precipitating day is reduced from 1 to $0.25\,\mathrm{mm\,day^{-1}}$, the median value of precipitation is reduced at all resolutions (by approximately 30%) but the impact on the extreme precipitation is minimal (not shown).

The HRES data show a marked decrease both in the median and extreme precipitation values at nearly all locations in comparison to the station data (Figs. 5a–d). This decrease in extreme precipitation with objective analysis is consistent with the results of Ensor and Robeson (2008). However, Ensor and Robeson (2008) did not find significant differences in the mean precipitation between the selected stations and their closest analyzed point. We find that the mean precipitation (not shown) behaves similarly to the median precipitation (Figs. 5a,c), suggesting that use of the median instead of the mean is not the cause of the discrepancy between our study and that of Ensor and Robeson (2008). One explanation is that Ensor and Robeson (2008) only compared stations that were in close proximity to grid points, and consequently measured the smallest errors possible between a station and an analyzed point. Their study also only included the Midwest, which our results show has smaller changes in the median than other regions of the United States.

In general, median and extreme precipitation are higher in the LRES-interp than in the LRES (Figs. 5e–h), with differences ranging from 0% to 30% (Fig. 6). These differences are solely attributed to the interpretation of a model grid box being a point estimate or area average respectively, since both low-resolution fields are derived from the same HRES data. These results are in agreement with Chen and Knutson (2008), who also examined the impact of interpolation and remapping on extreme values but at a lower resolution. They show that the 5-yr and 50-yr return period values of daily precipitation were smaller when using an interpolation method as opposed to a remapping method, where the return period is defined here as the daily amount of precipitation that is expected to occur only once every 5 and 50 years.

These results are important to consider when validating GCM output against station observations. Differences between the station value and the LRES median and extreme precipitation can be as large as 50%. This exemplifies why direct comparison between station data and GCM output is inappropriate because
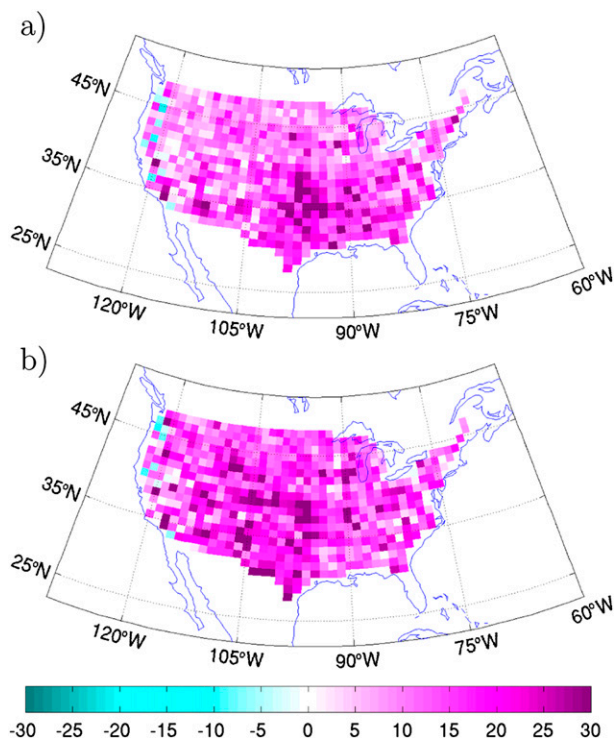

FIG. 6. Percent difference in the average annual (a) median and (b) extreme precipitation between the LRES-interp and LRES fields.

of the smoothing that occurs during the spatial transformation. The minimum value to define a precipitating day is also an issue across scales, as it is easier to attain at the station level than averaged over an entire grid box. Furthermore, any change in precipitation in a GCM could represent a much larger change at a point location. We also show the importance of the interpretation of model data as either a point value or an area average. In our subsequent analysis we will use the area averaged interpretation. As discussed previously, model precipitation is often parameterized and dependent on fluxes across grid boundaries, and as such we believe it is best represented as an area average within a grid box, in keeping with Chen and Knutson (2008).

b. Impacts of station density

In this section, we present the results of our experiment on the impact of reducing station density on the statistics of HRES and LRES precipitation fields. Distributions of climatological errors with respect to station density are produced by creating scatterplots of the percent climatological errors of all the stations within various regions of the United States. Climatological errors in the HRES and LRES data exhibit similar behavior, but in general the HRES data (Fig. 7) exhibit
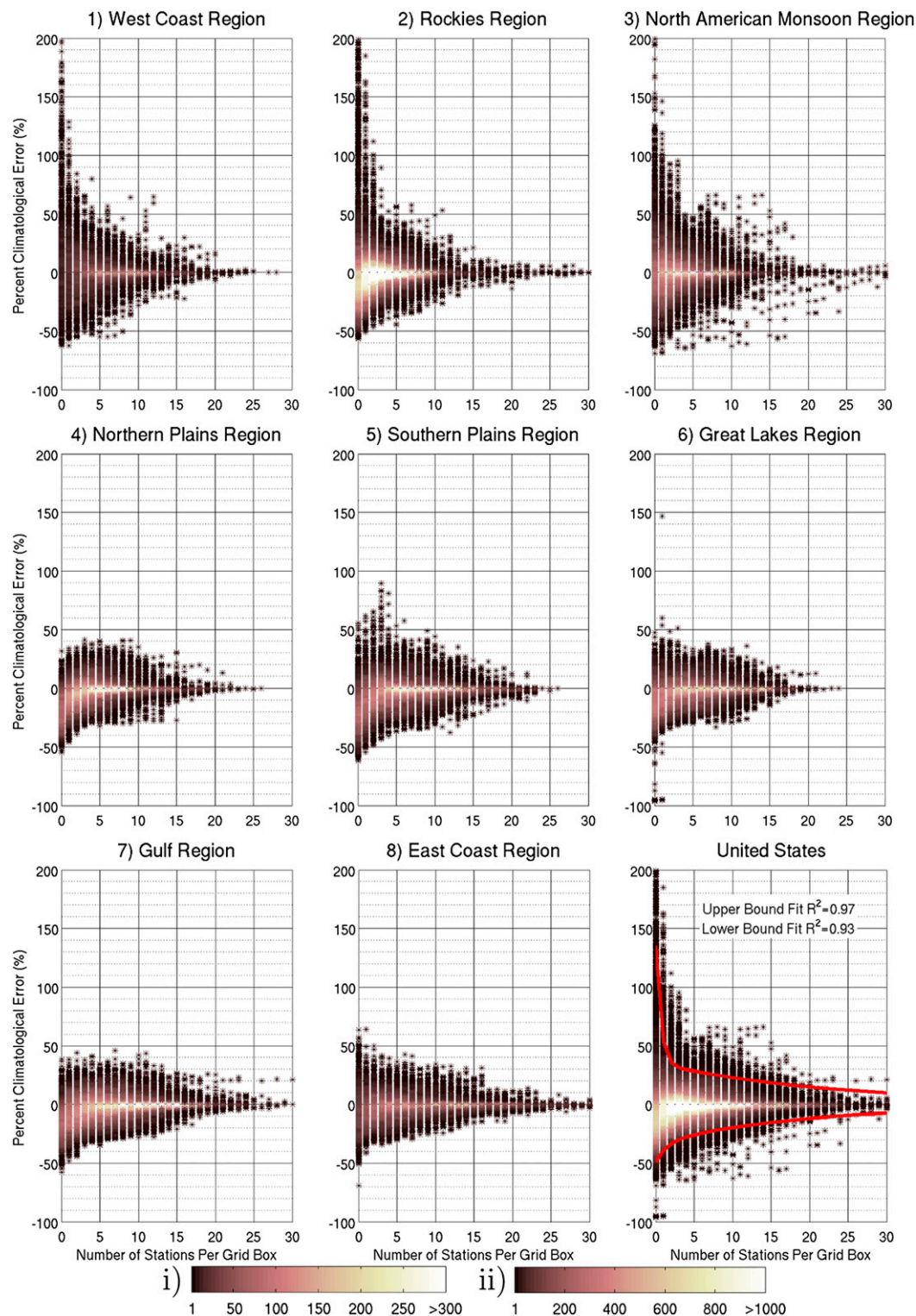
FIG. 7. Percent climatological error of annual extreme precipitation (1979–2003) for all HRES grid boxes in a region and removal steps, as a function of station density (number of stations per 0.9° × 1.25° box). The color of the symbols represents the concentration of climatological error points within 1% error bins, for a given station density. The corresponding color bars are for (i) the regions and (ii) the United States. Exponential fits are applied to the 1st and 99th percentiles of the U.S. distributions (red lines) and the coefficients of determination (R^2) of the fits are displayed.

larger percent errors than the LRES data (Fig. 8), as evidenced for the extreme precipitation errors. Unlike the errors in the LRES data, the HRES errors are often large even when the initial station density is high. This implies that there are many locations where the GHCN data do not have an adequate station density to represent extreme precipitation with the HRES product. The LRES gridded data, however, are less sensitive to data density due to area averaging. Results are similar for the median precipitation; however, the climatological errors are smaller than in the extreme precipitation for both HRES and LRES fields (not shown). The larger impact of station density on extreme precipitation than on median precipitation seen here is in keeping with observations in other studies that smoothing has a large impact on extreme values (Ensor and Robeson 2008; Hofstra et al. 2010; Chen and Knutson 2008).

The shapes of the climatological error distributions can be broadly separated into two categories. The first is characterized by errors that initially grow at higher station density but remain bounded at lower station density, hereinafter referred to as a bounded response to decreasing station density. This distribution shape is found in the central and eastern United States consisting of the northern and southern Great Plains, Great Lakes, Gulf, and East Coast regions (Figs. 7 and 8). The second distribution shape is an exponential increase with decreasing station density (exponential response), which is found in the western United States consisting of the West Coast, Rockies, and North American monsoon regions (Figs. 7 and 8). These responses to decreased station density are both prominent in the HRES (Fig. 7) and LRES (Fig. 8) data even with the smoothing involved in the LRES data. These results are consistent with Daly (2006), who suggests that regions such as the western United States where the coast or complex terrain influence precipitation will be more difficult to represent with objective analysis schemes. The two types of error distributions also have different seasonalities. For instance, the Gulf region has larger percent errors in the JA than the JF period, while the Rockies region shows the opposite seasonality (Fig. 9). The Gulf and Rockies regions are representative of all regions in the eastern and western United States respectively (not shown).

The shape and seasonality of the error distribution are further investigated using the decorrelation length scale of precipitation. The decorrelation lengths are longer in JF than in JA (Fig. 10). In the JF period, the decorrelation lengths are longer in the east and along the West Coast than in the central United States, ranging from approximately 500 to 200 km respectively. The longer decorrelation lengths coincide with regions that experience more synoptic-scale winter precipitation systems. The decorrelation lengths are generally shorter in the summer with longer lengths to the north (~250 km) than the south (~100 km). This is consistent with the northward movement of the storm track in the summer, resulting in more synoptic-scale systems to the north while the south is more prone to air mass convection. In general, the difference in spatial gradient in decorrelation lengths is smaller in JA than in JF.

Geographic differences in the decorrelation length scale were also noted by Osborn and Hulme (1997) in western Europe. For instance, they show that the decorrelation length scales in France (400–480 m) were 4 times that in northern Italy (80–160 m). Decorrelation lengths were found to be longer in the winter compared to the summer across all of Europe (Osborn and Hulme 1997; Hofstra and New 2009), which is in accord with results presented here for the eastern United States. This was attributed to the predominance of larger-scale precipitation systems in the winter and smaller-scale convective systems in the summer (Osborn and Hulme 1997; Hofstra and New 2009). Furthermore, Hofstra and New (2009) examined the relationship between synoptic typing and decorrelation length, which further demonstrated that the presence of synoptic-scale forcing leads to longer decorrelation lengths, consistent with this seasonal dependence.

Chen et al. (2008) examined the impact of station density on the relative biases in correlations between a set of withheld stations and gridded station datasets, using different objective analysis methods. They withheld 10% of the initial number of input stations for cross-comparison, while the remaining stations were gridded several times with systematic removals of input stations, using the different objective analysis methods. Each withheld station was then cross-compared to the nearest grid point in the analyses with decreasing input stations. They found that their biases increased as station density decreased, and that this effect was highest in the summer season. In our study, we see two different seasonal responses in precipitation statistics depending on the region of study, whereas they examine an average over the entire United States. Since their withheld stations are randomly chosen and there are significantly more stations located in the eastern United States, their verification set is likely biased toward the eastern United States. This would explain the agreement with our results for the eastern United States, as they likely saw a predominantly eastern U.S. response. Our results are also independent of the large differences in station density across the United States because we examine errors with respect to station density as opposed to percent input stations.
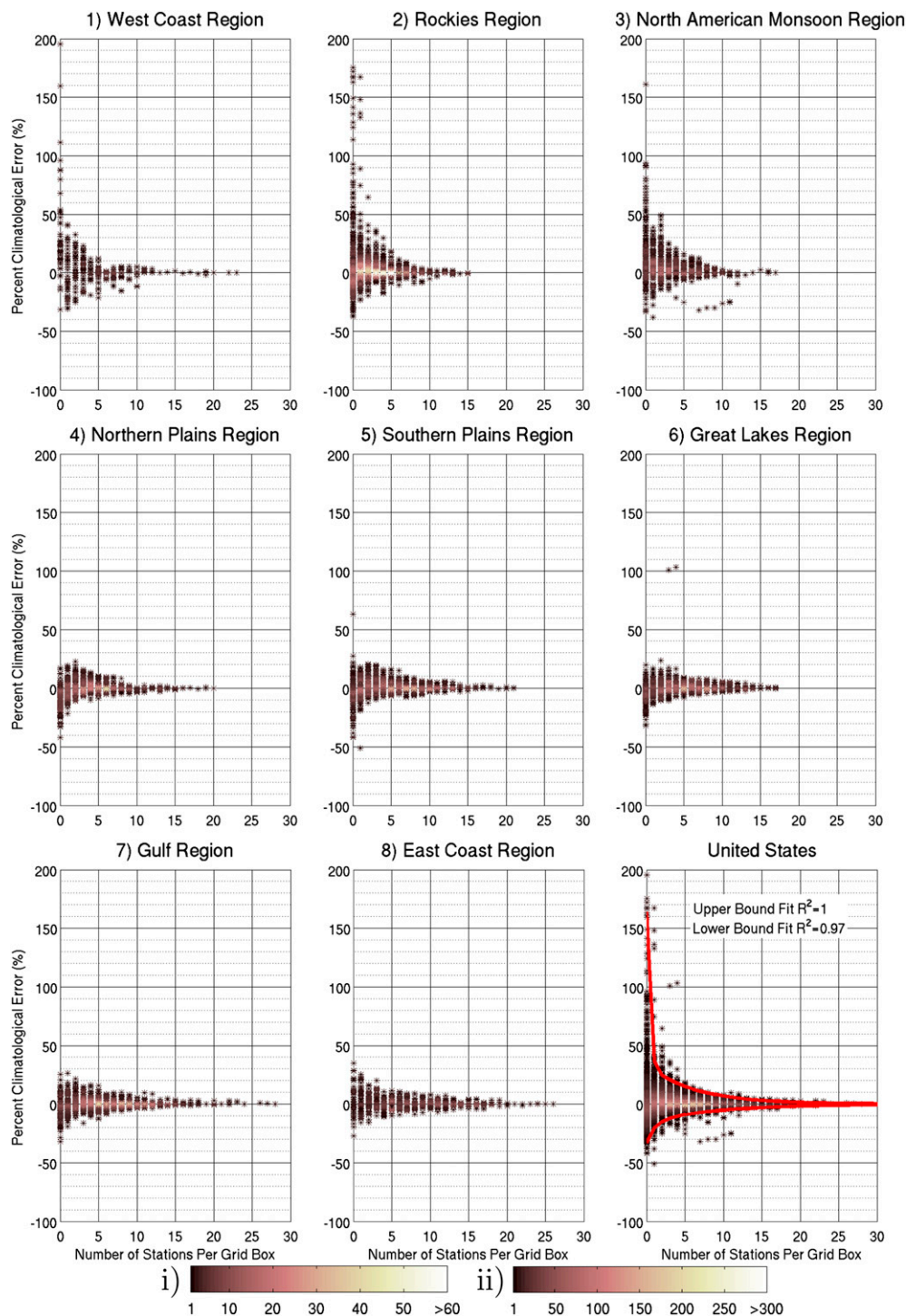
FIG. 8. Percent climatological error of annual extreme precipitation (1979–2003) for all LRES grid boxes in a region and removal steps, as a function of station density (number of stations per 0.9° × 1.25° box). The color of the symbols represents the concentration of climatological error points within 1% error bins, for a given station density. The corresponding color bars are for (i) the regions and (ii) the United States. Exponential fits are applied to the 1st and 99th percentiles of the U.S. distributions (red lines) and the coefficients of determination (R^2) of the fits are displayed.
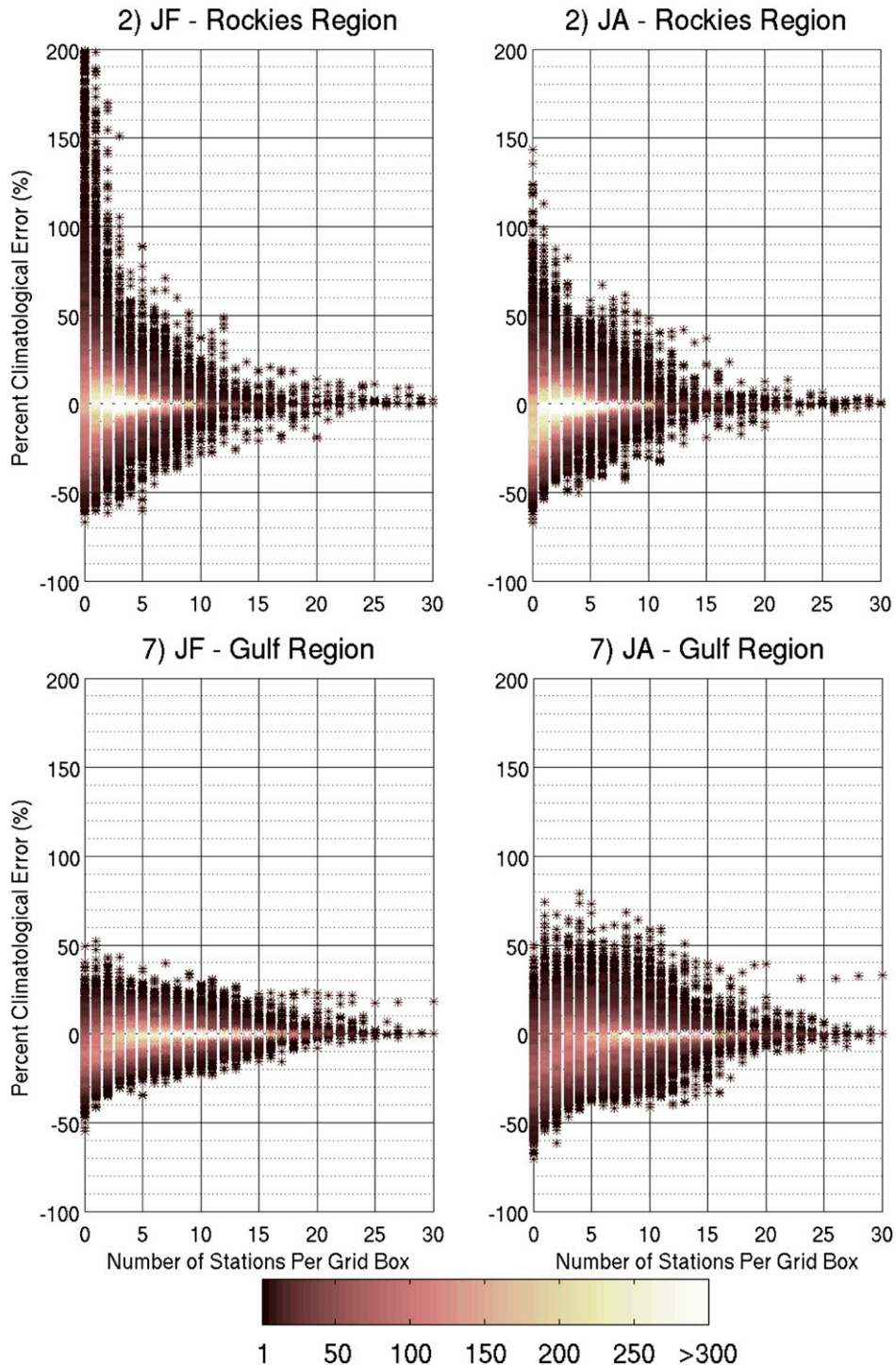
FIG. 9. Percent climatological error of JA and JF extreme precipitation (1979–2003) in the Rockies and Gulf regions, for all HRES grid boxes in a region and removal steps, as a function of station density (number of stations per $0.9° \times 1.25°$ box). The color of the symbols represents the concentration of climatological error points within 1% error bins, for a given station density.
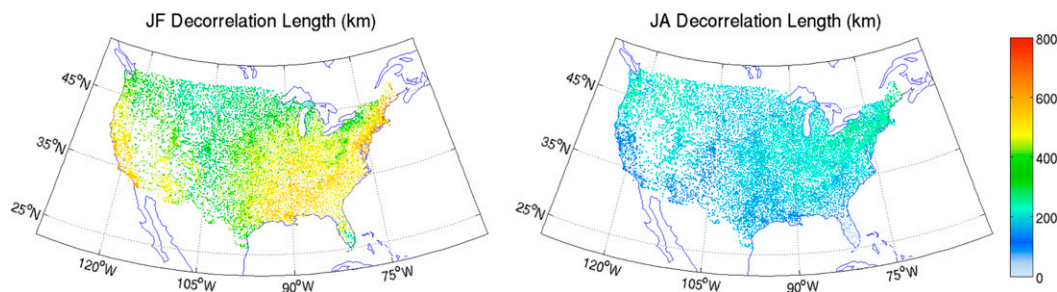
FIG. 10. Station decorrelation lengths (km) for all stations within the United States for both the JF and JA periods.

The central goal of the station density experiment was to determine the range of potential representativeness errors in gridded station data related to station density. Considering the more general case of the annual climatological error over the entire United States, we use exponential fits applied to the 1st and 99th percentiles of the error distributions (red line in Figs. 7 and 8, produced as described in section 3c) to obtain an estimate of the lower and upper error bounds versus station density respectively. A table of these values of the upper and lower bounds of percent error for given station density is provided for median and extreme precipitation, and for the HRES and LRES grids (Table 1). These results were duplicated using a lower minimum threshold to define a precipitating day of 0.25 mm day^{-1}. There are relatively small differences when using the 0.25 mm day^{-1} instead of the 1 mm day^{-1} threshold, with somewhat larger magnitudes of errors and similar behaviors of representativeness errors with respect to station density (Table 2).

Using the initial station density across the United States (Fig. 3), maps of the upper and lower error bounds at each grid box are created (Figs. 11 and 12). The median climatological errors, in both the HRES (Figs. 11a,b) and the LRES fields (Figs. 12a,b), are typically lower than the extreme climatological errors (Figs. 11c,d and 12c,d, respectively). In general, climatological errors in median and extreme precipitation are higher in the HRES (Fig. 11) than the LRES (Fig. 12) data. This is expected as the area averaging in the LRES data tends to reduce climatological errors. The magnitude of the upper bound of climatological errors (Figs. 11a,c and 12a,c) tends to be higher than the lower bound of climatological errors (Figs. 11b,d and 12b,d) at all resolutions and for both the median and extreme precipitation, indicating a tendency toward positive climatological errors in precipitation. Biases in LRES precipitation due to inadequate station density for the median (Figs. 12a,b) and extreme (Figs. 12c,d) precipitation can range from as low as 0% in well-sampled regions in the East to as high as 50% in the

TABLE 1. Table of upper and lower bounds of percent errors in median and extreme precipitation due to station density, over the entire United States. Values are taken from exponential fits applied to the outer limits of the distribution of errors with decreasing station density, for both the HRES and LRES grids. The fits for the extreme precipitation are shown as red curves in Figs. 7 and 8, for the HRES and LRES fields respectively. Station density is defined as the number of stations within a 0.9 × 1.25° grid box.

	HRES error (%)				LRES error (%)			
	Median		Extreme		Median		Extreme	
Station density	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
0	79	−34	135	−49	87	−18	164	−33
1	37	−30	54	−39	28	−14	36	−21
2	27	−28	36	−34	21	−12	25	−15
3	24	−26	31	−30	18	−10	21	−12
4	23	−25	30	−28	16	−8	18	−10
5	22	−24	28	−26	14	−7	15	−9
6	22	−22	27	−25	13	−6	13	−8
7	21	−21	26	−23	11	−5	11	−7
8	21	−20	25	−22	10	−5	9	−6
9	20	−19	24	−21	9	−4	8	−6
10	19	−18	23	−20	8	−4	7	−5
15	17	−14	18	−15	4	−2	3	−3
20	15	−11	15	−12	2	−1	1	−2
25	13	−8	12	−9	1	−1	1	−1
30	12	−7	10	−7	1	0	0	−1

TABLE 2. As in Table 1, but using a smaller threshold of 0.25 mm day^{-1} to define a precipitating day.

Station density	HRES error (%)				LRES error (%)			
	Median		Extreme		Median		Extreme	
	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower
0	99	−41	153	−50	99	−23	178	−34
1	51	−37	59	−40	39	−17	45	−21
2	37	−35	39	−35	31	−14	27	−15
3	33	−34	33	−31	27	−11	22	−12
4	31	−32	31	−29	24	−9	19	−10
5	30	−30	30	−27	21	−8	16	−9
6	29	−29	28	−26	19	−7	14	−8
7	28	−28	27	−24	17	−6	12	−7
8	27	−26	26	−23	15	−5	10	−7
9	27	−25	25	−22	13	−5	9	−6
10	26	−24	24	−21	11	−4	8	−5
15	22	−19	19	−16	6	−2	4	−3
20	19	−15	16	−12	3	−1	2	−2
25	16	−11	13	−9	2	−1	1	−1
30	14	−9	10	−7	1	−0	0	−1

poorly sampled Rocky Mountains. The lower and upper error bounds tend to be dominated by the larger errors found in the western United States at lower station density; however, this has a small impact on the results because the initial station density is higher in the eastern United States.

5. Discussion and conclusions

This study explores the representativeness errors of gridded precipitation data through the changes in precipitation statistics as station data are gridded. We observe a dramatic decrease in median and extreme precipitation as station data are upscaled to the high-resolution (HRES) objectively analyzed (OA), low-resolution interpolated (LRES-interp), and low-resolution remapped (LRES) fields. This implies that even if a GCM were to perfectly represent areal averaged precipitation within model grid boxes, its median and extreme precipitation would be lower than that of a station measurement due to representativeness errors. This is an important factor when using future climate predictions from GCMs to determine the societal implications of climate change, as society experiences precipitation at a point location as opposed to an area averaged region.

The interpretation of a model grid as a point value or an area average across a GCM grid box can have large impacts on the resulting precipitation statistics. The point value assumption generally leads to larger median and extreme values than the area average assumption, with differences reaching 30%. These results are consistent with Chen and Knutson (2008), but in this analysis it is repeated at a resolution typical of the current generation of GCMs. This has significant consequences for GCM validation,

demonstrating the importance of the methods used to upscale station data to GCM resolutions. We advocate objectively analyzing to a higher resolution followed by remapping to a lower resolution to upscale station data for comparison with model output, in agreement with others (Hewitson and Crane 2005; Chen and Knutson 2008). This is consistent with the area average view of a GCM precipitation output.

Climatological errors resulting from low station density are examined for different regions of the United States. Two characteristic climatological error responses to decreasing station density depending on the homogeneity of station precipitation distributions within the radius of influence are identified and can be broadly geographically separated into the eastern and western United States. Climatological errors in the eastern United States begin at higher station densities but do not grow exponentially and in general have a small negative bias. The error structure and seasonality in the western United States are different from those of the eastern United States. As station density decreases, the upper and lower bounds on climatological errors grow exponentially in both positive and negative directions. Furthermore, these two error responses exhibit differing seasonalities: in the eastern (western) United States the percent error is greater in the JA (JF) period.

In a previous study by Bussières and Hogg (1989), it has been shown that decreased distance between stations and OA grid points results in decreased OA errors. How this translates to climatological errors in precipitation distribution, however, is not straightforward. In an OA scheme there will always be an element of smoothing due to the influence of neighboring points. This smoothing is reduced as the proximity of stations to
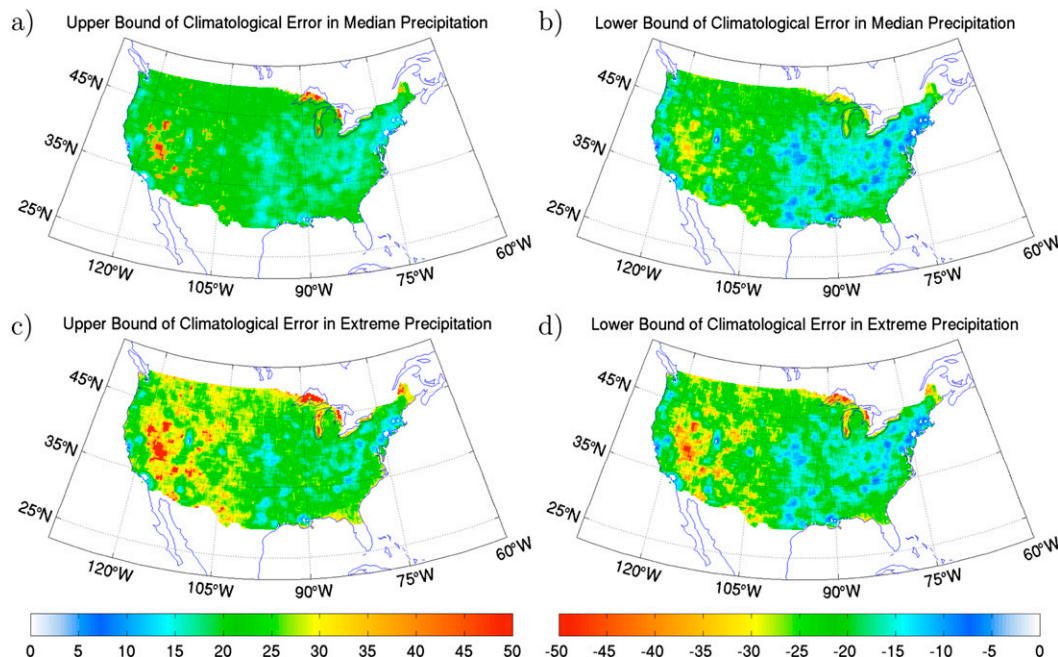
FIG. 11. (left) Upper and (right) lower bound on the percent climatological error in average annual (a),(b) median and (c),(d) extreme of precipitation (1979–2005) for HRES data using the exponential fits of the 99th and 1st percentiles. Note that the color scales are reversed between the upper and low bound maps such that the magnitudes of the color schemes are identical but in opposing directions.

the analysis point increases and the OA point is closer to the true precipitation field. We propose two conceptual frameworks to explain the observed impact of station density on the climatological average of OA precipitation statistics. The first will be applicable to the

entire United States and the second solely for the western United States.

In the first conceptual framework, we assume that the distribution of precipitation is relatively homogeneous. This is the case in the eastern United States, as evidenced
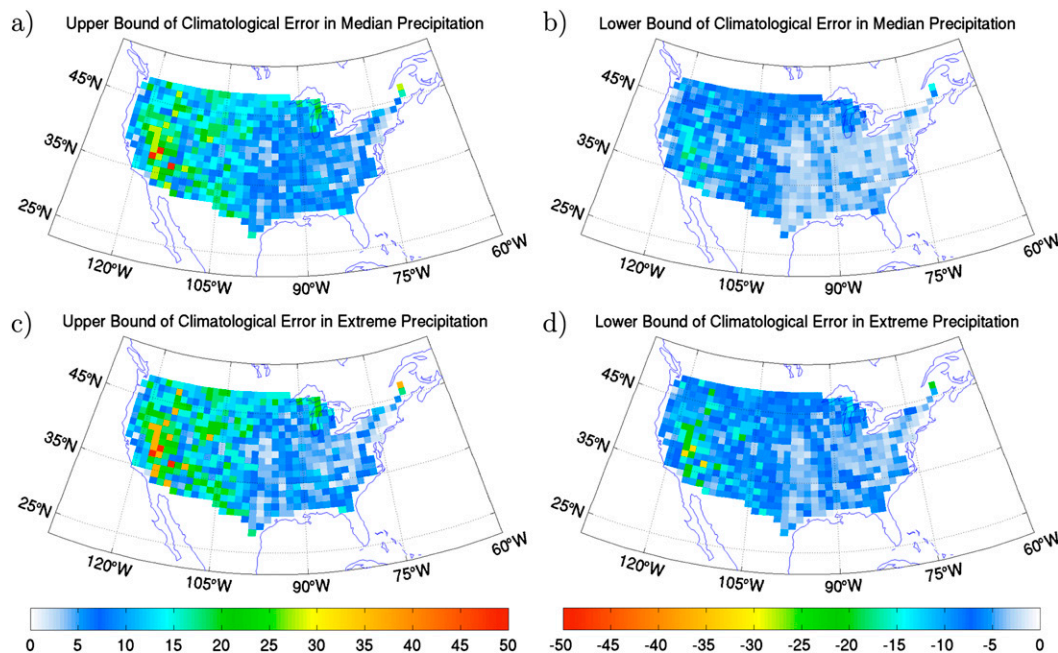

FIG. 12. As in Fig. 11, but for LRES data.

by the homogeneity in the median and extreme precipitation value in the east (Figs. 5a,b). The higher the station density, the closer the analysis is to the truth, and the lower the station density the greater the influence of more distant stations on the analysis point. In the case of homogeneous distributions, this implies that we will have a greater influence of stations with less shared variance (i.e., less correlated), but which have a similar distribution. As a result, the averaging of less shared variance biases the OA of precipitation toward lower climatological median and extreme values, as station density is decreased. This explains the small bias toward negative climatological errors observed in many of the eastern regions at both the annual and bimonthly averaging periods, for HRES and LRES fields (Figs. 7–9). The seasonality of climatological errors in this framework is impacted by the decorrelation length. As the decorrelation length decreases, the impact of stations with less shared variance on the analysis points will increase for the same search radius. This is consistent with the observation that climatological errors are higher when decorrelation lengths are shorter in JA relative to JF, in the eastern United States (Fig. 9).

These results are consistent with those found by Hofstra and New (2009) in western Europe, and disagree with those of Osborn and Hulme (1997) and Kursinski and Zeng (2006). We assert that the disparity between these sets of studies is due to differences in the gridding methods used and can be described within the first conceptual framework. The current study and that of Hofstra and New (2009) conduct an OA on the station data and then perform an area average. In Osborn and Hulme (1997) and Kursinski and Zeng (2006), a simple averaging of station data is applied to achieve a grid box averaged value. As discussed in Osborn and Hulme (1997), the method of simple station averaging allows for the application of the theory of randomly correlated variables to explain the impact of stations density on grid averaged variance. The theory relates grid averaged variance (S_n^2), averaged station variance (s_i^2), average correlation between stations ($\bar{r}$), and number of stations (n), through the following relationship:

$$S_n^2 = s_i^2 \left[\frac{1 + (n-1)\bar{r}}{n} \right]$$

(Osborn and Hulme 1997). The premise is that S_n^2 decreases as n increases because stations with less shared variance are averaged. In our method, it is the areal-averaging procedure that is analogous to an average of randomly correlated variables. The number of OA grid points is constant, so this is not where we see a dependence

on station density. It is instead the OA step that depends on the station density, and the areal-averaging step simply perpetuates these errors with some smoothing due to averaging. This implies that the simple averaging method is not expected to follow the first conceptual framework. We in fact expect the opposite response, which explains the disparity.

The second conceptual framework applies when the distribution of precipitation is inhomogeneous. In this case, as station density decreases, more distant stations with substantially different precipitation distributions have a larger impact on the OA point, resulting in large climatological errors. In the western United States, there is a predominance of orographically forced precipitation. This results in preferred regions for higher amounts of precipitation, as well as large contrasts between precipitation medians and extremes, depending on the specific location (Figs. 5a,b). Systematic errors in precipitation metrics can then result depending on the specific stations employed to conduct the analysis, making these regions more sensitive to station loss.

In the second framework, we may expect a similar relationship between the decorrelation length and errors that was seen in the first framework. However, the seasonality in the steepness of the inhomogeneity must be considered. We argue that in the western United States the preferred wet/dry season of heavy precipitation, driven by a stronger jet stream and more intense storm track in the winter, steepens the gradient in climatological median (not shown) and extreme precipitation (Fig. 13). This effect will not be apparent in the decorrelation length as the Kendall's tau rank method employed does not assume a linear relationship for the correlation. As such, a change in the steepness of the gradient in precipitation statistics will not necessitate a change in decorrelation length. The decorrelation lengths in the western United States are also longer in the winter than in the summer. This is more pronounced on the West Coast than in the Rockies or the North American monsoon regions (Fig. 10). Although there are also some increases in the errors at higher station densities in JA compared to JF in the Rockies region, consistent with the first conceptual framework, the overriding signal is an exponential increase in errors at lower station density that is higher in JF than in JA (Fig. 13). In the context of the second conceptual framework, this is thus explained based on the seasonality in the magnitude of the homogeneity in the western United States.

An envelope of potential upper and lower bounds of errors for all station densities is computed. Applying these boundaries to the actual station density at each grid point provides an estimate of the representativeness error, due to station density, across the United States.
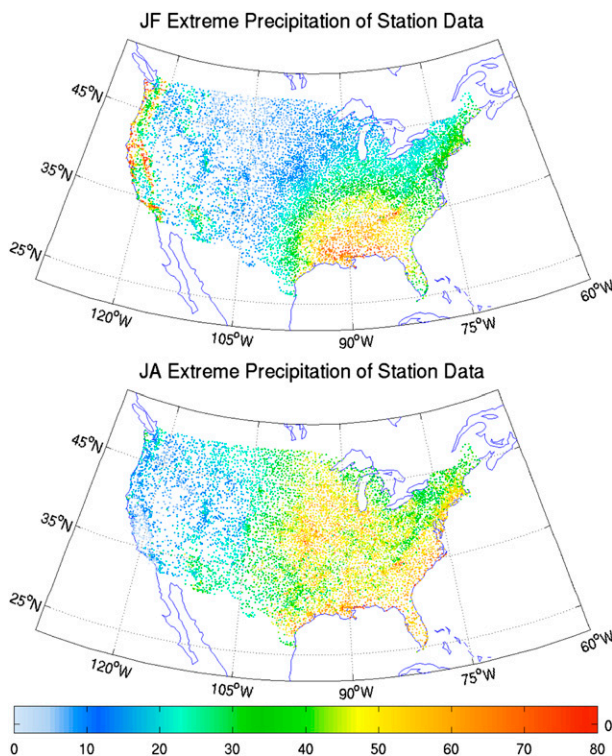
FIG. 13. Climatological station extreme precipitation (mm day^{-1}) for both the JF and JA periods.

These climatological errors are higher for the HRES field than the LRES field, and higher for extremes than median precipitation. Even within the United States, which is known for having a relatively dense network of stations, there are wide regions with the potential for large climatological errors in median and extreme precipitation. For the LRES field, much of the eastern United States has low values of potential errors, with upper bounds of 10%–15%, whereas in the western United States these climatological errors are often around 35%–45% (Fig. 12). When using the LRES field to validate a GCM, consideration of these errors is important for the interpretation of the model's ability to represent precipitation in the historical period.

REFERENCES

Accadia, C., S. Mariani, M. Casaioli, A. Lavagnini, and A. Speranza, 2003: Sensitivity of precipitation forecast skill scores to bilinear interpolation and a simple nearest-neighbor average method on high-resolution verification grids. *Wea. Forecasting,* **18,** 918–932, doi:10.1175/1520-0434(2003)018<0918:SOPFSS>2.0.CO;2.

Adam, J. C., and D. P. Lettenmaier, 2003: Adjustment of global gridded precipitation for systematic bias. *J. Geophys. Res.,* **108,** 4257, doi:10.1029/2002JD002499.

Bussières, N., and W. Hogg, 1989: The objective analysis of daily rainfall by distance weighting schemes on a mesoscale grid. *Atmos.–Ocean,* **27,** 521–541, doi:10.1080/07055900.1989.9649350.

Chen, C.-T., and T. Knutson, 2008: On the verification and comparison of extreme rainfall indices from climate models. *J. Climate,* **21,** 1605–1621, doi:10.1175/2007JCLI1494.1.

Chen, M., W. Shi, P. Xie, V. B. S. Silva, V. E. Kousky, R. Wayne Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.,* **113,** D04110, doi:10.1029/2007JD009132.

Cherry, J. E., L.-B. Tremblay, M. Stieglitz, G. Gong, and S. J. Déry, 2007: Development of the pan-Arctic snowfall reconstruction: New land-based solid precipitation estimates for 1940–99. *J. Hydrometeor.,* **8,** 1243–1263, doi:10.1175/2007JHM765.1.

Cressman, G., 1959: An operational objective analysis system. *Mon. Wea. Rev.,* **87,** 367–374, doi:10.1175/1520-0493(1959)087<0367: AOOAS>2.0.CO;2.

Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. *J. Climate,* **19,** 4605–4630, doi:10.1175/ JCLI3884.1.

Daly, C., 2006: Guidelines for assessing the suitability of spatial climate data sets. *Int. J. Climatol.,* **26,** 707–721, doi:10.1002/ joc.1322.

Ensor, L. A., and S. M. Robeson, 2008: Statistical characteristics of daily precipitation: Comparisons of gridded and point datasets. *J. Appl. Meteor. Climatol.,* **47,** 2468–2476, doi:10.1175/ 2008JAMC1757.1.

Gervais, M., J. Gyakum, E. Atallah, L.-B. Tremblay, and R. B. Neale, 2014: How well are the distribution and extreme values of daily precipitation over North America represented in the Community Climate System Model? A comparison to reanalysis, satellite, and gridded station data. *J. Climate,* **27,** 5219–5239, doi:10.1175/JCLI-D-13-00320.1.

Gober, M., E. Zsoter, and D. Richardson, 2008: Could a perfect model ever satisfy a naive forecaster? On grid box mean versus point verification. *Meteor. Appl.,* **15,** 359–365, doi:10.1002/ met.78.

Goodison, B., P. Louis, and D. Yang, 1998: WMO solid precipitation measurement intercomparison. WMO/TD-872, 211 pp.

Groisman, P., R. Knight, D. Easterling, T. Karl, G. Hegerl, and V. Razuvaev, 2005: Trends in intense precipitation in the climate record. *J. Climate,* **18,** 1326–1350, doi:10.1175/JCLI3339.1.

Hewitson, B., and R. Crane, 2005: Gridded area-averaged daily precipitation via conditional interpolation. *J. Climate,* **18,** 41–57, doi:10.1175/JCLI3246.1.

Hofstra, N., and M. New, 2009: Spatial variability in correlation decay distance and influence on angular-distance weighting interpolation of daily precipitation over Europe. *Int. J. Climatol.,* **29,** 1872–1880, doi:10.1002/joc.1819.

——, ——, and C. McSweeney, 2010: The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. *Climate Dyn.,* **35,** 841–858, doi:10.1007/s00382-009-0698-1.

Hutchinson, M. F., D. W. McKenney, K. Lawrence, J. H. Pedlar, R. F. Hopkinson, E. Milewska, and P. Papadopol, 2009: Development and testing of Canada-wide interpolated spatial models of daily minimum-maximum temperature and precipitation for 1961–2003. *J. Appl. Meteor. Climatol.,* **48,** 725–741, doi:10.1175/2008JAMC1979.1.

Jones, P., 1999: First- and second-order conservative remapping schemes for grids in spherical coordinates. *Mon. Wea. Rev.,* **127,** 2204–2210, doi:10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2.

Kursinski, A. L., and X. Zeng, 2006: Areal estimation of intensity and frequency of summertime precipitation over a midlatitude region. *Geophys. Res. Lett.,* **33,** L22401, doi:10.1029/2006GL027393.

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston, 2012: An overview of the Global Historical Climatology Network-Daily database. *J. Atmos. Oceanic Technol.,* **29,** 897–910, doi:10.1175/JTECH-D-11-00103.1.

Osborn, T. J., and M. Hulme, 1997: Development of a relationship between station and grid-box rainday frequencies for climate model evaluation. *J. Climate,* **10,** 1885–1908, doi:10.1175/1520-0442(1997)010<1885:DOARBS>2.0.CO;2.

Shiu, C.-J., S. C. Liu, C. Fu, A. Dai, and Y. Sun, 2012: How much do precipitation extremes change in a warming climate? *Geophys. Res. Lett.,* **39,** L17707, doi:10.1029/2012GL052762.

Sun, Y., S. Solomon, A. Dai, and R. Portmann, 2006: How often does it rain? *J. Climate,* **19,** 916–934, doi:10.1175/JCLI3672.1.

——, ——, ——, and R. W. Portmann, 2007: How often will it rain? *J. Climate,* **20,** 4801–4818, doi:10.1175/JCLI4263.1.

Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The changing character of precipitation. *Bull. Amer. Meteor. Soc.,* **84,** 1205–1217, doi:10.1175/BAMS-84-9-1205.

Tustison, B., D. Harris, and E. Foufoula-Georgiou, 2001: Scale issues in verification of precipitation forecasts. *J. Geophys. Res.,* **106,** 11 775–11 784, doi:10.1029/2001JD900066.

Wilks, D., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Academic Press, 627 pp.

Xie, P., M. Chen, S. Yang, A. Yatagai, T. Hayasaka, Y. Fukushima, and C. Liu, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.,* **8,** 607–626, doi:10.1175/JHM583.1.