

## Improved Historical Temperature and Precipitation Time Series for U.S. Climate Divisions

RUSSELL S. VOSE, SCOTT APPLEQUIST, MIKE SQUIRES, IMKE DURRE, MATTHEW J. MENNE,  
CLAUDE N. WILLIAMS JR., CHRIS FENIMORE, KARIN GLEASON, AND DEREK ARNDT

*NOAA/National Climatic Data Center, Asheville, North Carolina*

(Manuscript received 26 July 2013, in final form 7 January 2014)

### ABSTRACT

This paper describes an improved edition of the climate division dataset for the conterminous United States (i.e., version 2). The first improvement is to the input data, which now include additional station networks, quality assurance reviews, and temperature bias adjustments. The second improvement is to the suite of climatic elements, which now includes both maximum and minimum temperatures. The third improvement is to the computational approach, which now employs climatologically aided interpolation to address topographic and network variability. Version 2 exhibits substantial differences from version 1 over the period 1895–2012. For example, divisional averages in version 2 tend to be cooler and wetter, particularly in mountainous areas of the western United States. Division-level trends in temperature and precipitation display greater spatial consistency in version 2. National-scale temperature trends in version 2 are comparable to those in the U.S. Historical Climatology Network whereas version 1 exhibits less warming as a result of historical changes in observing practices. Divisional errors in version 2 are likely less than 0.5°C for temperature and 20 mm for precipitation at the start of the record, falling rapidly thereafter. Overall, these results indicate that version 2 can supersede version 1 in both operational climate monitoring and applied climatic research.

### 1. Introduction

The National Oceanic and Atmospheric Administration/National Climatic Data Center (NOAA/NCDC) issues State of the Climate reports on a monthly basis. These reports summarize recent conditions and long-term trends at a variety of spatial scales, the smallest being the climate division level. For reporting purposes, the conterminous United States is divided into 344 divisions (Fig. 1), the boundaries of which reflect multiple considerations (e.g., climatic conditions, county lines, crop districts, drainage basins) rather than strict climatic homogeneity (Guttman and Quayle 1996). The historical record for each division consists of temperature and precipitation averages for each month from 1895 to the present. Derived quantities such as degree days and drought indices are also available.

The divisional dataset has two primary strengths that have led to its widespread application in climatological

research. First, its long-term, serially complete record and its relatively modest size facilitate the rapid calculation of state, regional, and national averages for individual events, which can then be placed into a century-scale perspective. Second, its spatial coherence makes it useful in tracking large-scale features over extended periods; major events such as the extreme droughts of the 1930s and the cold winters of the 1970s, for example, are easy to discern. Consistent with these strengths, the divisional dataset has been extensively employed in climate change research (e.g., Grundstein 2009; Hidalgo et al. 2009; Casola et al. 2009; Miller and Piechota 2008; Rogers et al. 2007), drought assessments (e.g., Myoung and Nielsen-Gammon 2010; Plank and Shuman 2009; Seager et al. 2009; Quiring 2009; Goodrich 2007), precipitation studies (e.g., Anderson et al. 2010; Quiring and Kluver 2009; Goodrich and Ellis 2008; Grantz et al. 2007; McCabe et al. 2007), and other varied applications (e.g., Maugé et al. 2009; Stahle et al. 2009; Livezey et al. 2007; Preisler and Westerling 2007).

The divisional dataset also has four major weaknesses that render it suboptimal for certain applications, including, to some extent, the estimation of spatial means

Corresponding author address: Russell S. Vose, NOAA/National Climatic Data Center, 151 Patton Ave., Asheville, NC 28801.  
E-mail: russell.vose@noaa.gov



FIG. 1. Map of the 344 climate divisions in the conterminous United States. Divisions highlighted in gray are discussed in the paper.

and temporal trends. First, each divisional value from 1931 to the present is just the arithmetic average of the station data within it, a computational practice that results in a bias when a division is spatially undersampled in a month (e.g., because some stations did not report) or is climatologically inhomogeneous in general (e.g., due to large variations in topography). Second, all divisional values before 1931 stem from state averages published by the U.S. Department of Agriculture (USDA) rather than from actual station observations, producing an artificial discontinuity in both the mean and variance for 1895 to 1930 relative to 1931 to the present (Guttman and Quayle 1996). Third, many divisions experienced a systematic change in average station location and elevation during the twentieth century, resulting in spurious historical trends in some regions (Keim et al. 2003; Keim et al. 2005; Allard et al. 2009). Finally, none of the station-based temperature records contain adjustments for historical changes in observation time, station location, or temperature instrumentation—inhomogeneities that further bias temporal trends (Peterson et al. 1998).

This paper describes the construction of an improved divisional dataset addressing these weaknesses. The first improvement is to the input data, which now include additional station records and contemporary bias adjustments (Menne and Williams 2009). The second improvement is to the suite of climatic elements, which has been expanded to include both maximum and minimum temperatures. The final (and most extensive) improvement is to the computational methodology, which now addresses topographic and network variability via climatologically aided interpolation (Willmott and Robeson 1995). The outcome of these improvements is a new dataset, hereafter termed version 2, which maintains the strengths of its predecessor while providing more robust estimates of areal averages and long-term trends.

## 2. Station data

The Global Historical Climatology Network-Daily (GHCN-Daily) dataset (Menne et al. 2012) is the source of station data for version 2. GHCN-Daily contains several major observing networks in North America, six of which are used here. The primary networks include the Cooperative Observer (COOP) program and the Automated Surface Observing System (ASOS). Notably, the COOP and ASOS networks were the only source of data used in the original divisional dataset. To improve coverage in western states and along international borders, version 2 also includes the National Interagency Fire Center (NIFC) Remote Automatic Weather Station (RAWS) network, the USDA Snow Telemetry (SNOTEL) network, the Environment Canada (EC) network (south of 52°N), and part of Mexico's Servicio Meteorológico Nacional (SMN) network (north of 24°N). Note that version 2 does not use RAWS precipitation data because that network's tipping-bucket gauges are unheated, leading to suspect cold-weather data.

All GHCN-Daily stations are routinely processed through a suite of logical, serial, and spatial quality assurance reviews (Durre et al. 2010) to identify erroneous observations. For version 2, all such data were set to missing before computing monthly values, which in turn were subjected to additional serial and spatial checks to eliminate residual outliers (Lawrimore et al. 2011). Overall, these checks deemed less than 0.25% of the monthly data as being erroneous. Stations having at least 10 years of valid monthly data since 1950 were used in version 2. These criteria resulted in the exclusion of only a modest number of long-term (>50 yr) COOP sites in the first half of the twentieth century, that is, about 250 for temperature and 375 for precipitation.

GHCN-Daily temperature records do not contain adjustments for historical changes in observing practice. Consequently, bias adjustments were computed specifically for version 2 to account for changes in observation time, station location, temperature instrumentation, and siting conditions. The first step in this process entailed using the method of Karl et al. (1986) to address documented changes in observation time at COOP stations and to adjust the records to a midnight local standard time (LST) observation schedule (matching ASOS, RAWS, and SNOTEL). COOP station histories were obtained from the NCDC Historical Observing Metadata Repository (HOMR) and the U.S. Historical Climatology Network (HCN; Menne et al. 2009). The second step in the adjustment process involved using the “pairwise” method of Menne and Williams (2009) to address all other documented and undocumented changes at any

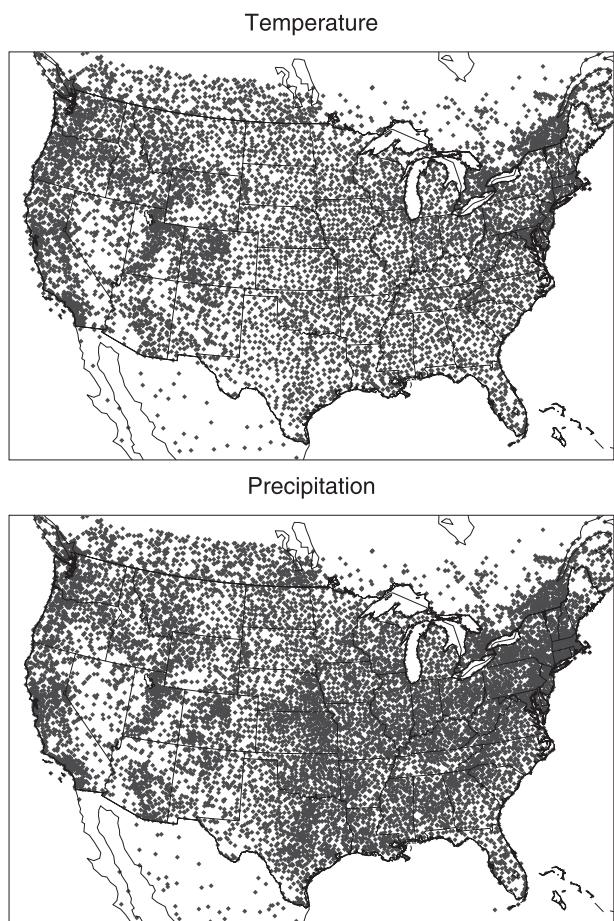


FIG. 2. Maps of stations used in the construction of version 2: (top) temperature and (bottom) precipitation.

station in any network. For example, the pairwise approach was applied to account for documented changes in station location and temperature instrumentation at COOP and ASOS stations, again using HOMR and HCN for station history information. Likewise, the pairwise approach was used to address undocumented changes in observation time, station location, and temperature instrumentation across all networks, including COOP and ASOS. Because the pairwise method largely accounts for local, unrepresentative trends that arise from changes in siting conditions (Menne et al. 2010; Hausfather et al. 2013), version 2 contains no separate adjustment in that regard.

Figure 2 depicts the final station network, which consists of 10 325 temperature and 14 702 precipitation stations. Except for northern Maine, station coverage is fairly uniform east of the Rocky Mountains. Coverage is more variable in the West (i.e., from the Pacific coast to the Rocky Mountains), with large gaps in Nevada. In general, there is a decline in station density going back in

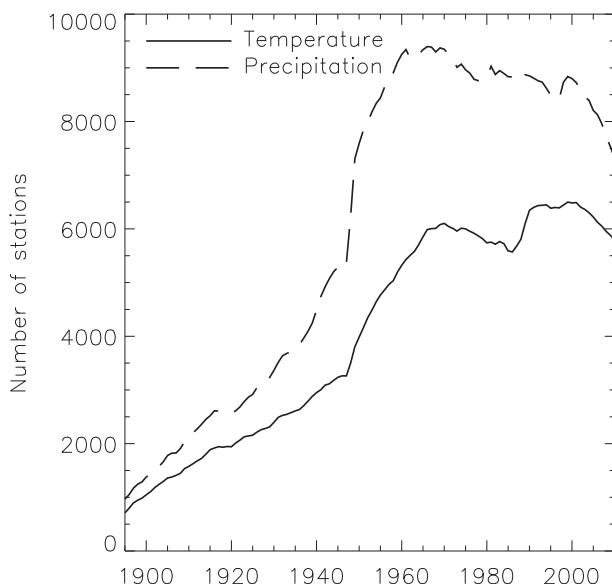


FIG. 3. Plot of the number of temperature and precipitation stations through time.

time; the losses are proportionately larger in the West, but not to an extreme degree. For instance, about a third of all precipitation stations are located in the West in the early twenty-first century, with the proportion falling to about a fifth by the late nineteenth century.

Figure 3 depicts the temporal evolution of the station network. For temperature, the network increases in size until the mid-1960s, declines slightly over the next two decades, increases abruptly again in the late 1980s (due to the growth of RAWS), and declines modestly in the last decade. For precipitation, the network increases in size until the late 1950s, declines gradually through the late 1990s and then more rapidly thereafter. There is a slight change in slope around 1948 for all elements attributable to the creation of the NCDC digital archive. From a monitoring perspective, data are usually available for at least 4000 stations in near-real time (i.e., for the previous month), increasing to roughly 5000 stations after a 2-month lag.

### 3. Gridding method

Climate division values in version 2 were derived from area-weighted averages of gridpoint estimates interpolated from station data. A nominal latitude-longitude spacing of 5 km was used to ensure that all divisions had sufficient gridpoint representation (only four small divisions had fewer than 100 points) and because the impact of elevation on precipitation is minimal below a spatial resolution of 5 km (Sharples et al. 2005). Station data were gridded via climatologically aided interpolation to

minimize biases resulting from topographic and network variability (Willmott and Robeson 1995; Hamlet and Lettenmaier 2005). In brief, this approach typically requires the creation of three grid types: first-guess grids of climate normals by month, anomaly grids (i.e., departures from normal) for each year and month, and composite normal-anomaly grids for each year and month. The normal grids capture finescale detail using all stations whereas the anomaly grids capture broad departures from average using only the sites available at that time. Because anomalies generally exhibit less spatial variability than the mean field, the approach is usually viable for sparse networks (New et al. 2000), though some local features (e.g., inversions) may not be resolved if they do not occur on a climatological basis (Daly 2006).

The remainder of this section presents the gridding approach in detail, the discussion consisting of three topics. The first section describes the computation of station-based climate normals, which were required for the creation of both the normal and anomaly grids. The next section introduces the thin-plate smoothing spline method, which was employed in generating the normal grids. The third section details the preparation of the composite grids, emphasizing the computation and interpolation of station anomalies via inverse-distance weighting.

#### a. Climate normals

The first step involved the computation of climate normals for each station, month, and element. Normals were computed using a base period of 1981–2010 to maximize the number of stations from the RAWS and SNOTEL networks, which expanded markedly during the period. If a station had a complete record in the base period, then the normals were simply averages of the 30 monthly values. If a station had an incomplete record in the base period, then estimates were generated for all missing months before computing the normals (to minimize the bias caused by excluding an unusually cold or dry year). If a station was missing more than two-thirds of its record in the base period, then two steps were employed to compute the normals: first, averages were calculated for a previous 30-yr span (e.g., 1971–2000) and, second, neighbor-based adjustments were applied to those averages, producing normals consistent with the 1981–2010 base period.

Most stations required estimates for missing values, with 25% of all sites lacking half of their base-period data. Estimates for missing values were created using least absolute deviation regression (Mielke and Berry 2001). For each regression model, the dependent variable was the time series at the target station, and the

independent variables were the series from up to five neighboring stations that were climatologically similar to the target. Similarity was quantified using the index of agreement  $d$  to capture both additive and proportional differences between series (Legates and McCabe 1999). Neighbors were included in a stepwise fashion until reaching a maximum of five or a decline in performance (i.e., a decrease in the  $d$  between the target and the model-predicted series). Each station and month had a unique regression model, and in computing a normal, each missing value estimate had the same weight as each observed value (i.e.,  $1/30$ ).

One-quarter of all sites lacked sufficient base-period data for the direct computation of normals. In such cases, normals were obtained by finding a previous 30-yr span with sufficient data, filling in missing values, computing averages by month, and then adjusting the averages to approximate 1981–2010. Six previous periods were considered (1976–2005, 1971–2000, 1966–95, 1961–90, 1956–85, and 1951–80), with about 4% of the sites falling into each period. The last step (i.e., the adjustment process) was performed separately by month and used up to five neighboring stations to estimate climatological differences between the periods. Similarity was again quantified using  $d$ , with the added constraint that each neighbor had both a base-period normal and an average for the same time as the target. An adjustment factor was calculated for each neighbor; for temperature, this factor was the base-period normal minus the average from the earlier period, and for precipitation, this factor was the base-period normal divided by the average. The neighbor-based adjustments factors were then composited, with each site receiving a weight proportional to its  $d$  value. Finally, the composite adjustment was applied to the target's average, producing a normal consistent with the base period.

#### b. Normal grids

Climatologically aided interpolation requires the creation of three grid types, the first being grids of station-based climate normals for each element and month. These “normal” grids were produced using the thin-plate smoothing spline method (Hutchinson 1995). As noted by Daly (2006), this method is well suited for a large domain such as the United States because the relationship between the dependent variable (e.g., temperature) and the predictor variables (e.g., elevation) can vary in space, facilitating the reconstruction of complex geographical patterns (e.g., Hijmans et al. 2005; Rehfeldt 2006). Notably, the approach generates a continuous surface rather than an exact interpolation through the data, reducing the chance that measurement error (e.g., from poor siting) leads to unrealistic spatial gradients.

The method itself is formally implemented in a software package called the Australian National University Splines (ANUSPLIN) package, which has become a leading technology for interpolation (McKenney et al. 2011). The basic elements of this implementation are described here; for further details, see Wahba (1990) and Hutchinson (2004).

The objective of the thin-plate smoothing spline method is to estimate an approximate function  $g(x_i, y_i)$  that matches the data values as closely as possible while also being as smooth as possible. This is accomplished by minimizing

$$\frac{1}{n} \sum_{i=1}^n [(z_i - g(x_i, y_i))^2 + \lambda J_m(g)],$$

where  $z_i$  is the value of the dependent variable at station  $i$  (e.g., its normal temperature in July),  $x_i$  and  $y_i$  denote the location of station  $i$  (e.g., its latitude and longitude), the summation term is the mean-square error;  $\lambda$  is a smoothing parameter,  $J_m(g)$  is a penalty function that quantifies surface roughness with high-order partial derivatives, and  $m$  is the order of the spline ( $m = 3$  in this study, meaning a cubic spline). The smoothing parameter  $\lambda$  balances the influence of the mean-square error and the roughness penalty; a smaller  $\lambda$  results in a lesser cost for roughness and thus a tighter fit through the data (i.e., a more variable surface).

The smoothing parameter is usually obtained by minimizing the generalized cross validation (GCV), a measure of predictive skill for the fitted surface. In essence, GCV is computed by removing a point, fitting a surface through the remaining points, estimating the value at the location of the withheld point, calculating the squared residual for that location, and then repeating the process for all other points (the GCV being the average of the residuals). To account for short-range correlation between the values, the GCV is often minimized for a subset of more evenly spaced stations (knots) rather than the full network (Bates and Wahba 1982). In this approach, all stations influence the shape of the surface, but the GCV itself is only based on stations that were selected as knots. The present investigation used 75% of the stations as knots (as in Sharples et al. 2005), and to increase computational efficiency, the conterminous United States was divided into three tiles that overlapped by  $5^\circ$  of longitude (i.e.,  $130^\circ$ – $100^\circ$ W,  $105^\circ$ – $85^\circ$ W, and  $90^\circ$ – $65^\circ$ W). Surfaces were fit separately to each tile and then merged, the weight of each tile being a linear function of the distance between a particular location and the edge of that tile.

All climate normals were modeled as a smooth function of latitude, longitude, and elevation because those

locational attributes explain much of the spatial variation in climate. Coordinates were scaled in decimal degrees while elevation was scaled in kilometers, effectively exaggerating its influence by a factor of 100 for consistency with the generally accepted horizontal and vertical distance scales of atmospheric dynamics (Sharples et al. 2005; Daley 1991). To supplement the coordinates and elevation, additional locational predictors were included for individual elements to improve skill in areas where thin-plate splines can have difficulty in simulating abrupt transitions (Daly 2006). In particular, a metric of coastal influence was included to model the damping effect of large water bodies on maximum and minimum temperatures. Likewise, a secondary metric of elevation was included to model the influence of atmospheric inversions on minimum temperature. Finally, metrics of slope and aspect were included to enhance the identification of windward and leeward exposures and their impact on precipitation. For a detailed description of these additional predictors, see the appendix.

The thin-plate smoothing spline method produces a spatially continuous surface of the dependent variable. As a result, it can be resolved to any desired grid by supplying an appropriate lattice of the predictor variables, usually in the form of a digital elevation model (DEM). The present application employed the Global 30 arc-s elevation dataset (GTOPO30) DEM (USGS 2013) in that regard, resampling from its native 1-km resolution to a nominal latitude-longitude spacing of 5 km using a focal median technique (i.e., by finding the median within a search radius of four grid cells).

### c. Composite grids

Climatologically aided interpolation requires the development of three grid types for each climatic element. The previous section described the preparation of the first type, that is, normal grids, which are designed to capture finescale topographic detail using all stations. This section describes the creation of the remaining grid types (i.e., anomaly and composite grids), which are designed to capture interannual variability using only those stations available during each year and month. The process itself consisted of three steps: computing anomalies from climate normals for each station, year, and month; creating anomaly grids via inverse-distance weighting; and merging the normal and anomaly grids into a composite field for each year and month.

Anomalies were employed in the gridding process to account for large gradients arising from differences in station elevation and location. For temperature, the anomaly for a given year and month was computed by subtracting a station's climate normal for the calendar month from the station's actual temperature in that year

and month. For precipitation, the anomaly for a given year and month was computed by dividing the station's actual total in that year and month by the station's climate normal for the calendar month. On average, anomalies exhibit less spatial variability than the original data, facilitating interpolation with sparse station networks (Mitchell and Jones 2005). However, anomaly fields still contain some roughness for a variety of reasons (e.g., station siting, coastal effects, topographic positioning); furthermore, increases in station density can result in increasingly complex anomaly patterns through time, particularly in high-elevation areas of the West. If an anomaly deviated significantly from its neighbor-based average, then it was smoothed to reduce its influence during interpolation, thus reducing the chance that a divisional average was heavily impacted by a (presumably) local-scale feature. Smoothing was accomplished using the interpolation residual-based index of Lui et al. (2001); specifically, if an index exceeded a  $z$  score of 3, then the value used in gridding was a weighted average of the original anomaly (20%) and the index-based prediction (80%). Approximately 1% of the anomalies were smoothed in this fashion.

Inverse-distance weighting was applied to interpolate the anomalies onto the 5-km grid. The approach used here was essentially that of Willmott et al. (1985), which was employed in the original version of climatologically aided interpolation described in Willmott and Robeson (1995). As with all distance-based algorithms, the approach estimates a value at each grid point using a small number (15–25) of nearby stations, their respective weights being proportional to the distance between the stations and the grid point. The algorithm performs interpolation in spherical rather than Cartesian coordinates to increase predictive skill, accounts for both the distance and angular relationships between stations and grid points, corrects for the directional isolation of a station relative to its neighbors, and permits grid points to take on values outside the range of the data via a limited extrapolation function. In this investigation, the distance between each station and grid point was artificially inflated by a modest 25 km (roughly five grid cells) to prevent the algorithm from being an exact interpolator, further minimizing the impact of local-scale features on gridded fields (and thus divisional averages). Slightly smaller and larger distances were also tested, but the resulting grids were generally comparable.

The creation of the composite grid for each year and month differed slightly by climatic element. For maximum and minimum temperatures, the composite grid for a given year and month was created by adding the normal grid for the calendar month to the anomaly grid for that year and month. Average temperature was

computed by taking the mean of maximum and minimum temperatures. For precipitation, the composite grid for a given year and month was created by multiplying the normal grid for the calendar month by the anomaly grid for that year and month.

#### 4. Divisional examples

Climate division values for each element were derived from the composite grids. More specifically, the value for each division, year, and month was computed as the area-weighted average of the composite gridpoint values whose centroids fell within the boundaries of that division in that month. This approach differs substantially from the methods used in the original divisional dataset, resulting in systematic differences between the two versions, particularly for divisions with large topographic variability. This section presents examples of climate divisions that had substantial differences in 2012, each case demonstrating the benefits obtained by using climatologically aided interpolation.

The San Joaquin drainage division in California (Fig. 4) illustrates the impact of elevation on average temperature at the divisional level. For example, during July 2012 the divisional value in version 1 (24.9°C) is 1.5°C higher than the value in version 2 (23.4°C). The version 1 value is higher because it is an arithmetic mean of station observations, the majority of which are in the San Joaquin Valley, which is warmer than the Sierra Nevada in the eastern third of the division. Notably, the composite grid depicts lower temperatures throughout the Sierras (16% of the composite grid points have elevations above 2000 m versus just one COOP station), contributing to a lower divisional value in version 2.

The Central Coast drainage division in California (Fig. 4) illustrates the impact of coastal effects on average temperature at the divisional level. For instance, during July 2012 the divisional value in version 1 (18.8°C) is 1.4°C lower than the value in version 2 (20.2°C). The version 1 value is lower because it is an arithmetic mean of station observations, 61% of which are within 25 km of the coast, which is generally cooler than areas just a short distance inland as a result of onshore flow in summer (Daly et al. 2002). The composite grid captures this relatively narrow inland penetration of marine air, particularly for maximum temperature, helping explain the higher divisional value in version 2.

The Rio Grande drainage division in Colorado (Fig. 5) illustrates the impact of atmospheric inversions on average temperature at the divisional level. For example, during January 2012 the divisional value in version 1 ( $-6.6^{\circ}\text{C}$ ) is 1.5°C lower than the value in version 2 ( $-5.1^{\circ}\text{C}$ ). The version 1 value is lower because it is an

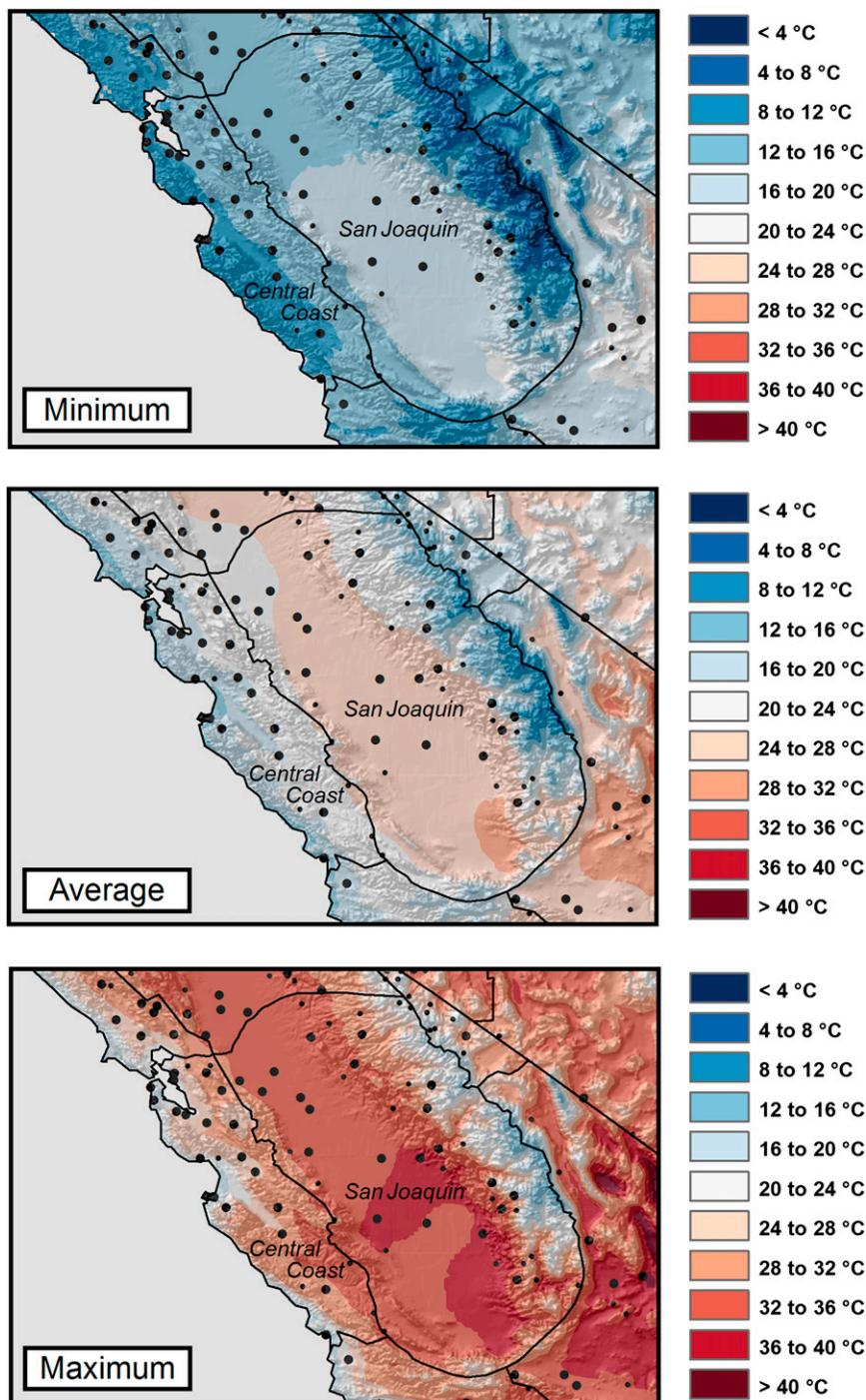


FIG. 4. Maps of minimum, average, and maximum temperatures during July 2012 in the San Joaquin and Central Coast divisions of southern California. Large dots denote the locations of ASOS and COOP stations while small dots denote the locations of RAWS and SNOTEL stations.

arithmetic mean of station observations, most of which are located in a basin that experiences cold-air drainage (and thus very low minimums) at night. The composite grid generally replicates the resulting inversion pattern

in minimum temperature, wherein the lowest values are at the lowest elevations and the highest values occur roughly at midslope, contributing to the higher divisional value in version 2.

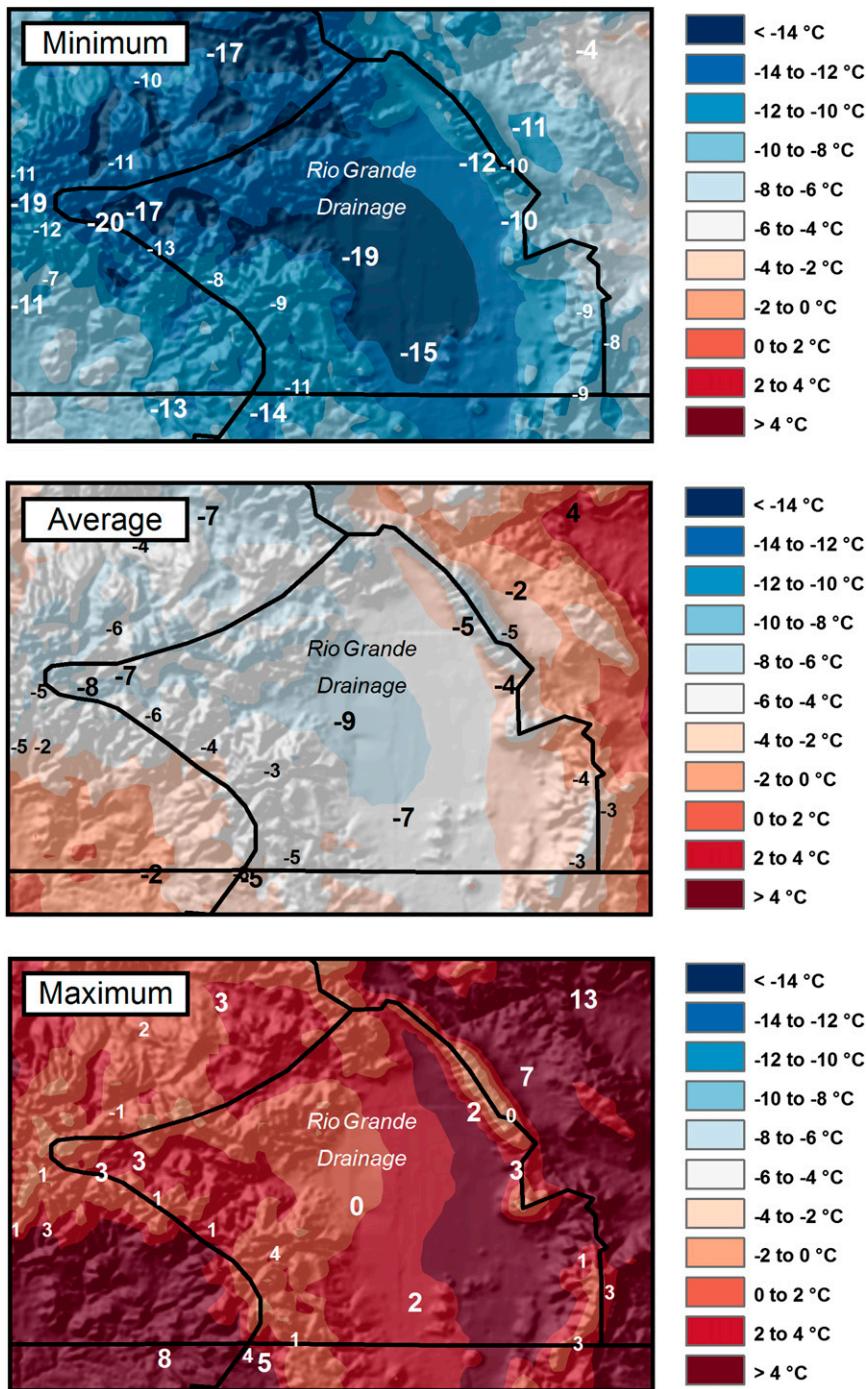


FIG. 5. Maps of minimum, average, and maximum temperatures during January 2012 in the Rio Grande drainage division of southern Colorado. Numbers denote temperatures ( $^{\circ}\text{C}$ ) at individual stations (values at ASOS and COOP stations are in boldface).

The Northeast division in Arizona (Fig. 6) illustrates the impact of differences in methodology on the spatial mean of precipitation. For example, during July 2012 the divisional value in version 1 (81.3 mm) is 25% wetter

than the value in version 2 (64.5 mm) even though station totals are generally consistent with the composite gridded field. The version 1 value is higher because it is an arithmetic mean of station totals, which, due to

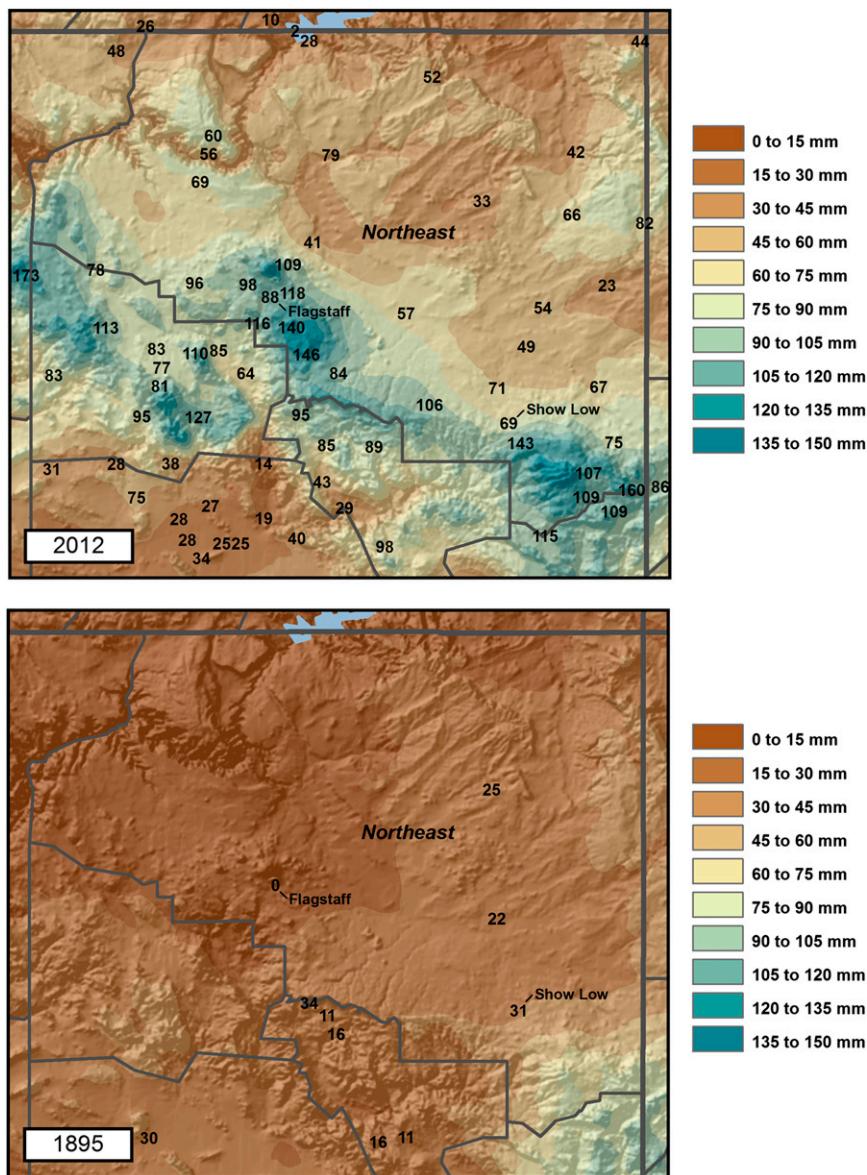


FIG. 6. Maps of precipitation in the northeast division of Arizona during July 2012 and July 1895. Numbers represent precipitation totals (mm) at individual stations.

orographic effects, are higher along the Mogollon Rim (the southern boundary of the division), where the station density is greater. The Northeast division also illustrates the suboptimal nature of version 1 from 1895 to 1930, a period when divisional values were estimated using regression models that employed USDA statewide averages as predictors. For instance, in July 1895 the regression estimate of version 1 (26.4 mm) is 31% wetter than the value in version 2 (20.1 mm). Ironically, the latter is nearly identical to the arithmetic mean of the station totals (19.5 mm), which again is consistent with the composite gridded field. Notably, both depict much

lower totals north of the Mogollon Rim in 1895 than in 2012 (as illustrated by the values near Flagstaff and Show Low).

## 5. Database differences

There are several prominent differences between version 2 and version 1 at both the divisional and the national levels. This section explores the major differences from 1895 to 2012, focusing on long-term means and temporal trends (computed using Kendall-Theil robust lines; Helsel and Hirsch 1991). The emphasis is on

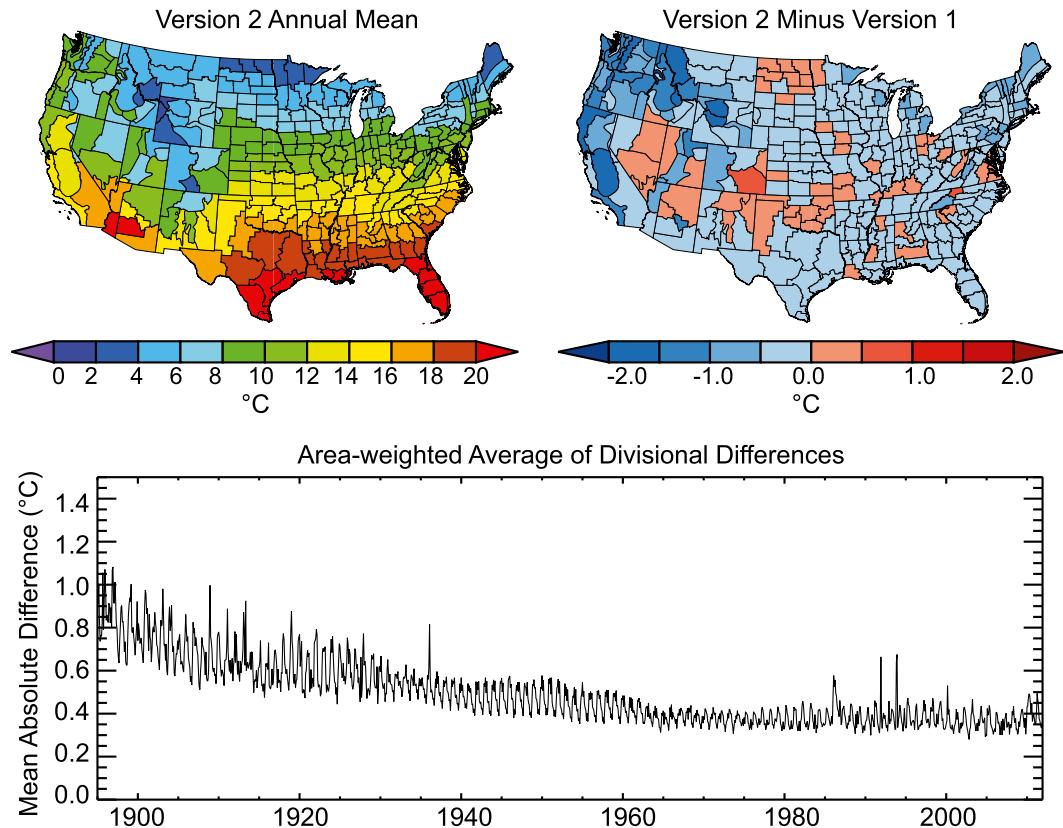


FIG. 7. Comparison of average temperatures in version 2 and version 1 for the period 1895–2012. (top left) The annual average in each division in version 2. (top right) The difference in the annual average (i.e., version 2 minus version 1) in each division. (bottom) The monthly time series of the area-weighted average of divisional differences.

large-scale differences at the annual and seasonal time scales, with a brief mention of select divisions having improved estimates in version 2. Temperature and precipitation are discussed separately.

#### a. Temperature

Figure 7 depicts the differences in average temperature at both the divisional and the national levels. From a divisional perspective, most differences in the annual average are less than 0.5°C, particularly in the East, and most differences in excess of 1.0°C are in the West. The latter are usually negative, indicating lower averages in version 2. These largely result from better sampling of cool, high-elevation areas in version 2, as exemplified by the San Joaquin drainage division described in the previous section. From a national perspective, the area-weighted difference between the datasets is  $-0.3^{\circ}\text{C}$ , indicating that version 2 is slightly cooler overall than version 1. Notably, the average gridpoint elevation in version 2 (790 m) is also higher than the average elevation of observing stations in version 1 (568 m). From a temporal perspective, differences generally decline until the

mid-1960s and remain fairly stable thereafter, with occasional increases in isolated months (caused by disparities in large divisions where coverage in version 2 is superior owing to RAWS and SNOTEL). From a seasonal perspective, differences are marginally larger in winter than summer, which is generally consistent with greater variability during the cold season.

Figure 8 depicts trends in annual average temperature at the divisional level over the period 1895–2012. Version 2 shows warming throughout the nation except for a relatively small area of cooling in the Southeast. Version 1 exhibits the same primary features, but the divisional trends are often smaller, the cooling area is more extensive, and the overall pattern is less consistent (e.g., the northern division of Maine exhibits cooling while the rest of New England is warming). In general, most of these trend differences are attributable to historical changes in station location, temperature instrumentation, observing practice, and siting conditions, none of which are addressed at the station level in version 1 (i.e., it contains no station-specific bias adjustments for such changes). In particular, the smaller trends and greater

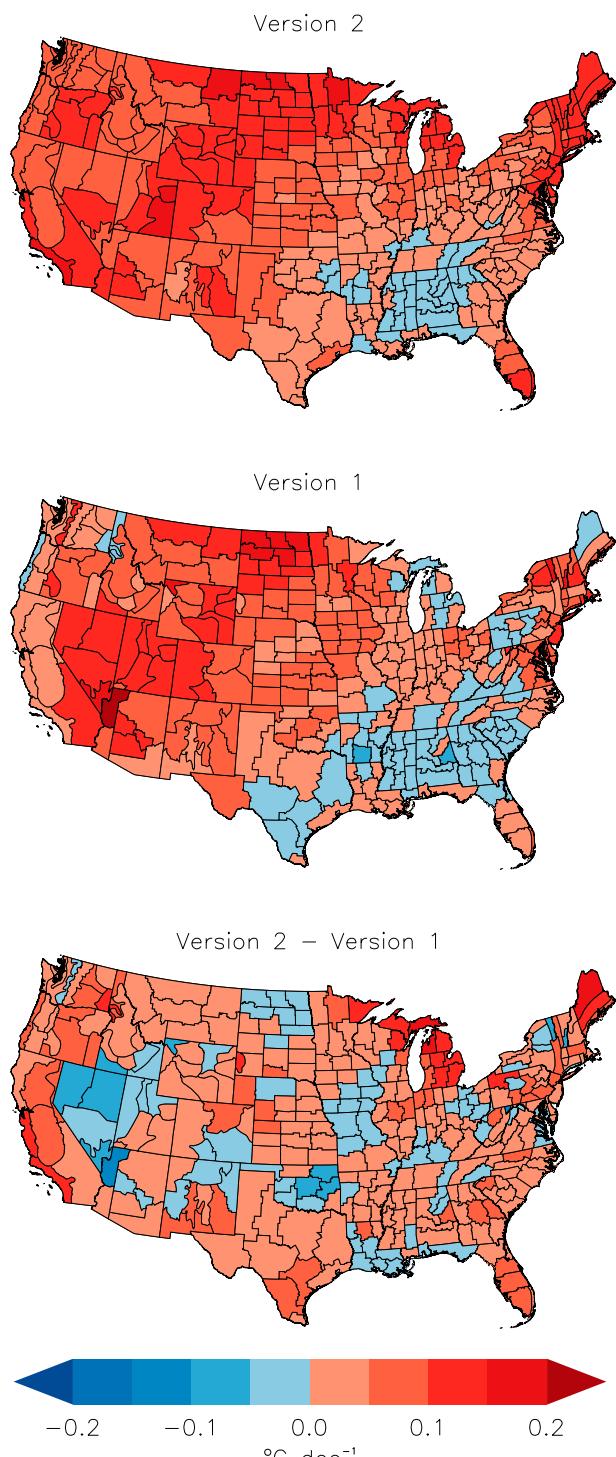


FIG. 8. Maps of annual average temperature trends ( $^{\circ}\text{C decade}^{-1}$ ) by division from 1895 to 2012: (top) version 2, (middle) version 1, and (bottom) version 2 – version 1.

TABLE 1. Area-average temperature trends ( $^{\circ}\text{C decade}^{-1}$ ) over the conterminous United States for the period 1895–2012.

Dataset	Annual	Winter	Summer
Avg			
Version 1	0.044	0.077	0.047
Version 2	0.069	0.120	0.055
HCN 2.5	0.067	0.106	0.065
Max			
Version 1	—	—	—
Version 2	0.066	0.119	0.035
HCN 2.5	0.064	0.098	0.050
Min			
Version 1	—	—	—
Version 2	0.076	0.122	0.071
HCN 2.5	0.073	0.112	0.077

cooling in version 1 reflect the systematic change from afternoon to morning observing times since the 1950s and the installation of the Maximum–Minimum Temperature System since the 1980s, both of which artificially cooled the U.S. temperature record (Williams et al. 2012; Vose et al. 2012). The lower spatial consistency in version 1 reflects these network-wide events, other changes at the local level (such as station location), and changes in station coverage through time, all of which are known to increase noise in trend patterns (Menne et al. 2009). Trend differences are also apparent at the state level, with Nevada and North Dakota being obvious examples. These differences reflect the use of the state-based regression models in version 1 for the period 1895–1930.

Table 1 shows national-scale trends over the period 1895–2012. Consistent with its lack of bias adjustments, version 1 has smaller average temperature trends than both version 2 and HCN, a bias-adjusted dataset that has been used by NOAA to monitor national trends for many years. Version 2 and HCN have nearly identical trends for all elements at the annual time scale, with the former exhibiting slightly more warming in winter ( $\sim 0.015^{\circ}\text{C decade}^{-1}$ ) and slightly less warming in summer ( $\sim 0.010^{\circ}\text{C decade}^{-1}$ ).

#### b. Precipitation

Figure 9 depicts the differences in average precipitation at both the divisional and the national levels. From a divisional perspective, most differences in average annual precipitation are less than 50 mm, particularly in the East, while most differences in excess of 100 mm are in the West. The latter are usually positive, indicating higher totals in version 2. These largely result from better sampling of upslope areas subject to orographic precipitation in version 2, as exemplified by the western division in Montana [an area in which precipitation is highly

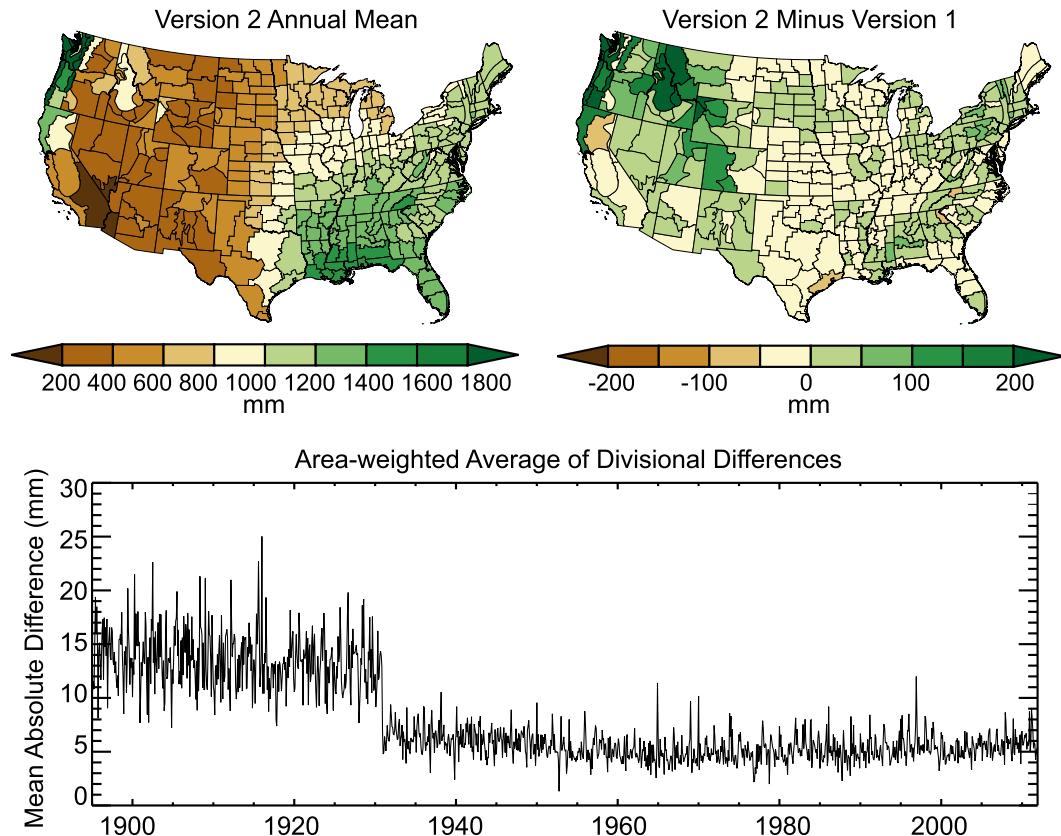


FIG. 9. As in Fig. 7, but for precipitation.

correlated with elevation; [Silverman et al. \(2013\)](#)]. From a national perspective, the area-weighted difference between the datasets is 20 mm, indicating that version 2 is slightly wetter overall than version 1. From a temporal perspective, the datasets exhibit much greater similarity starting precisely in 1931, reflecting the change in computational methods in version 1 at that time. There are otherwise no systematic changes in the national-scale difference series through time, nor is there any obvious seasonal signal.

[Figure 10](#) depicts trends in annual average precipitation at the divisional level over the period 1895–2012. Version 2 shows increases in most of the East (except for parts of the Southeast) and a mixture of increases and decreases in the West. Version 1 exhibits the same primary features, but the divisional trends are often larger, and the overall pattern is less consistent. Discrepancies are also evident at both the state and divisional levels; for instance, version 1 has larger increases in Alabama and South Dakota, and it exhibits drying in small areas that are mostly surrounded by increases (such as the northeast division in Arkansas). The lower spatial consistency in version 1 is primarily related to changes in station coverage through time. As with temperature, the

state-level differences reflect the use of the state-based regression models in version 1 for the period 1895–1930. While not shown here, version 2 exhibits only small trend differences at the divisional level with the Full Network Enhanced Precipitation (FNEP) dataset ([McRoberts and Nielsen-Gammon 2011](#)), a divisional database that was designed for the analysis of climate variability and change. In contrast, version 1 exhibits substantial differences with FNEP, as thoroughly documented by [McRoberts and Nielsen-Gammon \(2011\)](#).

Despite division-level differences, version 1, version 2, and FNEP have very similar trends at the national scale for the period 1895–2012. In particular, all of the datasets exhibit statistically significant increases in annual precipitation (4.0, 3.3, and 4.2 mm decade<sup>-1</sup>, respectively). These increases are driven primarily by changes in fall precipitation (2.3, 2.2, and 2.4 mm decade<sup>-1</sup>, respectively). Trends in all other seasons are <1 mm decade<sup>-1</sup>, and none are statistically significant.

## 6. Error estimates

Two analyses were performed to provide uncertainty estimates for version 2. In the first step, cross validation

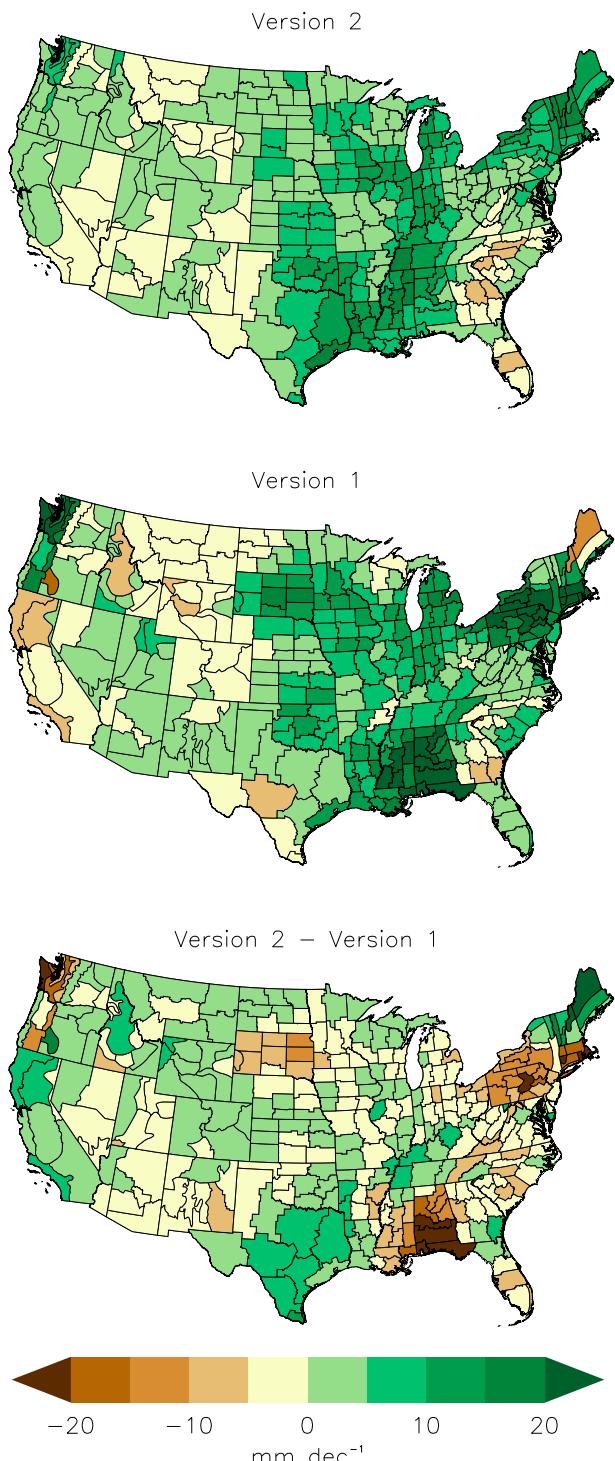


FIG. 10. As in Fig. 8, but for average annual precipitation trends ( $\text{mm decade}^{-1}$ ).

was used to quantify interpolation error over the conterminous United States as a whole. In the second step, Monte Carlo simulations were used to estimate error at the climate division level.

#### *a. Composite grids*

Interpolation error was quantified using a cross-validation exercise consisting of three steps for each element and calendar month. The first step involved computing a climate normal residual: the difference between the actual climate normal at each station and the normal predicted by the spline surface at that location. The second step entailed calculating a climate anomaly residual: the difference between the actual anomaly at each station in each year and the anomaly predicted by neighboring stations (i.e., interpolated via inverse-distance weighting). Finally, the normal and anomaly residuals were summed at the station level, converted to absolute differences, interpolated onto the 5-km grid, and then area averaged over the conterminous United States.

Figure 11 depicts area-averaged cross-validation errors over the period 1895–2012. For temperature, errors decrease rapidly until about 1905 and then gradually until about 1990, remaining relatively stable thereafter. Errors exceed  $1.0^\circ\text{C}$  early in the record and attain present-day minima of about  $0.60^\circ\text{C}$  for maximum temperature and  $0.75^\circ\text{C}$  for minimum temperature. From a seasonal perspective, temperature errors are about a tenth of a degree smaller in summer than winter. For precipitation, errors decrease gradually until about 1950 and then generally level off thereafter, with substantial seasonal and interannual variability throughout the record. Errors in summer exceed 25 mm in the late nineteenth century, falling to about 20 mm in recent years. Errors in winter are about half as large as in summer.

#### *b. Divisional averages*

Divisional error was quantified using a Monte Carlo exercise that entailed four general steps for each element and calendar month. In essence, the objective was to determine how well historical, low-density networks reproduced divisional averages based upon contemporary, high-density networks. The first step entailed selecting a recent year with a high-density network to serve as a baseline (e.g., 2000). The second step involved reducing the density of this network such that it mimicked the station density in an earlier year (e.g., 1895). The third step required computing climate anomalies, composite grids, and divisional values for the baseline year using only those stations in the reduced network. The final step involved calculating differences between reduced- and baseline-network divisional values. The last three steps were performed 100 times, stations for each simulation being selected in a stratified random fashion using  $5^\circ \times 5^\circ$  grid boxes as a sampling guide (e.g.,

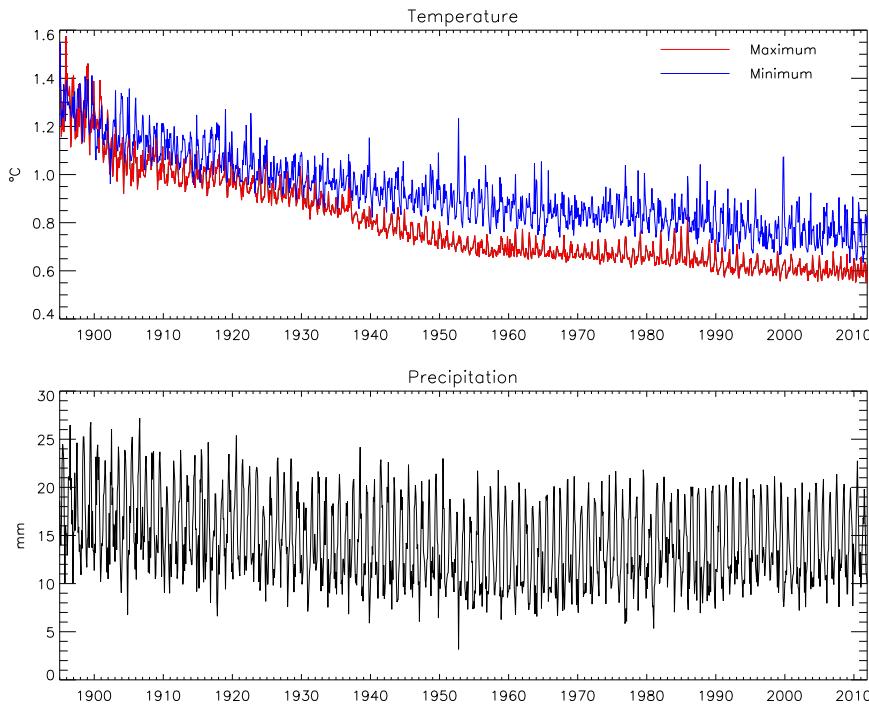


FIG. 11. Plot of area-averaged cross-validation errors versus time: (top) temperature and (bottom) precipitation.

by determining the number of stations in each box in 1895, then randomly selecting stations from the network in 2000 such that each box had the same number of stations as in 1895).

Figure 12 summarizes the differences between reduced- and baseline-network divisional values on a quinquennial basis using 2000 (i.e., the year with the most stations) as a baseline. The median difference is zero throughout the period for all elements and months, and there is a decline in the differences through time. For temperature, approximately 95% of the divisional differences are less than  $0.50^{\circ}\text{C}$  by 1900 and less than  $0.25^{\circ}\text{C}$  by 1925; for precipitation, the corresponding differences are 20 and 15 mm, respectively. From a seasonal perspective, temperature differences exhibit slightly greater spread in January, which is consistent with larger variability in winter. In contrast, precipitation differences exhibit considerably greater spread in July, which is consistent with the spotty nature of convective rainfall during summer.

Figure 13 depicts median differences at the divisional level using 2000 for the baseline network and 1895 (i.e., the year with the fewest stations) for the reduced network. For temperature, most divisions east of the Rocky Mountains have differences less than  $0.25^{\circ}\text{C}$ . Differences in the West are about twice as large, with a few topographically varied divisions exceeding  $0.50^{\circ}\text{C}$ . From

a seasonal perspective, temperature differences are slightly larger in January than in July. For precipitation, most divisional differences are less than 10 mm, with slightly larger differences in July in the East (reflecting the region's higher seasonal totals). In absolute terms, the largest precipitation differences generally correspond to areas with the highest average totals (e.g., the Pacific Northwest in January, Florida in July). As a percent of normal, however, the largest precipitation differences usually align with areas having low totals on a climatological basis (such as the Southwest in winter and the West in summer) or low totals in this particular year (as in the case of Texas).

## 7. Summary

This paper described an improved edition of the climate division dataset for the conterminous United States. The first improvement was to the input data, which now include additional station networks, quality assurance reviews, and temperature bias adjustments. The second improvement was to the suite of climatic elements, which now includes both maximum and minimum temperatures. The third improvement was to the computational approach, which now employs climatologically aided interpolation to address topographic and network variability.

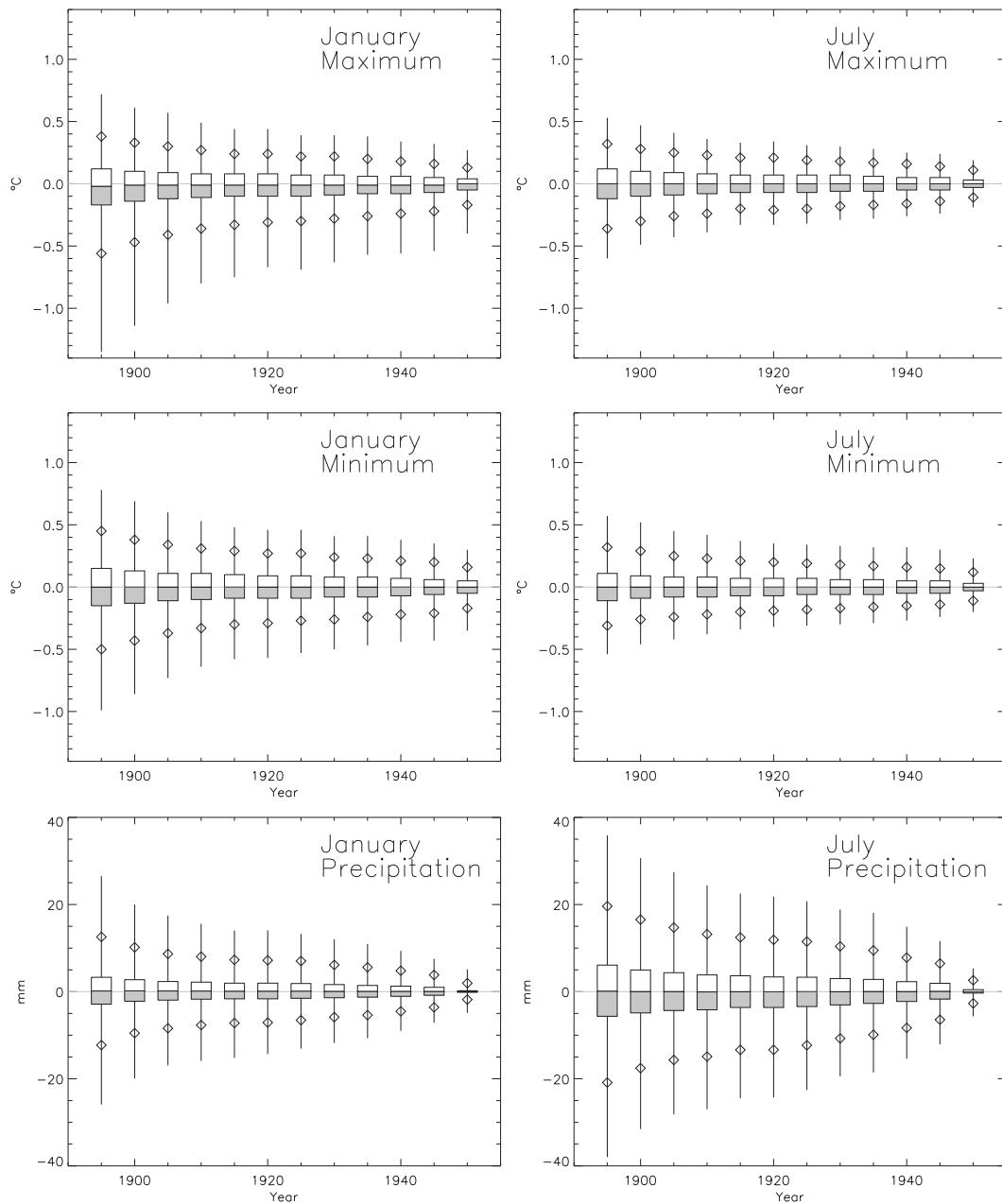


FIG. 12. Plots of differences between reduced- and baseline-network divisional values. Each year depicts 34 400 differences (i.e., 344 divisions times 100 Monte Carlo simulations). The boxes represent the 25th, 50th, and 75th percentiles; the diamonds represent the 95th percentiles; and the lines represent the 99th percentiles. The baseline year was 2000 (i.e., the year with the largest number of stations).

Version 2 exhibits substantial differences from version 1 over the period 1895–2012. For example, divisional averages in version 2 tend to be cooler and wetter, particularly in mountainous areas of the western United States. Division-level trends in temperature and precipitation display greater spatial consistency in version 2. National-scale temperature trends in version 2 are comparable to

those in HCN, whereas version 1 exhibits less warming as a result of historical changes in observing practice. Divisional errors in version 2 are likely less than 0.5°C for temperature and 20 mm for precipitation at the start of the record, falling rapidly thereafter; in contrast, methodological considerations precluded the estimation of divisional uncertainty in version 1. Overall, these results

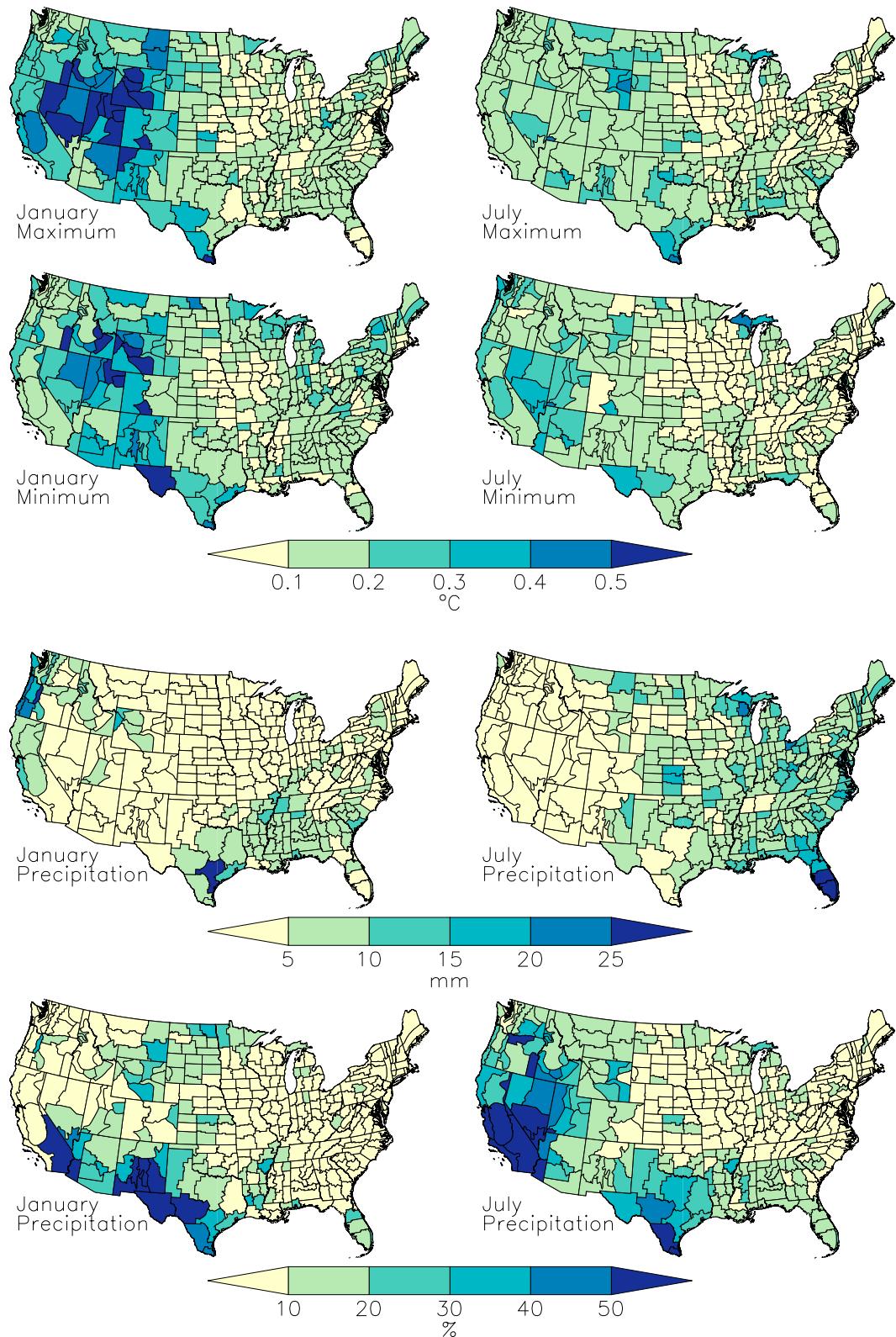


FIG. 13. Maps of median differences between reduced-network divisional values and baseline-network divisional values. Medians are based on 100 Monte Carlo simulations using the year 1895 for the reduced network and the year 2000 for the baseline network.

indicate that version 2 can supersede version 1 in both operational climate monitoring and applied climatic research.

*Acknowledgments.* The authors thank the anonymous reviewers, whose comments substantially improved this manuscript.

## APPENDIX

### Additional Predictors

The normals for all climatic elements were gridded via thin-plate smoothing splines that used coordinates and elevation as predictors. Additional predictors were also included for individual elements to improve skill in topographic environments that could impact divisional averages. In particular, a coastal influence index was used for maximum–minimum temperatures, an inversion index was used for minimum temperature, and slope-aspect were used for precipitation. The following is a brief description of the relatively simple indices used in this study; for more elaborate alternatives, see [Daly et al. \(2008\)](#).

#### a. Coastal index

Coastal influence was modeled using an index that simulates the impact of distance and terrain on marine damping of temperature. In essence, the farther a location is from the shore, the smaller its index; likewise, the greater its topographic separation from the shore (e.g., leeward exposure), the smaller its index. A simplified representation is given by

$$c_i = s_i - t_i,$$

where  $c_i$  is the coastal influence index for location  $i$  ( $c_i = 1$  implies maximum coastal effects),  $s_i$  is a distance-decay term, and  $t_i$  is a terrain effects term. Distance-decay effects can exist near any large water body whereas terrain effects mainly exist on the Pacific coast, where mountain ranges limit the inland penetration of marine air. The distance-decay term  $s_i$  is calculated as

$$s_i = \exp(\sqrt[3]{2.5/d_i} - 1),$$

where  $d_i$  is the distance to the coast (km). The exponential function imposes a steep decline in coastal influence within  $\sim 10$  km of the shore, while full marine effects are assigned within half a grid cell of the coast (i.e.,  $s_i = 1$  if  $d_i \leq 2.5$  km). The terrain effects term  $t_i$  is computed as

$$t_i = s_i \left[ l_i \left( 1 - \frac{d_i}{100} \right) \right]^2,$$

where  $l_i$  is the least-cost path to the coast, cost being the accumulation of gridded slope values traversed in reaching the shore. The least-cost path approaches 0.5 in extreme terrain; thus, the penalty can be one-quarter of the size of the distance-decay effect. The penalty has its largest impact near the shore, gradually losing its influence inland and disappearing entirely at 100 km.

#### b. Inversion index

Atmospheric inversions were modeled using an index that simulates the vertical profile of nocturnal temperature. This index is zero if the location is above the inversion or if no climatological inversion exists. If the location is below the inversion, then the index is the difference in elevation (km) between the location and the inversion top. Basically, the index implies that locations with small differences in elevation are warmer than those with large differences, the former reflecting close proximity to the inversion, and the latter indicative of depressions such as valleys. Stated formally, the inversion index  $\Delta_i$  is

$$\Delta_i = \begin{cases} 0; & z_i > h_i \\ h_i - z_i; & z_i < h_i \end{cases}$$

where  $z_i$  is the elevation of location  $i$  and  $h_i$  is the elevation of the inversion top. The latter is

$$h_i = b_i + w_i,$$

where  $b_i$  is the smoothed “base” elevation for location  $i$  (as in [Daly et al. 2008](#)) and  $w_i$  is the height of the “warm point” above the ground. The base elevation  $b_i$  was obtained by finding the minimum of the gridded elevations within a 10-km radius, then averaging the minima within a 20-km radius (simple thresholds being applied to exclude false minima from canyons in adjacent areas). The warm point  $w_i$  was obtained using temperature profiles derived from radiosonde data at 119 stations ([Durre et al. 2006](#)). For each 1200 UTC sounding, temperature was interpolated in 10-m increments above the surface using mandatory and significant levels as predictors, and then a normal was computed by month for each increment. The inversion top was defined as the warmest point in the climatological profile ([Kahl 1990](#)) or the level at which the lapse rate fell below a critical threshold (for isothermal profiles). Warm-point surfaces were then created by month via thin-plate smoothing splines that used latitude and longitude as predictors, the

height of the warm point  $w_i$  being obtained by interrogating the appropriate monthly surface at location  $i$ .

### c. Slope–aspect indices

The impact of slope and aspect on precipitation was modeled using eastings and northings (Hutchinson 1998). In essence, the larger the indicator is, the steeper the slope in that direction, with flat areas having values of zero. The eastings  $p_i$  and northings  $q_i$  are given by

$$p_i = \cos(\alpha_i) \sin(\theta_i) \quad \text{and}$$

$$q_i = \sin(\alpha_i) \cos(\theta_i),$$

where  $\theta_i$  and  $\alpha_i$  are the slope and aspect angles for location  $i$ , respectively. As in Hutchinson (1998),  $p_i$  and  $q_i$  were computed on a smoothed grid (50 km) to approximate meso-beta-scale phenomena such as enhanced upslope precipitation and concomitant rain shadows. Smoothing was performed using a focal median technique and a 32-cell search radius.

### d. Interpolation skill

Cross validation was employed to estimate the improvement resulting from the additional predictors. Specifically, normals for each station were estimated twice: once using a simplified model (i.e., only coordinates and elevation) and once using the full model (i.e., including the additional predictors). Improvements are illustrated here by examining states where the full models are more impactful (e.g., California, Nevada, and Washington). In California, which has cool onshore flow in summer, the coastal influence index reduces the mean absolute error (MAE) for maximum temperature in July by 0.17°C (from 1.45° to 1.28°C) for coastal stations (i.e.,  $0.0 \leq d_j \leq 5.0$  km). Estimated normals have a warm bias of 0.77°C without the index, indicating the simplified model underestimates the marine influence. In the full model, the estimated normals have a slight cool bias of 0.21°C. In Nevada, where inversions are common in winter amid the basin and range topography, the inversion index reduces the MAE for minimum temperature in January by 0.14°C (from 1.04° to 0.90°C) for stations just below the inversion (i.e.,  $0.0 \leq h_i \leq 0.2$  km). Estimated normals have a cool bias of 0.34°C without the index, indicating the simplified model underestimates the temperature increase near inversions. In the full model, the bias falls to 0.11°C. In Washington, where orographic lifting is common, the slope–aspect indices reduce the MAE for precipitation in January by 8.6 mm (from 16.6 to 8.0 mm) for sites on the steepest 20% of slopes (i.e.,  $p_i$  or  $q_i \geq 0.0125$ ). As a percent of normal, this is a drop from 10.0% to 4.5%.

Notably, Figs. 4–6 contain patterns directly attributable to the additional predictors, that is, features that are nonexistent or muted on maps based on simplified models (not shown). For instance, Fig. 4 depicts a narrow band of cooler maximum temperatures near the Pacific resulting from the coastal influence index. Similarly, Fig. 5 depicts an increase in minimum temperature with height up to midslope resulting from the inversion index. Finally, Fig. 6 shows stronger orographic effects along the Mogollon Rim resulting from the slope/aspect indices.

## REFERENCES

- Allard, J., B. D. Keim, J. E. Chassereau, and D. Sathiaraj, 2009: Spuriously induced precipitation trends in the southeast United States. *Theor. Appl. Climatol.*, **96**, 173–177, doi:10.1007/s00704-008-0021-9.
- Anderson, B. T., J. Wang, G. Salvucci, S. Gopal, and S. Islam, 2010: Observed trends in summertime precipitation over the southwestern United States. *J. Climate*, **23**, 1937–1944, doi:10.1175/2009JCLI3317.1.
- Bates, D., and G. Wahba, 1982: Computational methods for generalized cross-validation with large data sets. *Treatment of Integral Equations by Numerical Methods*, C. Baker and G. Miller, Eds., Academic Press, 283–296.
- Casola, J. H., M. T. Stoeling, J. M. Wallace, L. Cuo, B. Livneh, D. P. Lettenmaier, and P. W. Mote, 2009: Assessing the impacts of global warming on snowpack in the Washington Cascades. *J. Climate*, **22**, 2758–2772, doi:10.1175/2008JCLI2612.1.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.
- Daly, C., 2006: Guidelines for assessing the suitability of spatial climate data sets. *Int. J. Climatol.*, **26**, 707–721, doi:10.1002/joc.1322.
- , W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris, 2002: A knowledge-based approach to the statistical mapping of climate. *Climate Res.*, **22**, 99–113, doi:10.3354/cr022099.
- , M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.*, **28**, 2031–2064, doi:10.1002/joc.1688.
- Durre, I., R. S. Vose, and D. B. Wuertz, 2006: Overview of the Integrated Global Radiosonde Archive. *J. Climate*, **19**, 53–68, doi:10.1175/JCLI3594.1.
- , M. J. Menne, B. E. Gleason, T. G. Houston, and R. S. Vose, 2010: Comprehensive automated quality assurance of daily surface observations. *J. Appl. Meteor. Climatol.*, **49**, 1615–1633, doi:10.1175/2010JAMC2375.1.
- Goodrich, G. B., 2007: Influence of the Pacific decadal oscillation on winter precipitation and drought during years of neutral ENSO in the western United States. *Wea. Forecasting*, **22**, 116–124, doi:10.1175/WAF983.1.
- , and A. W. Ellis, 2008: Climatic controls and hydrologic impacts of a recent extreme seasonal precipitation reversal in Arizona. *J. Appl. Meteor. Climatol.*, **47**, 498–508, doi:10.1175/2007JAMC1627.1.
- Grantz, K., B. Rajagopalan, M. Clark, and E. Zagona, 2007: Seasonal shifts in the North American monsoon. *J. Climate*, **20**, 1923–1935, doi:10.1175/JCLI4091.1.

- Grundstein, A., 2009: Evaluation of climate change over the continental United States using a moisture index. *Climatic Change*, **93**, 103–115, doi:10.1007/s10584-008-9480-3.
- Guttman, N. B., and R. G. Quayle, 1996: A historical perspective of U.S. climate divisions. *Bull. Amer. Meteor. Soc.*, **77**, 293–303, doi:10.1175/1520-0477(1996)077<0293:AHPOUC>2.0.CO;2.
- Hamlet, A. F., and D. P. Lettenmaier, 2005: Production of temporally consistent gridded precipitation and temperature fields for the continental United States. *J. Hydrometeor.*, **6**, 330–336, doi:10.1175/JHM420.1.
- Hausfather, Z., M. J. Menne, C. N. Williams Jr., T. Masters, R. Broberg, and D. Jones, 2013: Quantifying the impact of urbanization on U.S. Historical Climatology Network temperature records. *J. Geophys. Res.*, **118**, 481–494, doi:10.1029/2012JD018509.
- Helsel, D. R., and R. M. Hirsch, 1991: *Statistical Methods in Water Resources*. Techniques of Water-Resources Investigations of the United States Geological Survey, Book 4: Hydrologic Analysis and Interpretation, United States Geological Survey, 510 pp.
- Hidalgo, H. G., and Coauthors, 2009: Detection and attribution of streamflow timing changes to climate change in the western United States. *J. Climate*, **22**, 3838–3855, doi:10.1175/2009JCLI2470.1.
- Hijmans, R., S. E. Cameron, J. Parra, P. Jones, and A. Jarvis, 2005: Very high resolution interpolated climate surface for global land areas. *Int. J. Climatol.*, **25**, 1965–1978, doi:10.1002/joc.1276.
- Hutchinson, M. F., 1995: Interpolating mean rainfall using thin plate smoothing splines. *Int. J. Geogr. Inf. Syst.*, **9**, 385–403, doi:10.1080/02693799508902045.
- , 1998: Interpolation of rainfall data with thin plate smoothing splines—Part II: Analysis of topographic dependence. *J. Geogr. Inf. Dec. Anal.*, **2**, 152–167.
- , 2004: ANUSPLIN version 4.3. Centre for Resource and Environmental Studies, Australian National University, Canberra, ACT, Australia. [Available online at <http://fennerschool.anu.edu.au/research/products/anusplin-vrsn-44>.]
- Kahl, J. D., 1990: Characteristics of the low-level temperature inversion along the Alaskan Arctic coast. *Int. J. Climatol.*, **10**, 537–548, doi:10.1002/joc.3370100509.
- Karl, T. R., C. N. Williams Jr., P. J. Young, and W. M. Wendland, 1986: A model to estimate the time of observation bias associated with monthly mean maximum, minimum, and mean temperature for the United States. *J. Climate Appl. Meteor.*, **25**, 145–160, doi:10.1175/1520-0450(1986)025<0145:AMTETT>2.0.CO;2.
- Keim, B. D., A. M. Wilson, C. P. Wake, and T. G. Huntington, 2003: Are there spurious temperature trends in the United States Climate Division Database? *Geophys. Res. Lett.*, **30**, 7, doi:10.1029/2002GL016295.
- , M. R. Fischer, and A. M. Wilson, 2005: Are there spurious precipitation trends in the United States Climate Division Database? *Geophys. Res. Lett.*, **32**, L04702, doi:10.1029/2004GL021985.
- Lawrimore, J., M. Menne, B. Gleason, C. Williams Jr., D. Wuertz, R. Vose, and J. Rennie, 2011: An overview of the Global Historical Climatology Network monthly mean temperature dataset, version 3. *J. Geophys. Res.*, **116**, D19121, doi:10.1029/2011JD016187.
- Legates, D. R., and G. J. McCabe Jr., 1999: Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model evaluation. *Water Resour. Res.*, **35**, 233–241, doi:10.1029/1998WR900018.
- Livezey, R. E., K. Y. Vinnikov, M. M. Timofeyeva, R. Tinker, and H. van del Dool, 2007: Estimation and extrapolation of climate normals and climate trends. *J. Appl. Meteor. Climatol.*, **46**, 1759–1776, doi:10.1175/2007JAMC1666.1.
- Lui, H., K. C. Jezek, and M. O’Kelly, 2001: Detecting outliers in irregularly distributed spatial data sets by locally adaptive and robust statistical analysis and GIS. *Int. J. Geogr. Inf. Syst.*, **15**, 721–741.
- Maquet, S., J. Zhang, and J. Ko, 2009: The value of ENSO forecast information to dual-purpose winter wheat production in the U.S. southern high plains. *J. Appl. Meteor. Climatol.*, **48**, 2100–2117, doi:10.1175/2009JAMC2018.1.
- McCabe, G. J., L. E. Hay, and M. P. Clark, 2007: Rain-on-snow events in the western United States. *Bull. Amer. Meteor. Soc.*, **88**, 319–328, doi:10.1175/BAMS-88-3-319.
- McKenney, D. W., and Coauthors, 2011: Customized spatial climate models for North America. *Bull. Amer. Meteor. Soc.*, **92**, 1611–1622, doi:10.1175/2011BAMS3132.1.
- McRoberts, D. B., and J. W. Nielsen-Gammon, 2011: A new homogenized climate division precipitation dataset for analysis of climate variability and change. *J. Appl. Meteor. Climatol.*, **50**, 1187–1199.
- Menne, M. J., and C. N. Williams Jr., 2009: Homogenization of temperature series via pairwise comparisons. *J. Climate*, **22**, 1700–1717, doi:10.1175/2008JCLI2263.1.
- , —, and R. S. Vose, 2009: The United States Historical Climatology Network monthly temperature data—Version 2. *Bull. Amer. Meteor. Soc.*, **90**, 993–1007, doi:10.1175/2008BAMS2613.1.
- , —, and M. A. Palecki, 2010: On the reliability of the U.S. surface temperature record. *J. Geophys. Res.*, **115**, D11108, doi:10.1029/2009JD013094.
- , I. Durre, B. G. Gleason, T. Houston, and R. S. Vose, 2012: An overview of the Global Historical Climatology Network Daily dataset. *J. Atmos. Oceanic Technol.*, **29**, 897–910, doi:10.1175/JTECH-D-11-00103.1.
- Mielke, P. W., and K. J. Berry, 2001: *Permutation Methods: A Distance Function Approach*. Springer-Verlag, 352 pp.
- Miller, W. P., and T. C. Piechota, 2008: Regional analysis of trend and step changes observed in hydroclimatic variables around the Colorado River basin. *J. Hydrometeor.*, **9**, 1020–1034, doi:10.1175/2008JHM988.1.
- Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high resolution grids. *Int. J. Climatol.*, **25**, 693–712, doi:10.1002/joc.1181.
- Myoung, B., and J. W. Nielsen-Gammon, 2010: The convective instability pathway to warm season drought in Texas. Part I: Free-tropospheric modulation of convective inhibition. *J. Climate*, **23**, 4461–4473, doi:10.1175/2010JCLI2946.1.
- New, M., M. Hulme, and P. D. Jones, 2000: Representing twentieth century space-time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *J. Climate*, **13**, 2217–2238, doi:10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2.
- Peterson, T. C., and Coauthors, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.*, **18**, 1493–1517, doi:10.1002/(SICI)1097-0088(19981115)18:13<1493::AID-JOC329>3.0.CO;2-T.
- Plank, C., and B. Shuman, 2009: Drought-driven changes in lake areas and their effects on the surface energy balance of Minnesota’s lake-dotted landscape. *J. Climate*, **22**, 4055–4065, doi:10.1175/2009JCLI1978.1.

- Preisler, H. K., and A. L. Westerling, 2007: Statistical model for forecasting monthly large wildfire events in western United States. *J. Appl. Meteor. Climatol.*, **46**, 1020–1030, doi:10.1175/JAM2513.1.
- Quiring, S. M., 2009: Developing objective operational definitions for monitoring drought. *J. Appl. Meteor. Climatol.*, **48**, 1217–1229, doi:10.1175/2009JAMC2088.1.
- , and D. B. Kluver, 2009: Relationship between winter/spring snowfall and summer precipitation in the northern Great Plains of North America. *J. Hydrometeor.*, **10**, 1203–1217, doi:10.1175/2009JHM1089.1.
- Rehfeldt, G. E., 2006: A spline model of climate for the western United States. RMRS General Tech. Rep. RMRS-GTR-165, USDA Forest Service, 21 pp.
- Rogers, J. C., S. Wang, and J. S. M. Coleman, 2007: Evaluation of a long-term (1882–2005) equivalent temperature time series. *J. Climate*, **20**, 4476–4485, doi:10.1175/JCLI4265.1.
- Seager, R., A. Tzanova, and J. Nakamura, 2009: Drought in the southeastern United States: Causes, variability over the last millennium, and the potential for future hydroclimate change. *J. Climate*, **22**, 5021–5045, doi:10.1175/2009JCLI2683.1.
- Sharples, J. J., M. F. Hutchinson, and D. R. Jellet, 2005: On the horizontal scale of elevation dependence of Australian monthly precipitation. *J. Appl. Meteor. Climatol.*, **44**, 1850–1865, doi:10.1175/JAM2289.1.
- Silverman, N. L., M. P. Maneta, S. H. Chen, and J. T. Harper, 2013: Dynamically downscaled winter precipitation over complex terrain of the Central Rockies of western Montana, USA. *Water Resour. Res.*, **49**, 458–470, doi:10.1029/2012WR012874.
- Stahle, D. W., and Coauthors, 2009: Cool- and warm-season precipitation reconstructions over western New Mexico. *J. Climate*, **22**, 3729–3750, doi:10.1175/2008JCLI2752.1.
- USGS, cited 2013: GTOPO30 Global Digital Elevation Model. EROS Data Center, U.S. Geological Survey, Sioux Falls, SD. [Available online at <http://www.temis.nl/data/gtopo30.html>.]
- Vose, R. S., S. Applequist, M. J. Menne, and C. N. Williams Jr., 2012: An intercomparison of temperature trends in the U.S. Historical Climatology Network and recent atmospheric reanalyses. *Geophys. Res. Lett.*, **39**, L10703, doi:10.1029/2012GL051387.
- Wahba, G., 1990: *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics, Book 59, Society for Industrial and Applied Mathematics, 180 pp.
- Williams, C. N., M. J. Menne, and P. W. Thorne, 2012: Benchmarking the performance of pairwise homogenization of surface temperatures in the United States. *J. Geophys. Res.*, **117**, doi:10.1029/2011JD016761.
- Willmott, C. J., and S. M. Robeson, 1995: Climatologically aided interpolation (CAI) of terrestrial air temperature. *Int. J. Climatol.*, **15**, 221–229, doi:10.1002/joc.3370150207.
- , C. M. Rowe, and W. D. Philpot, 1985: Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *Amer. Cartogr.*, **12**, 5–16, doi:10.1559/152304085783914686.