

Statistical Downscaling Using Localized Constructed Analogs (LOCA)*

DAVID W. PIERCE

*Division of Climate, Atmospheric Sciences, and Physical Oceanography, Scripps Institution of Oceanography,
University of California, San Diego, La Jolla, California*

DANIEL R. CAYAN

*Division of Climate, Atmospheric Sciences, and Physical Oceanography, Scripps Institution of Oceanography,
University of California, San Diego, and U.S. Geological Survey, La Jolla, California*

BRIDGET L. THRASHER

Climate Analytics Group, Menlo Park, California

(Manuscript received 16 April 2014, in final form 7 July 2014)

ABSTRACT

A new technique for statistically downscaling climate model simulations of daily temperature and precipitation is introduced and demonstrated over the western United States. The localized constructed analogs (LOCA) method produces downscaled estimates suitable for hydrological simulations using a multiscale spatial matching scheme to pick appropriate analog days from observations. First, a pool of candidate observed analog days is chosen by matching the model field to be downscaled to observed days over the region that is positively correlated with the point being downscaled, which leads to a natural independence of the downscaling results to the extent of the domain being downscaled. Then, the one candidate analog day that best matches in the local area around the grid cell being downscaled is the single analog day used there. Most grid cells are downscaled using only the single locally selected analog day, but locations whose neighboring cells identify a different analog day use a weighted combination of the center and adjacent analog days to reduce edge discontinuities. By contrast, existing constructed analog methods typically use a weighted average of the same 30 analog days for the entire domain. By greatly reducing this averaging, LOCA produces better estimates of extreme days, constructs a more realistic depiction of the spatial coherence of the downscaled field, and reduces the problem of producing too many light-precipitation days. The LOCA method is more computationally expensive than existing constructed analog techniques, but it is still practical for downscaling numerous climate model simulations with limited computational resources.

1. Introduction

A changing climate will have numerous impacts on society, such as increasing peak energy demand, altering freshwater availability, changing ecosystems, modifying growing degree days and pests relevant to agriculture,

and increasing heat-related health risks (e.g., Parry et al. 2007). Properly anticipating many of these impacts requires climate change information on a spatial scale on the order of 10 km, yet global climate models (GCMs) still rarely have a spatial resolution finer than on the order of 100 km, leading to a disconnect between the information available from GCMs and that needed to inform climate change adaptation strategies. This is particularly true for hydrological simulations, which are sensitive to elevation, local soil properties, topography, and slope orientation, and so can benefit from higher spatial resolutions than global models provide.

This difference in scales is generally addressed using downscaling, that is, techniques that infer smaller-spatial-scale structure using the original coarse-resolution fields

* Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JHM-D-14-0082.s1>.

Corresponding author address: David W. Pierce, Division of Climate, Atmospheric Sciences, and Physical Oceanography, Scripps Institution of Oceanography, University of California, San Diego, Mail Stop 0224, La Jolla, CA 92093-0224.
E-mail: dpierce@ucsd.edu

along with finer-scale observations, topography, and dynamics. There are two main types of downscaling: dynamical, which uses regional climate models driven at the domain boundaries by global climate model output, and statistical, which uses the historic relationships between large- and small-scale conditions. Dynamical methods can capture nonstationary relationships between the large- and finescale that may develop in the future and can produce more variables than is practical for statistical methods, but are orders of magnitude more computationally expensive than the statistical methods, require large volumes of global climate model output, and typically need to be bias corrected before use in applications anyway. Our purpose is to provide an improved downscaling technique for impact studies that have a limited computational budget but wish to sample a range of different global model results, so we explore statistical downscaling in this work.

There are many different forms of statistical downscaling [reviews can be found in Wilby et al. (2004), Fowler et al. (2007), and Maraun et al. (2010); see also Hwang and Graham (2014)]. These methods include stochastic weather generators (e.g., Wilby et al. 1998; Wilks 2012); various approaches that use large-scale fields as predictors for fine-resolution fields through regression or artificial neural networks (e.g., von Storch et al. 1993; Schoof and Pryor 2001; Chen et al. 2014); weather-type methods that use observed associations between characteristic recurring large-scale weather patterns and local responses typically seen when that weather type is present (e.g., Goodess and Palutikof 1998); and even simple “delta methods,” which typically add the time-mean model-predicted change to the sequence of observations. Many of these methods use different climate variables for the coarse- and fine-resolution fields (e.g., using large-scale atmospheric circulation to predict local precipitation).

Two statistical downscaling methods in particular have been widely applied over the western United States: bias correction with spatial disaggregation (BCSD; Wood et al. 2004) and “constructed analog”-based techniques such as the original constructed analogs (CA; Hidalgo et al. 2008), bias correction with constructed analogs (BCCA; Maurer et al. 2010), and multivariate adapted constructed analogs (MACA; Abatzoglou and Brown 2012). Unlike many of the statistical downscaling methods noted above, these procedures use a coarse-resolution depiction of a climate variable to produce a fine-resolution version of the same climate variable. BCSD has proven to be an effective technique (e.g., Maurer et al. 2010), but as traditionally implemented it is a monthly method, which limits the possible changes in the shape of daily distributions and does not preserve the GCM’s daily

weather sequences [although Piani et al. (2010) and Thrasher et al. (2012) give extensions to the daily time scale, which is evaluated in Hwang and Graham (2014)]. Our interests include impacts that are affected by both the distribution and specific sequences of daily extremes and how those might change in the future, such as energy use during heat waves (e.g., California electricity use rises after the third consecutive day of hot weather) and floods generated by sequential days of heavy precipitation, so we focus on improving the constructed analog process rather than changes to BCSD.

Constructed analogs are inspired by analog weather forecasting techniques (van den Dool 1994), where the best matching historical occurrence of a target pattern is identified and the forecast assumes that the weather will evolve the same way it did before. A difficulty with analog techniques is that the current pattern never looks exactly like the historical analog. Constructed analogs address this by identifying the N_a (typically 30) best matching analog days, then optimally combining the N_a analog days to best reproduce the target pattern.

When applied to the downscaling problem, the “forecast” is of finescale conditions given a coarse-scale (i.e., GCM) predictor field. The observations are first coarsened to the GCM grid, and then the N_a -coarsened historical analog days that best match the climate model day to be downscaled are identified. The N_a optimal weights of those coarse analog days that best reproduce the model day to be downscaled are computed and then applied to the original finescale observed fields to form a weighted average of N_a days. The result is the downscaled field in CA and BCCA; MACA adds an additional postdownscaling bias correction step (Hildago et al. 2008; Maurer et al. 2010; Abatzoglou and Brown 2012).

The constructed analog method is computationally efficient because all weights for the downscaling are found by operations on the coarse GCM grid. However, forming a weighted average of analog days has some drawbacks. First, it smears finescale spatial features and increases the spatial coherence of the final downscaled field. This affects flooding, which is influenced by the spatial coherence of the precipitation field (the same problem has been found in BCSD; Zhang and Georgakakos 2012; Hwang and Graham 2014). Second, averaging tends to reduce the temporal variance of the final result (e.g., von Storch 1999). Model output statistics (MOS) correctors can partially take this into account but have problems of their own, particularly when the spatial coherence is artificially large (Maraun 2013). Third, averaging tends to produce too much drizzle when downscaling precipitation, because an original global model day that has precipitation in one area but is clear in another area can easily be matched to analog days that have

precipitation in both areas. The weighted sum of the analog days will then have low levels of precipitation in both areas. This issue was noted in [Hidalgo et al. \(2008\)](#), who found that skill of the CA technique for daily precipitation increased as the domain size decreased.

Here, we introduce the localized constructed analogs (LOCA) technique, which was developed to address these issues. Rather than forming the downscaled field as a weighted sum of the N_a analog days, LOCA constructs the downscaled field, point by point, from the single analog day in the pool of N_a analog days that best matches weather in the local region around the point being considered. By avoiding averaging, many of the problems noted above are diminished.

This work evaluates LOCA through its downscaling of daily precipitation and temperature. Daily precipitation is of obvious importance to hydrological simulations. Daily maximum temperature can also have a limited effect on evapotranspiration and is of primary importance to nonhydrological applications such as energy use and human health. The domain used is the western United States, which provides a rich test set of hydrological conditions to evaluate the method, ranging from the arid southwestern deserts to the rain forest of the Olympic peninsula. However, tests over the conterminous United States indicate that the downscaling performs similarly over that broader domain.

Our experimental design, and the data required to implement it, are described in [section 2](#). The LOCA technique is described in [section 3](#), with details in the [appendix](#). Results for LOCA downscaling daily maximum temperature and precipitation are provided in [sections 4](#) and [5](#), respectively. An example of LOCA multivariate downscaling to determine daily minimum temperature is given in [section 6](#). A summary and conclusions are presented in [section 7](#).

2. Data and experimental design

An entire downscaling system based on constructed analogs potentially consists of multiple steps, including regridding data from multiple climate models to a common grid, bias correction, and the spatial downscaling step. MACA adds an additional bias correction after the downscaling. In this work, we are primarily interested in the spatial downscaling step alone; bias correction will be treated in a separate work. That is, we focus on the question of how well a fine-resolution field can be computed given an unbiased but coarse-resolution input field, and we do not address the problem of producing an unbiased input field from a climate model. We believe that taking the opportunity to distinguish between errors arising from downscaling versus those arising from

residual biases in the fields being downscaled can add clarity to the analysis.

Accordingly, we examine LOCA downscaling observations over the period 1940–69 that have been aggregated to a $1^\circ \times 1^\circ$ grid. Observations over the period 1970–2005 are used as the training data. There are two advantages to downscaling the coarsened observations for our purposes. First, the role of bias correction is minimized, because the statistical properties of the data over the period 1940–69 are similar to those found in the training data, more so than would likely be found when downscaling data from a climate model. Nonetheless, the two periods differ, primarily from natural internal climate variability. So, analogous to a climate model, bias correction still needs to be applied to the 1940–69 data. We use the quantile mapping approach from [Wood et al. \(2004\)](#), but the effect of bias correction on our analysis is small, and the same bias correction method is used for both LOCA and BCCA (to which we compare our results). Second, downscaling coarsened observations means we have the original, fine-resolution dataset available for comparison.

We follow the example of the U.S. Bureau of Reclamation (USBR) phases 3 and 5 of the Coupled Model Intercomparison Project (CMIP3 and CMIP5, respectively) archive website (http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/dcpInterface.html) by downscaling starting from a common $1^\circ \times 1^\circ$ latitude–longitude grid (the southwesternmost grid cell is centered at 29.5°N , 125.5°W ; the northeasternmost grid cell is centered at 52.5°N , 104.5°W). All references to the coarse grid will therefore refer to this $1^\circ \times 1^\circ$ grid.

We use the Livneh observationally based gridded product ([Livneh et al. 2013](#)) on a $1/16^\circ$ latitude–longitude grid over the western United States (west of 105°W) for both our coarsened data to be downscaled and our training data. We use anomalies when downscaling temperature and absolute values when downscaling precipitation (the original CA scheme used anomalies for precipitation, while BCCA and MACA use absolute values after bias correction, as is done here). Results shown here are produced using version 1.0 of the Livneh et al. observational dataset, which was the latest available when this study was undertaken. We later checked the LOCA downscaling results using version 1.2 of the Livneh et al. dataset and found them little changed.

3. LOCA technique

a. Overview

The basic physical assumption of constructed analog downscaling techniques is that meteorological processes

produce cyclostationary statistical relationships between large-scale, area-averaged values (0.5° – 2° latitude–longitude) and finer-scale estimates of a climatological field. Global climate model outputs are then considered to be estimates of the large-scale averages, and the observed relationships between area-averaged coarse and fine scales are used to infer a plausible distribution of values on a fine-resolution grid ($1/16^{\circ}$ here).

To downscale a model day using LOCA, the 30 observed days that best match the model day in the wider region around the point being downscaled are found, then the single one of those 30 days that best matches the model day in the local neighborhood of the point being downscaled is identified. This multiscale matching is a key aspect of LOCA and ensures that the final downscaled field is consistent with the day being downscaled on both local and synoptic length scales. The final downscaled value is the value from the single best matching observed day, scaled so that its amplitude matches the model day being downscaled (additively for temperature, multiplicatively for precipitation).

b. LOCA procedure

A higher-level, conceptual view of the LOCA downscaling process is given in the main text; details can be found in the [appendix](#). The steps for LOCA downscaling are as follows.

- 1) Select the locations at which the 30 analog days will be chosen [[section 3b\(1\)](#); [appendix section a](#)]. This differs from existing constructed analog techniques, which use the same analog days at all points in the domain. This step is only done once; the following steps are repeated for each day being downscaled.
- 2) Select the pool of 30 analog days at each location identified in step 1 [[section 3b\(2\)](#); [appendix section b](#)].
- 3) Out of the pool of 30 analog days, find the one that best matches the model field being downscaled in the local neighborhood of the point to which it is being downscaled [[section 3b\(3\)](#); [appendix section c](#)].
- 4) Construct the final downscaled field by scaling the observed day to match the model day [[section 3b\(4\)](#); [appendix section d](#)].

1) ANALOG POOL LOCATIONS AND SPATIAL MASKS

Existing constructed analog techniques use a single pool of N_a analog days for downscaling everywhere in the domain. Those analog days are determined by computing, over the entire domain, the N_a best-fit days between the model field and the coarsened observations. In LOCA, different locations can use different pools of analog days. We refer to the locations where the pool of

analog days is chosen as the analog pool points, and we select them from the coarse grid [[Fig. 1](#) (top left) shows the analog pool points used here]. Our sensitivity tests show little influence of the exact positioning or number of analog pool points on the final result. Although every coarse-scale grid point could be selected, this is too computationally expensive, so we only choose a limited number of locations.

At every analog pool point, we use a spatial mask to limit the region over which the model field is compared to the coarsened observations when determining the fit between the two. The mask is computed from the coarsened observations; a different mask is used for each combination of variable and season [December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON)]. The mask is set to 1 at locations where the Pearson correlation with the time series at the analog pool point is greater than 0. This masking, which is not included in the existing constructed analog techniques, allows us to choose the analog days based on a limited region of influence. It serves little purpose to require the analog days to be able to describe precipitation in locations so far apart that the precipitation patterns are unrelated. Limiting the analog pool domain renders the downscaling insensitive to changes in the domain beyond a certain region of influence; for example, analogs for summer precipitation in Arizona are not influenced by precipitation in Montana [[Fig. 1](#) (bottom left)]. This feature allows LOCA to be applied on a very large, even global domain.

[Figure 1](#) (top right), (bottom left), and (bottom right) shows the precipitation masks for different analog pool points and seasons. The masks tend to be smaller in summer than winter (by about 9.5% on average), owing to the change to more convective systems in summer rather than the large-scale planetary wave–driven precipitation in winter. For more details on selecting the analog pool locations, see [section a](#) of the appendix.

2) SELECTING THE ANALOG DAYS AT THE REGIONAL SCALE

Our method of selecting analog days follows that used in [Hidalgo et al. \(2008\)](#), although in LOCA, this is done at all analog pool points rather than once for the entire domain.

At each analog pool location and day to be downscaled, we choose the pool of N_a ($=30$ here) analog days to use for downscaling as those observed days that have the smallest root-mean-square (RMS) difference with the model field being downscaled (taking the domain mask into account). This matching is done on the coarse grid. We further require that the day of year of an analog day must be within 45 days of the day of year of the

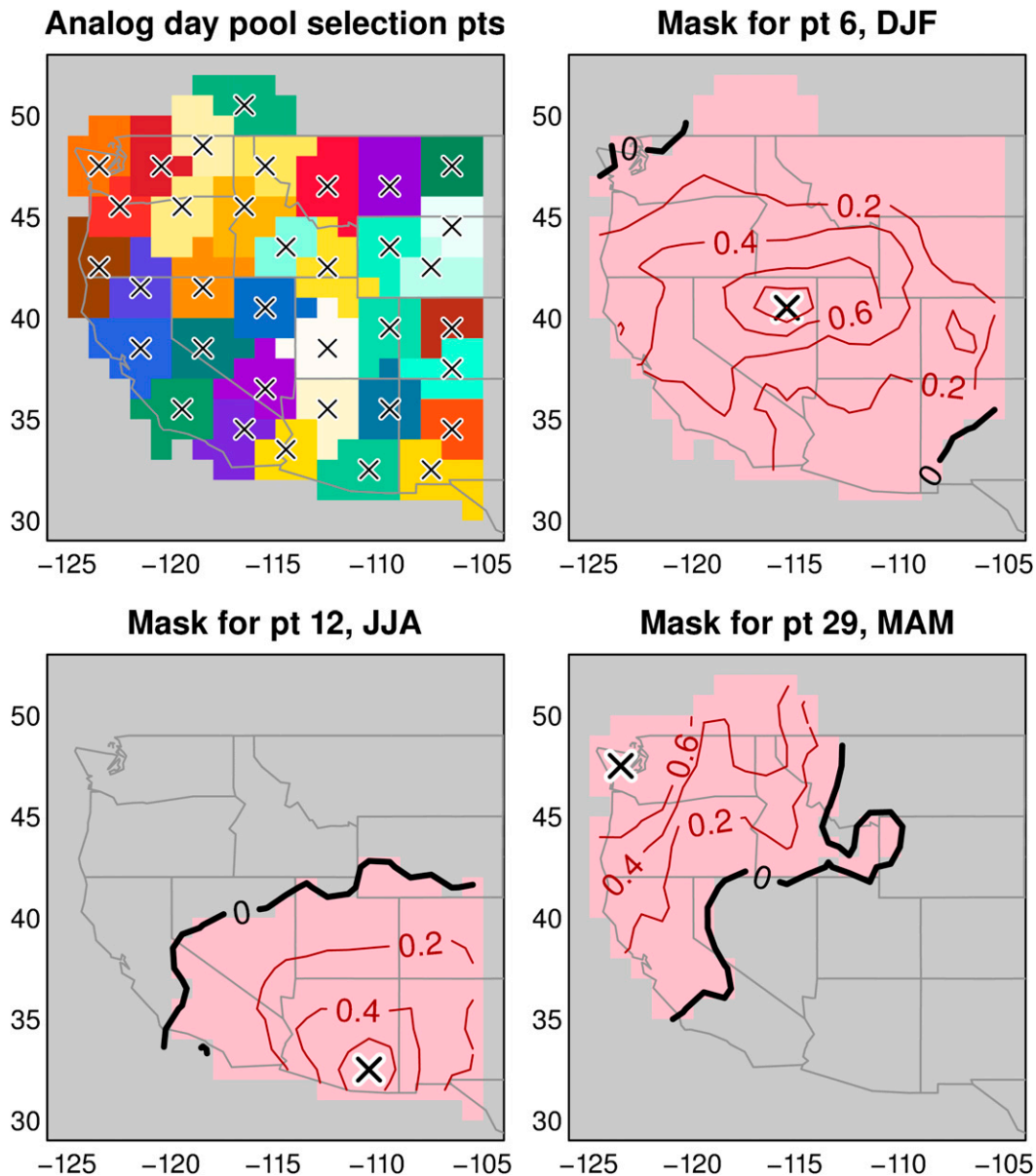


FIG. 1. (top left) Crosses indicate the analog day pool selection points; colors indicate the domain over which a point is used. (top right), (bottom left), (bottom right) Example weighting masks [Γ in Eq. (A1)] for precipitation in three of the points and seasons. Pink areas show where the mask is 1; gray areas show where the mask is 0. Contours show the temporal correlation between the daily time series of values (for the indicated season only) at the analog pool selection point, indicated by a cross, and every other location (plotted for positive values only).

model day being downscaled, as done in [Hidalgo et al. \(2008\)](#). Downscaling at each point on the fine-resolution grid then draws from the pool of N_a analog days obtained at the closest analog pool point. This approach allows us to retain the desirable constructed analog property that downscaled results at the fine grid cells are constructed exclusively from analog days that are the best match to large-scale conditions, yet reduces the sensitivity of the downscaled results to the chosen

domain size. For more details on selecting the pools of analog days, see [section b](#) of the appendix.

3) FINDING THE ONE BEST MATCHING ANALOG DAY AT THE LOCAL SCALE

Once the pools of 30 analog days are obtained at all the analog pool locations, then at each location on the fine grid we select from that pool the single best matching analog day in the immediate neighborhood of the point to

which it is being downscaled. Because this matching must be implemented on the finescale grid, we first interpolate the model field to the fine grid using bicubic interpolation. We do not, however, locally match the interpolated model field to the observations at each location because the latter has significant information content at the finest spatial scales that the former lacks. Instead, we bicubic interpolate the coarsened observed data of the analog days back to the fine grid. This yields a representation of the analog days on the fine-resolution grid that have their finescale spatial information reduced in the same way as the model, so the two can be sensibly compared.

At each fine-resolution location, we find the single analog day that minimizes the RMS difference between the interpolated model field and interpolated analog day in a square region of size $2r + 1$ finescale grid cells around the center grid point being downscaled. In this work, we used $r = 10$. [Figure 2](#) (top left) shows, as an example, the analog day number selected when downscaling precipitation for 1 January 1940 (the first day of the period analyzed in [section 4](#)). For more details on finding the best match at the local scale, see [section c](#) of the appendix.

4) CONSTRUCTING THE FINAL DOWNSCALED FIELD

Once the best matching single analog day is found at each point on the fine-resolution grid, the final downscaled field can be constructed. We still need to match the amplitude of the values on the analog day to the amplitude of the model fields, similar to how BCCA calculates the optimal weights to best match the combination of analog days to the model day being downscaled. We do this amplitude matching after selecting the analog days so that, for example, an unusually warm model day is matched to an unusually warm observed day, not to an average observed day that has been scaled to match the model amplitude.

When downscaling precipitation, we scale the selected analog day by the ratio of the interpolated model field to the interpolated analog day. When downscaling temperature, we add to the selected analog day the difference between the interpolated model field and the interpolated analog day.

To reduce the chance of discontinuities between regions with different analog days, any grid cell adjacent to a cell being downscaled with a different analog day is treated as an “edge cell.” The final downscaled value at edge cells is the weighted sum of the values computed using the multiple analog days in question, where the weight is determined by the number of adjacent grid cells using the different analog day. Typically, $\sim 30\%$ of the points are edge cells. [Figure 2](#) (top right) shows the

edge points found when downscaling 1 January 1940. [Figure 2](#) (bottom left) shows the final downscaled field and [Fig. 2](#) (bottom right) shows the original observations on that day, using the experimental procedure described in [section 3](#). Many more examples of individual downscaled days, and maps of the correlation with the time series at a point, are shown in the supplemental materials (Figs. S5–S8). For details on constructing the final downscaled field, see [section d](#) of the appendix.

c. Multivariate downscaling

[Abatzoglou and Brown \(2012\)](#) and [Zhang and Georgakakos \(2012\)](#) have noted the importance of downscaling some quantities using information from multiple variables simultaneously. This is easily accomplished in LOCA. There are two steps at which a best-fit selection of analog days is performed: during the choosing of the pool of analog days [[section 3b\(2\)](#)] and when selecting the best single analog day in the local region around the fine-resolution grid point being downscaled [[section 3b\(3\)](#)]. When performing multivariate downscaling, in both these steps ranks are assigned to the candidate analog days for each variable such that rank 1 is the best match. Then, the weighted average rank of each candidate analog day across all the variables being downscaled is computed, and the day with the minimum weighted average rank is taken as the best match. The best matches 2–30 are chosen similarly. Weights are specified by the user and can be set so that all variables contribute equally (in which case the same analog day is used for all variables), one variable dominates the others, or only information from one variable is considered at all (in which case the result is the same as univariate downscaling). An example of multivariate downscaling is given in [section 6](#).

d. Changes in climatology

Because temperature is downscaled as an anomaly, the fine-resolution climatology is needed to construct the final downscaled field. For the historical period, observed climatology from the training data is used. However for future climate projections, the fine-resolution change in climatology (from historical to future) must be calculated. Different downscaling techniques handle this problem differently. BCSD removes the climatology from the coarse grid values, downscales the field to the fine grid, and then adds back in the bilinearly interpolated, model-predicted change in climatology ([Wood et al. 2004](#)). This is effective but means the change in climatology has no downscaled information, because it is determined purely by the global model. CA and BCCA do not treat the change in climatology in any special way; so, for example, a heat wave at the end of the century is merely

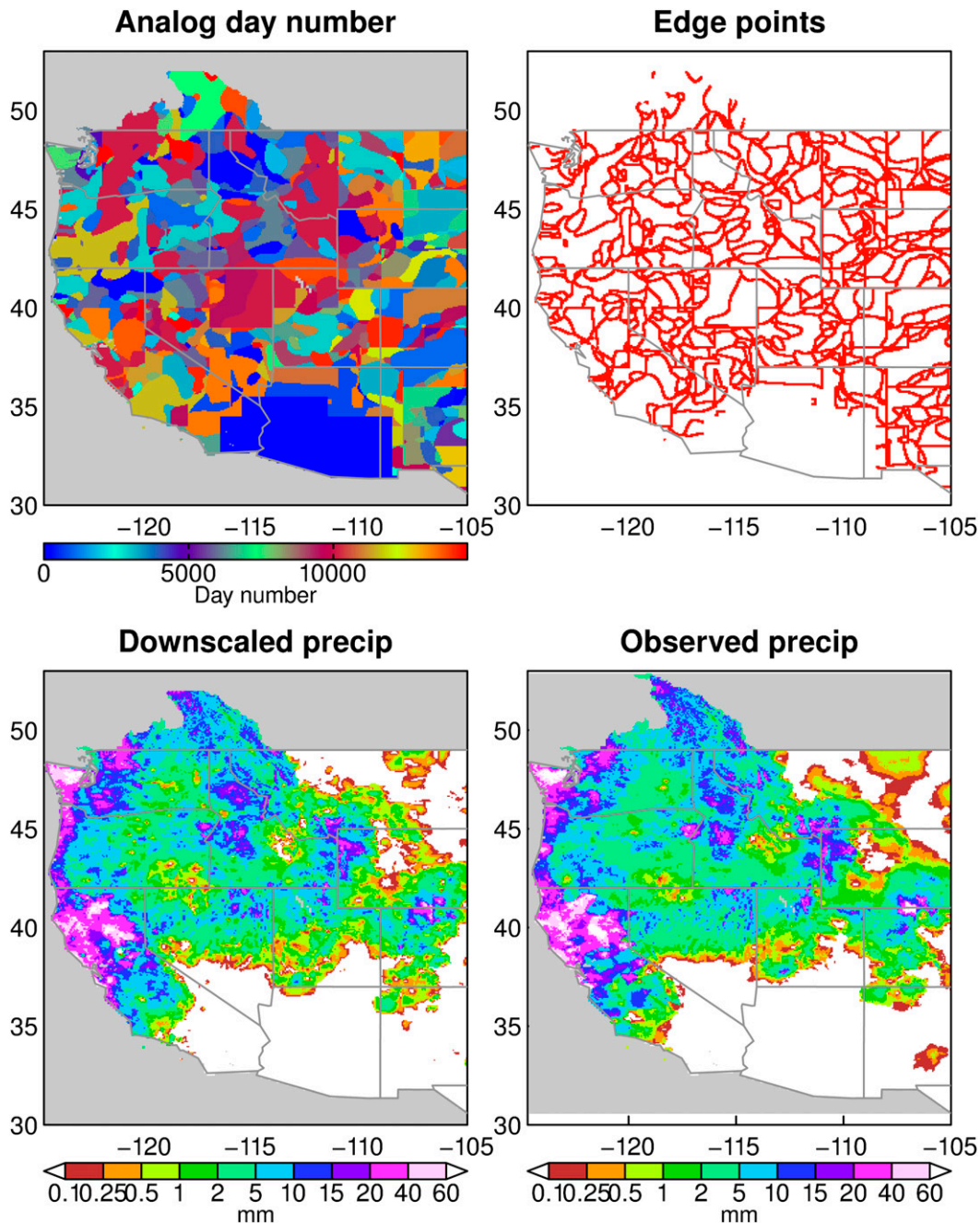


FIG. 2. Example fields when downscaling precipitation, illustrated for 1 Jan 1940. (top left) The selected analog day number (integer day number starting 1 Jan 1970 and ending 31 Dec 2010). (top right) Edge points (those adjoining a region with a different selected analog day number). (bottom left) Downscaled and (bottom right) observed precipitation field for that day (mm).

considered to have very large anomalies with respect to the historical era (Hildago et al. 2008; Maurer et al. 2010). The concern with that approach is that future anomalies might fall outside the range of historical variability, complicating the process of finding good historical analogs. MACA addresses this by removing the change in climatology, then

reintroducing it after the constructed analogs are obtained (Abatzoglou and Brown 2012).

LOCA follows the approach used in Pierce and Cayan (2013). The model-predicted change in climatology, which is available on the coarse grid, is simply considered another anomaly field to be downscaled and so is itself

downscaled using LOCA. However, the change in climatology represents, by definition, an average change across many individual days. Using a single analog day to represent the change in average conditions is inappropriate, not unlike how some of BCCA's shortcomings arise from representing a single day as the average of numerous analog days. Therefore, when downscaling the model-predicted change in climatology, the downscaling is repeated for all N_a analog days and the results are averaged together to produce the change in climatology.

This procedure means that the downscaled change in temperature climatology starts with the interpolated model change in temperature climatology, as is used in BCSD. However, in LOCA, this is additionally modified by the mean difference between the 30 original fine-resolution analog days and the coarse-resolution version of the 30 analog days. So, if the analog days suggest there is a consistent, small-spatial-scale feature in the change in climatology, it can be recovered by this process instead of only the interpolated model field being used. See [section e](#) of the appendix for additional details on calculating the change in climatology.

4. LOCA results: Daily maximum temperature

a. Monthly mean daily maximum temperature

[Figure 3](#) shows the monthly mean daily maximum temperature for selected months from the (left) observations and errors in (middle) BCCA and (right) LOCA. Averaged over the domain, the mean errors in LOCA's downscaled fields are small ($<0.01^\circ\text{C}$), while RMS errors are on the order of 0.3°C . BCCA's mean and RMS errors are larger than those found in LOCA, with a mean error of $\sim 0.03^\circ\text{C}$ and an RMS error of $\sim 0.5^\circ\text{C}$. LOCA's errors do not show any particular spatial pattern, which is a desirable characteristic as it implies that regional users of the downscaled product are unlikely to run into biases. The "spotty" error field seen in LOCA (both here and in later figures) may be influenced by two factors associated with the observed dataset as noted in [Gutmann et al. \(2014\)](#): 1) observed statistics tend to be slightly different at points with stations compared to points without stations and 2) the spots may be associated with stations that have moved or otherwise changed between the historical and training periods.

b. Hottest day in 1 and 20 years

Climate change is likely to increase extreme daily maximum temperatures as greenhouse gases accumulate in the atmosphere (e.g., [Parry et al. 2007](#)). This affects the peak electricity demand that the electrical system has to be designed for, poses health risks, and could affect wildfires, agriculture, and ecosystems. Here,

we look at the single hottest day per 1 year and per 20 years. The 1-day-in-1-yr extreme was obtained by calculating the highest value of maximum daily temperature in each year, then averaging across years. The 1-day-in-20-yr extreme was estimated using a bootstrap method, picking 20 random years from the dataset at a time, taking the maximum across the 20 years, then averaging across the 1000 trials. (Using a sliding 20-yr window gave similar results.) The short period of record means that the estimate is subject to large sampling uncertainty.

[Figure 4](#) shows the hottest day in 1 year and 20 years from (left) the observations and the error using (middle) BCCA and (right) LOCA. On average, LOCA values are within 0.1°C of the observed values and the RMS errors are $\sim 0.5^\circ\text{C}$. BCCA mean and RMS errors are both larger than those found with LOCA (with a mean error $\sim 0.5^\circ\text{C}$ and RMS error $\sim 1^\circ\text{C}$). More notable is that BCCA's error pattern displays a distinct cold bias along the West Coast, particularly in central California, Northern California, and Oregon. This may be due to problems representing the marine layer using domain-wide analog patterns, because the marine layer is confined to the coast. As much of California's population is concentrated along the coast, this cold bias could degrade simulations of peak energy use in the state.

c. Spatial variability in daily maximum temperature

Downscaled fields should realistically represent the spatial variability of the original observed field. To remove the strong effects of topography and other spatially varying land surface influences on daily maximum temperature, we first compute anomalies with respect to the daily climatological mean value, then calculate the spatial standard deviation in a 3×3 array of grid cells around the center point. Spatial variability is a function of the length scale over which it is evaluated, so we compute the variability progressively on data that have been aggregated to $1/8^\circ$ and $1/4^\circ$ (i.e., by a factor of 2 and 4) in each direction.

[Figure 5](#) shows the spatial variability ($^\circ\text{C}$) for an initial spatial coarsening of the temperature field by a factor of (top) 2 and (bottom) 4 for the (left) observations and the error with respect to the observations using (middle left) and (middle right) BCCA and (right) LOCA. LOCA captures the observed pattern accurately for the finest spatial scales (coarsening of 2) and reasonably well for the longer spatial scales. Averaged over the domain, the spatial standard deviations obtained from the downscaled data are within 0.02°C of those from the observations for a coarsening of 2 and 0.1°C for a coarsening of 4. BCCA downscaling has considerably larger errors than LOCA in both the mean and RMS error for both coarsening levels, with lower spatial variability in the downscaled result than observed.

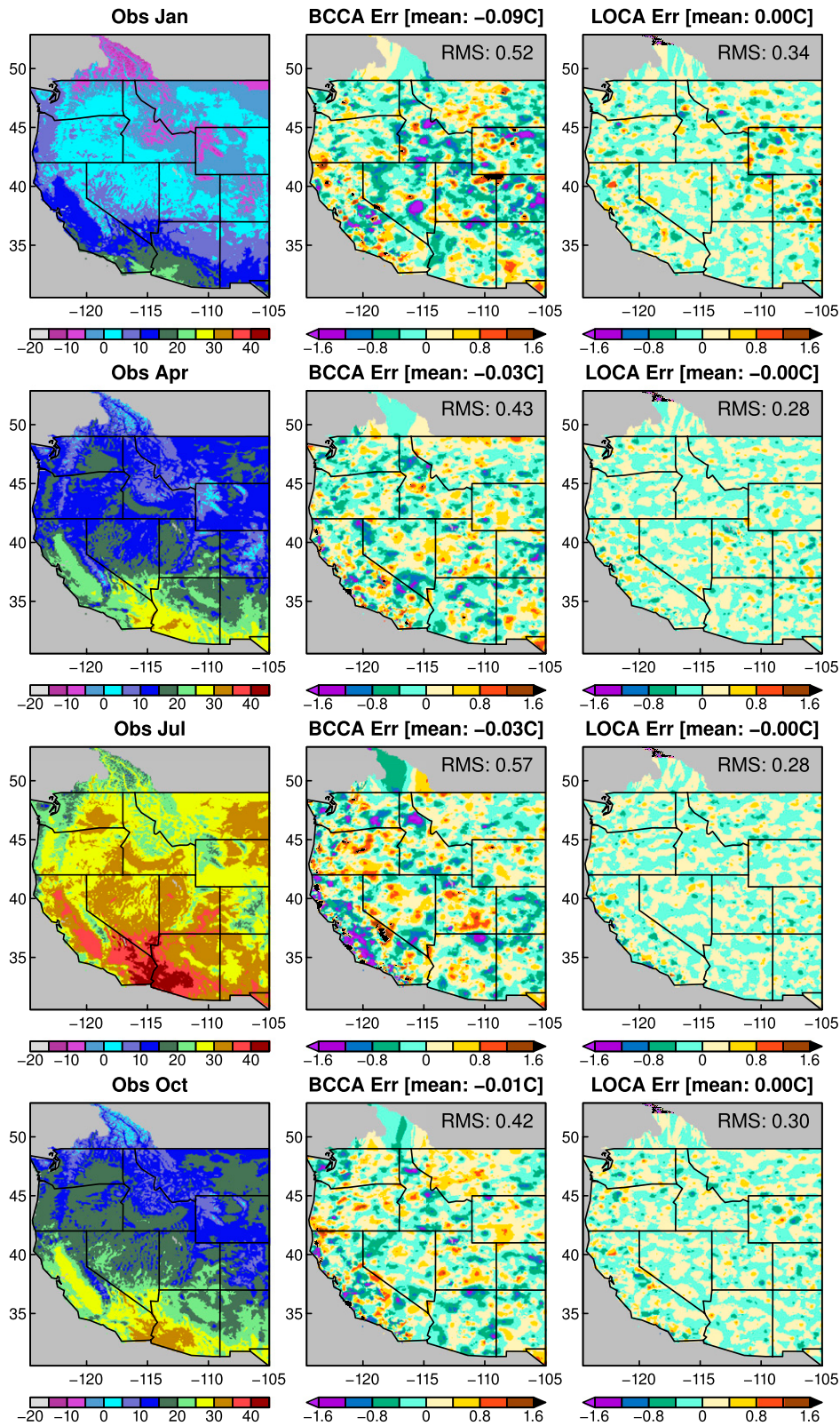


FIG. 3. Monthly mean daily max temperature ($^{\circ}\text{C}$) for selected months from (left) observations (1970–2010) and the error ($^{\circ}\text{C}$) with respect to observations after downscaling with (middle) BCCA and (right) LOCA.

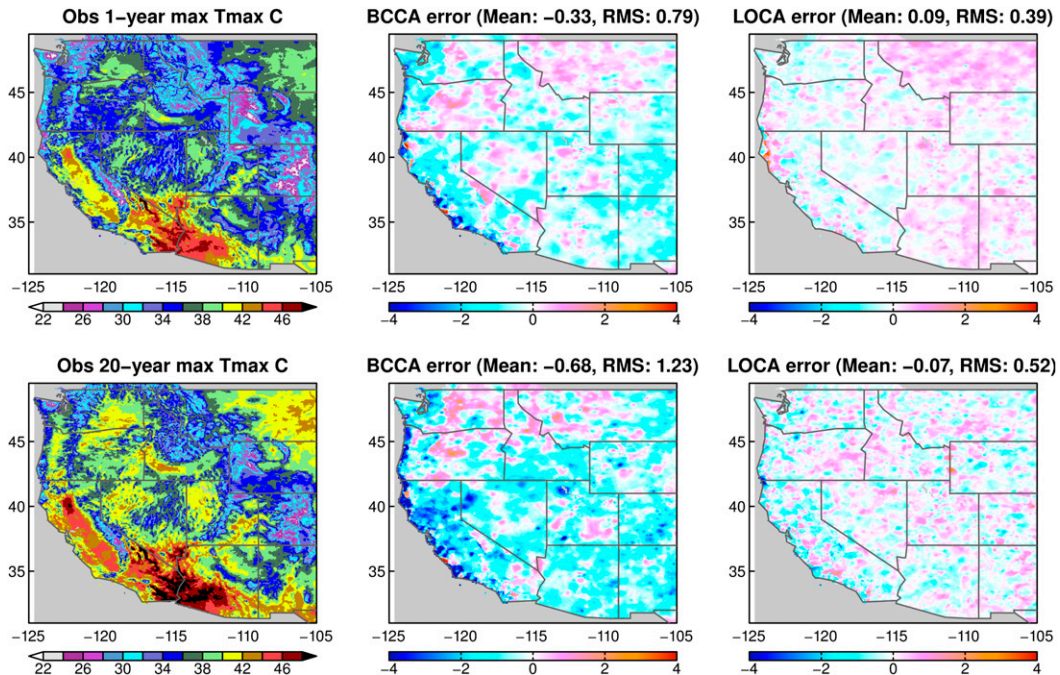


FIG. 4. Average highest daily max temperature ($^{\circ}\text{C}$) from (left) observations and the error ($^{\circ}\text{C}$) with respect to observations after downscaling with (middle) BCCA and (right) LOCA. Results are over (top) a year and (bottom) 20 years.

As mentioned in the introduction, much of this unrealistically large spatial coherence arises from the averaging of multiple analog days in BCCA. This is emphasized by Fig. 5 (middle right), which shows the error in spatial variability after a final bias correction step has been added after the BCCA downscaling process (similar to what MACA does). Even with postdownscaling bias correction, the spatial variance is lower than that found in either the observations or LOCA.

d. Temporal variability in daily maximum temperature

The downscaled field should properly represent temporal variability in the daily maximum temperature. This variability includes extremes, described above, but here we consider a broader representation of anomalous variation represented by the temporal standard deviation. Figure 6 shows the seasonally averaged temporal standard deviation computed at every point from the (left) observations and the error with respect to observations after downscaling with (middle) BCCA and (right) LOCA. Most of the year the standard deviation is highest over Montana and lowest along the West Coast, although this pattern changes to a north–south gradient in summer. LOCA slightly but consistently underestimates the variability, with a mean error of $\sim 0.05^{\circ}\text{C}$ on values that range from 3° to 8°C . The errors in LOCA have little spatial pattern, instead appearing evenly distributed across the

domain. BCCA has a considerably larger mean error than LOCA and is systematically biased low, as well as having larger RMS errors than LOCA. BCCA also has a deficit of variability adjacent to the west coast.

Another important aspect of the temporal behavior of daily maximum temperature is the duration of heat waves, which affects energy use, agriculture, ecosystems, and public health. Figure 7 (left) shows percentiles in the observed distribution of heat wave length (days), defined as consecutive days with daily maximum temperature $\geq 35^{\circ}\text{C}$. Values are only plotted in locations where there are at least 20 heat waves, except for the median (50th percentile) value, which shows values at all locations. Errors in the downscaled fields with respect to observations are shown in Fig. 7 (middle) and (right) for BCCA and LOCA, respectively. LOCA generally captures the distribution quite well, with a slight bias that rises to ~ 1 day shorter than observed at the 95th percentile, averaged across the domain. BCCA is similar to LOCA except in the desert region near the border of California and Arizona, where BCCA does notably worse than LOCA, especially in the 95th percentile values.

e. Correlation of downscaled daily maximum temperature with observations

Because we are downscaling coarsened observations, we can compute the temporal correlation between the downscaled field and original observations at every

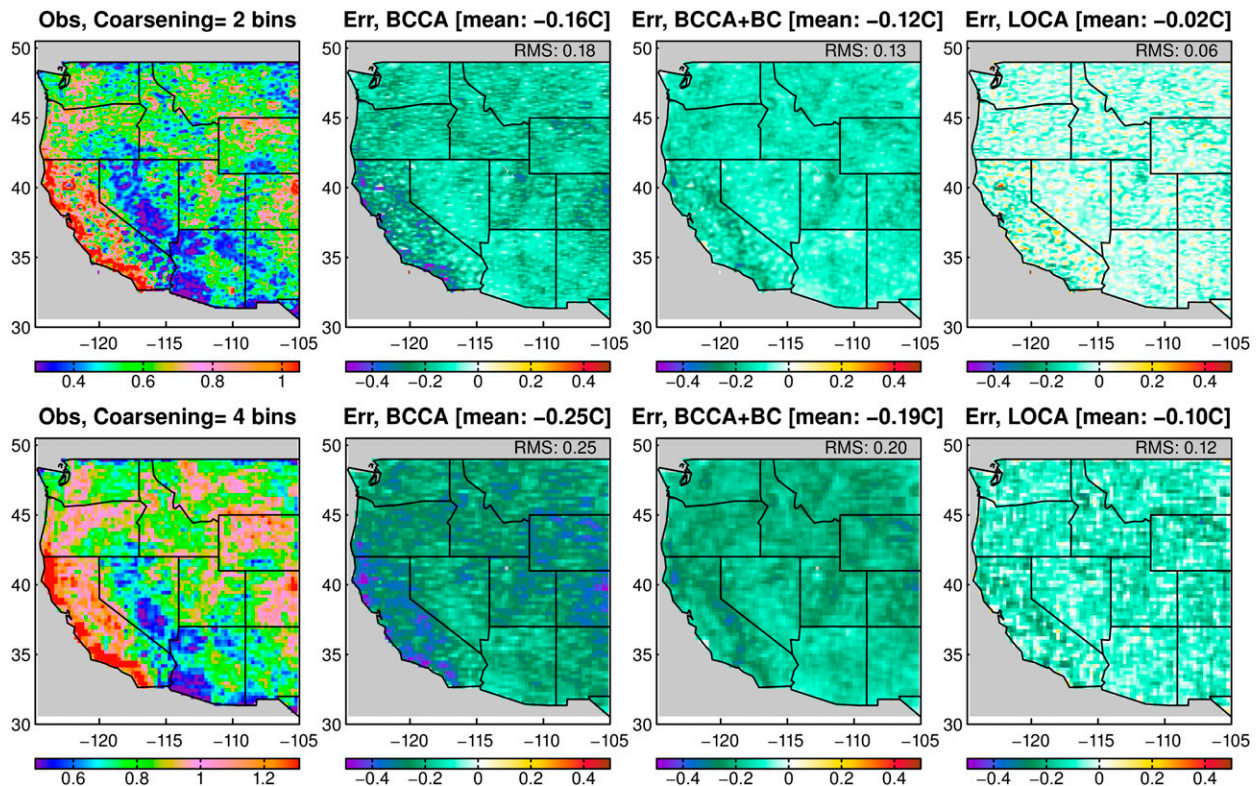


FIG. 5. The spatial variability of daily max temperature anomalies ($^{\circ}\text{C}$), computed as the spatial std dev in a 3×3 array around the center grid cell, from (left) observations and the error ($^{\circ}\text{C}$) with respect to observations after downscaling with (middle left) BCCA, (middle right) BCCA (after an additional postdownscaling bias correction is added to BCCA), and (right) LOCA. Results are shown for the daily max temperature anomaly field coarsened by a factor of (top) 2 and (bottom) 4 in each direction (lat, lon).

point. The correlation is shown in Fig. 8 for both (left) BCCA and (right) LOCA, using three ways of averaging the data. The correlation of the time series of daily anomalies is shown in Fig. 8 (top); for LOCA the average value is 0.95, although values are slightly lower along the coast than in the interior. This may be due to the coarse-scale field being unable to preserve information about coastal marine layer processes that are important to daily maximum temperature along the coast. Monthly averaged anomalies [Fig. 8 (middle)] are represented incrementally better than daily anomalies, while yearly averaged anomalies are somewhat worse [Fig. 8 (bottom)]. In all three cases, BCCA gives modestly lower correlation values than LOCA.

5. LOCA results: Precipitation

a. Annual and monthly average precipitation

Figure 9 compares the LOCA and BCCA results when downscaling January daily precipitation over the western United States, evaluating in terms of both percentage and actual value. As noted by Gutmann et al. (2014), BCCA

underpredicts the precipitation amount (averaged over the domain, the result is about 15% weaker than observed), particularly in the driest areas. In the wettest regions of the Sierra Nevada and Oregon Cascades the errors are slightly positive, so the errors appear to be inversely related to the amount of precipitation. The LOCA January mean, by contrast, is only 0.8% weaker than observed, shows a much reduced signature of the magnitude of the error being linked to the magnitude of precipitation, and has little spatial coherence. In both absolute (cm yr^{-1}) and relative (%) terms, the average LOCA error is ~ 20 times smaller than the BCCA error.

In July (Fig. 10), BCCA has a widespread region where downscaled precipitation is less than half that observed. Of course, in much of California, July is a very dry month, and LOCA has trouble accurately reproducing these extremely small values as well. However, the errors in LOCA for both January and July have the appearance of being mostly random noise or sampling errors, unlike the spatially coherent error fields found in BCCA. The tendencies seen here for January and July are also found in the other months and annual average (Figs. S1 and S2 in the supplemental materials).

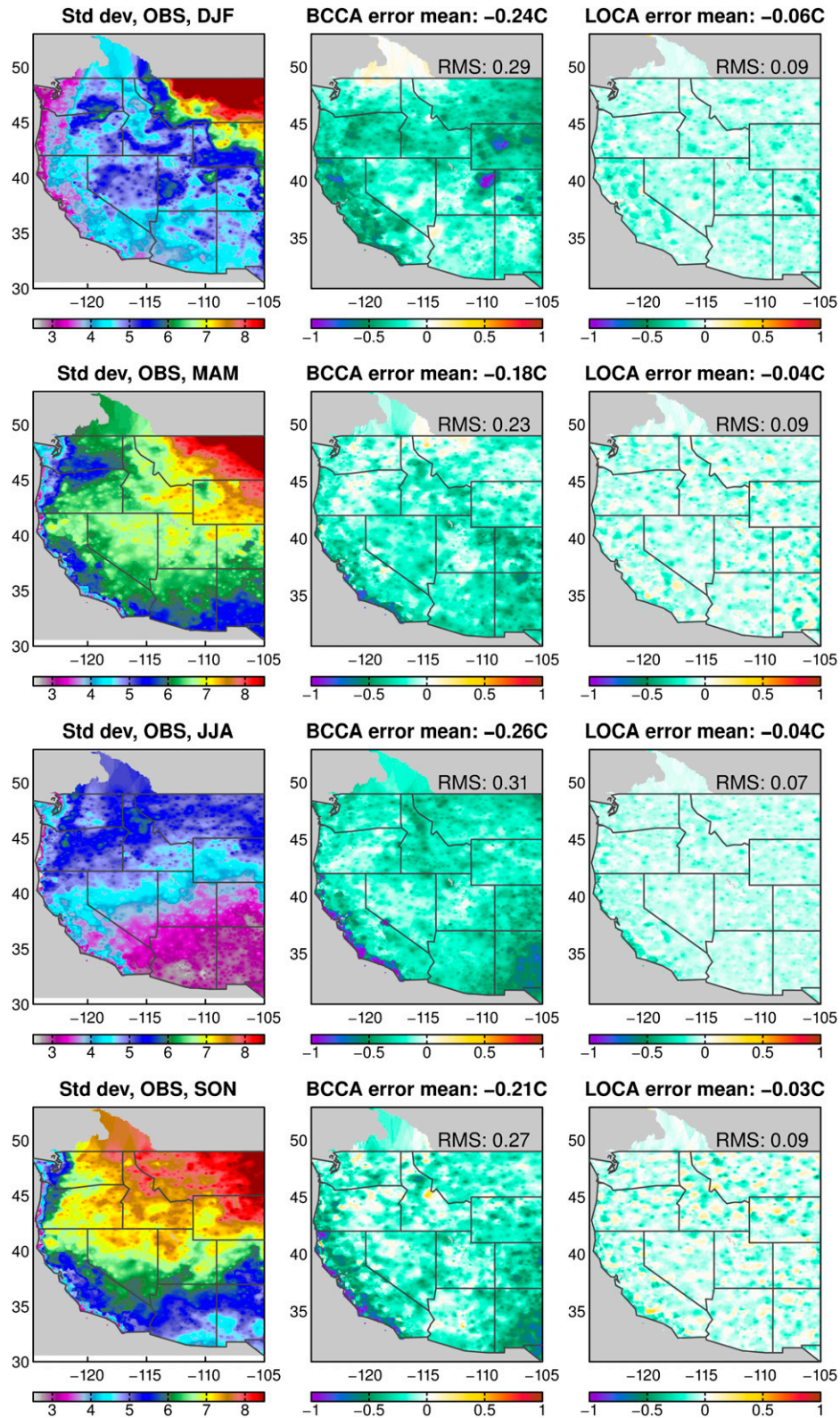


FIG. 6. Temporal std dev ($^{\circ}\text{C}$) of daily max temperature, by season, for (left) observations and the error ($^{\circ}\text{C}$) with respect to observations using (middle) BCCA and (right) LOCA.

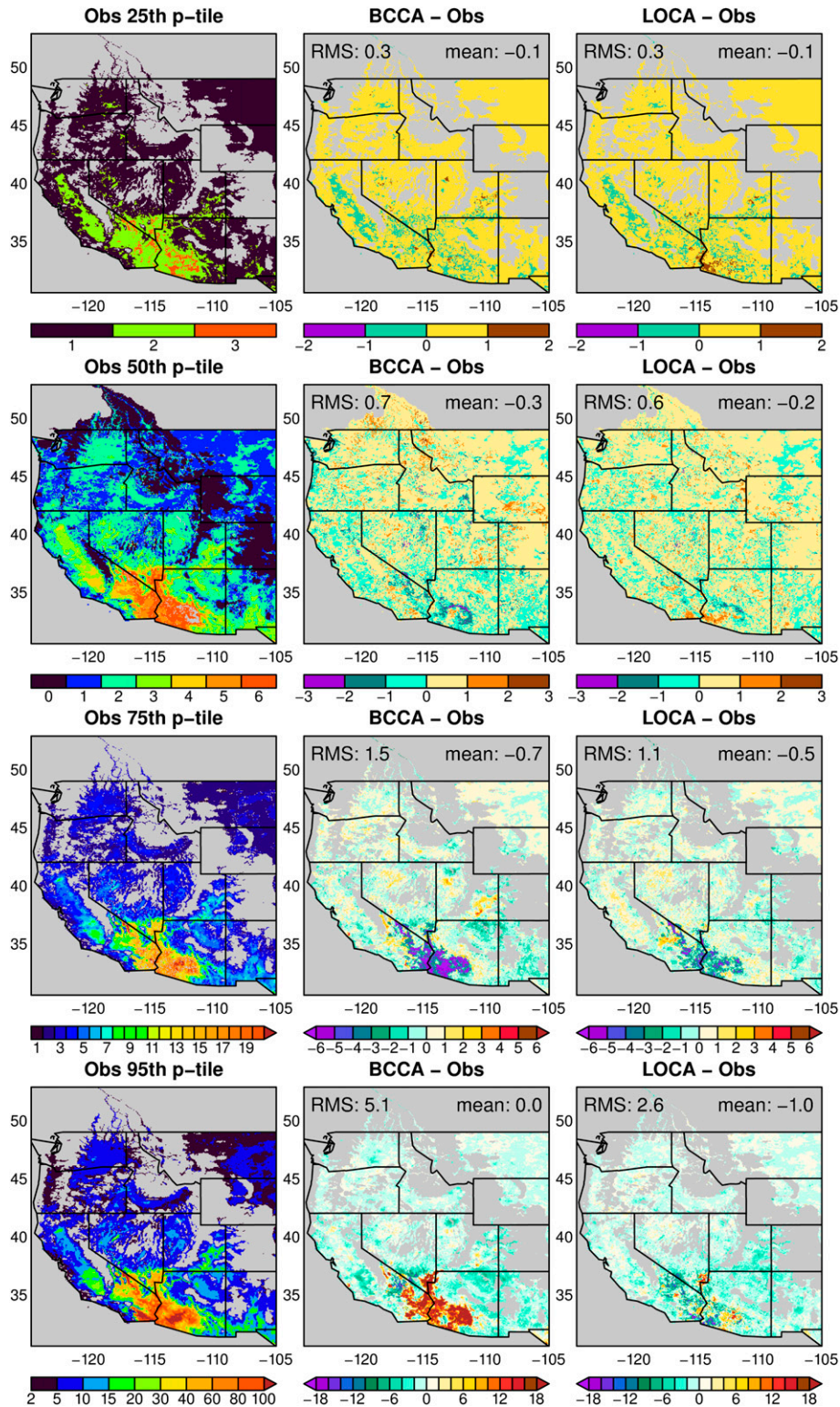


FIG. 7. Percentiles in the distribution of heat wave duration (days), where a heat wave is defined as a consecutive sequence of days with $T_{max} \geq 35^{\circ}\text{C}$, for (left) observations and the error (days) with respect to observations when downscaled using (middle) BCCA and (right) LOCA. For example, (upper middle) shows that the median (50th percentile) length of a heat wave in the border region between California and Arizona is 6 days.

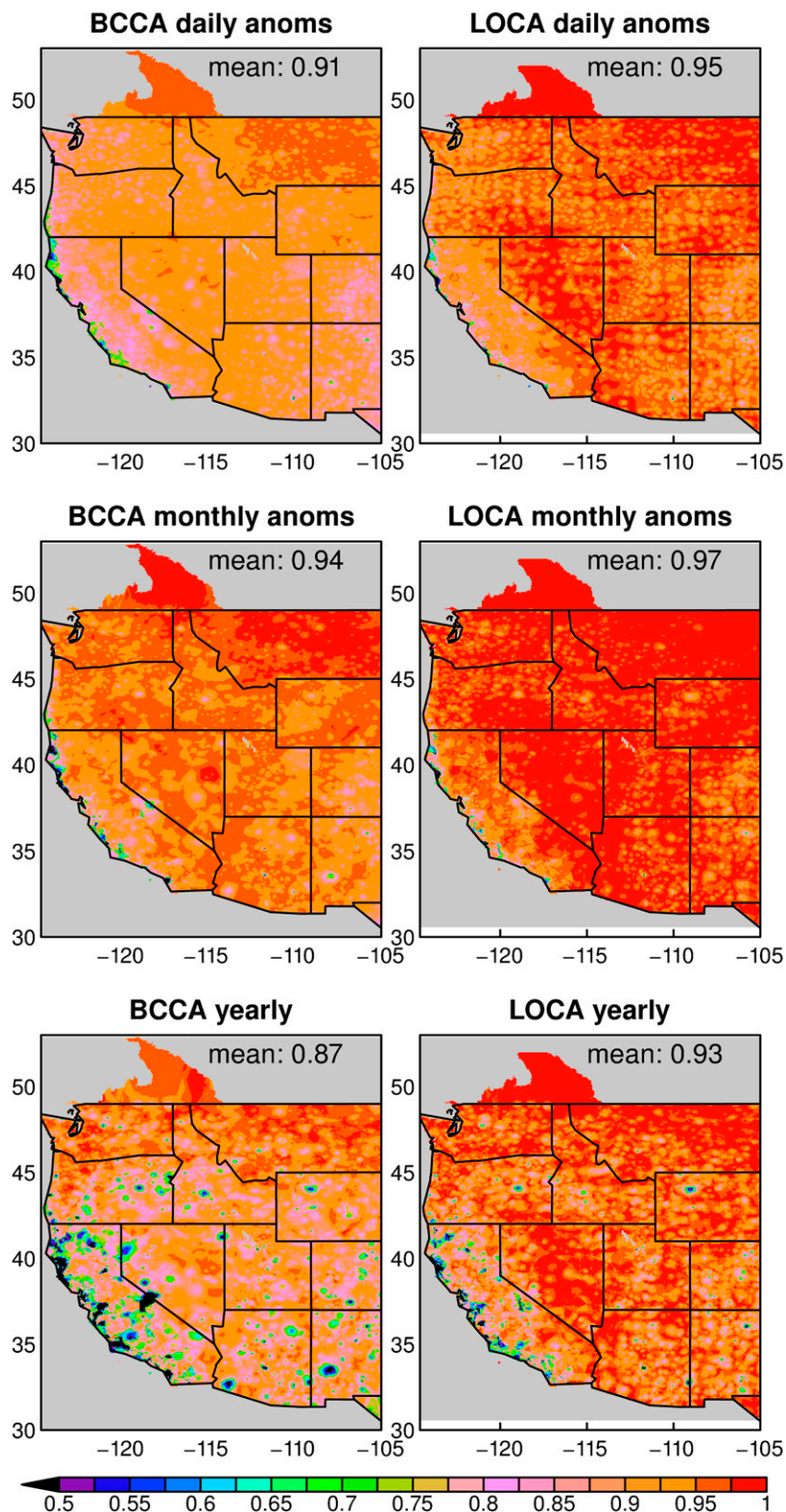


FIG. 8. Temporal correlation between observations and downscaled daily max temperature for (left) BCCA and (right) LOCA for (top) daily, (middle) monthly, and (bottom) yearly anomalies.

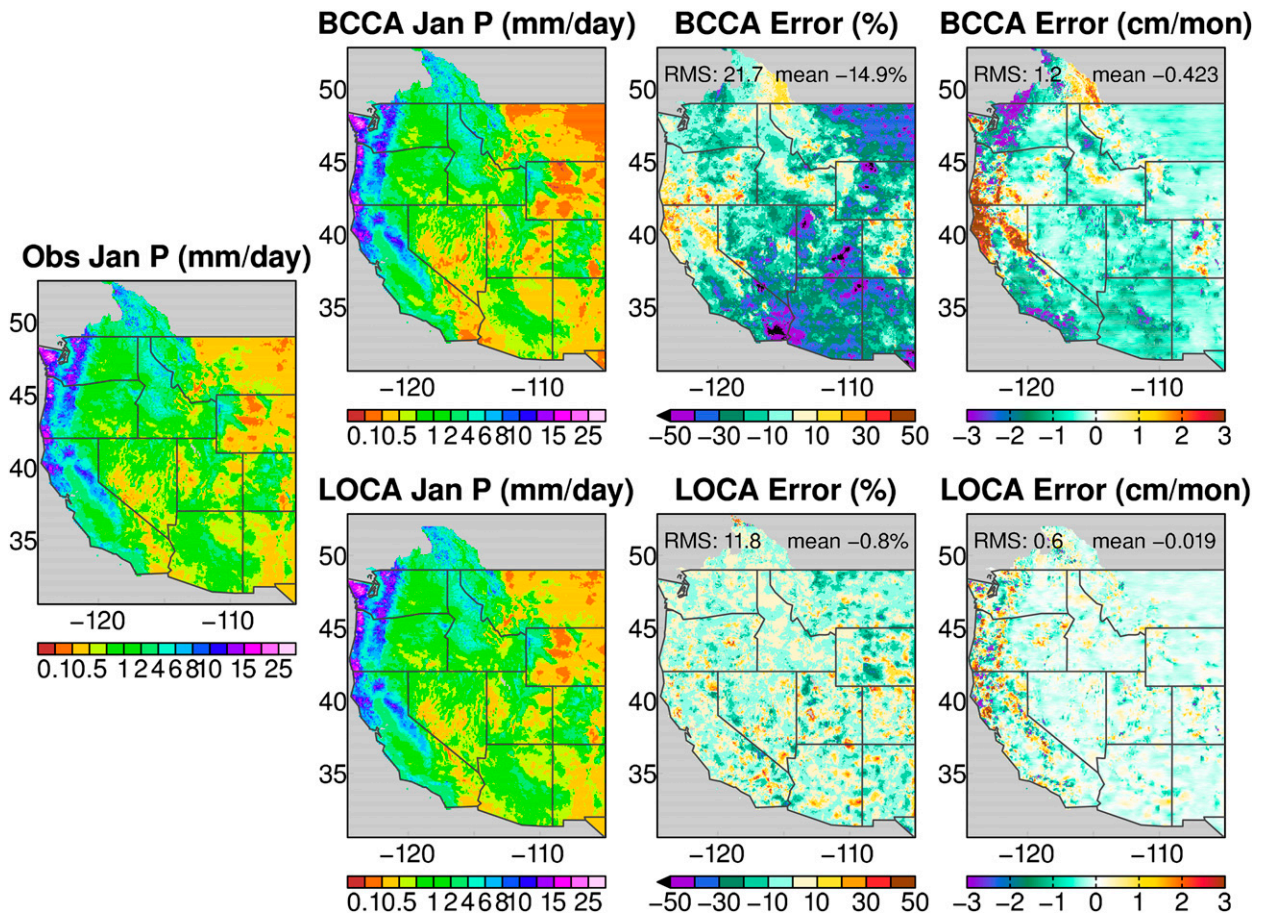


FIG. 9. A comparison of LOCA and BCCA downscaling January average precipitation (mm day^{-1}) over the western United States.

b. Precipitation extremes

Climate change has the potential of changing extreme high precipitation (e.g., Parry et al. 2007), which in some locations can dominate the overall annual precipitation change (Pierce et al. 2013). Also, climate change could increase flooding in mountainous watersheds as more winter storms drop rain instead of snow, and what snow there is melts earlier in the year (Das et al. 2013). This is a significant concern; for example, even in the present climate, flood damage has generated the largest total agricultural insurance payout in California over the past 15 years (Lobell et al. 2009). Here, we look at the single day with the greatest amount of precipitation typically found in 1 and 20 years. The methodology is the same as used for temperature extremes (section 3c).

Figure 11 shows the average yearly maximum precipitation in a 1-yr period. BCCA tends to underestimate the yearly maximum, by about 23% averaged over the domain. Similar to the pattern seen for monthly and annual mean precipitation, the error pattern has spatial structure, with overestimates in the high-precipitation

regions of the Sierra Nevada of California, the Oregon Cascades, and the Olympic peninsula and severe underestimates in the dry Mojave Desert region. By contrast, LOCA errors are much smaller and show little spatial coherence.

The average maximum daily precipitation experienced in 20 years exhibits similar patterns, only with more noise, which is not surprising considering the difficulty of estimating this quantity from only 30 years of data (not shown). BCCA continues to show systematic errors that are mildly positive over the Sierra Nevada, Oregon Cascades, and Olympic peninsula and strongly negative over dry areas in the Southwest. LOCA's average error is much smaller (mean of -4.2% versus -24.4% for BCCA) and shows minimal spatial patterns.

c. Precipitation spatial variability

Another key feature of a downscaled field that influences hydrological phenomena, including flooding, is the spatial variability of the precipitation field. If the downscaled field has too little spatial variability, then

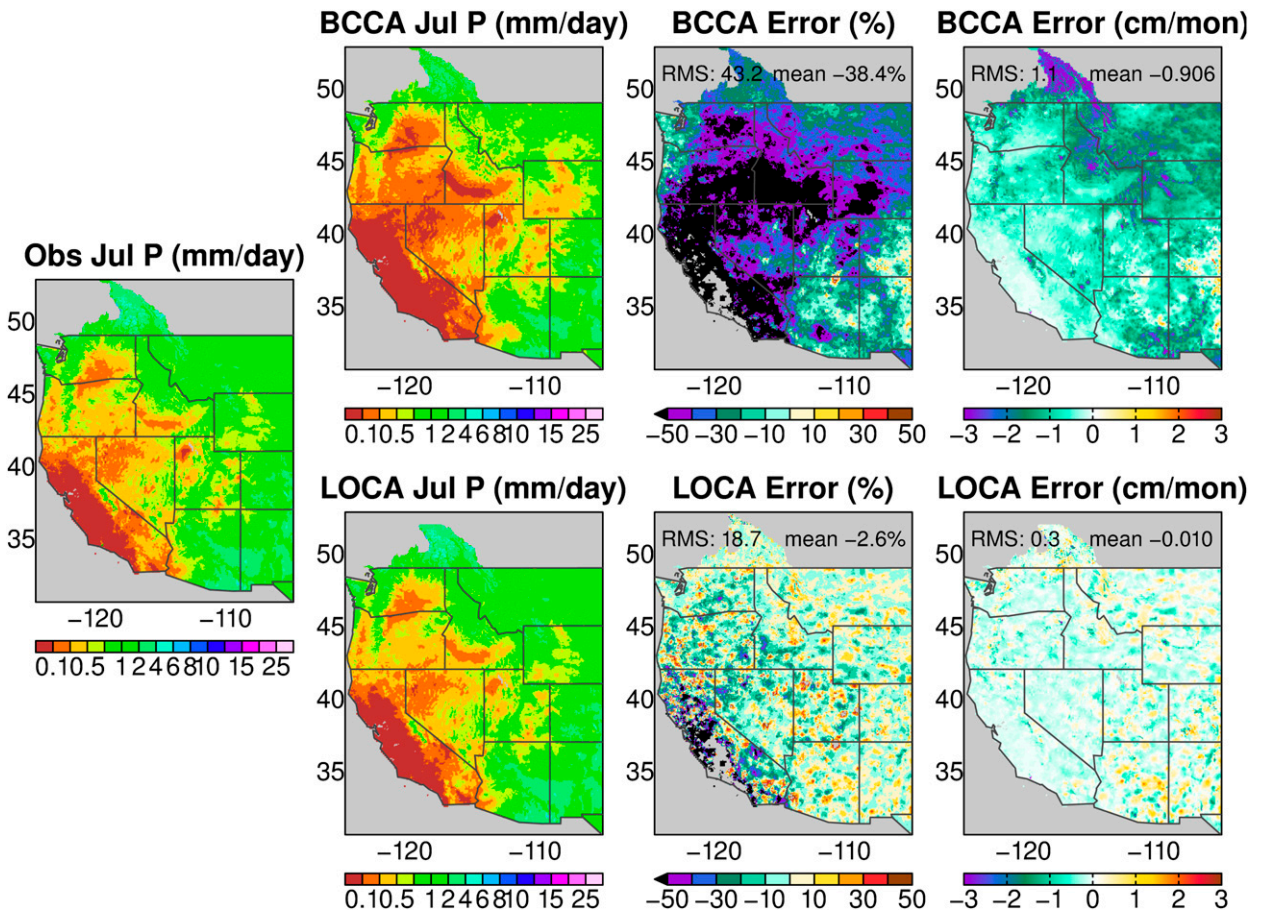


FIG. 10. As in Fig. 9, but for July.

adjoining catchments are more likely to receive high precipitation simultaneously, artificially inflating flooding probabilities.

We calculate the coefficient of spatial variability for precipitation the same as for daily maximum temperature (section 3d), with the exceptions that actual precipitation is used for the calculation instead of anomalies and that the spatial standard deviation is normalized by the precipitation value in the center grid point to produce a coefficient of variability (temperature anomalies cannot be normalized this way because values can be zero). We only include days when the center gridpoint precipitation is $\geq 2.5 \text{ mm day}^{-1}$ to avoid excessively inflating the value under conditions of little precipitation.

Figure 12 shows results for an initial spatial coarsening by a factor of (top) 2 and (bottom) 4. BCCA has overly weak spatial variability, by about 16% on average. LOCA has only one-fifth this error, and the sense is in the opposite direction (slightly too much spatial variability) for a coarsening of 2. Similar results are found for a spatial coarsening of 4 in each direction; the BCCA

results are noisier but otherwise little changed, while LOCA tends to slightly (3%) underestimate the spatial variability over almost all the domain.

As discussed above, averaging multiple analog days in BCCA increases the spatial coherence, so even bias correcting the overly weak BCCA precipitation fields after downscaling (like MACA does) leaves too much spatial coherence [Fig. 12 (middle right)]. The post-downscaling bias correction improves the representation of spatial coherence, but it still has roughly twice the error in the mean and RMS as does LOCA.

d. Precipitation temporal variability

A proper representation of precipitation’s temporal variability is key to impacts relating to flooding, drought, and agriculture. Figure 13 shows the temporal standard deviation of daily values at each point for all winter (DJF) days. BCCA underpredicts the values with a mean error of 21%, while LOCA is much closer, underpredicting the values with a mean error of 2%. In BCCA’s case, this reflects the negative bias in mean

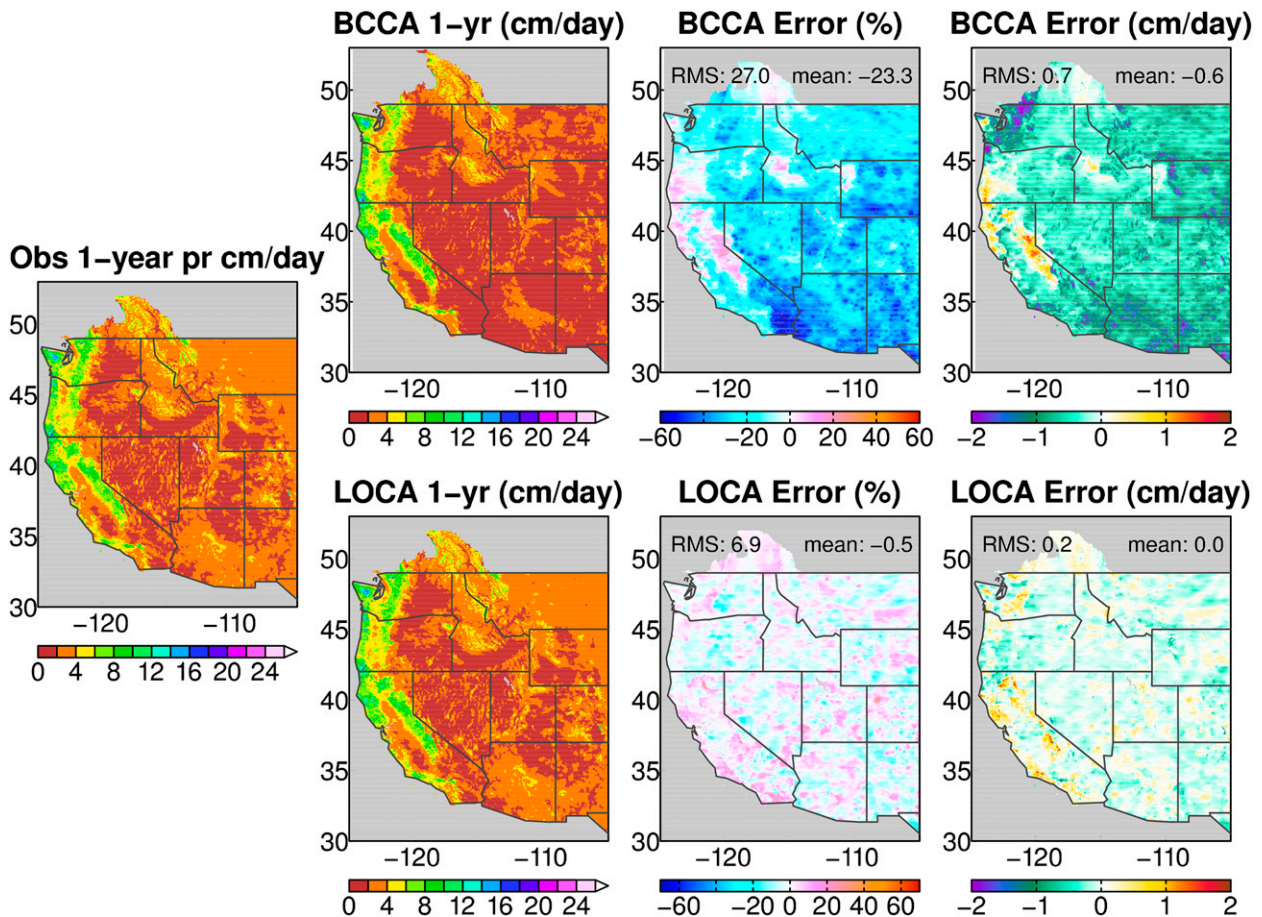


FIG. 11. The most precipitation that falls in a single day, on average over a year (cm day^{-1}).

precipitation (e.g., Figs. 9, 10). As seen in previous fields, BCCA tends to have errors that mirror the actual values of the field; for example, the errors are positive over the regions of large precipitation in the Sierra Nevada, Northern California, and southern coastal Oregon. By contrast, LOCA has a pattern of errors that appears to be closer to random sampling variability.

In summer (Fig. 14), the temporal standard deviation of observed precipitation drops to very low values over California, which is a challenge for both methods to represent. BCCA nonetheless underpredicts the standard deviation over virtually the entire domain, with a mean value of -30% . LOCA has appreciable errors only in the very lowest-valued parts of the region, and the errors appear to be noisy, with both positive and negative values. Averaged over the domain, the LOCA standard deviation is $\sim 2.8\%$ too small.

A critical aspect of precipitation's temporal behavior is the persistence of heavy-precipitation events, because it has bearing on floods (for multiday rain events) or transportation (for multiday snow events). Percentiles in

the distribution of the length of wet spells are shown in Fig. 15 for a threshold of 10 mm day^{-1} . LOCA generally captures the duration of wet spells quite well. Results using a threshold of 25 mm day^{-1} are similarly well represented (not shown).

An important point to emphasize here is that LOCA is a spatial downscaling scheme; every time step is downscaled independently, without regard to any other time step. The method has no “knowledge” of the temporal characteristics of either the observations or coarse-resolution GCM data. So if, for example, a coarse climate model precipitation field had a poor representation of wet-spell length, then LOCA, being a spatial downscaling scheme, would not mitigate this temporal bias. In an entire downscaling system that used LOCA as the spatial downscaling step, such a bias in wet-spell length would have to be addressed in the bias correction step.

e. Fraction of zero-precipitation days

One drawback of BCCA is that it tends to produce drizzle in locations where the original field being downscaled

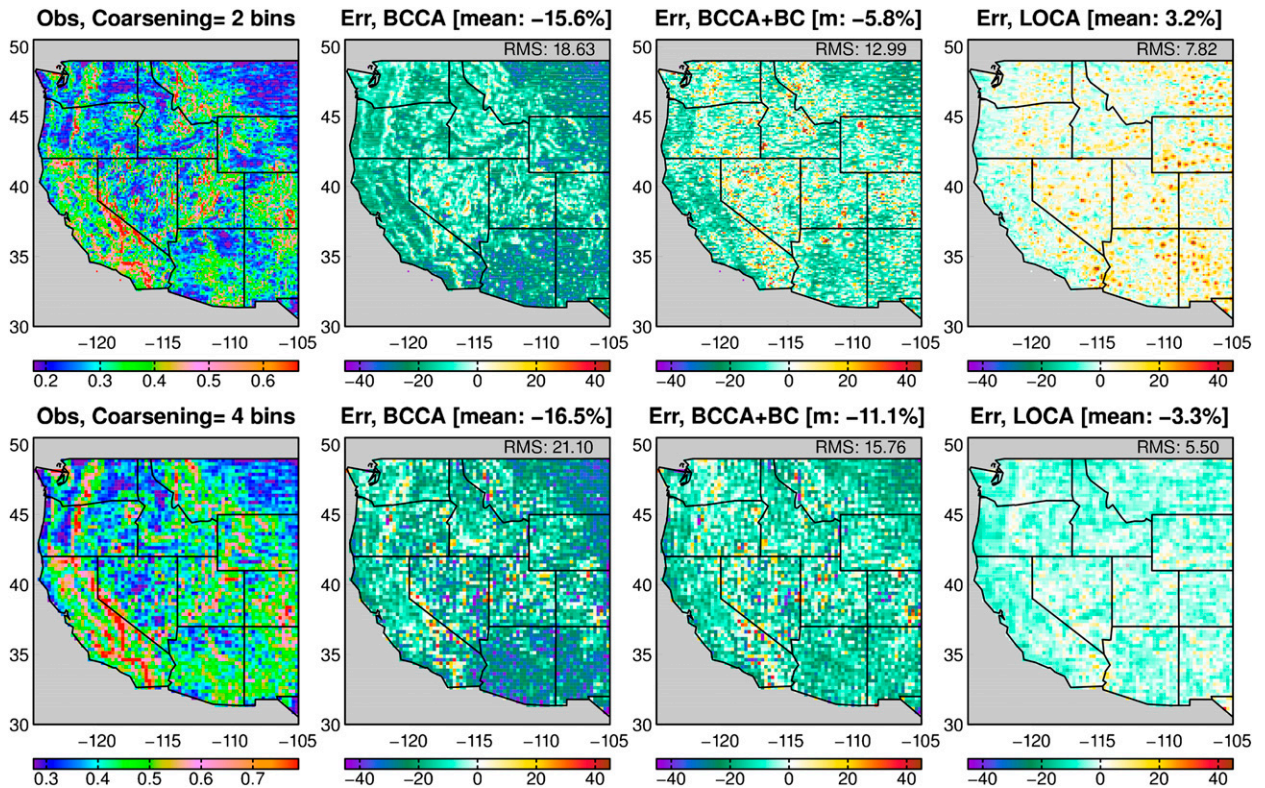


FIG. 12. The coefficient of spatial variability (nondimensional) for daily precipitation, computed as the spatial std dev in a 3×3 array divided by the value at the center.

is dry, because there is a reasonable chance that some of the 30 analog days that are weighted into the downscaled estimate have precipitation where the field being downscaled has none. This is illustrated in Fig. 16, which shows the fraction of winter (DJF) days that have precipitation less than 0.1 mm day^{-1} . This threshold is chosen to avoid pointlessly exaggerating the problem in the BCCA results; if a threshold of zero had been used, BCCA’s depiction would look substantially worse, but such precipitation values are so small as to be of little importance for most applications. (One possible exception might be if a cloud or solar insolation parameterization is affected by the presence of nonzero precipitation.) As expected, BCCA depicts fewer days than are seen in the observations, by about 11% on average. The errors are not random, having a distinct spatial pattern that mimics that of the overall zero-precipitation day rate. LOCA produces about 2% more days than observed, with only a weak spatial pattern.

The fraction of $<0.1 \text{ mm day}^{-1}$ precipitation days in summer (JJA; Fig. 17) shows similar results. The area-averaged BCCA shortfall is nearly the same as in winter (-11.3%), and the spatial pattern of errors again reflects the mean field, although the pattern itself is different

from the winter pattern. The errors in LOCA during the summer are about -2% , on average.

Compared to the traditional CA approach, LOCA generates less drizzle in locations where the coarse input field is dry because LOCA only needs one analog day to be dry in the local region being downscaled to produce a dry downscaled field. By contrast, traditional CA may produce drizzle if any of the 30 analog days has precipitation in the location being downscaled, depending on how exactly the combination of 30 analog weights multiplied by the precipitation values cancels to zero in the dry location. However, it should be understood that LOCA is a spatial downscaling technique and does not “know” if the coarse field being downscaled has drizzle because the day really did have drizzle or because it is from a GCM that is systematically biased toward having too much drizzle. Bias correction that produces a realistic precipitation distribution must be applied to a GCM field before it is downscaled with LOCA to avoid excess drizzle in the downscaled result. Likewise, a training dataset that has too much drizzle would produce downscaled fields with too much drizzle, as the LOCA scheme assumes (like other statistical methods) that the training dataset is accurate. This may be an issue with the training

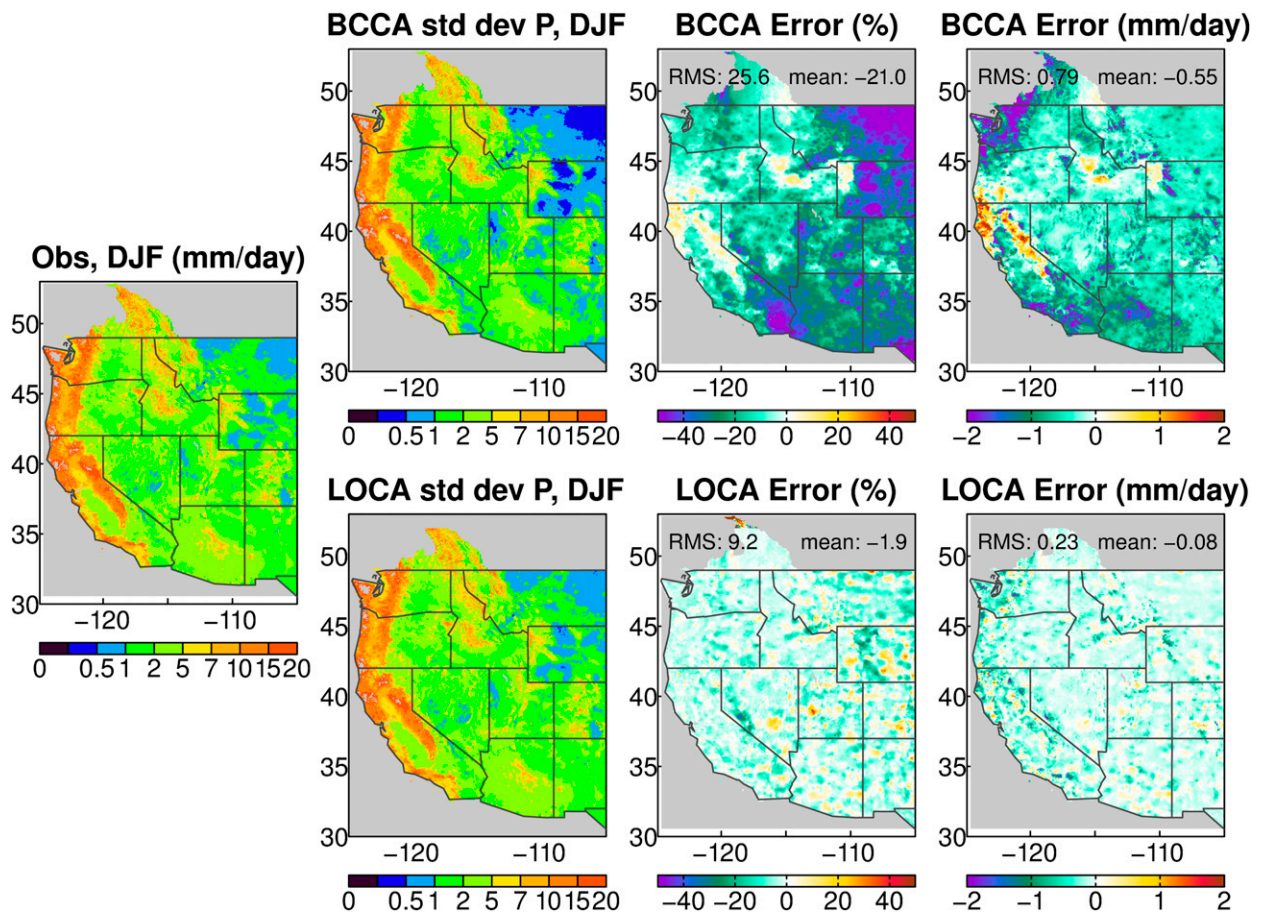


FIG. 13. Temporal std dev of daily precipitation in winter (DJF).

data used here, where the time of observation is used to split measured precipitation amounts across 2 days (Livneh et al. 2013).

f. Correlation of downscaled precipitation with original observations

Figure 18 (top) shows the temporal correlation between the precipitation time series at each point from the observations and downscaled data. LOCA shows higher values than BCCA over the majority of the domain and a higher average correlation overall. The subsequent rows of Fig. 18 show the analogous correlation fields for the daily anomalies, monthly anomalies, and yearly anomalies. LOCA generally has higher skill in reproducing the original fields than BCCA. LOCA daily anomaly correlations with observations, averaged over the western U.S. domain, are 0.83 compared to 0.76 for BCCA. Skill increases at the longer time scales, particularly from daily to monthly aggregation.

LOCA's correlations tend to be slightly higher where the finescale grid box falls in the center of the original

coarse $1^\circ \times 1^\circ$ data that were downscaled, as can be seen in Fig. 18. This is likely because at the center of a coarse grid cell, LOCA's local spatial matching is covering approximately the same region as was aggregated to produce the coarse-scale data in the first place.

6. LOCA results: Multivariate downscaling for daily minimum temperature

Generating downscaled fields from multiple climate variables should preserve important aspects of their joint distribution (e.g., Zhang and Georgakakos 2012). For example, Thrasher et al. (2012) recommend calculating bias-corrected daily minimum temperature T_{min} as the difference between daily maximum temperature T_{max} and diurnal temperature range DTR, because this prevents $T_{min} > T_{max}$. This approach was implemented in LOCA by calculating downscaled T_{min} as the difference between downscaled T_{max} and downscaled DTR. However, when T_{max} and DTR are downscaled independently, the derived T_{min} has unrealistically high

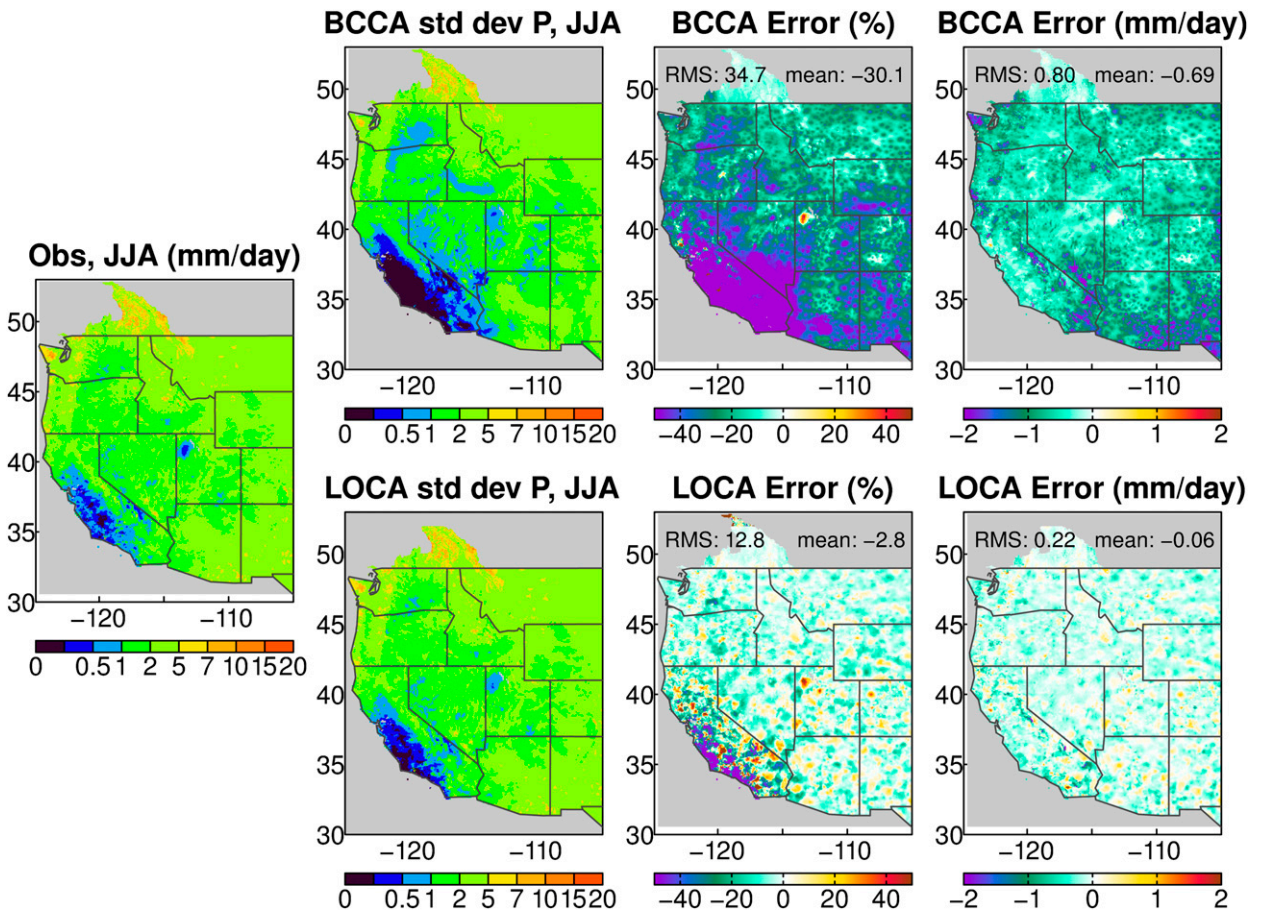


FIG. 14. As in Fig. 13, but for summer (JJA).

variability along the West Coast even though downscaled Tmax and DTR themselves do not show excessive variability. This is likely because Tmin adjacent to the coast is influenced by ocean surface temperatures (which are less variable than land temperatures), so that a high Tmax along the coast is associated with a high DTR—that association is not properly captured when Tmax and DTR are downscaled independently.

As described in section 3c, LOCA implements multivariate downscaling by choosing analog days on the basis of weighted ranks across all the variables being downscaled. If w is the weighting parameter and two variables are being downscaled, then the final multivariate rank of a day when downsampling variable 1 along with information from variable 2 is $wR_1 + (1 - w)R_2$ where R_n is the rank of the observed day for variable n . If $w = 0.5$, the variables are weighted equally, while if $w = 1.0$, then the two variables are downscaled independently, exactly as if a univariate downscaling were being performed.

The procedure is illustrated when computing Tmin as the difference between downscaled Tmax and downscaled DTR in Fig. 19. Plotted values are the error in the

simulated daily Tmin standard deviation in summer as a function of the weight w used in the multivariate downscaling (indicated in titles). The RMS error increases monotonically as the weight increases from (top left) 0.5, indicating equal weighting of Tmax and DTR, to (bottom right) 1.0, indicating independent downscaling of Tmax and DTR. These results show that multivariate downscaling is necessary to improve the simulation of daily Tmin computed as $T_{max} - DTR$ and should be evaluated for other circumstances and variables. In particular, Abatzoglou and Brown (2012) indicate the importance of multivariate downscaling when computing relative humidity.

As a test, we recomputed Fig. 8 (the point-by-point temporal correlation between the downscaled and observed Tmax fields) with Tmax from the multivariate downscaling and found the results to be nearly indistinguishable. So using the multivariate scheme does not appear to degrade the quality of the Tmax downscaled field.

Temperature/precipitation relationships and snow

K. Hegewisch and J. Abatzoglou (2013, personal communication) have pointed out the importance of

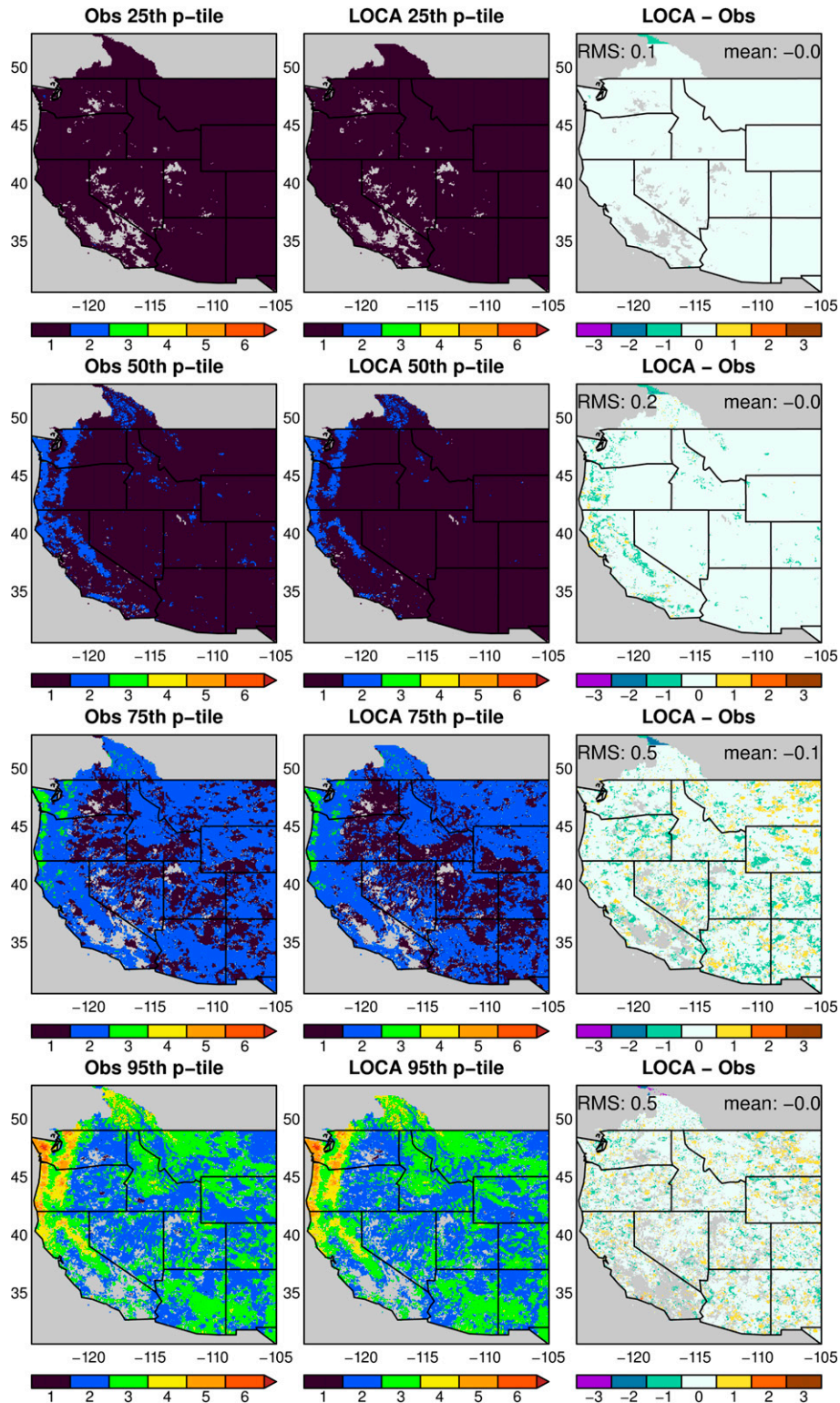


FIG. 15. Percentiles in the distribution of wet-spell length, where a wet spell is defined as a string of consecutive days with precipitation $\geq 10 \text{ mm day}^{-1}$: (left) observations, (middle) LOCA, and (right) the difference (LOCA minus observations). Values are only plotted in locations that experience at least 20 wet spells, except for the median (50th percentile), which is plotted in all locations with at least one value.

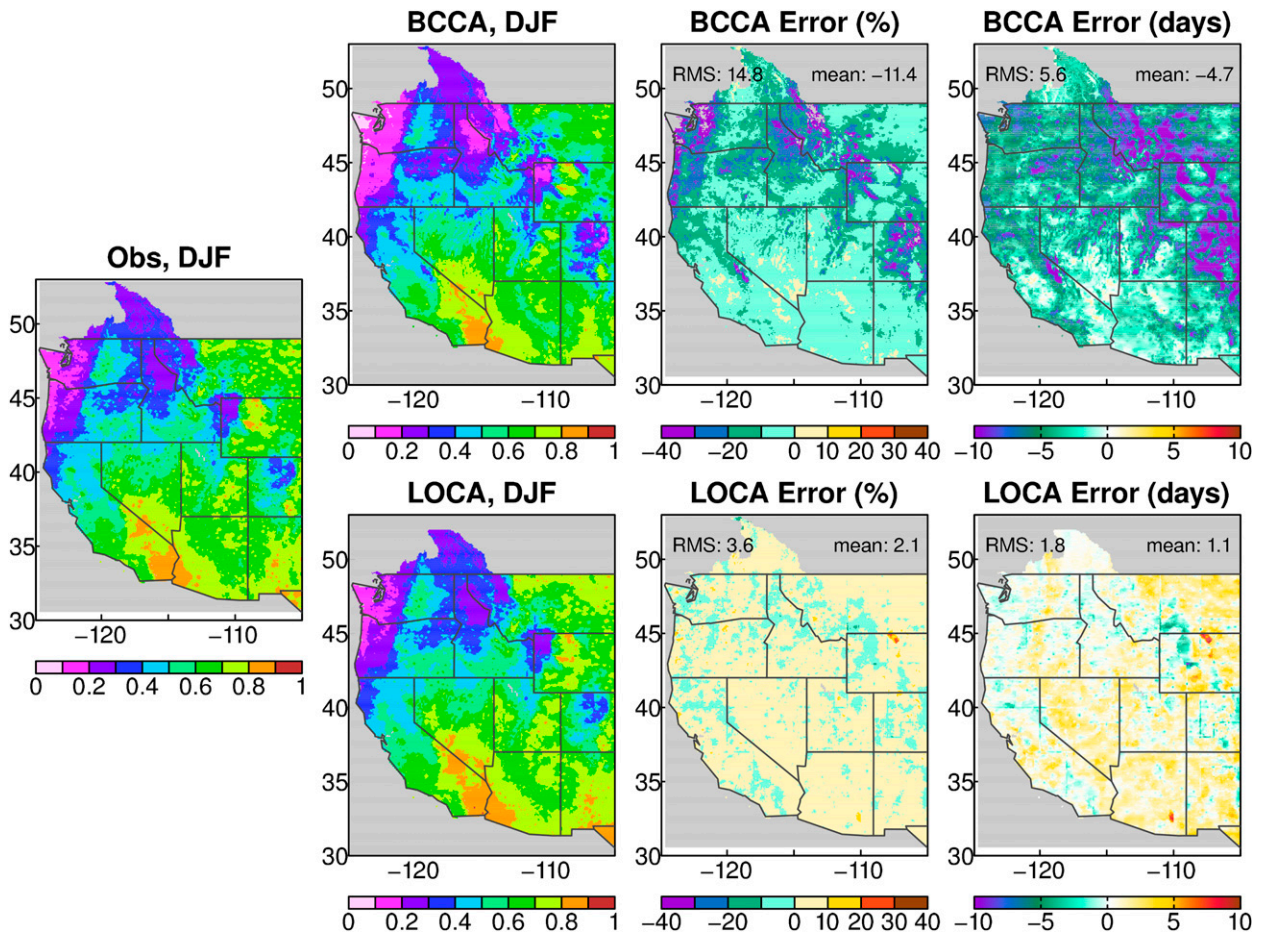


FIG. 16. Fraction of days with precipitation $>0.1 \text{ mm day}^{-1}$ (referred to as zero-precipitation days) for winter (DJF).

preserving observed relationships between temperature and precipitation in downscaled datasets used for hydrological applications. In many locations and seasons, observations show that wet days are colder than dry days. If a downscaling scheme fails to preserve this difference, the simulated precipitation might arrive as rain instead of snow and thus run off quickly rather than being preserved in the snowpack. We evaluate this by comparing the seasonal distribution of T_{max} found on wet days to that found on dry days, using the observations, downscaled LOCA, and BCCA data. Overall, LOCA does a good job of preserving the wet/dry day temperature difference; BCCA does nearly as well if a 1 mm day^{-1} precipitation threshold is used in the analysis to reduce the impact of BCCA's too numerous drizzle days (Fig. S3 in the supplemental materials).

A similar question is whether LOCA, which downscales based on temperature anomalies, does as well reproducing temperature on below-freezing days as it does on above-freezing days. Selection of the analog days does not take into account whether or not the

temperature is below freezing, yet snow might preferentially be present on such days, perhaps leading to characteristically different spatial patterns of temperature. We evaluated this by calculating the temporal correlation between LOCA downscaled T_{max} anomalies and observed T_{max} anomalies as a function of actual T_{max} (i.e., not the anomaly) for the day. Results (Fig. S4 in the supplemental materials) show that correlations tend to be highest for temperatures near the middle of the distribution and fall off at the extreme temperatures, but there is no systematic change in correlation as the freezing point is crossed.

7. Summary and conclusions

We have introduced a new technique for statistical downscaling called localized constructed analogs (LOCA). Existing constructed analog techniques such as CA, BCCA, and MACA (Hidalgo et al. 2008; Maurer et al. 2010; Abatzoglou and Brown 2012) average together multiple days to construct the final downscaled field.

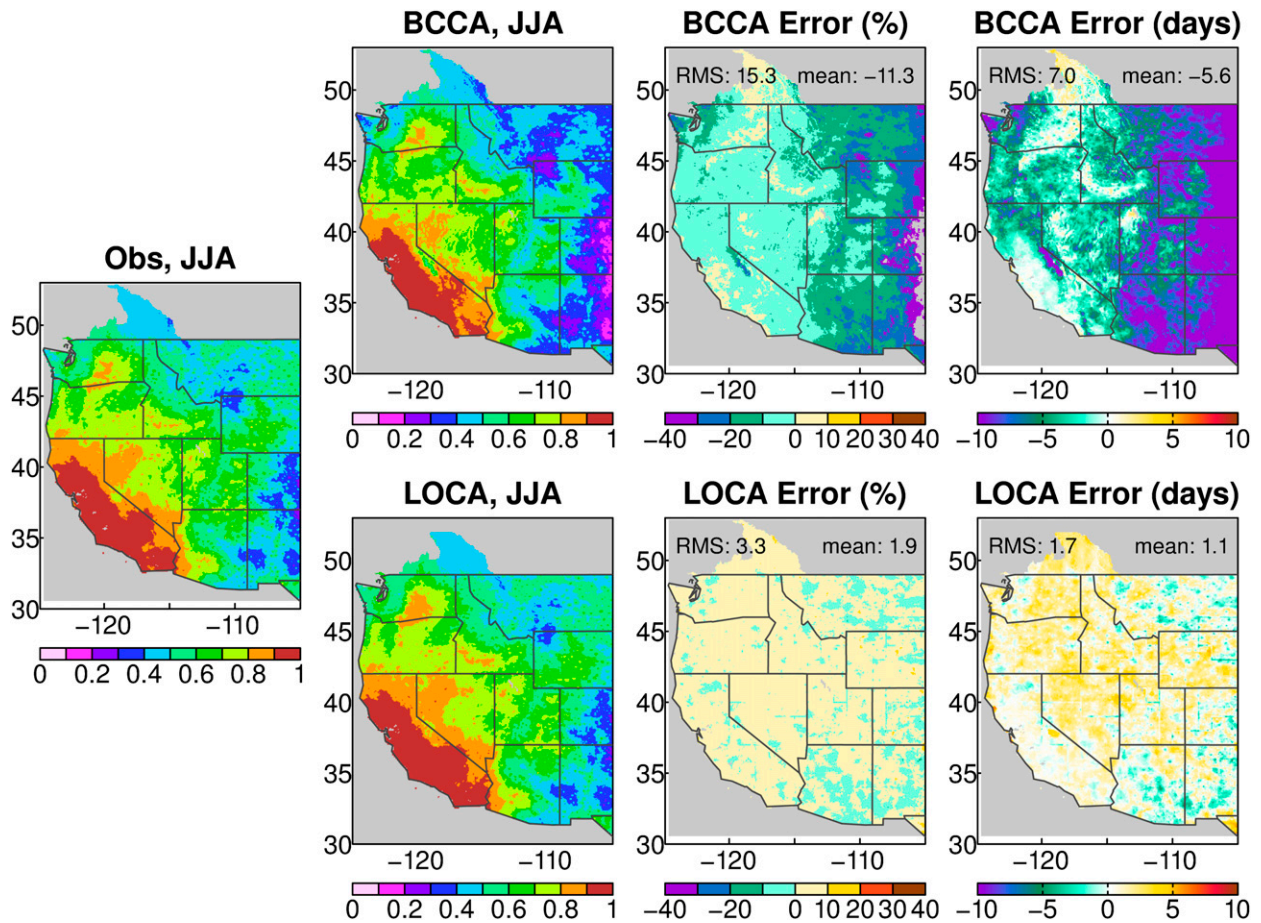


FIG. 17. As in Fig. 16, but for summer (JJA).

This averaging increases the spatial coherence of the downscaled fields, dilutes extremes, and produces drizzle in nonprecipitating areas of the original model. LOCA avoids these problems by selecting the single best matching analog day in a local region about the point being downscaled.

Results from downscaling daily maximum temperature and precipitation over the western United States illustrate that LOCA reproduces the extremes in downscaled summer maximum daily temperature and winter daily precipitation quite well. LOCA provides a respectable representation of the overall variability, with temporal correlations against observations that are superior to BCCA over a range of time scales. LOCA also realistically captures the spatial variability of the original field and avoids the production of drizzle in nonprecipitating areas.

Unlike the other constructed analog methods, there is a natural domain independence to LOCA. Starting from a point being downscaled, increasing the domain size past the distance where the correlation between the

distant point and the starting point drops to zero does not affect the downscaling results at the starting point. Like all other statistical methods, though, LOCA is trained on historical observations, and it therefore assumes that this spatial relationship will remain largely unchanged in the future climate. Likewise, LOCA assumes that the characteristic relationship between local and area-averaged climate fields will not appreciably change in the future. Such assumptions could be tested with targeted dynamical downscaling simulations.

Multivariate downscaling is straightforward to incorporate into the LOCA framework with user-specifiable weights for the multiple variables. The results here show the utility of multivariate downscaling when computing daily minimum temperature as the difference between daily maximum temperature and the diurnal temperature range.

Recently, BCCA has been found to significantly underestimate monthly averaged precipitation as well as spatial and temporal variability and extreme events (Gutmann et al. 2014). Postdownscaling bias correction,

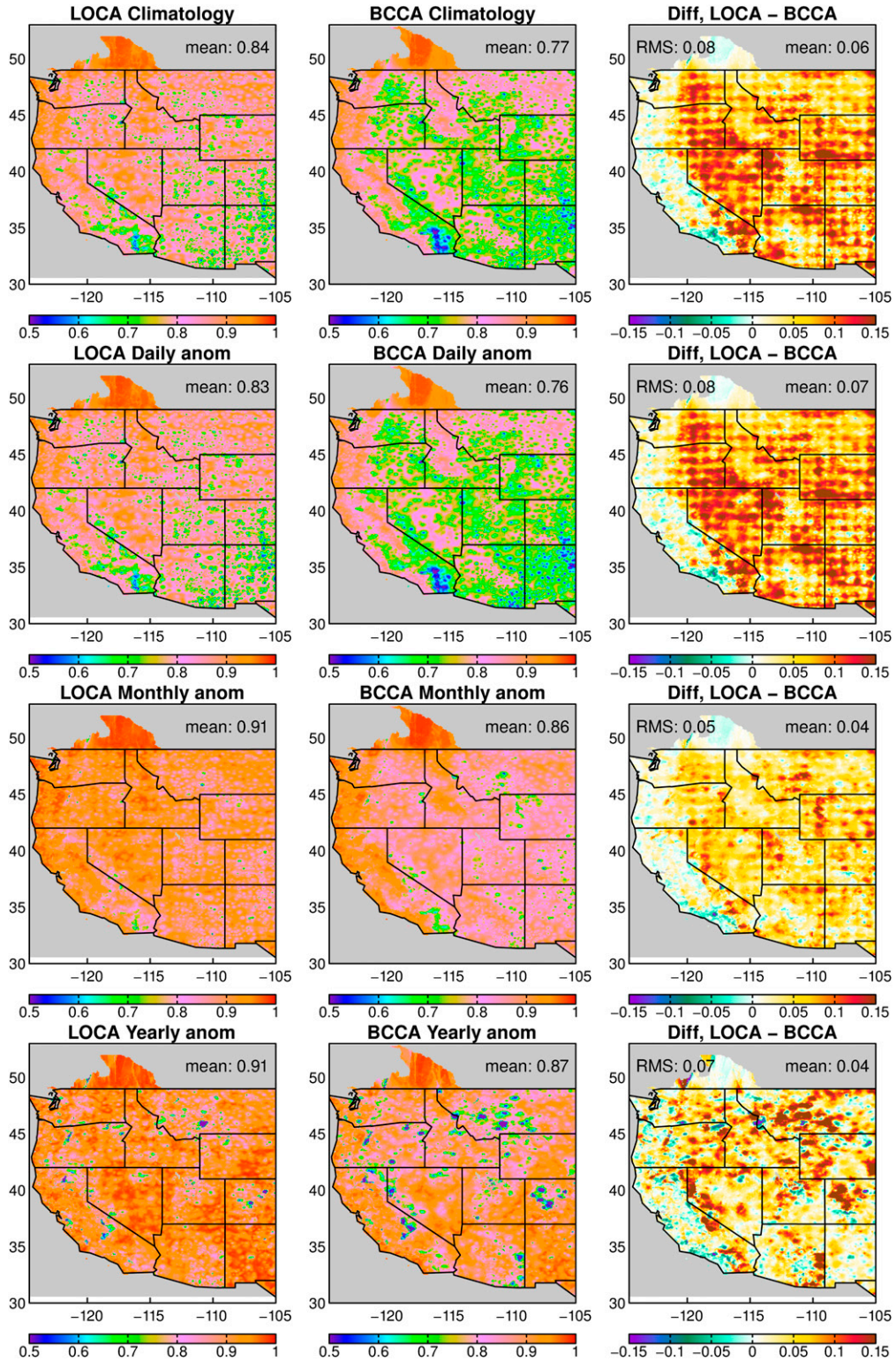


FIG. 18. Correlation between time series of precipitation at each point from the original observations and downscaled results for (left) LOCA, (middle) BCCA, and (right) their difference for (from top to bottom) the entire period without taking anomalies first, daily anomalies, monthly anomalies, and yearly anomalies.

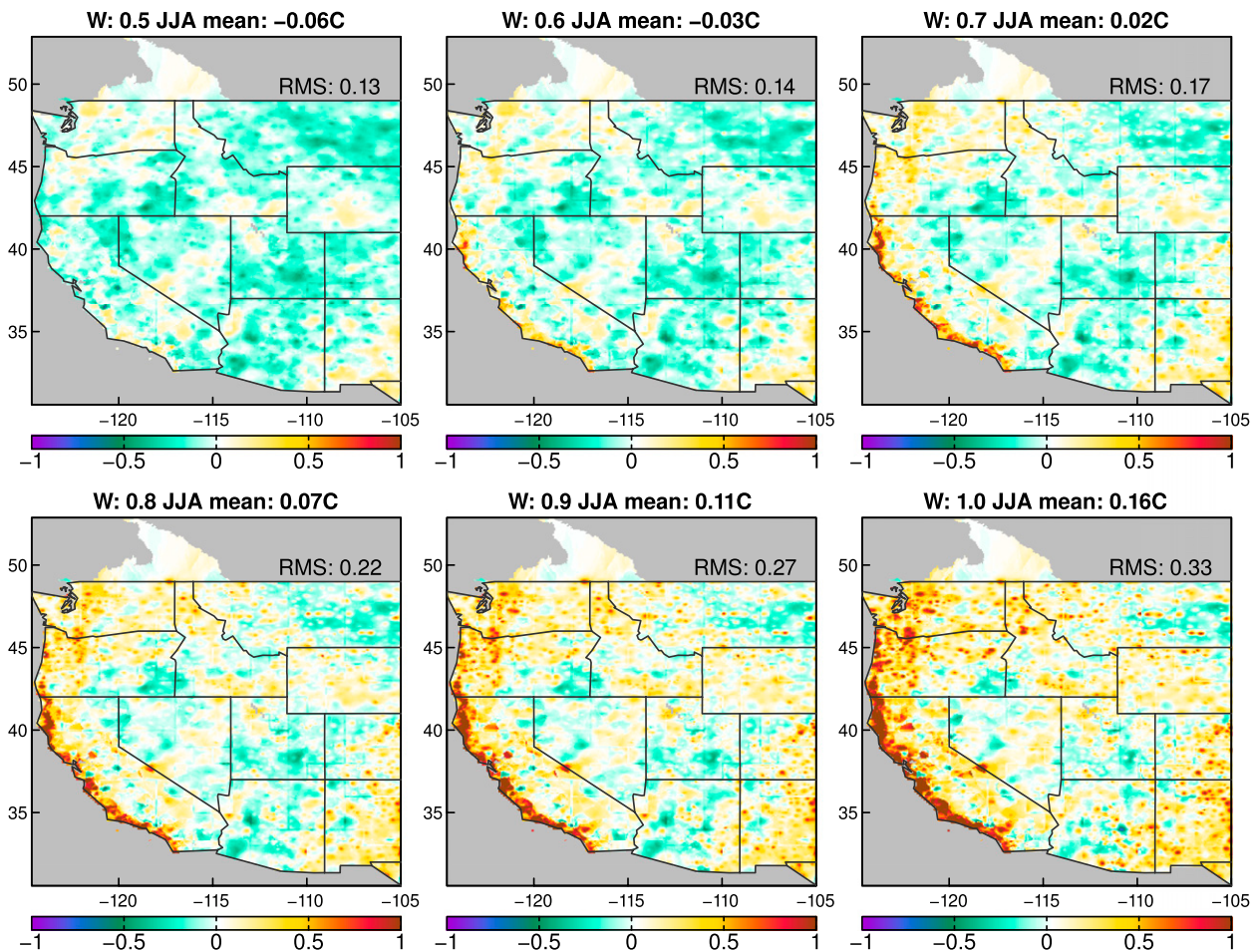


FIG. 19. Error ($^{\circ}\text{C}$) with respect to observations in the std dev of summer (JJA) daily temperature min, computed as the difference between downscaled Tmax and downscaled DTR. Each is computed using a different weight in the multivariate downscaling procedure. A weight of (top left) 0.5 means that the two variables (Tmax and DTR) are considered equally in the multivariate downscaling procedure; a weight of (bottom right) 1.0 means that the two variables are downscaled independently. See text for details.

such as used in MACA (Abatzoglou and Brown 2012), reduces many of these errors but is unable to repair all of them. For example, we have shown that the representation of spatial variability is improved but not completely fixed by an additional bias correction step added after BCCA. Also, different fine-resolution grid cells are bias corrected independently, so the spatial structure of an event (say, a storm) will be modified by this process. A related question is whether it is better to use a downscaling scheme that produces less bias in the first place, versus producing downscaled fields that initially have larger biases and then correcting them with another processing step. Also, some applications may be intrinsically better suited to the smoothed spatial characteristics of the BCCA downscaled fields. Although the answers to these questions depend on a host of factors such as the application being considered and the

processing time available, we believe that the LOCA scheme represents a useful addition to the arsenal of downscaling tools currently available.

Acknowledgments. The authors thank John Abatzoglou, Edwin Maurer, Hugo Hidalgo, Ethan Gutmann, and two other anonymous reviewers for helpful comments and suggestions that improved this work. This project was sponsored by the California Energy Commission under Grant CEC-500-10-041. Additional support for D.W.P. and D.C. came from the USGS through the Southwest Climate Science Center and from NOAA through the California Nevada Climate Applications Project (CNAP) Regional Integrated Science Applications (RISA) program. We thank the University of Washington Land Surface Hydrology Research Group for supplying the Livneh et al. (2013) $1/16^{\circ}$ dataset.

APPENDIX

Details of the LOCA Procedure

Let $O^F(x, y, t)$ be a time series of daily observations, which are a function of longitude x , latitude y , and day t on a fine-resolution grid (here, $1/16^\circ$). When downscaling temperature, anomalies are calculated with respect to day of year mean values estimated using cubic spline interpolation of monthly mean values in order to reduce the sampling errors generated by calculating day of year mean values directly with only 41 years of data. Leap days are ignored, with the mean for 29 February taken as the mean of the values on 28 February and 1 March.

Let $M^C(\chi, \psi, t')$ be the model field on the coarse-resolution grid (we indicate the coarse-grid longitude and latitude by χ and ψ and the fine grid resolution by x and y to emphasize their difference), and t' is a model day to be downscaled. Let $D^Y(t)$ be the day of the year of day t .

a. The analog pool locations: Finding matches at the regional scale

We refer to the locations where the pool of analog days are chosen as the analog pool points \mathbf{x}^P and select them from the coarse grid. For each \mathbf{x}^P , we use a spatial mask $\Gamma(\chi, \psi)$ to limit the region over which the model day is compared to the coarsened observations when determining the fit between the two. Variable Γ is set to 1 where the Pearson correlation between the observed time series at locations \mathbf{x}^P and $O^C(\chi, \psi) > 0$ over the historical training period, where $O^C(\chi, \psi, t)$ is the observations aggregated to the coarse model grid. Variable Γ is set to 0 elsewhere. As a sensitivity test, we also tried setting Γ to 1 where the correlation is significantly positive at the 95% level; with ~ 30 years of daily training data by season, this yields a significance value of ~ 0.07 (assuming a synoptic autocorrelation time scale of 5 days), which made little difference to the final result.

To determine the locations of the \mathbf{x}^P in a systematic way, we modeled them as points that are repelled from both each other and the domain edges. We initially placed 35 points randomly in the domain (providing approximately one \mathbf{x}^P per $3^\circ \times 3^\circ$ latitude–longitude square) and then let them evolve forward in time, moving according to their mutual repulsion from both the domain edges and all the other points, until the velocity of the points dropped to near zero. The \mathbf{x}^P were then chosen as the closest coarse-grid locations to the final point positions. This process only needs to be done once for a given domain. In sensitivity tests, we found the points to generally end up near the same final locations no matter where they started, except in a few limited regions where the points would alternate between

different near-stable configurations. We also tried using 20 and 70 \mathbf{x}^P points; although the results differed in details, the overall statistics of the final downscaled field were little changed. We conclude that the exact placement of the points does not matter as long as the total number of points is unchanged and they are distributed approximately equally across the domain.

For picking \mathbf{x}^P locations in a new domain, the key point is that the spacing between analog pool points should be smaller than the extent of the Γ masks. This allows a gradual change in the composition of the analog day pool as the domain is traversed, rather than an abrupt transition going from one \mathbf{x}^P to another. It also ensures that every fine-resolution grid cell falls inside the mask used to compute analog days for that grid cell. So, for example, if the Γ masks have a typical extent of $10^\circ \times 10^\circ$ in longitude–latitude (Fig. 1), placing an \mathbf{x}^P point every 3° or finer is appropriate.

b. Selecting the analog days at the regional scale

After determining the \mathbf{x}^P locations for analog day selection, we choose the analog days by finding those t such that

$$\left(\sum_{\chi, \psi} \{ \Gamma(\chi, \psi, \mathbf{x}^P, \hat{t}) [O^C(\chi, \psi, t) - M^C(\chi, \psi, t')] \}^2 / n \right)^{1/2} \tag{A1}$$

is minimized, subject to the constraint that $D^Y(t)$ and $D^Y(t')$ are within 45 days of each other. This procedure is like that used in Hidalgo et al. (2008), but calculated at all analog pool locations and with a spatial mask applied. Here, n is the number of active points on the coarse grid, and, as described above, $\Gamma(\chi, \psi, \mathbf{x}^P, \hat{t})$ is a spatial mask that is different at each \mathbf{x}^P point and is a function of season \hat{t} . The applicable season is determined by t' . When conducting cross-validated testing (i.e., downscaling coarsened observations over the same period as the training period), we additionally require that the potential analogs must be well removed from the downscaled target day ($|t - t'| > 320$ days), so that no information from any day near the day being downscaled is used to train the model.

Equation (A1) is computationally expensive [for each location (χ, ψ) there are over 3000 values of t in our training dataset that have to be checked for each t'], which is why we choose only a limited number of \mathbf{x}^P locations.

c. Finding the one best matching analog day at the local scale

As noted in the main text, finding the best match between the model and observations must be implemented

on the finescale grid. We therefore first interpolate $M^C(\chi, \psi, t')$ to $M^F(x, y, t')$ using bicubic interpolation. However, $O^F(x, y, t)$ has significant information content at the finest spatial scales that $M^F(x, y, t')$ lacks. To address this, we bicubic interpolate $O^C(\chi, \psi, i)$ from the N_a analog days (where i is the analog day number) in the pool of analog days back to the fine grid to produce $\hat{O}^F(x, y, i)$, which can sensibly be compared to $M^F(x, y, t')$.

At each fine-resolution grid location (x, y) we find the single analog day τ from the pool of analog days that minimizes

$$\left\{ \sum_{\substack{x'=x+r, y'=y+r \\ x'=x-r, y'=y-r}} [\hat{O}^F(x', y', \tau) - M^F(x', y', t')]^2 / n \right\}^{1/2}. \quad (\text{A2})$$

This single analog day selection is done locally (hence the name, localized constructed analogs), in a square region of size $2r + 1$ grid cells around the center grid point being downscaled. It is desirable to have r as small as practical, because the overall downscaling time increases with r , but not so small as to produce downscaled results that are spatially incoherent. We used $r = 10$ for both temperature and precipitation. We also tried $r = 20$ and found that the computational time increased with no net benefit in the final result (some measures were better, others worse).

d. Constructing the final downscaled field

To construct the final downscaled field we define a scale factor S . When downscaling absolute values (i.e., precipitation), we use a multiplicative scale factor:

$$S = \begin{cases} 0 & \text{if } \hat{O}^F(x, y, \tau) = 0, \\ \frac{M^F(x, y, t')}{\hat{O}^F(x, y, \tau)} & \text{otherwise} \end{cases}. \quad (\text{A3})$$

To avoid having S blow up for small values of $\hat{O}^F(x, y, \tau)$, we limit it to 2. When downscaling anomalies (i.e., temperature), we use an additive scale factor:

$$S = M^F(x, y, t') - \hat{O}^F(x, y, \tau). \quad (\text{A4})$$

The downscaled value $d(x, y)$ is then

$$d(x, y) = \begin{cases} SO^F(x, y, \tau) & \text{for absolute values} \\ S + O^F(x, y, \tau) & \text{for anomalies} \end{cases}. \quad (\text{A5})$$

At edge cells using a different analog day (τ'), Eqs. (A3)–(A5) are recomputed at (x, y) using τ' . The final downscaled value at edge cells is then the weighted sum

of the values computed using τ and τ' , where the weight is determined by the number of adjacent grid cells using τ' versus τ .

e. Calculating the change in future climatology

LOCA downscales temperature as an anomaly, so when downscaling a climate projection, the systematic future change in climatology toward warmer conditions must be addressed. CA and BCCA do not treat the change in climatology in any special way; for example, a hot day at the end of the century is simply considered an extremely hot day compared to the historical training period.

In LOCA, the change in climatology is calculated in separate 30-yr periods. The selection of 30-yr climatologies is motivated by the recommendation of the World Meteorological Organization (WMO) that climatological normals be calculated over 30-yr periods [a brief history of climatological normals can be found in Trewin (2007)], a recommendation followed by the National Oceanic and Atmospheric Administration (NOAA) and the National Climatic Data Center (NCDC; www.ncdc.noaa.gov/oa/climate/normals/usnormals.html). For example, imagine we are downscaling the future period 2040–69. We calculate the mean over the historical period and the mean over the future period (for each day, since LOCA is a daily method), and take the difference. The result is simply an anomaly field representing the change in climatology from the historical period to the future period, which is itself downscaled using LOCA as described in the main text. The result is a fine-resolution representation of the change in climatology, which is added to the historical climatology to produce the climatology of the future period. The downscaled temperature anomaly fields LOCA generates are then added to this future climatology. If we were also downscaling a different future period, say 2070–99, then the process is repeated for that future period.

One advantage to calculating the change in climatology this way is that it becomes easier to find historical analogs for temperature in the face of global warming. CA and BCCA must find historical analogs even for days during heat waves at the end of this century, which could be challenging given the likely temperature changes. By contrast, LOCA's anomalies are only calculated with respect to the future 30-yr climatology and so are easier to match with historical analogs.

Consider how a model's representation of decadal variability such as the Pacific decadal oscillation (PDO) would be affected by this procedure. Assume for simplicity that the model's PDO was high over the period 2040–69 and low over the period 2070–99. Then (by construction) the mean of the downscaled field over 2040–69 would match the mean of the GCM over 2040–69,

and similarly for the 2070–99 period. So, the downscaled product would retain the original GCM's PDO signal on those long time scales. On shorter time scales (i.e., within the 2040–69 period), assume that the model PDO was trending toward a peak. That trend would also be preserved by LOCA because the temperatures would start out cold at the beginning of the 2040–69 period and trend to warm at the end of the 2040–69 period, and LOCA preserves the model-predicted temperature evolution (e.g., Fig. 8). Thus, the overall result is that the GCM's PDO will be reproduced in the downscaled field.

REFERENCES

- Abatzoglou, J. T., and T. J. Brown, 2012: A comparison of statistical downscaling methods suited for wildfire applications. *Int. J. Climatol.*, **32**, 772–780, doi:10.1002/joc.2312.
- Chen, J., F. P. Brissette, and R. Leconte, 2014: Assessing regression-based statistical approaches for downscaling precipitation over North America. *Hydrol. Processes*, **28**, 3482–3504, doi:10.1002/hyp.9889.
- Das, T., E. P. Maurer, D. W. Pierce, M. D. Dettinger, and D. R. Cayan, 2013: Increases in flood magnitudes in California under warming climates. *J. Hydrol.*, **501**, 101–110, doi:10.1016/j.jhydrol.2013.07.042.
- Fowler, H. J., S. Blenkinsop, and C. Tebaldi, 2007: Linking climate change modeling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.*, **27**, 1547–1578, doi:10.1002/joc.1556.
- Goodess, C. M., and J. P. Palutikof, 1998: Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling. *Int. J. Climatol.*, **18**, 1051–1083, doi:10.1002/(SICI)1097-0088(199808)18:10<1051::AID-JOC304>3.0.CO;2-1.
- Gutmann, E. D., T. Pruitt, M. P. Clark, L. Brekke, J. R. Arnold, D. A. Raff, and R. M. Rasmussen, 2014: An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.*, **50**, 7167–7186, doi:10.1002/2014WR015559.
- Hidalgo, H. G., M. D. Dettinger, and D. R. Cayan, 2008: Downscaling with constructed analogues: Daily precipitation and temperature fields over the United States. CEC PIER Project Rep. CEC-500-2007-123, 48 pp. [Available online at www.energy.ca.gov/2007publications/CEC-500-2007-123/CEC-500-2007-123.PDF.]
- Hwang, S., and W. D. Graham, 2014: Assessment of alternative methods for statistically downscaling daily GCM precipitation outputs to simulate regional streamflow. *J. Amer. Water Resour. Assoc.*, **50**, 1010–1032, doi:10.1111/jawr.12154.
- Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. Andreadis, E. P. Maurer, and D. P. Lettenmaier, 2013: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Updates and extensions. *J. Climate*, **26**, 9384–9392, doi:10.1175/JCLI-D-12-00508.1.
- Lobell, D. B., A. Torney, and C. B. Field, 2009: Climate extremes in California agriculture. California Energy Commission Rep. CEC-500-2009-040-F, 8 pp. [Available online at www.energy.ca.gov/2009publications/CEC-500-2009-040/CEC-500-2009-040-F.PDF.]
- Maraun, D., 2013: Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *J. Climate*, **26**, 2137–2143, doi:10.1175/JCLI-D-12-00821.1.
- , and Coauthors, 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, **48**, RG3003, doi:10.1029/2009RG000314.
- Maurer, E. P., H. G. Hidalgo, T. Das, M. D. Dettinger, and D. R. Cayan, 2010: The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrol. Earth Syst. Sci.*, **14**, 1125–1138, doi:10.5194/hess-14-1125-2010.
- Parry, M. L., O. F. Canziani, J. P. Palutikof, P. J. van der Linden, and C. E. Hanson, Eds., 2007: *Climate Change 2007: Impacts, Adaptation, and Vulnerability*. Cambridge University Press, 976 pp.
- Piani, C., J. Haerter, and E. Coppola, 2010: Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.*, **99**, 187–192, doi:10.1007/s00704-009-0134-9.
- Pierce, D. W., and D. R. Cayan, 2013: The uneven response of different snow measures to human-induced climate warming. *J. Climate*, **26**, 4148–4167, doi:10.1175/JCLI-D-12-00534.1.
- , and Coauthors, 2013: The key role of heavy precipitation events in climate model disagreements of future annual precipitation changes in California. *J. Climate*, **26**, 5879–5896, doi:10.1175/JCLI-D-12-00766.1.
- Schoof, J. T., and S. C. Pryor, 2001: Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks. *Int. J. Climatol.*, **21**, 773–790, doi:10.1002/joc.655.
- Thrasher, B., E. P. Maurer, C. McKellar, and P. B. Duffy, 2012: Technical note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrol. Earth Syst. Sci.*, **16**, 3309–3314, doi:10.5194/hess-16-3309-2012.
- Trewin, B. C., 2007: The role of climatological normals in a changing climate. WCDMP-61/WMO-TD-1377, 45 pp. [Available online at www.wmo.int/dataset/documents/WCDMPNo61_1.pdf.]
- van den Dool, H. M., 1994: Searching for analogues, how long must we wait? *Tellus*, **46A**, 314–324, doi:10.1034/j.1600-0870.1994.t01-2-00006.x.
- von Storch, H., 1999: On the use of “inflation” in statistical downscaling. *J. Climate*, **12**, 3505–3506, doi:10.1175/1520-0442(1999)012<3505:OTUOH>2.0.CO;2.
- , E. Zorita, and U. Cubasch, 1993: Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *J. Climate*, **6**, 1161–1171, doi:10.1175/1520-0442(1993)006<1161:DOGCE>2.0.CO;2.
- Wilby, R. L., T. M. L. Wigley, D. Conway, P. D. Jones, B. C. Hewiston, J. Main, and D. S. Wilks, 1998: Statistical downscaling of general circulation model output: A comparison of methods. *Water Resour. Res.*, **34**, 2995–3008, doi:10.1029/98WR02577.
- , S. P. Charles, E. Zorita, B. Timbal, P. Whetton, and L. O. Mearns, 2004: Guidelines for use of climate scenarios developed from statistical downscaling methods. IPCC Doc., 27 pp. [Available online at www.ipcc-data.org/guidelines/dgm_no2_v1_09_2004.pdf.]
- Wilks, D. S., 2012: Stochastic weather generators for climate-change downscaling, part II: Multivariable and spatially coherent multisite downscaling. *Wiley Interdiscip. Rev.: Climate Change*, **3**, 267–278, doi:10.1002/wcc.167.
- Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **62**, 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.
- Zhang, F., and A. P. Georgakakos, 2012: Joint variable spatial downscaling. *Climatic Change*, **111**, 945–972, doi:10.1007/s10584-011-0167-9.