2021

The Perils of Regridding: Examples Using a Global Precipitation Dataset

Rajulapati, Chandra Rupa

American Meteorological Society (AMS)

10.1175/JAMC-D-20-0259.1

The Perils of Regridding: Examples Using a Global Precipitation Dataset

CHANDRA RUPA RAJULAPATI,[a,b] SIMON MICHAEL PAPALEXIOU,[a,b,c] MARTYN P. CLARK,[a,b] and JOHN W. POMEROY[a,b]

[a]*Centre for Hydrology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada*
[b]*Global Institute for Water Security, University of Saskatchewan, Saskatoon, Saskatchewan, Canada*
[c]*Department of Civil, Geological and Environmental Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada*

(Manuscript received 12 November 2020, in final form 5 October 2021)

ABSTRACT: Gridded precipitation datasets are used in many applications such as the analysis of climate variability/change and hydrological modeling. Regridding precipitation datasets is common for model coupling (e.g., coupling atmospheric and hydrological models) or comparing different models and datasets. However, regridding can considerably alter precipitation statistics. In this global analysis, the effects of regridding a precipitation dataset are emphasized using three regridding methods (first-order conservative, bilinear, and distance-weighted averaging). The differences between the original and regridded dataset are substantial and greatest at high quantiles. Differences of 46 and 0.13 mm are noted in high (0.95) and low (0.05) quantiles, respectively. The impacts of regridding vary spatially for land and oceanic regions; there are substantial differences at high quantiles in tropical land regions, and at low quantiles in polar regions. These impacts are approximately the same for different regridding methods. The differences increase with the size of the grid at higher quantiles and vice versa for low quantiles. As the grid resolution increases, the difference between original and regridded data declines, yet the shift size dominates for high quantiles for which the differences are higher. While regridding is often necessary to use gridded precipitation datasets, it should be used with great caution for fine resolutions (e.g., daily and subdaily), because it can severely alter the statistical properties of precipitation, specifically at high and low quantiles.

SIGNIFICANCE STATEMENT: Regridding has a substantial impact on the statistical properties of precipitation. The impacts of regridding vary spatially as well as at different quantiles. Regridding should be used with great caution.

KEYWORDS: Precipitation; Data processing; Interpolation schemes

1. Introduction

Regridding or remapping is the process of interpolating from one grid to another. This may involve both temporal and spatial (horizontal) interpolations, yet typically, regridding refers to spatial interpolation. Regridding is necessary in model coupling (e.g., coupling atmospheric and hydrological models) and comparing different models and datasets (Gautam et al. 2018; Sun et al. 2018). Regridding is commonly used in climate change impact studies (Alexander et al. 2006), hydrological modeling (Chen and Brissette 2017; Li et al. 2019), forecast applications (Lavers et al. 2009), moisture transport studies (Ghodichore et al. 2019), trend analysis (Harrison et al. 2019), and wildfire prediction (Yue et al. 2014). The effects of regridding vary with the variable under consideration. Precipitation is a critical variable for hydrology and is especially sensitive to regridding due to its high spatial variability and intermittency. Though regional studies have explored the effects of regridding (Accadia et al. 2003; Berndt and Haberlandt 2018; Diaconescu et al. 2015; Ensor and Robeson 2008; Rauscher et al. 2010), a global perspective is still lacking as these effects vary with region. It is important to understand changes in precipitation at different latitudes and for different quantiles to distinguish regions and situations with low or high changes.

Global gridded precipitation products are developed using different data sources and assimilation techniques (Ashouri et al. 2015; Chen et al. 2008, 2017; Fuka et al. 2014; Saha et al. 2014; Weedon et al. 2014). Given several datasets, there is always interest to assess differences among them in various regions (Akinsanola et al. 2017; Dinku et al. 2008; Donat et al. 2014; Hu et al. 2018; Kidd et al. 2012; Sun et al. 2014; Zhang et al. 2013). However, the spatial resolution of datasets can be very different, and in some cases even if the grid resolution is same, the grids are unaligned. For example, the Multi-Source Weighted-Ensemble Precipitation (MSWEP) and the National Centers for Environmental Prediction Climate Forecast System Reanalysis (NCEP–CFSR) products have the same resolution (0.5° × 0.5°), yet the grids start at (89.75°, −179.75°) and (90°, −179.50°), respectively. Therefore, comparing two products for quantitative assessment, necessitates regridding them to a unified resolution and grid placement (Henn et al. 2018; Sun et al. 2018).

Typically, to assess the frequency and severity of extreme events, several climate indices are used in the literature and gridded products are used to estimate such indices (e.g., the Expert Team on Climate Change Detection and Indices

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JAMC-D-20-0259.s1.

Papalexiou's current affiliations: Department of Civil Engineering, University of Calgary, Calgary, Alberta, Canada and Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Prague, Czech Republic.

Corresponding author: Chandra Rupa Rajulapati, chandra.rajulapati@usask.ca

DOI: 10.1175/JAMC-D-20-0259.1

(ETCCDI), Alexander et al. 2006). Indices such as R95 (Annual total precipitation from days greater than 95th percentile), R20 (the number of days with precipitation greater than 20 mm), consecutive dry days, etc., are cross compared among the gridded products (Contractor et al. 2015). However, most studies do not pay attention to the effects of regridding the datasets on the analysis, results, and conclusions. For example, one regional study that highlight the effects of regridding in the North Atlantic basin documents a 4% increase in low precipitation values and a corresponding decrease in zero precipitation (Booth et al. 2018). However, the effects are not constant in space and vary with respect to the magnitude of the precipitation. There have been no studies at the global scale highlighting the discrepancies and/or the effects of regridding precipitation data products. Also, the effects of grid resolution and the size of the shifted grid while regridding have not been explored. Here, the effects of regridding precipitation using different techniques including bilinear, conservative and distance-weighted averaging remapping, are quantified using one global gridded product as an example. Specifically, the aim is to understand the effects of 1) regridding across different latitudes and quantiles, 2) the regridding method used, 3) the size of the grid shift, and 4) the spatial resolution of the gridded product.

2. Data and methods

A global precipitation gridded dataset, the MSWEP, version 2.0 (Beck et al. 2017), was chosen to evaluate the effects of regridding. The widely used MSWEP dataset is formed by optimally merging data from various sources (including gauge observations, satellite, and reanalysis). The dataset is available at two spatial resolutions, 0.5° × 0.5° and 0.1° × 0.1° with daily temporal resolution from 1969 to 2016; it covers both land and ocean. Here, four different analyses were performed in line with the objectives: 1) a simple shift of 0.25° in the 0.5° × 0.5° grid was made so that the starting coordinates (89.75°, −179.75°) are shifted to (89.50°, −179.50°). The regridded data still have the original 0.5° × 0.5° spatial resolution (Fig. 1a), 2) different regridding methods are used to understand the effects of the method, 3) to check the effect of the size of the shift, multiple small shifts were made in the 0.5° × 0.5° data and regridded back to the original 0.5° × 0.5° (Figs. 1b,c), and 4) a simple shift of 0.05° was made in the 0.1° × 0.1° data (Fig. 1). In this study, the grid shift is considered as a diagonal shift, that is, shift in both latitude and longitude.

Numerous grid interpolation methods exist; simple classical methods such as bilinear, nearest-neighbor approaches, weighting based on nearest neighbors, and applications of splines (see, e.g., Accadia et al. 2003; Berndt and Haberlandt 2018; Hofstra et al. 2008), as well as more sophisticated approaches such as conservative and patch-based methods (see, e.g., Fischer et al. 2014; McGinnis et al. 2010). Regridding precipitation is always challenging due to the high spatial variability and intermittency. Here three regridding techniques, first-order conservative, bilinear interpolation, and distance-weighted approach, were used. The first-order conservative regridding method preserves the integral of the source field (precipitation in this case) across grids (Jones 1999). The bilinear approach uses a linear interpolation in two directions considering the four nearest grids of original data. The distance-weighted average or the inverse distance weighting approach is based on the distance of the surrounding four points in the original grid that are closest to each point in the target grid. The Climate Data Operator (CDO) was used to regrid the data (Schulzweida 2019). CDO is a commonly used platform to manipulate and analyze gridded data (for more information, see https://code.mpimet.mpg.de/projects/cdo/). The nearest-neighbor method, a most widely used method for precipitation (Accadia et al. 2003; Berndt and Haberlandt 2018; Hofstra et al. 2008; Shen et al. 2001), typically considers the grid to the nearest target grid, that is, the grid is just shifted with same precipitation time series; that is, data with coordinates (89.75°, −179.75°) are shifted to (89.50°, −179.50°). The nearest-neighbor method is highly sensitive to the distance between the target and the source, and for larger grid shifts this usually causes undesired results. Therefore, this method is not considered in this study. Also, effects due to the spatial reference systems including geographic coordinate systems and projected coordinate systems/map projections (e.g., Lambert equal-area azimuthal, equidistant azimuthal, Albers equal-area conic, equidistant conic, and

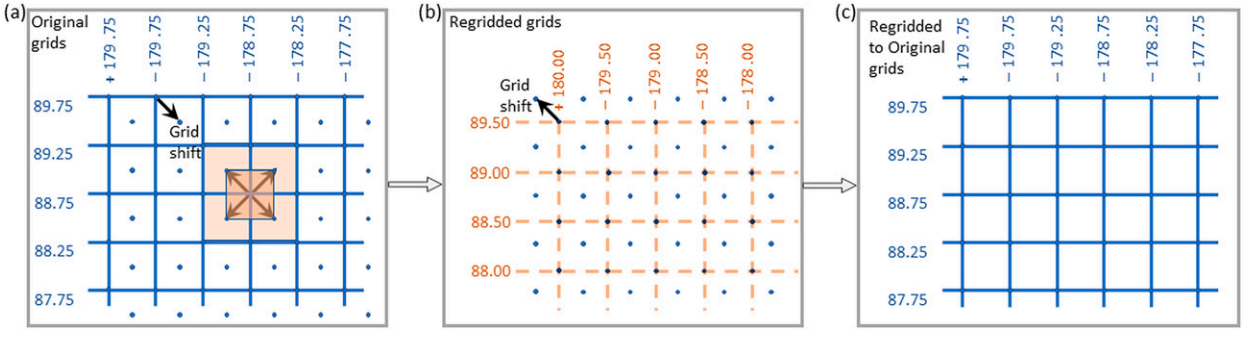

FIG. 1. Grids of the original (blue) and regridded (orange) data: (a) grid shift in original grids, (b) grid shift in regridded grids, and (c) regridded-to-original grids. In (a), the highlighted blue lines show area represented by grid (88.75, −178.75), and the highlighted orange points and area represent the surrounding grid points of the grid.

Lambert conformal conic) are not considered in this work, which have effects on the grids at high latitudes because the grid is distorted (Battersby 2009; Li et al. 2011; Robeson 1997; Tissot 1881; Willmott and Robeson 1995).

In the example provided in Fig. 1, the area represented by a given original grid comprises 1/4 of the area of each of the surrounding four regridded grids. Specifically, the area represented by grid point (88.75°, −178.75°) contains parts of the four nearest regridded points (Fig. 1a) centered at (89.00°, −179.00°), (88.50°, −179.00°), (89.00°, −178.50°), and (88.50°, −175.50°). Therefore, the precipitation of a given original grid with each one of the four surrounding grids is compared at different quantiles. For a given grid i and a given quantile Q, the percentage difference $\mathrm{PD}_{Q,i}$ between original and regridded datasets is estimated as:

$$\mathrm{PD}_{Q,i} = \left(\frac{\mathrm{OP}_{Q,i}}{\mathrm{RP}_{Q,i}} - 1\right) \times 100\%, \tag{1}$$

where $\mathrm{OP}_{Q,i}$ and $\mathrm{RP}_{Q,i}$ is the precipitation at Q for original and regridded grids, respectively. Note that there are four different regridded grids neighboring the original grid and thus four comparisons can be made. Figure S1 in the online supplemental material shows the maps of PD_{50} in the four directions using the conservative regridding method. No considerable difference is noted among the four directions (four maps), either in spatial patterns or in the percentage difference. Therefore, results for one neighboring grid (southeast) is shown for comparison.

The PD, however, is deceptive in some cases for understanding the degree of variation in precipitation between gridded products. At low values, a small variation in precipitation can lead to a large PD. To examine this, in addition to the PD, the actual differences d between the original and regridded precipitation ($d_{Q,i} = \mathrm{OP}_{Q,i} - \mathrm{RP}_{Q,i}$) at different quantiles are calculated.

3. Results

a. *Regridding 0.5° × 0.5° data*

The results show a considerable shift in the global statistics of precipitation between the original and regridded datasets for all regridding methods considered (see Table 1). There is a consistent reduction in all statistics of the regridded dataset, including the standard deviation, relative to the original data, and this reduction is consistent among the three methods considered. Figure 2 shows the mean precipitation and standard deviation for the original and regridded data. While visual inspection of the spatial patterns does not discern a large change, a decrease of 12% and 8% in the mean and standard deviation, respectively, is observed at global scale in shifting from original to regridded datasets using all the methods considered. Globally statistics of regridded precipitation using different methods are approximately the same (Table 1). The maximum precipitation observed in the original dataset drops by 483 mm (about 28%) in the regridded data.

Regridding smooths the precipitation, especially the extremes; however, smoothing may not be consistent for all quantiles. A range of quantiles were considered and precipitation at each quantile Q was compared between the original and regridded data (Fig. 3). The precipitation at high quantiles is greatly reduced relative to the low quantiles, in terms of absolute differences. At 5th and 25th quantiles, a considerable shift is found in the histograms. The magnitude of reduction is large at high quantiles, thus underestimating extreme precipitation and potentially affecting studies related to extreme events, including flooding (Slinskey et al. 2019). Yet the fractional reduction is higher in low quantiles than in high quantiles. There is also a consistent increase in the number of wet days in the regridded dataset (though a threshold is typically considered in characterizing a wet day, here precipitation greater than zero is considered). Globally, the average probability or fraction of zero precipitation days P_0 is reduced from 0.514 to 0.446 (Fig. S2 in the online supplemental material shows the histograms of P_0) in all the methods considered. Accordingly, the average global days of low precipitation increased by approximately 1.5 times, and the average days doubled in some grids. This is much higher than found in regional studies (Booth et al. 2018), where about 4% has been reported for the Arctic region). This may have a great impact on drought studies where the interest lies in understanding the changes in low quantiles (Raziei et al. 2011), leading in miscalculating indices related to droughts such as the standard precipitation index. The estimated annual count of days for which precipitation is greater than 10 mm (R10) and 20 mm (R20) indices (e.g., see Alexander et al. 2006) also shows considerable difference between the datasets (Fig. S3 in the online supplemental material).

TABLE 1. Mean value, standard deviation (std dev) and quantiles Q of original, regridded, and regridded-to-original-grids data using first-order conservative, bilinear, and distance-weighted regridding techniques.

		First-order conservative		Bilinear		Distance weighted	
Statistic (mm)	Original	Regridded	Regridded to original	Regridded	Regridded to original	Regridded	Regridded to original
Mean	4.860	4.273	3.893	4.273	3.893	4.273	3.893
Std dev	10.459	9.564	8.999	9.564	8.999	9.564	8.999
Q_5	0.030	0.005	0.0025	0.005	0.0025	0.005	0.0025
Q_{25}	0.180	0.097	0.058	0.098	0.058	0.097	0.058
Q_{50}	1.190	0.872	0.671	0.873	0.671	0.872	0.671
Q_{75}	4.900	4.170	3.673	4.170	3.673	4.170	3.672
Q_{95}	21.710	19.634	18.300	19.633	18.290	19.632	18.300
Q_{99}	49.600	45.347	42.667	45.345	42.665	45.344	42.661

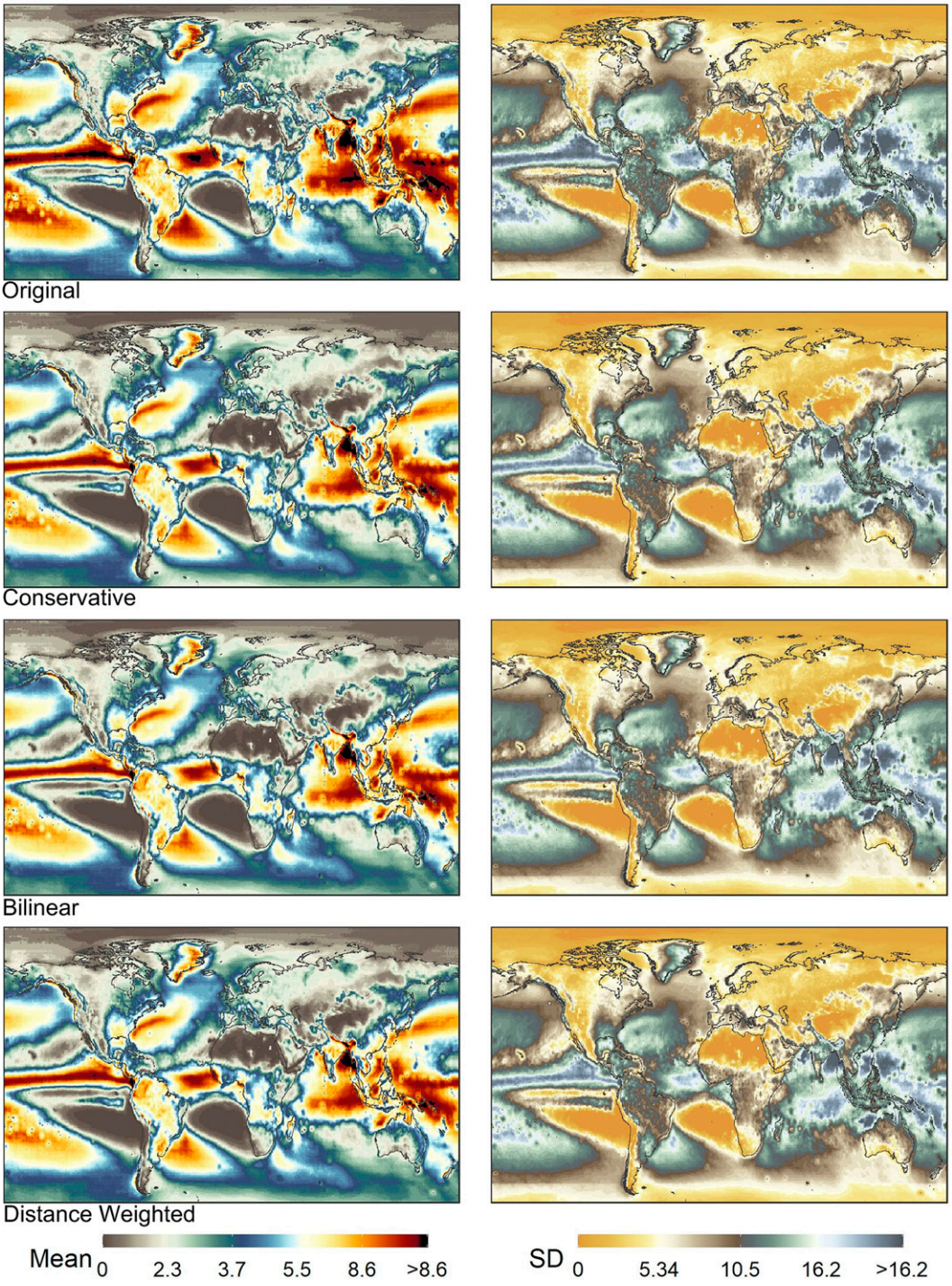

FIG. 2. Spatial pattern of (left) mean and (right) standard deviation of (top) original and regridded using (top middle) conservative, (bottom middle) bilinear, and (bottom) distance-weighted average techniques.

The absolute differences d between the original and regridded precipitation at different quantiles are mapped; see Fig. 4 for 5th and 95th quantiles and Fig. S4 in the online supplemental material for 25th and 75th quantiles. For high quantiles, high d (>12 mm) is observed in tropical oceans, and low values (<1 mm) near polar regions in all regridding methods; in fact, the method of regridding does not have a major impact on the spatial patterns except in the bilinear method at

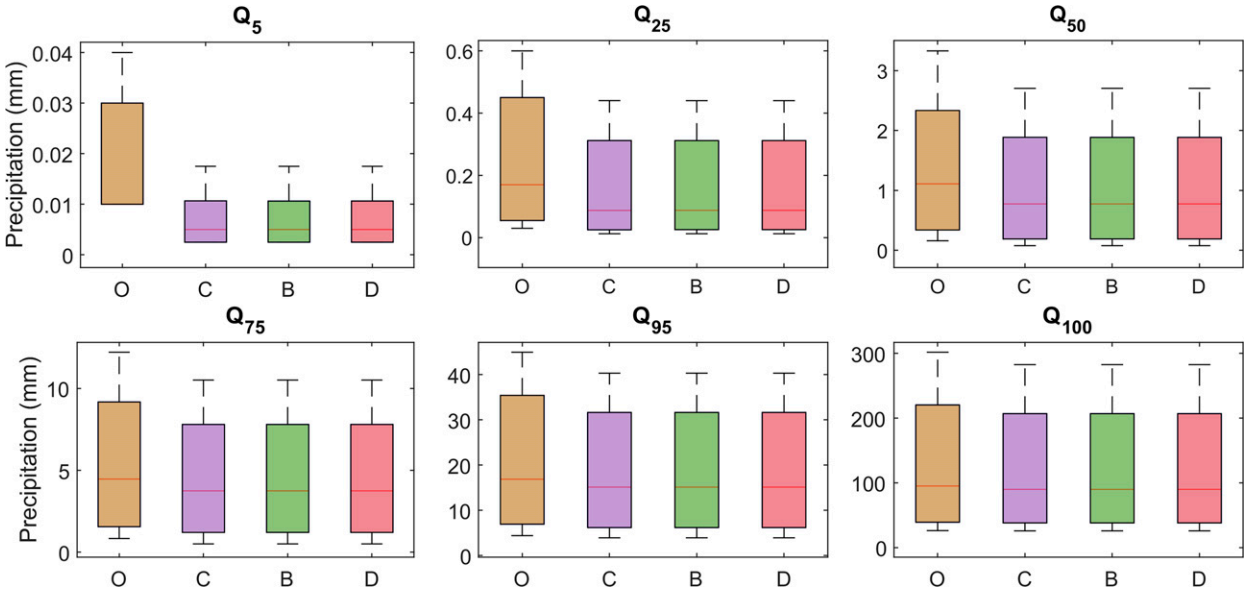

FIG. 3. Boxplots of original (O) and regridded data using conservative (C), bilinear (B), and distance-weighted (D) methods at different quantiles at global scale. For the boxplots, whiskers denote the 95% empirical confidence interval. Plots were formed from 2162000 grids, covering both ocean and land.

low quantiles. Large differences were noted in high quantiles in comparison with low quantiles. Maps of PD, comparing the original grid and its southeast neighbor for different quantiles demonstrate a substantial spatial variation (Figs. S5–S7 in the online supplemental material) irrespective of the regridding method. Especially large variations between the original and regridded precipitation are evident near the tropics. For low quantiles, there are high (~150%) PD over the tropical oceans, yet PD is negative or small (~14%) over tropical land areas. In contrast, high quantiles consistently correspond to high PD (~30%) values over both oceans and land areas. While the spatial patterns of PD and d do not match at low quantiles, they match well at high quantiles, except in the South Pacific Ocean near South America, South Atlantic Ocean near Africa, and the desert interior of Africa. A few areas, the northwestern coast of Australia, western coast of the continental United States and Canada, and southern parts of South America, have both high d and PD, including at the low quantiles. It should be noted that PDs are misleading for low quantiles, that is, a high PD refers to a small value of d.

To test if the process is reversible, the data were regridded by shifting back the coordinates of the regridded dataset to those of the original dataset. This regridded-to-original dataset used the same method as the initial regridding. Considerable differences between the original and regridded-to-observations data are evident in Fig. 5 and Table 1 (see also Figs. S8–S11 in the online supplemental material). Notably, the PD are higher at low quantiles than at high quantiles and vice versa for d. P_0 is further reduced to 0.393 at the global scale (Fig. 6). Further, the cross correlations between the original and regridded-to-observations quantile maps are 0.59, 0.85, 0.93, 0.96, 0.98, 0.99, and 0.98 at 5%, 25%, 50%, 75%, 95%, 99%, and 100% quantiles, respectively. Figure 6 shows boxplots of P_0 at all grids for the original (O), regridded (R), and regridded-to-original (RO) grids for all the methods considered in this study. A reduction of about 76% is noted in P_0 on an average between the O and RO. This shows that the process is irreversible, with the greatest absolute differences for high quantiles and percentage differences for low quantiles. This may be due to a systematic bias.

b. *Effect of shift size*

In the previous analysis, considerable changes in precipitation statistics were noted with a shift of 0.25°. To understand if the size of the shift has any effect on these changes, three shifts were made, that is, 0.0625° (one-eighth)—shift-1 (labeled S1), 0.125° (one-fourth)—shift-2 (labeled S2) and 0.375° (three-fourths of the grid size)—shift-4 (labeled S4) in addition to the shift of 0.25° (labeled S3). The shifts were made and regridded back to the original grid. Given the similarity of results among the three methods in section 3a, for checking the effect of grid shift, only first-order conservative regridding method is used. Results show that as the grid shift decreases, the regridded data are close to the original data, and the mean values of mean, standard deviation, and high quantiles are close to those of the original data. However, there is no effect of the shift size on P_0. This is due to the fact that if any grid has a nonzero precipitation and its surrounding grids have a zero precipitation, the regridded grid will have a precipitation value depending on the shift size, and the P_0 is constant irrespective of the shift size. Another implication of this effect is the low quantiles were greatly reduced for lower shifts (Fig. 7). Yet, at high quantiles, the smaller the shift size the closer to the original data. The grid shift of

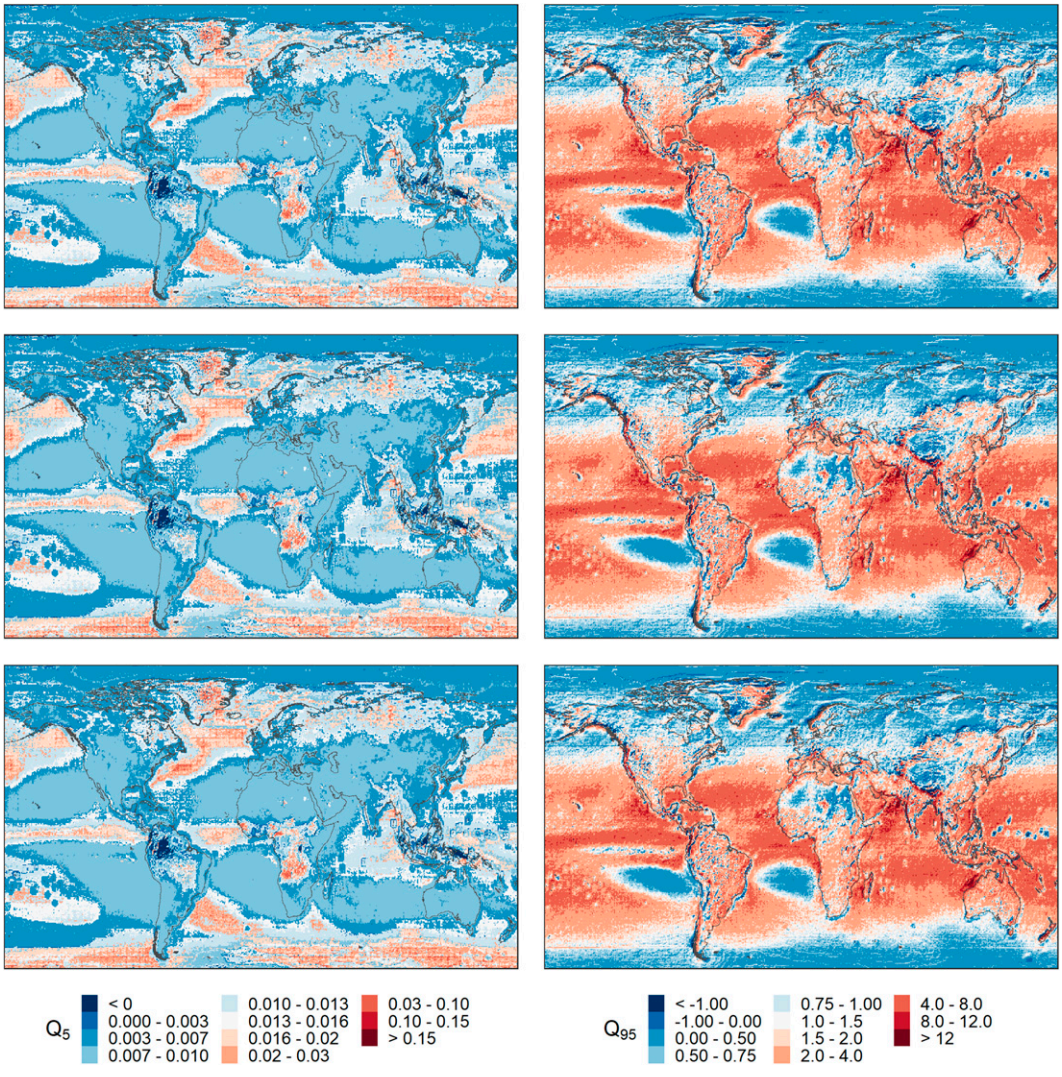

FIG. 4. Spatial patterns of differences (mm) between the original and regridded data at (left) 5th and (right) 95th quantiles using (top) conservative, (middle) bilinear, and (bottom) distance-weighted average techniques.

0.125° is same as the 0.375° shift, that is, the effect of the shift reverses after 0.25°. It is important to note that though the regridded data are close to the original data as the shift size reduces, there are considerable differences from the original, unless otherwise a very minor shift, of order one-hundredth, is made. However, such small shifts are not applied in practice. Spatially, the differences at high quantiles are higher than at low quantiles (Fig. 8) as seen in the boxplots.

c. *Effect of grid resolution*

To check the effect of grid resolution on regridding, the 0.1° data are considered and a shift size of 0.05° is made. Again, given the similarity of the results using different regridding methods, only the first-order conservative regridding method is used and the regridded data are shifted back to the original grids. Though visual inspection of the spatial patterns does not discern a large change (see Fig. 9, top and middle rows), a decrease of 15% and 18% in the mean and standard deviation, respectively, is observed at global scale. These decreases are higher than the 0.5° resolution (12% and 8%), specifically for standard deviation. This is the due to higher mean, standard deviation, and other quantiles in the 0.1° relative to the 0.5° (see Fig. S12 in the online supplemental material). Typically, at lower resolution, precipitation refers to averaged values over a smaller area in comparison with high resolutions and therefore, high values as well as high variations, including high P_0 at 0.1°. The spatial patterns of differences between the original and regridded-to-original data are approximately similar as those of 0.5°, yet the differences are higher at low quantiles and lower at high quantiles (Figs. 9 and 10). The low differences at high quantiles in the 0.1° as compared with the 0.5° is noted, which is due to the shift size.

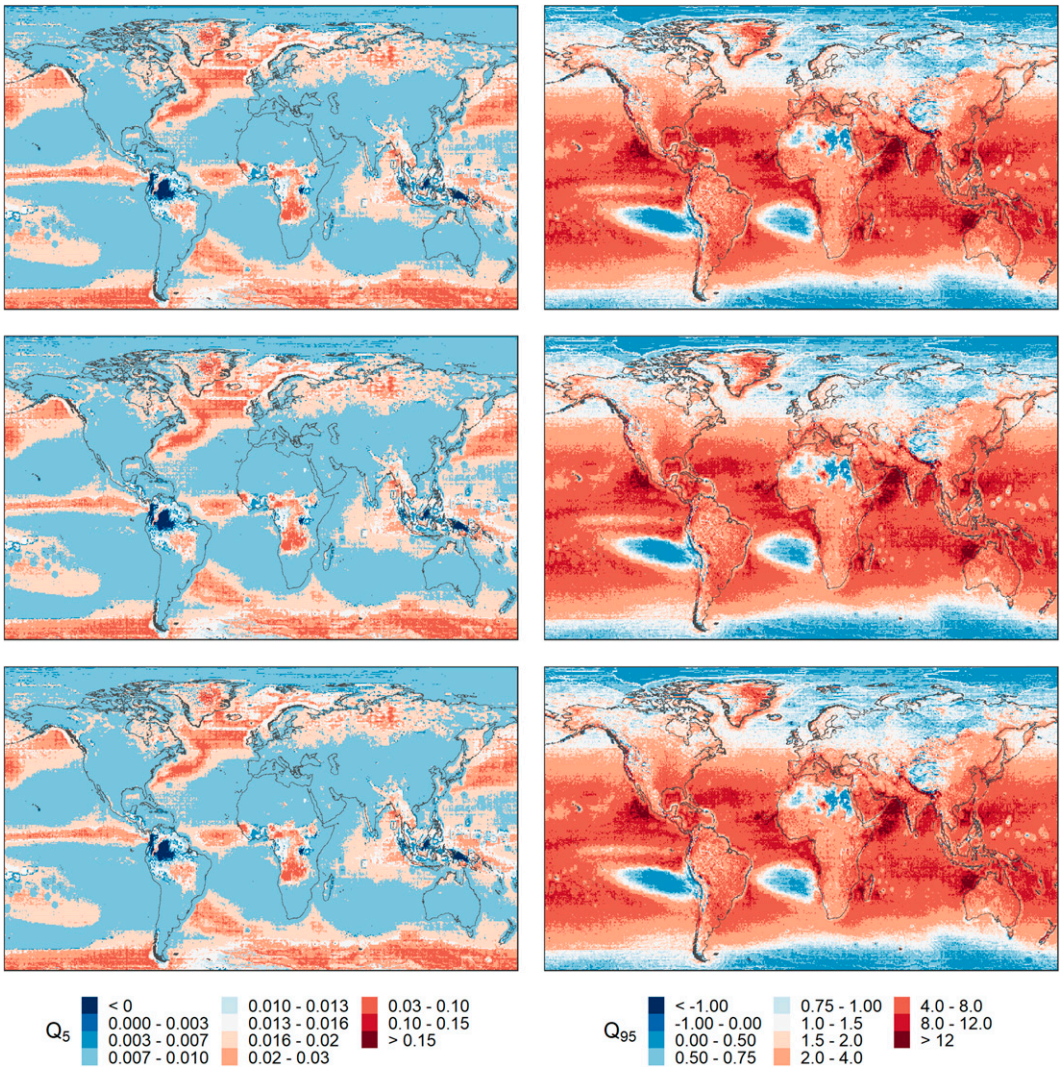

FIG. 5. As in Fig. 4, but between the original data and regridded-to-original-grids data.

4. Discussion

It is well known that regridding precipitation introduces a systematic bias, yet this analysis shows that the precipitation at high quantiles is more impacted by regridding than at low quantiles in terms of absolute differences, and vice versa if percentage changes are considered. Regridded points generally having a value greater than zero if the surrounding three grids have no precipitation and the fourth one has precipitation. Such a regridded point will, when converted back to the original grid, further smooth the precipitation, that is, averaging the averages for a simple averaging technique. The reason for the increase in wet days is that the zero precipitation grids are given a positive value in regridding due to at least one positive value in the surrounding grids. This is also the reason for much lower precipitation values as the shift size is reduced (Fig. 7).

Spatially, patterns of precipitation at different quantiles and their corresponding differences patterns are similar except at a few grids, that is, grids with high precipitation have high absolute differences and vice versa. However, it is to be noted that due to regridding, precipitation do not decrease (due to averaging), there are a few grids where precipitation increases. For example, in the Fig. 5, there are grids with negative absolute differences where the regridded precipitation is higher than the original, specifically in the tropical lands. This is again due to huge variation in the precipitation among the neighboring grids. Precipitation in the ocean grids on either side of South America is considerably lower relative to the land, therefore, the regridded precipitation is higher than the original in these land grids.

Spatially, the patterns of differences (both absolute and percentage) between the original and regridded data are

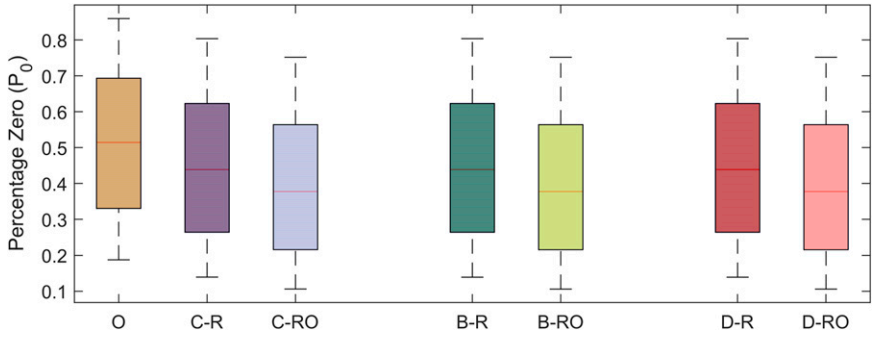

FIG. 6. Boxplots of the probability zero or the percentage of zero precipitation (P_0) at global scale for original (O), regridded (R), and regridded-to-original grid (RO) data using conservative (C), bilinear (B), and distance-weighted average (D) techniques. Whiskers denote 95% empirical confidence interval.

approximately similar for both the resolutions, specifically at high quantiles. Typically, precipitation refers to the average precipitation of the grid, and as the grid resolution reduces, precipitation statistics increase due to reduced area. In general, fine resolution grids (of order 4 km) have best agreement with point observations (Lee et al. 2017). As a result of high values in the 0.1° grids, the effects of regridding are also high at low quantiles, yet the size of the shift overcomes at high quantiles. In fact, precipitation at low quantiles is greatly affected at both the resolutions, yet the effect is high (approximately 4 times) at 0.1°resolution relative to the 0.5°. Similar effects are noted in the literature when station data are converted to grid data (Ensor and Robeson 2008). It is also noted in the literature that the shape of the distribution and autocorrelations are not greatly affected by regridding (Booth et al. 2018; Rauscher et al. 2010).

Therefore, a careful assessment of the effects of regridding should be made when comparing or coupling gridded products. Nevertheless, it is to be expected that different datasets will have different spatial and temporal resolutions. Resolution depends on the computational efficiency, computational capabilities, included physical processes and physical constraints of the underlying models used to develop the gridded products. To overcome this, a few datasets are provided at various spatial and temporal resolutions by the developers (Haylock et al. 2008). For example, the MSWEP dataset is also available at two temporal resolutions (3 h and daily). However, it is impractical to provide datasets at a wide range of spatial and temporal resolutions and so regridding to a common resolution is needed for comparison and coupling purposes. Unfortunately, the results of this study show that regridding precipitation may introduce a severe bias. To avoid this bias when comparing calculations derived from precipitation data provided on differing grids, it may be better to compute the climate index (to be compared) first at each grid and then to interpolate the index to a common resolution instead

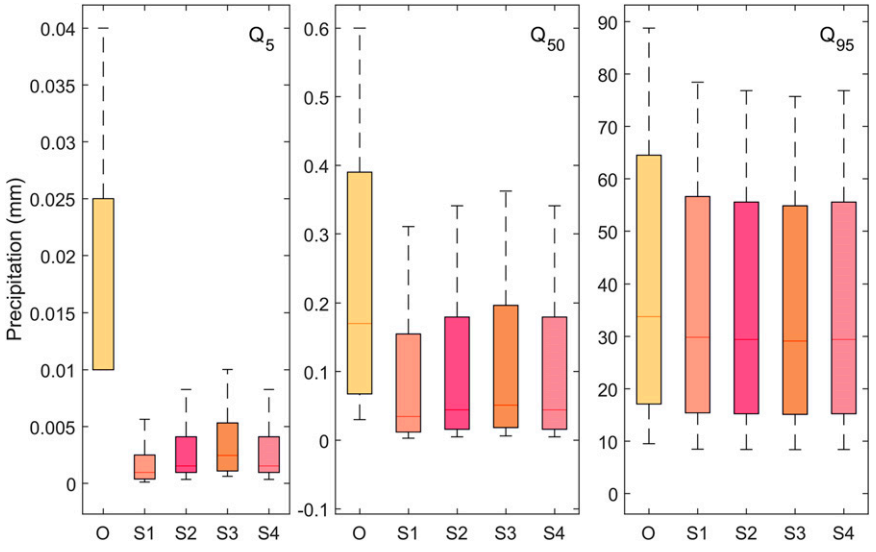

FIG. 7. Boxplots of precipitation at different quantiles for original (O) and regridded-to-original grid for a shift size of 0.0625° (S1), 0.125° (S2), 0.25° (S3), and 0.375° (S4) using the conservative regridding technique. Whiskers denote 95% empirical confidence interval.

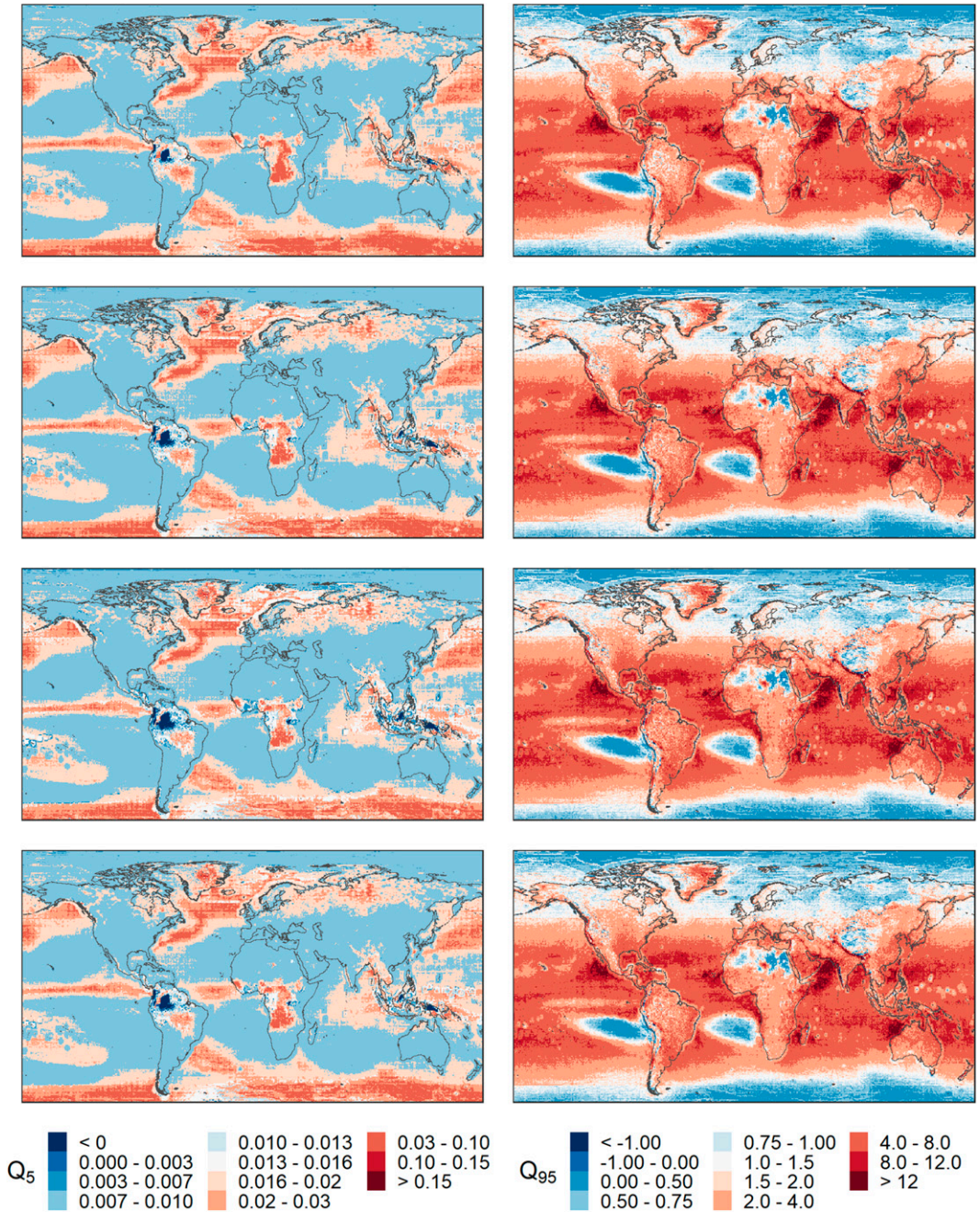

FIG. 8. Spatial patterns of differences (mm) between the original data and regridded-to-original-grids data at (left) 5th and (right) 95th quantiles for a shift size of (top) 0.0625°, (top middle) 0.125°, (bottom middle) 0.25°, and (bottom) 0.375°.

of using regridding/remapping the precipitation time series first and then calculating the index on the regridded data. This method of remapping (interpolating the index) has proven successful in minimizing the errors, irrespective of regridding methods (Diaconescu et al. 2015; Rajulapati et al. 2020; Risser et al. 2019). In the case of coupling, a careful assessment of biases is required, and the precipitation should be adjusted using covariates such as elevation, latitude, and longitude (Berndt and Haberlandt 2018; Demaria et al. 2013).

5. Conclusions

Regridding or remapping smooths precipitation fields. However, such smoothing is not consistent and varies spatially as well as for different quantiles. The effects of remapping are shown here using the example of the MSWEP dataset, regridded using the first-order conservative, bilinear, and distance-weighted average methods. Specifically, effects of regridding precipitation across different quantiles and regions,

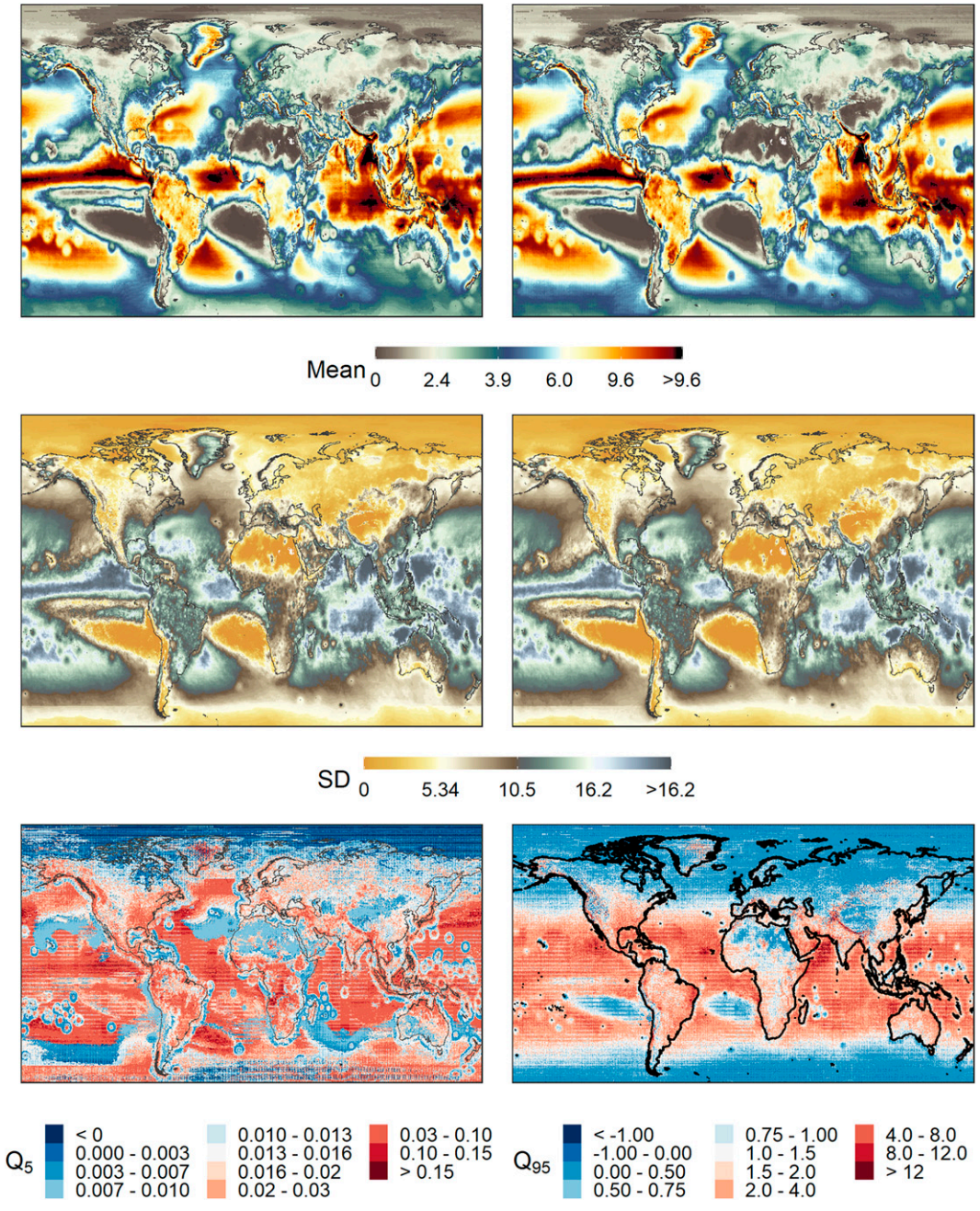

FIG. 9. Spatial patterns of (top) mean and (middle) standard deviation for both (left) original and (right) regridded-to-original data at 0.1° resolution. (bottom) differences between the original data and regridded-to-original-grids data at (left) 5th and (right) 95th quantiles for a shift size of 0.05°.

and the type of regridding method used are examined. The effect of the shift size and grid resolution were also demonstrated. Although a substantive percentage difference was found between original and regridded precipitation in low quantiles as compared with high quantiles, in terms of absolute differences, high quantiles are greatly affected. The frequency of dry days or days without precipitation was reduced by 30% and consequently the number of wet days increased in the regridded data. Precipitation was underestimated in the polar regions and overestimated in the tropics at high quantiles; at low quantiles, it was greatly overestimated over tropical land. The effects of regridding are irreversible, making them persistent in subsequent spatial data analysis. At the global scale, the choice of the regridding method has no effect on precipitation statistics, that is, the three methods resulted in approximately same differences. Selecting an appropriate

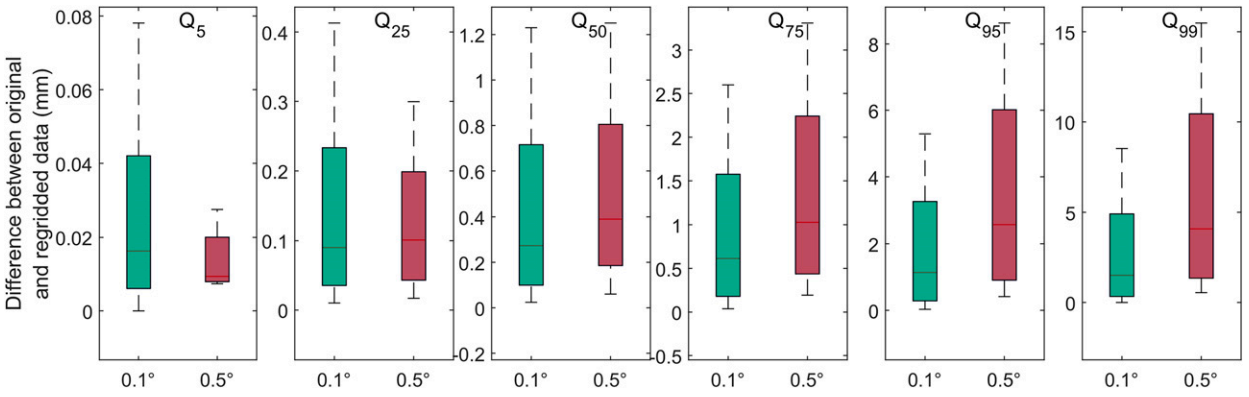

FIG. 10. Boxplots of differences between original and regridded-to-original-resolution data at different quantiles for 0.1° and 0.5° resolution. Whiskers denote the 95% empirical confidence interval.

interpolation method has been much discussed (Gervais et al. 2014; Hofstra et al. 2008); therefore, the aim of this study was not to suggest any specific method for regridding precipitation but to point out the spatial variability of differences between the original and regridded precipitation and the effects of regridding for different quantiles, specifically for the extremes that lead to floods and droughts. The size of the shift also effects the statistics of precipitation; as the grid shift decreases, the mean values of mean, standard deviation, and high quantiles are close to those of the original data. However, there is no effect of the shift size on P_0. And as the grid resolution becomes finer, average precipitation is further reduced, with the standard deviation declining faster than for low-resolution data. Effects of regridding are also high for low quantiles, yet the size of the shift dominates the effects at high quantiles.

It is challenging to regrid precipitation due to its high spatial variation across short distances and shifting distribution with temporal and spatial scales. At short time steps (e.g., daily and hourly) the precipitation distribution is J shaped and at long time steps (e.g., annual), it is bell shaped (Papalexiou et al. 2013; Papalexiou and Koutsoyiannis 2016). At fine spatial scales (e.g., 1 km), the precipitation distribution is J shaped, and at coarse spatial scales (e.g., 100 km), it is bell shaped because the precipitation is averaged over the area of the grid. Further research is required to select a regridding method based on grid size. For example, when the grid size is small, the nearest-neighbor method could be suitable for regridding to a fine resolution, but this may not be a good choice when this is to be regridded to a coarse resolution as the precipitation distribution varies. In addition, understanding the impacts of the shift size on precipitation scales and distribution shapes could be studied that are important in assessing impacts associated with changes in precipitation (Schoof 2012; Wilson and Toumi 2005). Also, the effects of projected coordinate systems or map projections on precipitation, specifically at high latitudes, due to different spatial interpolation methods could be explored. Though it was noted that distance-weighted interpolation has no effect for different projections (Jiang et al. 2014), this is to be examined for other techniques and grid sizes as well.

Overall, regridding should be used with caution as it can alter the statistical properties of precipitation to a great extent and adds uncertainty to analysis of combined precipitation products. It is preferable to calculate indices from each precipitation product and then to interpolate the derived indices rather than attempt to calculate an index from a regridded precipitation product formed from coupling of original data.

Acknowledgments. The authors acknowledge the financial contributions of the Canada First Research Excellence Fund's Global Water Futures program, the Natural Sciences and Engineering Research Council of Canada, the Canada Research Chairs program, and the Pacific Institute for Mathematical Studies for support of this research. The authors thank the one eponymous reviewer (Dr. Justin Schoof), the two anonymous reviewers, and the editor for their constructive comments and suggestions that helped to improve the paper. Color maps were inspired by ColorBrewer, version 2.0 (Harrower and Brewer 2003), and ColorMoves (Samsel et al. 2018).

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, https://doi.org/10.1175/1520-0434(2003)018<0918:SOPFSS>2.0.CO;2.

Akinsanola, A. A., K. O. Ogunjobi, V. O. Ajayi, E. A. Adefisan, J. A. Omotosho, and S. Sanogo, 2017: Comparison of five gridded precipitation products at climatological scales over West Africa. *Meteor. Atmos. Phys.*, **129**, 669–689, https://doi.org/10.1007/s00703-016-0493-6.

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.

Ashouri, H., K.-L. Hsu, S. Sorooshian, D. K. Braithwaite, K. R. Knapp, L. D. Cecil, B. R. Nelson, and O. P. Prat, 2015:

PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Amer. Meteor. Soc.*, **96**, 69–83, https://doi.org/10.1175/BAMS-D-13-00068.1.

Battersby, S. E., 2009: The effect of global-scale map-projection knowledge on perceived land area. *Cartographica*, **44**, 33–44, https://doi.org/10.3138/carto.44.1.33.

Beck, H. E., and Coauthors, 2017: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.*, **21**, 6201–6217, https://doi.org/10.5194/hess-21-6201-2017.

Berndt, C., and U. Haberlandt, 2018: Spatial interpolation of climate variables in northern Germany—Influence of temporal resolution and network density. *J. Hydrol. Reg. Stud.*, **15**, 184–202, https://doi.org/10.1016/j.ejrh.2018.02.002.

Booth, J. F., C. M. Naud, and J. Willison, 2018: Evaluation of extratropical cyclone precipitation in the North Atlantic basin: An analysis of ERA-Interim, WRF, and two CMIP5 models. *J. Climate*, **31**, 2345–2360, https://doi.org/10.1175/JCLI-D-17-0308.1.

Chen, J., and F. P. Brissette, 2017: Hydrological modelling using proxies for gauged precipitation and temperature. *Hydrol. Processes*, **31**, 3881–3897, https://doi.org/10.1002/hyp.11304.

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

Chen, T., and Coauthors, 2017: Comparison of spatial interpolation schemes for rainfall data and application in hydrological modeling. *Water*, **9**, 342, https://doi.org/10.3390/w9050342.

Contractor, S., L. V. Alexander, M. G. Donat, and N. Herold, 2015: How well do gridded datasets of observed daily precipitation compare over Australia? *Adv. Meteor.*, **2015**, 325718, https://doi.org/10.1155/2015/325718.

Demaria, E. M. C., E. P. Maurer, J. Sheffield, E. Bustos, D. Poblete, S. Vicuña, and F. Meza, 2013: Using a gridded global dataset to characterize regional hydroclimate in central Chile. *J. Hydrometeor.*, **14**, 251–265, https://doi.org/10.1175/JHM-D-12-047.1.

Diaconescu, E. P., P. Gachon, and R. Laprise, 2015: On the remapping procedure of daily precipitation statistics and indices used in regional climate model evaluation. *J. Hydrometeor.*, **16**, 2301–2310, https://doi.org/10.1175/JHM-D-15-0025.1.

Dinku, T., S. J. Connor, P. Ceccato, and C. F. Ropelewski, 2008: Comparison of global gridded precipitation products over a mountainous region of Africa. *Int. J. Climatol.*, **28**, 1627–1638, https://doi.org/10.1002/joc.1669.

Donat, M. G., J. Sillmann, S. Wild, L. V. Alexander, T. Lippmann, and F. W. Zwiers, 2014: Consistency of temperature and precipitation extremes across various global gridded in situ and reanalysis datasets. *J. Climate*, **27**, 5019–5035, https://doi.org/10.1175/JCLI-D-13-00405.1.

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, https://doi.org/10.1175/2008JAMC1757.1.

Fischer, R., S. Nowicki, M. Kelley, and G. A. Schmidt, 2014: A system of conservative regridding for ice–atmosphere coupling in a general circulation model (GCM). *Geosci. Model Dev.*, **7**, 883–907, https://doi.org/10.5194/gmd-7-883-2014.

Fuka, D. R., M. T. Walter, C. MacAlister, A. T. Degaetano, T. S. Steenhuis, and Z. M. Easton, 2014: Using the Climate Forecast System Reanalysis as weather input data for watershed models. *Hydrol. Processes*, **28**, 5613–5623, https://doi.org/10.1002/hyp.10073.

Gautam, S., C. Costello, C. Baffaut, A. Thompson, B. M. Svoma, Q. A. Phung, and E. J. Sadler, 2018: Assessing long-term hydrological impact of climate change using an ensemble approach and comparison with global gridded model—A case study on Goodwater Creek Experimental Watershed. *Water*, **10**, 564, https://doi.org/10.3390/w10050564.

Gervais, M., L. B. Tremblay, J. R. Gyakum, and E. Atallah, 2014: Representing extremes in a daily gridded precipitation analysis over the United States: Impacts of station density, resolution, and gridding methods. *J. Climate*, **27**, 5201–5218, https://doi.org/10.1175/JCLI-D-13-00319.1.

Ghodichore, N., C. T. Dhanya, and R. Vinnarasi, 2019: Examination of mean precipitation and moisture transport in reanalysis products over India. *ISH J. Hydraul. Eng.*, **25**, 51–61, https://doi.org/10.1080/09715010.2017.1364983.

Harrison, L., C. Funk, and P. Peterson, 2019: Identifying changing precipitation extremes in sub-Saharan Africa with gauge and satellite products. *Environ. Res. Lett.*, **14**, 085007, https://doi.org/10.1088/1748-9326/ab2cae.

Harrower, M., and C. A. Brewer, 2003: ColorBrewer.org: An online tool for Selecting colour schemes for maps. *Cartogr. J.*, **40**, 27–37, https://doi.org/10.1179/000870403235002042.

Haylock, M. R., N. Hofstra, A. M. G. K. Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.*, **113**, D20119, https://doi.org/10.1029/2008JD010201.

Henn, B., A. J. Newman, B. Livneh, C. Daly, and J. D. Lundquist, 2018: An assessment of differences in gridded precipitation datasets in complex terrain. *J. Hydrol.*, **556**, 1205–1219, https://doi.org/10.1016/j.jhydrol.2017.03.008.

Hofstra, N., M. Haylock, M. New, P. Jones, and C. Frei, 2008: Comparison of six methods for the interpolation of daily, European climate data. *J. Geophys. Res.*, **113**, D21110, https://doi.org/10.1029/2008JD010100.

Hu, Z., Q. Zhou, X. Chen, J. Li, Q. Li, D. Chen, W. Liu, and G. Yin, 2018: Evaluation of three global gridded precipitation data sets in central Asia based on rain gauge observations. *Int. J. Climatol.*, **38**, 3475–3493, https://doi.org/10.1002/joc.5510.

Jiang, W., J. Li, and Geoscience Australia, 2014: The effects of spatial reference systems on the predictive accuracy of spatial interpolation methods. Research Data Australia, accessed 6 August 2021, https://researchdata.edu.au/effects-spatial-reference-interpolation-methods/688287.

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

Kidd, C., P. Bauer, J. Turk, G. J. Huffman, R. Joyce, and D. Braithwaite, 2012: Intercomparison of high-resolution precipitation products over northwest Europe. *J. Hydrometeor.*, **13**, 67–83, https://doi.org/10.1175/JHM-D-11-042.1.

Lavers, D., L. Luo, and E. F. Wood, 2009: A multiple model assessment of seasonal climate forecast skill for applications. *Geophys. Res. Lett.*, **36**, L23711, https://doi.org/10.1029/2009GL041365.

Lee, H., D. E. Waliser, R. Ferraro, T. Iguchi, C. D. Peters-Lidard, B. Tian, P. C. Loikith, and D. B. Wright, 2017: Evaluating hourly rainfall characteristics over the U.S. Great Plains in dynamically downscaled climate model simulations using

NASA-Unified WRF. *J. Geophys. Res. Atmos.*, **122**, 7371–7384, https://doi.org/10.1002/2017JD026564.

Li, J., A. D. Heap, A. Potter, and J. J. Daniell, 2011: Application of machine learning methods to spatial interpolation of environmental variables. *Environ. Modell. Software*, **26**, 1647–1659, https://doi.org/10.1016/j.envsoft.2011.07.004.

Li, R., J. Shi, D. Ji, T. Zhao, V. Plermkamon, S. Moukomla, K. Kuntiyawichai, and J. Kruasilp, 2019: Evaluation and hydrological application of TRMM and GPM precipitation products in a tropical monsoon basin of Thailand. *Water*, **11**, 818, https://doi.org/10.3390/w11040818.

McGinnis, S. A., L. R. McDaniel, and L. O. Mearns, 2010: Effects of spatial interpolation algorithm choice on regional climate model data analysis. *2010 Fall Meeting*, San Francisco, CA, Amer. Geophys. Union, Abstract GC43F-1016, http://www.narccap.ucar.edu/doc/pubs/mcginnis-agu-poster-2010.pdf.

Papalexiou, S. M., and D. Koutsoyiannis, 2016: A global survey on the seasonal variation of the marginal distribution of daily precipitation. *Adv. Water Resour.*, **94**, 131–145, https://doi.org/10.1016/j.advwatres.2016.05.005.

——, ——, and C. Makropoulos, 2013: How extreme is extreme? An assessment of daily rainfall distribution tails. *Hydrol. Earth Syst. Sci.*, **17**, 851–862, https://doi.org/10.5194/hess-17-851-2013.

Rajulapati, C. R., S. M. Papalexiou, M. P. Clark, S. Razavi, G. Tang, and J. W. Pomeroy, 2020: Assessment of extremes in global precipitation products: How reliable are they? *J. Hydrometeor.*, **21**, 2855–2873, https://doi.org/10.1175/JHM-D-20-0040.1.

Rauscher, S. A., E. Coppola, C. Piani, and F. Giorgi, 2010: Resolution effects on regional climate model simulations of seasonal precipitation over Europe. *Climate Dyn.*, **35**, 685–711, https://doi.org/10.1007/s00382-009-0607-7.

Raziei, T., I. Bordi, and L. S. Pereira, 2011: An application of GPCC and NCEP/NCAR datasets for drought variability analysis in Iran. *Water Resour. Manage.*, **25**, 1075–1086, https://doi.org/10.1007/s11269-010-9657-1.

Risser, M. D., C. J. Paciorek, M. F. Wehner, T. A. O'Brien, and W. D. Collins, 2019: A probabilistic gridded product for daily precipitation extremes over the United States. *Climate Dyn.*, **53**, 2517–2538, https://doi.org/10.1007/s00382-019-04636-0.

Robeson, S. M., 1997: Spherical methods for spatial interpolation: Review and evaluation. *Cartogr. Geogr. Inf. Syst.*, **24**, 3–20, https://doi.org/10.1559/152304097782438746.

Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, https://doi.org/10.1175/JCLI-D-12-00823.1.

Samsel, F., S. Klaassen, and D. H. Rogers, 2018: ColorMoves: Real-time interactive colormap construction for scientific visualization. *IEEE Comput. Graphics Appl.*, **38**, 20–29, https://doi.org/10.1109/MCG.2018.011461525.

Schoof, J. T., 2012: Scale issues in the development of future precipitation scenarios. *J. Contemp. Water Res. Educ.*, **147**, 8–16, https://doi.org/10.1111/j.1936-704X.2012.00399.x.

Schulzweida, U., 2019: CDO user guide (version 1.9.8). Zenodo, 222 pp., https://zenodo.org/record/4246983#.YXAVkhrMJPY.

Shen, S. S. P., P. Dzikowski, G. Li, and D. Griffith, 2001: Interpolation of 1961–97 daily temperature and precipitation data onto Alberta polygons of ecodistrict and soil landscapes of Canada. *J. Appl. Meteor. Climatol.*, **40**, 2162–2177, https://doi.org/10.1175/1520-0450(2001)040<2162:IODTAP>2.0.CO;2.

Slinskey, E. A., P. C. Loikith, D. E. Waliser, and A. Goodman, 2019: An extreme precipitation categorization scheme and its observational uncertainty over the continental United States. *J. Hydrometeor.*, **20**, 1029–1052, https://doi.org/10.1175/JHM-D-18-0148.1.

Sun, Q., D. Kong, C. Miao, Q. Duan, T. Yang, A. Ye, Z. Di, and W. Gong, 2014: Variations in global temperature and precipitation for the period of 1948 to 2010. *Environ. Monit. Assess.*, **186**, 5663–5679, https://doi.org/10.1007/s10661-014-3811-9.

——, C. Miao, Q. Duan, H. Ashouri, S. Sorooshian, and K.-L. Hsu, 2018: A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Rev. Geophys.*, **56**, 79–107, https://doi.org/10.1002/2017RG000574.

Tissot, A., 1881: *Mémoire sur la Représentation des Surfaces et les Projections des Cartes Géographiques.* Gauthier-Villars, 337 pp.

Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo, 2014: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resour. Res.*, **50**, 7505–7514, https://doi.org/10.1002/2014WR015638.

Willmott, C. J., and S. M. Robeson, 1995: Climatologically aided interpolation (CAI) of terrestrial air temperature. *Int. J. Climatol.*, **15**, 221–229, https://doi.org/10.1002/joc.3370150207.

Wilson, P. S., and R. Toumi, 2005: A fundamental probability distribution for heavy rainfall. *Geophys. Res. Lett.*, **32**, L14812, https://doi.org/10.1029/2005GL022465.

Yue, X., L. J. Mickley, and J. A. Logan, 2014: Projection of wildfire activity in Southern California in the mid-twenty-first century. *Climate Dyn.*, **43**, 1973–1991, https://doi.org/10.1007/s00382-013-2022-3.

Zhang, Q., H. Körnich, and K. Holmgren, 2013: How well do reanalyses represent the southern African precipitation? *Climate Dyn.*, **40**, 51–962, https://doi.org/10.1007/s00382-012-1423-z.