

Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods

E. P. Maurer¹ and H. G. Hidalgo²

¹Civil Engineering Department, Santa Clara University, Santa Clara, CA, USA

²Scripps Institution of Oceanography, Div. of Climate Atmospheric Sciences & Physical Oceanography, La Jolla, CA, USA

Received: 16 August 2007 – Published in Hydrol. Earth Syst. Sci. Discuss.: 12 September 2007

Revised: 20 December 2007 – Accepted: 5 February 2008 – Published: 13 March 2008

Abstract. Downscaling of climate model data is essential to local and regional impact analysis. We compare two methods of statistical downscaling to produce continuous, gridded time series of precipitation and surface air temperature at a 1/8-degree (approximately 140 km² per grid cell) resolution over the western U.S. We use NCEP/NCAR Reanalysis data from 1950–1999 as a surrogate General Circulation Model (GCM). The two methods included are constructed analogues (CA) and a bias correction and spatial downscaling (BCSD), both of which have been shown to be skillful in different settings, and BCSD has been used extensively in hydrologic impact analysis. Both methods use the coarse scale Reanalysis fields of precipitation and temperature as predictors of the corresponding fine scale fields. CA downscales daily large-scale data directly and BCSD downscales monthly data, with a random resampling technique to generate daily values. The methods produce generally comparable skill in producing downscaled, gridded fields of precipitation and temperatures at a monthly and seasonal level. For daily precipitation, both methods exhibit limited skill in reproducing both observed wet and dry extremes and the difference between the methods is not significant, reflecting the general low skill in daily precipitation variability in the reanalysis data. For low temperature extremes, the CA method produces greater downscaling skill than BCSD for fall and winter seasons. For high temperature extremes, CA demonstrates higher skill than BCSD in summer. We find that the choice of most appropriate downscaling technique depends on the variables, seasons, and regions of interest, on the availability of daily data, and whether the day to day correspondence of weather from the GCM needs to be repro-

duced for some applications. The ability to produce skillful downscaled daily data depends primarily on the ability of the climate model to show daily skill.

1 Introduction

Climate models are the primary tool to evaluate the projected future response of the atmosphere-land-ocean system to changing atmospheric composition (MacCracken et al., 2003; Stocker et al., 2001), and they underpin most climate change impacts studies (Wilby and Harris, 2006). However there is a mismatch between the grid resolution of current climate models (generally hundreds of kilometers), and the resolution needed by environmental impacts models (typically ten kilometers or less). Downscaling is the process of transforming information from climate models at coarse resolutions to a fine spatial resolution. Downscaling is necessary, as the underlying processes described by the environmental impact models are very sensitive to the nuances of local climate, and the drivers of local climate variations, such as topography, are not captured at coarse scales.

There are two broad categories of downscaling: dynamic (which simulates physical processes at fine scales) and statistical (which transforms coarse-scale climate projections to a finer scale based on observed relationships between the climate at the two spatial resolutions) (Christensen et al., 2007). Dynamic downscaling, nesting a fine scale climate model in a coarse scale model, produces spatially complete fields of climate variables, thus preserving some spatial correlation as well as physically plausible relationships between variables. However, dynamic downscaling is very computationally intensive, making its use in impact studies limited, and essentially impossible for multi-decade simulations with



Correspondence to: E. P. Maurer
(emaurer@enr.scu.edu)

different global climate models and/or multiple greenhouse gas emission scenarios. Thus, most impacts studies rely on some form of statistical downscaling, where variables of interest can be downscaled using historical observations. There has been extensive work developing and intercomparing statistical downscaling techniques for climate impact studies (Giorgi et al., 2001; Wilby and Wigley, 1997).

Statistical downscaling is typically used to predict one variable at one site, though some techniques for simultaneous downscaling to multiple sites for precipitation have been developed (Harpham and Wilby, 2005; Wilks, 1999). However, for studies of some climate impacts such as river basin hydrology, it is important to downscale simultaneous values of multiple variables (such as precipitation and temperature) over large, heterogeneous areas, while maintaining physically plausible spatial and temporal relationships, though few downscaling techniques have been developed to do this.

In this study we compare two methods of statistical downscaling to produce gridded time series of precipitation and surface air temperature at a fine resolution over a large spatial domain. These two methods are termed constructed analogues (CA, Hidalgo et al., 2008; van den Dool, 1994) and bias correction and spatial downscaling (BCSD, Wood et al., 2004). The CA method has been shown to have significant skill in reproducing the variability of daily precipitation and temperature over the contiguous United States (U.S.), in particular in the western coast (Hidalgo et al., 2008). The BCSD method has been shown to provide downscaling capabilities comparable to other statistical and dynamical methods in the context of hydrologic impacts (Wood et al., 2004).

The main conceptual difference between the two methods compared here is that the daily correspondence of the coarse resolution and the fine resolution patterns is maintained in the CA method, while in the BCSD the monthly patterns are conserved but daily (intra-month) patterns are resampled randomly, and therefore the correspondence between downscaled and historic daily observations is not conserved. In this way, CA is designed to use the simulated daily sequences from a climate model (at a coarse spatial resolution) and downscales each simulated day, while BCSD downscales monthly simulated climate model output and randomly generates daily sequences to match the monthly values. While randomly resampling daily sequencing within a month has been shown to have a negligible impact for monthly and seasonal river basin hydrologic statistics (Wood et al., 2002), for impacts related to shorter-term extremes (e.g. heat waves, air quality episodes, flood peaks), changes in daily sequencing may be important. Where a climate model exhibits skill in simulating daily variability, CA would capture that skill, while BCSD would reflect climatological intra-month variability. Thus, for daily statistics, the two methods will be expected to distinguish themselves only inasmuch as the large-scale climate model exhibits skill at the daily time scale. By using these two techniques, and by using a Reanalysis as a surrogate for a best-case (unbiased) GCM, we aim in

this study to explore the potential gain in downscaled daily skill by downscaling daily, as opposed to monthly, GCM output. While there are other candidate methods that could have been utilized in an intercomparison exercise, these two methods were selected since they have been implemented over the Western United States, and will be included in future work on climate change impacts in California (Cayan et al., 2008).

2 Data sources and methods

2.1 Data sources

Daily precipitation, maximum and minimum temperature at $1/8^\circ$ resolution (approximately 140 km^2 per grid cell) were obtained from the University of Washington Land Surface Hydrology Research group (<http://www.hydro.washington.edu>), the development of which is described in Maurer et al. (2002). The data are daily station observations interpolated onto a regular grid, with temperature lapsed to grid cell center elevations and with precipitation adjusted for compatibility with the Parameter-elevation Regressions on Independent Slopes (PRISM, Daly et al., 1994) dataset, following Widmann and Bretherton (2000). This constitutes the main dataset used in the calibration and evaluation of the performance of the downscaling processes in this study.

We use the National Center of Environmental Prediction and the National Center of Atmospheric Research (NCEP/NCAR) reanalysis (hereinafter reanalysis, Kalnay et al., 1996) as a surrogate for a General Circulation Model (GCM), similar to the application of Widmann et al. (2003), which is then downscaled and compared to observations. Reanalysis data are available on a T62 Gaussian grid (approximately 1.9° square), a resolution comparable to current generation of GCMs. A favorable characteristic of reanalysis data is the availability of daily precipitation and temperature data, which is often not archived for long, climate change simulations by modeling groups.

While the objective of these methods is to downscale actual GCM output, intercomparing these two techniques using GCM output would be problematic since the skill assessment could not distinguish between biases related to downscaling and biases related to GCM process simulation. Thus, using reanalysis data provides a more objective measure of the skill of the downscaling techniques.

In the NCEP/NCAR reanalysis, observations of air temperature are assimilated into the model, resulting in reanalysis temperatures that are close to observations. Precipitation, on the other hand, does not benefit from assimilation of observations, and is a product of the reanalysis model, and can thus exhibit substantial regional biases (Maurer et al., 2001; Widmann and Bretherton, 2000; Wilby et al., 2000). Arguably, due to the assimilation of atmospheric observations, reanalysis represents the best possible simulation capability of a GCM. Because reanalysis temperature is strongly

connected to observations, the comparisons of temperature skill will reflect differences almost exclusively in the downscaling techniques; because reanalysis precipitation is not related to observations, the intercomparison will reflect differences between the downscaling methods, plus some influence of the reanalysis precipitation bias. The precipitation and temperature daily variability in the reanalysis has been shown to be plausible in some locations in the Western U.S. (Widmann and Bretherton, 2000), and the existence of skill in daily statistics of GCM output will be a major factor distinguishing the downscaling methods compared in this study.

2.2 “Observed” and “projected” time period definitions

We used 1950–1976 reanalysis precipitation and temperature as the period representing the “observations,” and 1977–1999 as “projections,” similar to past studies (e.g. Salathé, 2003; Wilby et al., 2000). These two periods have differing characteristics, with the second period reflecting the temperature increase of recent decades, as well as a phase shift in the Pacific Decadal Oscillation (PDO, Mantua et al., 1997) from cool phase (through 1976) to warm phase (1977 through at least the mid-1990s) (Mantua and Hare, 2002). The PDO influences North American climate in a similar manner to the El Niño Southern Oscillation (ENSO), with discernable differences in precipitation, temperature, and streamflow, under different PDO phases (Cayan, 1996; Hamlet and Lettenmaier, 1999). By contrast with ENSO, however, PDO persists for decades. The magnitude of observed warming trends in the Western U.S. of 1–3°C over the second half of the 20th century are non-uniform through the region and are not fully explained by the PDO shift (Stewart et al., 2005). Precipitation trends over recent decades are even more non-uniform spatially and variable through time (Mote et al., 2005), though a large-scale wetting trend is evident through the last few decades of the 20th century (Groisman et al., 2004). For the spatial domain used in this study, the latter period is warmer by 0.2°C and wetter by 7%, with the means of the two periods differing with high confidence (>90%, based on a 1-tailed t-test). In this way, while not dramatically warmer, the period used as projections in this study serves as a proxy for a changed climate from the one used to train the downscaling methods.

2.3 Bias-correction and spatial downscaling (BCSD)

The bias-correction and spatial downscaling (BCSD) method of Wood et al. (2004) is an empirical statistical technique in which the monthly precipitation and temperature output from a GCM are downscaled. The method was originally developed for adjusting GCM output for long-range streamflow forecasting (Wood et al., 2002) and was later adapted for use in studies examining the hydrologic impacts of climate change in the Western U.S. (Christensen et al., 2004; Payne

et al., 2004; Van Rheenen et al., 2004). The technique uses a quantile-based mapping (Panofsky and Brier, 1968) of the probability density functions for the monthly GCM precipitation and temperature onto those of gridded observed data, spatially aggregated to the GCM scale. This same mapping is applied to future GCM projections. This allows the mean and variability of a GCM to evolve in accordance with the GCM simulation, while matching all statistical moments between the GCM and observations for the base period. This technique has compared favorably to different statistical and dynamic downscaling techniques (Wood et al., 2004) in the context of hydrologic impact studies. The method is computationally efficient and has thus been applied to studies downscaling multiple, extended GCM simulations for hydrologic and other impact studies (Cayan et al., 2008; Hayhoe et al., 2004; Maurer and Duffy, 2005; Maurer, 2007). Recently 112 150-year GCM projections were downscaled over much of North America using the BCSD method (Maurer et al., 2007), demonstrating its applicability to downscaling large ensembles of GCM simulations.

To recover daily values historical months are selected at random and each day in the selected month is rescaled identically (using a multiplicative factor for precipitation and an additive factor for temperature) to match the projected monthly total precipitation and average temperature. Following Wood et al. (2004), screening of scale values for precipitation were applied to avoid unreasonable precipitation values. For example, if the randomly selected month has two or fewer days with precipitation and the scale factor is greater than three, another year with more wet days is randomly selected to avoid creating unreasonably intense precipitation. As discussed by Wood et al. (2002), while in principle daily GCM output could be used directly, the BCSD method was developed to streamline the translation of GCM output into fine-scale land surface meteorological forcing, and the assumption of climatological daily variability has proven reasonable for hydrological applications. In this way the BCSD method, as applied in this study, does not account for changes in the statistics of climate variability at scales less than monthly that may be projected by a GCM, and is not expected to exhibit skill at projecting statistics of daily extremes above simply assuming climatological daily variability. In other, more spatially limited settings, adjusting the random selection of the historic sequence used in rescaling based on climate similarity has been used (Salathé, 2005). However, applying that conditioning technique requires the ability to characterize the entire domain by mean monthly precipitation, which is only possible on much smaller domains than that used in this study.

This method is in essence a model output statistics (MOS) approach (Glahn and Lowry, 1972; Wilks, 2006) at the monthly level, providing a post-processing of model output to correct for model biases relative to observations, where biases may due, for example, to imperfect model parameterization of physical processes or inadequate topographic

description in the model. The types of biases that can be accommodated by the BCSD approach are restricted by the method's use of large scale precipitation and temperature, as opposed to, for example, remote predictors such as geopotential heights that could in theory be used to correct for model biases in large scale circulation. However, Widmann et al. (2003) note that using local predictors such as precipitation can perform adequately compared to other downscaling methods. When applied to GCM output in a climate change context, the BCSD method, as with any MOS-type approach, is model dependent and assumes that the relation between model scale and disaggregated, fine scale climate is constant; the method will be less effective where the relationship changes.

2.4 Constructed analogues

The CA method, described in detail by Hidalgo et al. (2008), is essentially a “perfect prog” method (Glahn and Lowry, 1972; Klein et al., 1959), and is thus fundamentally different from the MOS-type BCSD approach (Widmann et al., 2003; Wilks, 2006). In particular, CA, as with any perfect prog approach, makes no adjustment to correct for biases, but rather relates model-simulated variables (in this application, anomalies of daily precipitation and temperature) to predicted variables, using relationships established with observations. As discussed in Sect. 2.1 above, this distinction bears on the interpretation of the intercomparison of CA with BCSD, which statistically corrects for model biases, since where Reanalysis exhibits substantial biases, attributing the differences to these biases or downscaling approaches may be problematic. Since CA is based not on absolute simulated values but on anomalies, the primary biases that would affect the downscaling results would be related to spatial and temporal variability, for which no adjustment is made to account for differences between the simulated (reanalysis) and observed values.

The application of the CA technique begins with a library of observed coarse-resolution and corresponding high-resolution climate anomaly patterns, where the library contains the variable to be downscaled. During development of the CA method, many candidate predictors were assessed (including geopotential heights of 700 and 500 hPa levels, 500 hPa divergence, sea level pressure, precipitation and temperature), and the greatest skill was found using precipitation and temperature as predictors, despite the documented biases in reanalysis precipitation. Thus, for downscaling precipitation (temperature) the library is composed of precipitation (temperature) anomaly patterns at both coarse and fine scales. In our case the fine-resolution ($1/8^\circ$) library patterns are based on the daily climate patterns from the Maurer et al. (2002) data from 1950–1976. The coarse-resolution library patterns are composed of the same data aggregated to the resolution of the Reanalysis (T62). Anomalies are used, rather than absolute values, to avoid concentrating patterns in

areas with high values, and to avoid the influence of absolute biases in large-scale simulations. Just as BCSD in principle could be applied on daily scales, CA could theoretically be applied at a monthly scale. However, CA was developed to enable to direct downscaling of daily large-scale model output to capture the evolution of daily precipitation and temperature statistics as simulated by a model. Furthermore, since constructing analogues requires a large library from which to draw similar patterns, as discussed below, an observational record longer than the 1950–1999 period used in this study would be needed to support a robust monthly application of CA.

To downscale a daily pattern from the 1977–1999 Reanalysis (the “target pattern”) a subset of 30 patterns (predictors) are found from the library that show similitude at the coarse-resolution with the pattern to be downscaled (based on the pattern root mean square error, RMSE). A linear combination of the *coarse-resolution* version of the predictors is used to produce a linear estimate of the target pattern, called the constructed analogue (which is at the coarse-resolution scale, T62 in this case). The spatially downscaled estimate pattern ($1/8^\circ$) is then created by applying the same regression coefficients from the construction of the analogue, to the *high-resolution* versions of the predictors.

The estimation of the target pattern was constructed using as predictors the “best” 30 predictors (based on the pattern RMSE) selected from a window of potential patterns within ± 45 days from the target date. Precipitation was transformed by taking the square root to make its statistical distribution more Gaussian (Box and Cox, 1964; Dettinger et al., 2004; Panofsky and Brier, 1968).

Mathematically, for each day and variable (temperature and precipitation) to be downscaled, if we define $Z_{\text{analogues}}$ as the matrix of 30 best predictors from the 1950–1976 library at the coarse resolution (predictors) and define Z_{obs} as the target pattern, an analogue of that pattern (\hat{Z}_{obs}) can be constructed as a linear combination of the (preferred 30-member most-suitable subset of) predictor patterns, according to:

$$Z_{\text{obs}} \approx \hat{Z}_{\text{obs}} = Z_{\text{analogues}} A_{\text{analogues}} \quad (1)$$

where $Z_{\text{analogues}}$ is a matrix of the column vectors comprising the most-suitable subset of coarse-resolution patterns identified above specifically for Z_{obs} , and $A_{\text{analogues}}$ is a column vector of fitted least-squares estimates of the regression coefficients that are the linear proportions of the contributions of each column of $Z_{\text{analogues}}$ to the constructed analogue. The dimensions of the Z_{obs} matrix are $p_{\text{coarse}} \times 1$, where p_{coarse} is the number of considered gridpoints contained in each coarse-resolution weather pattern; that is, Z_{obs} is a column vector. The dimensions of $Z_{\text{analogues}}$ are $p_{\text{coarse}} \times n$, where n is the number of patterns in the most suitable predictors subset (i.e. 30). The dimension of $A_{\text{analogues}}$ is $n \times 1$. Assuming $Z_{\text{analogues}}$ has full rank (n) and using the definition

of the pseudo-inverse (Moore-Penrose inverse), $A_{\text{analogues}}$ is obtained from Eq. (1) by:

$$A_{\text{analogues}} = \left[\left(Z'_{\text{analogues}} Z_{\text{analogues}} \right)^{-1} Z'_{\text{analogues}} \right] Z_{\text{obs}} \quad (2)$$

where the ' superscript denotes the transpose of the matrix. The inversion of the matrix was performed using singular value decomposition routine (Press et al., 2007) in which small values of the decomposition were set equal to zero to avoid near-singular matrices.

To downscale the Z_{obs} pattern, the coefficients $A_{\text{analogues}}$ from Eq. (3) are applied to the high-resolution weather patterns corresponding to the same days as the coarse-resolution predictors $Z_{\text{analogues}}$, according to:

$$\hat{P}_{\text{downscaled}} = P_{\text{analogues}} A_{\text{analogues}} \quad (3)$$

From Eq. (2):

$$\hat{P}_{\text{downscaled}} = P_{\text{analogues}} \left[\left(Z'_{\text{analogues}} Z_{\text{analogues}} \right)^{-1} Z'_{\text{analogues}} \right] Z_{\text{obs}} \quad (4)$$

where $\hat{P}_{\text{downscaled}}$ is a constructed high-resolution analogue (e.g. a precipitation pattern on the $1/8^\circ$ grid) and $P_{\text{analogues}}$ is the set of high-resolution historical patterns corresponding to the same days as the $Z_{\text{analogues}}$. The dimension of the $\hat{P}_{\text{downscaled}}$ vector is $p_{\text{VIC}} \times 1$, and the dimension of the $P_{\text{analogues}}$ matrix is $p_{\text{VIC}} \times 1$, where p_{VIC} is the number of gridpoints in the high-resolution weather patterns. Note that the matrix, $Z'_{\text{analogues}} Z_{\text{analogues}}$, inverted with each application of the procedure is only of dimension $n \times n$, and therefore the computational resources needed to downscale the weather patterns are determined by the number of the patterns included in the most-suitable subset (in this case $n=30$) and can be quite small.

2.5 Comparison of methods

First, we assess the ability of the different methods to simulate average monthly precipitation and temperature. Second we compare both downscaling methods using metrics of daily P and temperature extremes.

2.5.1 Monthly and annual assessment

The monthly skill is characterized using correlations between the monthly averages of the downscaled estimates and the monthly averages of the Maurer et al. (2002) data. In addition, the biases in the climatological precipitation and temperature were computed. Scatter plots for different locations in California are also produced for assessment of the performance of the methods at point scales.

Table 1. Indices of daily extremes used in this study.

Index Name	Description
pq20	20th percentile of rainday amounts
pq90	90th percentile of rainday amounts
tavq10	10th percentile of daily average temperature
tavq90	90th percentile of daily average temperature
pxcdd	Maximum number of consecutive dry days
pxc wd	Maximum number of consecutive wet days

2.5.2 Comparison based on daily precipitation and temperature indices

To characterize precipitation and temperature at the daily scale, we use a subset of the indices that were developed as part of the Statistical and Regional dynamical Downscaling of Extremes for European regions (STARDEX) project, which provides standard diagnostics that have been used for the systematic inter-comparison of different downscaling methods (e.g., Harpham and Wilby, 2005; Haylock et al., 2006; Schmidli et al., 2006). Table 1 lists the indices used in this study, which are from the STARDEX framework, with the exception of the two temperature indices, which are tailored here to apply to daily average values. Statistics were computed on a seasonal (December–February; March–May; June–August; September–November) and annual level at each $1/8^\circ$ grid cell in the western United States. In computing the statistics (for the projection period of 1977–1999) for each grid cell, if fewer than 15 years were available for calculation of the statistic (such as many occurrences of zero precipitation amounts), that index was excluded for that grid cell.

Correlations were calculated for the years 1977–1999 between downscaled (CA or BCSD) and the gridded observed data (Maurer et al., 2002) for each statistic. Correlations are computed for daily data for each season. For plotting, the square of the correlation coefficient r^2 is used. To test the hypotheses that the correlation at each grid cell was zero, a Fisher's transform was applied to the Pearson correlation coefficients and a p -value (the probability that a non-zero correlation was reported when the downscaled and observed data are actually uncorrelated) was computed. A similar approach was used to test the hypotheses that the correlations produced by the two downscaling techniques are statistically the same.

3 Results and discussion

3.1 Monthly and seasonal skill

The correlation between monthly averages of downscaled precipitation and temperature and the Maurer et al. (2002)

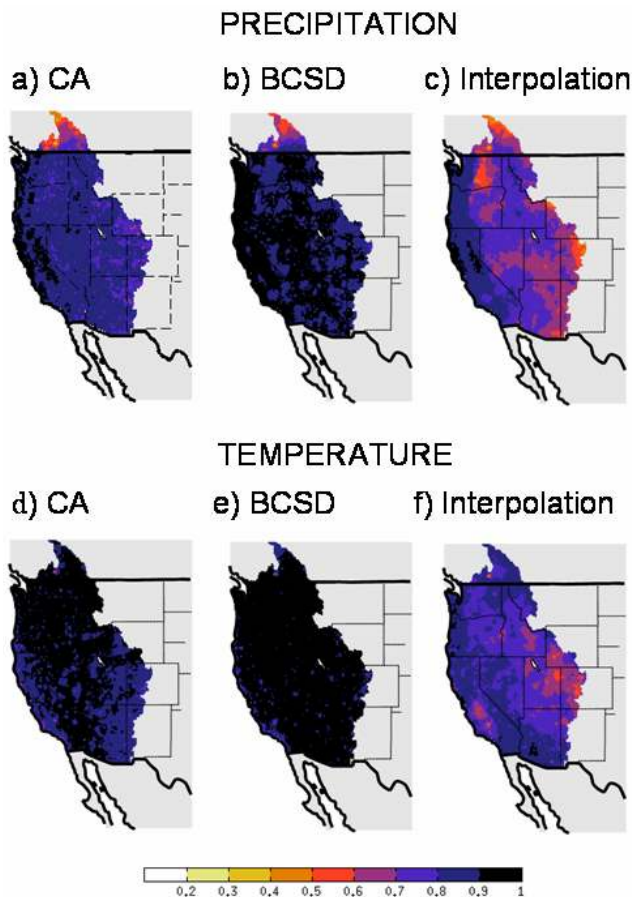


Fig. 1. Month to month correlation coefficients between observations from Maurer et al. (2002) dataset and downscaling estimates using three different approaches: the CA (left panel), the BCSD (middle panel) and a cubic interpolation of the large scale field.

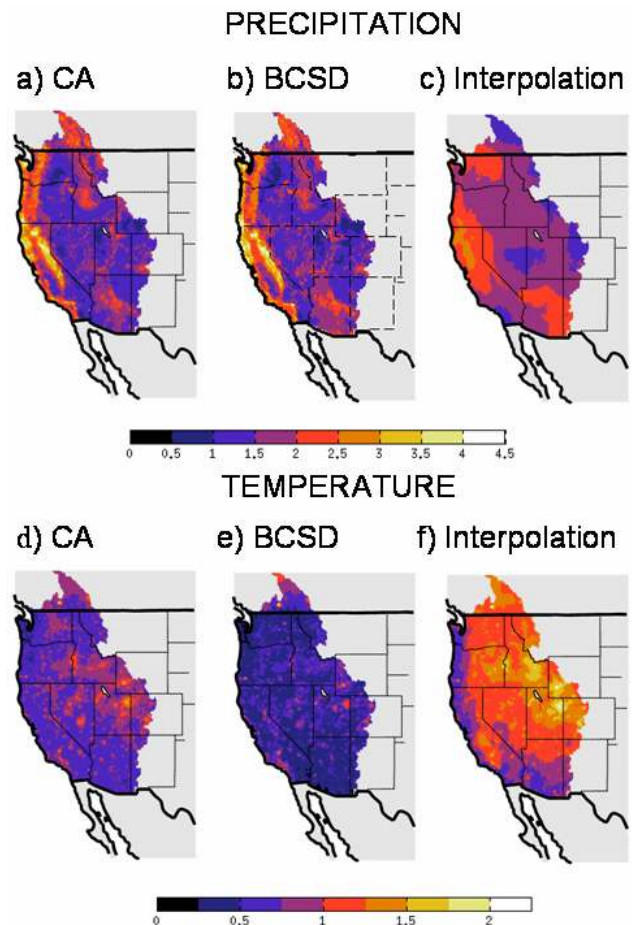


Fig. 2. Same as Fig. 1 but for the RMSE. Values in $(\text{mm day}^{-1})^{1/2}$ for precipitation and $^{\circ}\text{C}$ for temperature.

observations is shown in Fig. 1. An interpolation of the re-analysis data to the fine scale ($1/8^{\circ}$) grid is also shown as a reference or as a third “method” of downscaling the coarse scale data. For precipitation, the BCSD shows a larger area with very strong correlations, but the BCSD and CA downscaling methods are generally comparable when contrasted with the lower skill of the interpolated reanalysis. Figure 2 shows the root mean square error (RMSE) for the BCSD and CA are comparable for precipitation, and the BCSD method has lower RMSE over a larger region than CA for temperature. However, both methods exhibit much lower RMSE than the interpolated reanalysis, indicating that both downscaling methods provide substantial increases in skill at generating local climate features at the monthly scale.

Three points were selected to provide an example of the performance of the methods for different climate regimes: snow-controlled regions, snow-free regions and arid regions. For the Mojave Desert gridpoint, Fig. 3a shows that the correlations for precipitation and temperature are comparable for the CA and BCSD methods, with both methods

slightly underestimating precipitation and temperature but with high correlations. For the Madera gridpoint (Fig. 3b), located in California’s Central Valley, the CA method shows weaker correlations than the BCSD for precipitation and slightly stronger correlations for temperature, though again the two methods are very similar. For Madera and Yosemite (Fig. 3c), the CA method generally underestimates precipitation, while the BCSD generally underestimates temperature. Comparing the three sites, it appears that skillful downscaling to dry areas is more difficult than to wet locations.

A plot of the biases in precipitation and temperature can be found in Fig. 4. In general the precipitation biases are of similar magnitude for BCSD and CA, with larger biases occurring in similar locations for both methods (both generally along prominent mountain ranges), highlighting the difficulty in downscaling large-scale precipitation in areas of complex terrain. BCSD underestimates temperature to a greater degree than CA for the Upper Colorado River Basin, California’s San Joaquin Valley and the Canadian portion of the Columbia River Basin, though there is some

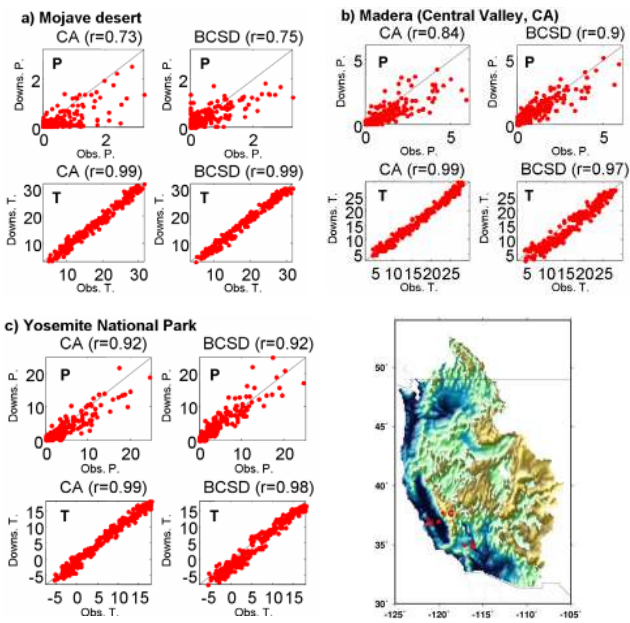


Fig. 3. Scatter plot of observed versus downscaled precipitation (P) and air temperature (T) for three grid cells. Locations of the three points are shown in the relief map in the lower right panel.

spatial correspondence in the regions with over- and under-estimation of temperature in both methods.

The overall similarity of skills between the CA and BCSD techniques for downscaled monthly precipitation and temperature should be interpreted in light of the distinctive properties of the methods. As noted above, the “perfect prog” CA approach makes no adjustment for biases in the large-scale spatial and temporal variability of reanalysis precipitation and temperature, while the “MOS” BCSD explicitly corrects any biases at the monthly level. Thus, the similarity in skills for downscaling monthly temperature and precipitation, where daily variability is not yet considered, indicates that the biases in reanalysis at this scale are not large enough to affect in a substantial way the downscaling skill.

Trends produced by a GCM are not explicitly corrected toward observations with either the CA or BCSD methods. Thus, as illustrated in Fig. 5, the large-scale temperature trends (which are not large for the projected period) are translated to the fine scale without generating fine-scale detail that may be present in the observations. BCSD, by extracting the temperature trend prior to bias-correction and replacing it afterward, exactly reproduces the large scale trends, while CA has a tendency to somewhat suppress them. The differences between large-scale trends simulated by reanalysis and observed station trends have been extensively explored (Kalnay and Cai, 2003; Kalnay et al., 2006). These differ due to many factors, notably because reanalysis does not reflect impacts of land use changes as well as other local and regional changes to clouds, snow, soil moisture, or instrumen-

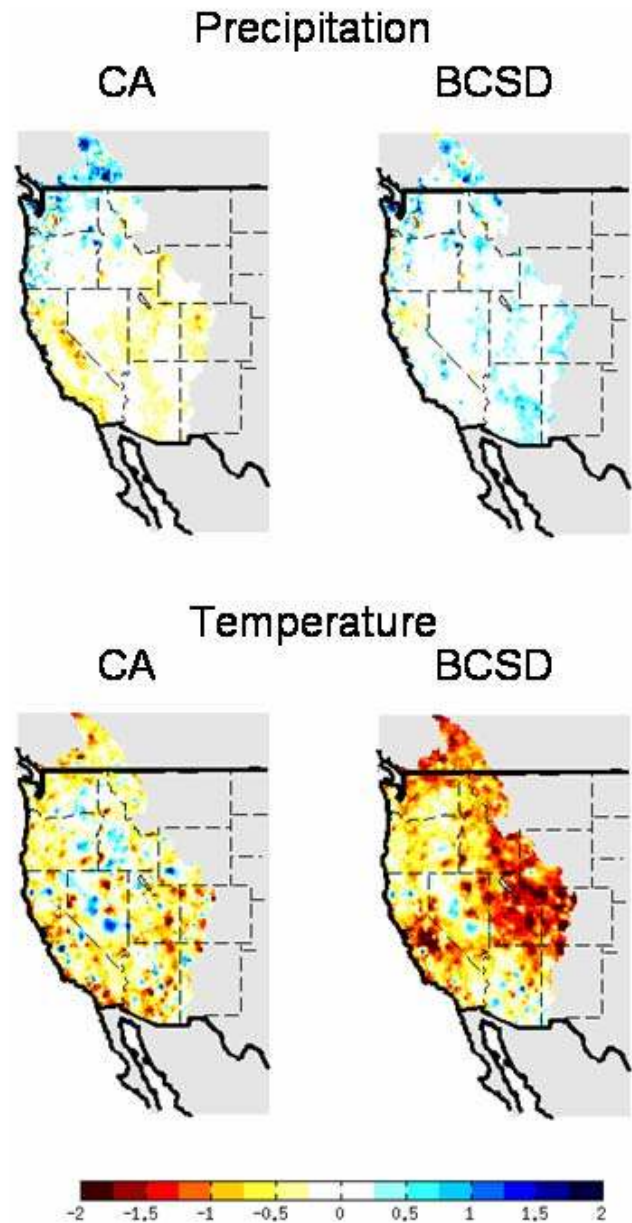


Fig. 4. Biases in mean precipitation (mm day^{-1}) and temperature ($^{\circ}\text{C}$) using CA and BCSD methods, based on monthly data for 1977–1999.

tal changes (Trenberth, 2004; Vose et al., 2004). Regardless, in general, trend simulation by a coupled GCM during the 20th century is not directly comparable to observed trends, since low-frequency natural oscillations can masquerade as trends (Knowles et al., 2006), and the phase of oscillations in an unconstrained GCM simulation would not be expected to mimic observations. Thus, correcting trends in a GCM toward observed trends would be a questionable practice.

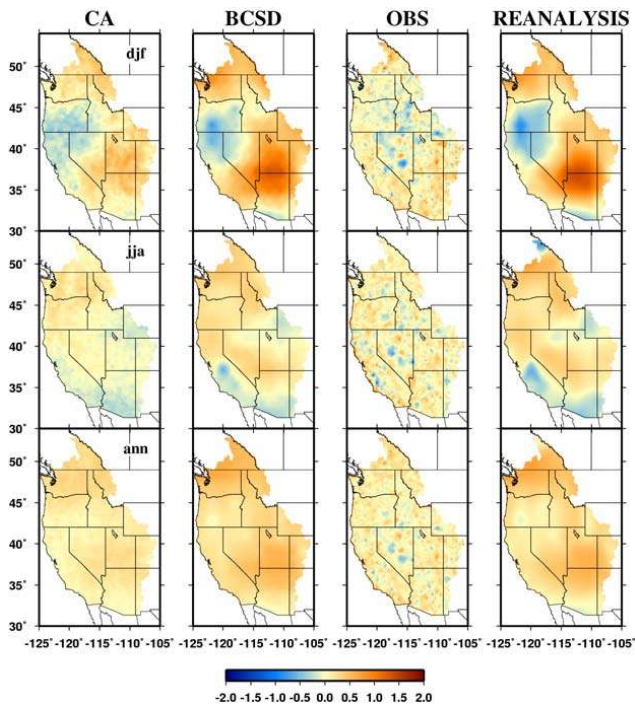


Fig. 5. Comparison of trends produced by downscaling with CA, BCSD, observed trends, and trends in the interpolated reanalysis data set. Values are in $^{\circ}\text{C}/\text{decade}$ for the projected period of 1977–1999.

3.2 Daily skill

There is only modest skill with either the CA and BCSD method for dry (20th percentile) daily precipitation extremes in winter (Fig. 6), and this limited skill is generally focused in coastal areas of the Pacific Northwest. Other seasons show lower skills, with much greater extent of area with insufficient data to calculate the statistics (as described in Sect. 2.5.2 above). There is no statistically confident difference between the methods for this measure. For wet (90th percentile) daily precipitation conditions both methods show some skill in winter, when most precipitation occurs (Fig. 7). As with Fig. 6, in general most of the domain has insufficient data for calculating this statistic during seasons other than winter. The CA method exhibits slightly higher correlations over certain regions such as the Sierra Nevada in California, but as with dry daily extremes, there is no statistically significant difference in the skills exhibited by the two methods. The lack of statistical significance to the differences suggests there is limited skill for extreme precipitation anomalies in the reanalysis, and neither method can recover daily skill. Where skill is exhibited, it is thus attributable to skill at longer time scales (monthly and seasonal), equivalently captured with both methods.

The stationarity of the transfer scheme is the key to the success of either the BCSD or CA method in translating

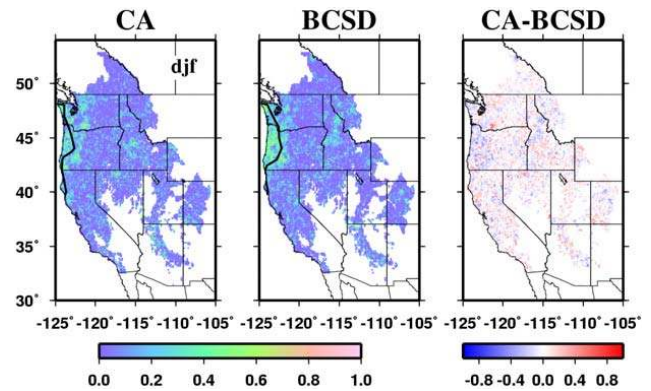


Fig. 6. r^2 values between observations and CA (left panel) and BCSD (center panel) for the 20th percentile (dry) daily precipitation statistic for winter season (as indicated in left panel). Right panel shows the difference between the two. The contour line delineates regions where the r^2 values achieve 90% confidence. Areas are absent if they have an inadequate number of years to compute the statistic.

large-scale climate to a fine scale. While the skill in simulating fine-scale dry or wet daily extremes (shown in Figs. 6 and 7) is generally low with either method, it is evident that wet extremes are captured more accurately for both methods, as also had been exhibited at the monthly level in Fig. 3. This suggests that the relationship established for the spatial transfer scheme, based on 1950–1976, holds to a greater degree under wet conditions than dry for the latter period of 1977–1999. One possible factor influencing such a change could be the PDO phase shift from negative in the earlier period to positive in the latter. The PDO has been shown to have a strong hydroclimatic teleconnection to the Pacific Northwest (Mantua and Hare, 2002), especially in its negative phase (Maurer et al., 2004), and the phase of PDO strongly influences the effect of ENSO on the region, amplifying dry extremes under a PDO warm phase (Hamlet and Lettenmaier, 1999). While many factors beside PDO influence this relationship, the differing skill in spatially downscaling precipitation under wet and dry extremes indicates that the transfer scheme may be disrupted more for dry extremes than wet for the periods used in this study.

In Fig. 8 the r^2 values between observations and the two downscaling methods for simulating the maximum number of consecutive dry days per season is shown. Most prominent is that, in contrast to the plots of precipitation skill, most of the domain (for both methods) shows significant r^2 values. The starkest difference is in winter, where in the southern half of the domain the CA downscaling technique better captures dry day sequences than the BCSD. In approximately 22% of the grid cells where CA has better skill than BCSD, the difference is also statistically significant ($p < 0.1$). For seasons other than winter, the CA technique still shows marginally better skill than BCSD overall, though there are

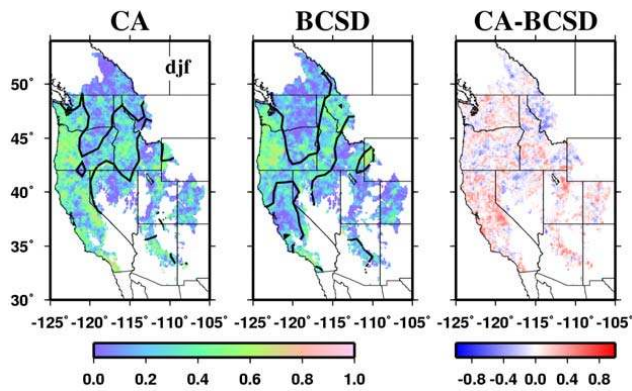


Fig. 7. Same as Fig. 6 but for the 90th percentile (wet) daily precipitation statistic.

more grid cells where BCSD shows better skill and the difference between the methods loses significance over the domain. Overall, at an annual level the two techniques are statistically indistinguishable, with only 5% of the grid cells showing differences in correlation between observations and each of the two methods that are statistically significant at the $p=0.1$ level, far fewer than would be expected by chance. This shows that temporal aggregation of daily extreme statistics can mask seasonal skill differences.

The skill of the methods at simulating the observed maximum number of consecutive wet days in each season, while not shown, has results similar to those of Fig. 8, where the highest skill and the greatest difference between the two methods is in winter, and in the Southern half of the domain. In winter, 23% of the grid cells exhibit statistically significant differences between the skill levels of the two methods. Again, at the annual level, the skill at reproducing observed patterns of maximum consecutive numbers of wet days is much less statistically distinguishable than in the Winter.

Many others have identified the connection of precipitation patterns to large-scale circulation patterns (e.g., Leung et al., 2003; Robertson and Ghil, 1999; Widmann et al., 2003); CA, because it utilizes domain-wide patterns can capture these relationships, while BCSD, defining only the spatial relationship between one large-scale grid box and the fine-scale grid boxes within it, is less able to. Reanalysis, due to its assimilation of observed atmospheric state variables, is known to represent better the large-scale atmospheric variables than derived quantities such as grid-cell precipitation (Widmann et al., 2003). Thus, the higher correlations exhibited by the CA method for dry- and wet-day sequencing in some locations may be a manifestation of the ability to relate these domain-wide patterns to local precipitation, at least to the extent that these effects are reflected accurately in the reanalysis.

In Fig. 9 the skill at reproducing extreme low temperature statistics, expressed as the 10th percentile daily temperature in each season is shown. In winter and fall, the CA method

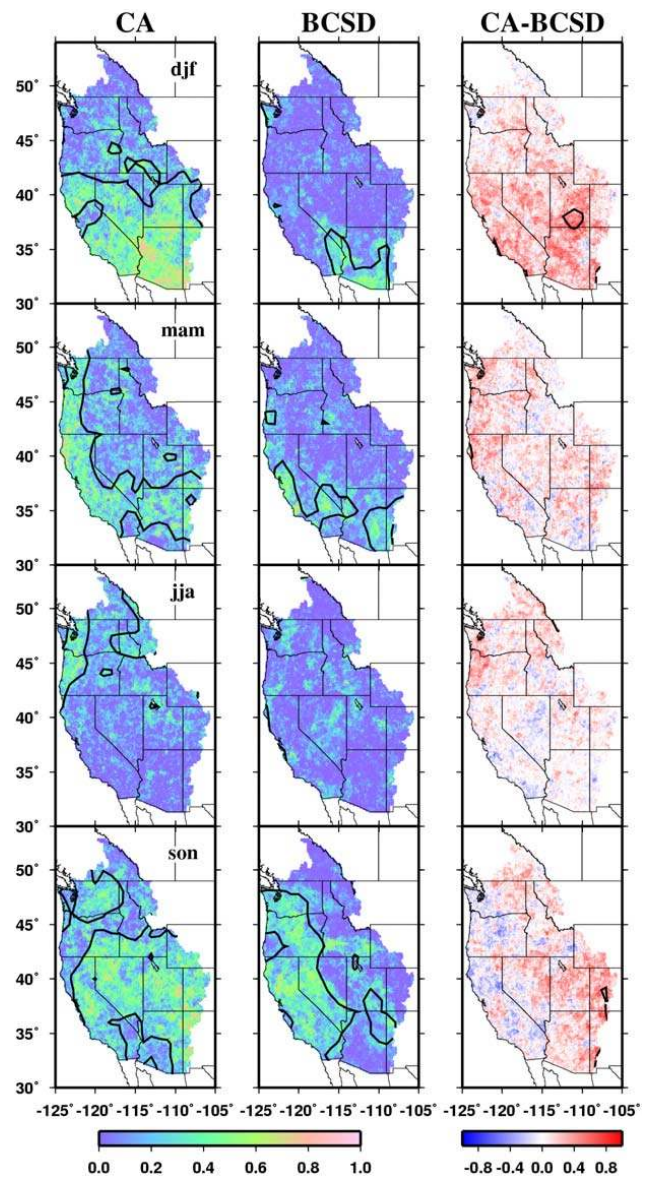


Fig. 8. Same as Fig. 6, but for maximum number of consecutive dry days per season. Each row corresponds to the season indicated in the right panel of that row.

has higher skill than BCSD, with 30% of the grid cells showing statistically significant differences between the methods. In the North, roughly corresponding to the Columbia River basin, the difference is most apparent. In this same region, however, the BCSD method shows greater skill in spring. Thus, the choice of most appropriate downscaling technique may depend not only on the statistic being analyzed, but also the region and season of focus.

In Fig. 10 the downscaling skill for reproducing observed daily warm anomalies, expressed as the 90th percentile temperature is illustrated. As was demonstrated above, the skill of the downscaling for daily temperature extremes exceeds

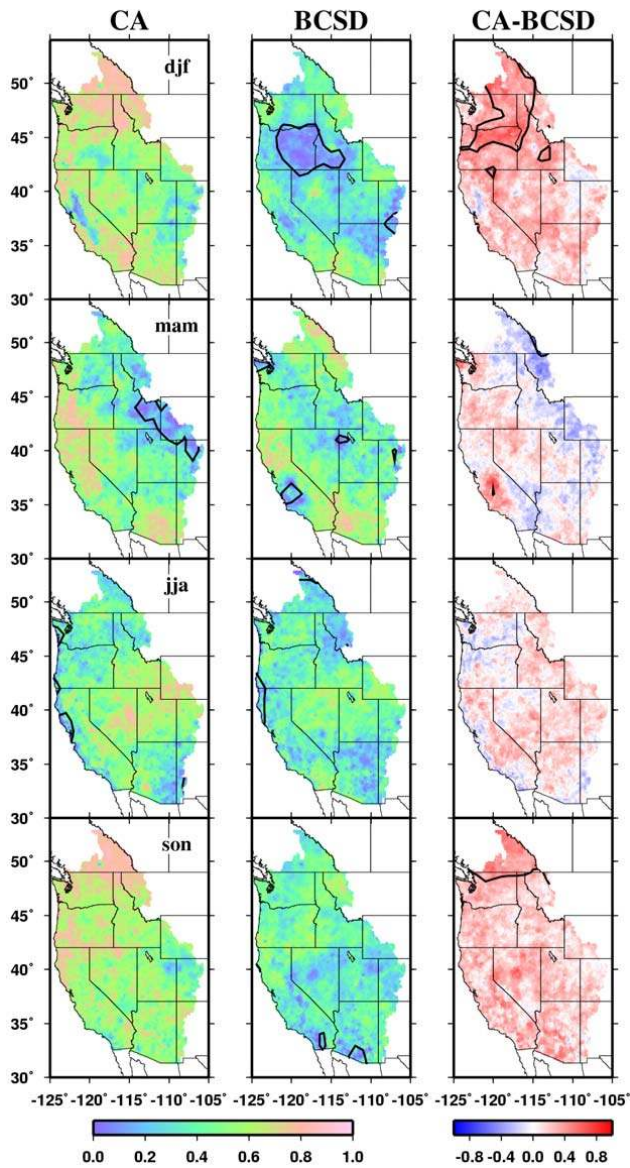


Fig. 9. Same as for Fig. 8, but for 10th percentile (cool) daily temperature in each season.

that for precipitation extremes. While the downscaling of average seasonal temperatures for the two downscaling methods was shown to be comparable, high temperature extremes are better simulated with the CA downscaling, most notably in Central California and the Great Basin in summer. For seasons other than summer, less than 7% of the grid cells showed differences between the skills of the two downscaling techniques, much less than would be expected by chance at 90% confidence, indicating the methods are statistically indistinguishable for these seasons.

As discussed in Sect. 2.1, since temperature observations are assimilated into reanalysis, there is much lower bias in reanalysis temperature than in reanalysis precipitation, a

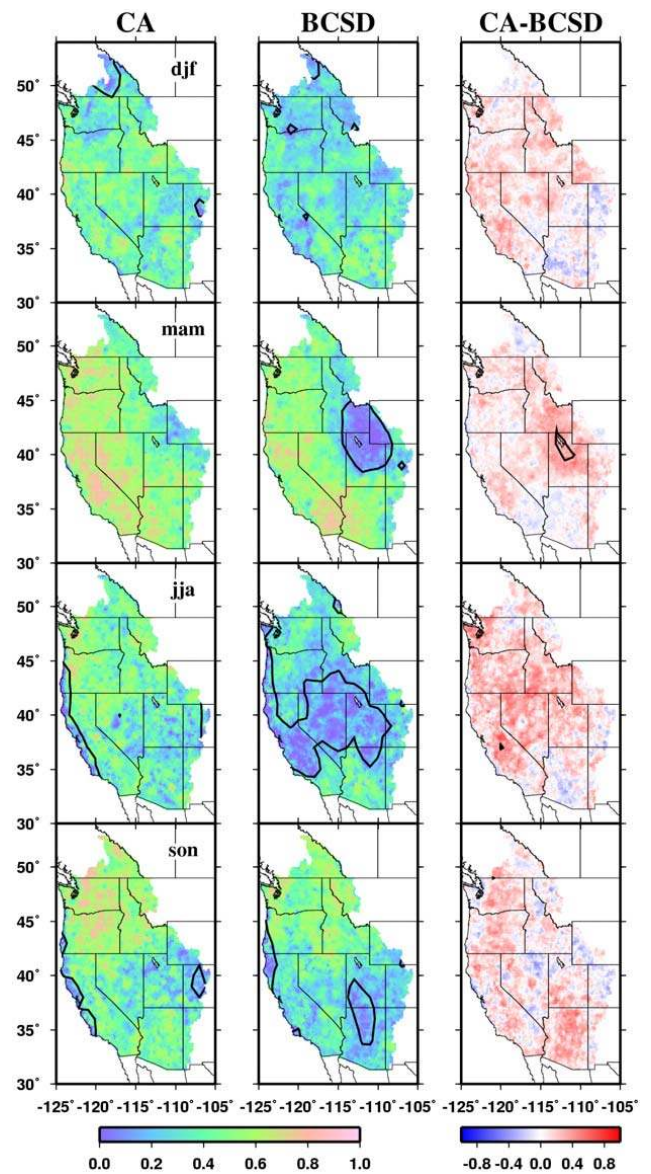


Fig. 10. Same as for Fig. 8, but for 90th percentile (warm) daily temperature in each season.

model-derived variable. Thus, downscaled daily temperature extremes, specifically high extremes in summer and low extremes in winter and fall, show improved skill with CA as compared to BCSD, while correlation at the monthly level was comparable (or even higher for BCSD) (Figs. 1 and 2), reflecting the translation of reanalysis daily skill to the fine-scale. In contrast, similarly to the monthly level, daily downscaled precipitation is generally indistinguishable between the two downscaling techniques, with the exception of some statistics related to sequencing of daily precipitation, showing that as large-scale daily simulations become less skillful, the downscaling based on daily model output will be less distinguishable from a random daily distribution.

It is interesting to note in Fig. 10 that the lack of significant skill with either method along large portions of the coast in the summer and fall. This shows that the assumptions of stationarity embedded in either statistical downscaling method (where large-scale weather patterns are related to historically observed fine scale observations) at the scale used in this study may not be valid for the coastal climate in this region, where local effects due to sea breeze and coastal upwelling affect extremes, and the relationship between large scale and fine scale climate may be changing (Lebassi et al., 2007).

4 Conclusions

At a monthly time scale, the two downscaling methods considered here, CA and BCSD, produce comparable skills in producing downscaled, gridded fields of precipitation and temperatures given coarse-scale reanalysis data as a surrogate GCM. The skill for temperature downscaling is considerably greater than that for precipitation, with precipitation showing much greater spatial variability in skill level.

Considering daily precipitation, both methods exhibit some skill in reproducing observed wet and dry extremes, generally in the Pacific Northwest, and the difference between the methods is not significant. This reflects the general low skill in daily precipitation variability in the large-scale reanalysis data over the domain, thus neither method can generate the skill absent in the large-scale signal. For reproducing fine scale observed consecutive sequences of wet and dry days, the CA method shows greater skill in winter in the Southwest, reflecting the presence of some reanalysis skill in simulating these daily precipitation characteristics. For other seasons and in other regions the methods are in general not statistically different.

The skill in downscaling daily temperature extremes exceeds that for precipitation extremes, which is not surprising given that temperature observations are assimilated in the reanalysis product, but precipitation is simulated. For low temperature extremes, the CA method produces greater downscaling skill than BCSD for fall and winter seasons. For high temperature extremes, CA demonstrates higher skill than BCSD in summer, though for other seasons differences are not significant. Contrasting the results from daily temperature and precipitation downscaling with these two methods, it is clear that as model (GCM) daily skill declines, the difference between using a downscaling technique based on daily model output versus applying a random daily distribution becomes less evident. As daily GCM skill, especially in regard to extreme events, continues to be assessed (Kharin et al., 2007; Tebaldi et al., 2006), the extent to which the daily GCM output exhibits skill in a region of interest will determine the utility of incorporating daily GCM output in a downscaling technique.

The choice of most appropriate downscaling technique depends in part on the variables, seasons, and regions of in-

terest. For precipitation, and impacts driven predominantly by precipitation, there is little distinction between the two methods, and the general lack of skill at a daily timescale in the large-scale reanalysis-simulated climate provides little incentive to favor either downscaling method. The presence of skill in the daily reanalysis temperature data allows the CA method to show superior skill compared to BCSD at reproducing local temperature extremes in some seasons and locations.

One drawback to using the CA technique with daily precipitation and temperature fields from a GCM is that the biases in the spatial and temporal variance exhibited by the GCM will be reconstructed in the downscaled fields. While the reanalysis data used here as a surrogate GCM can be considered a best possible GCM, since it assimilates observed temperature data, there are still substantial biases in some surface variables, and in particular for this study, precipitation. Actual GCM output reproduces extremes less reliably than reanalyses (Kharin et al., 2005), which will reduce the skill of the CA method, as applied here. Although the CA method works with anomalies and therefore biases in the mean of the GCM are not transferred to the fine scale results, some kind of bias correction is needed to remove biases in the variance of the GCM when the CA is to be applied to actual GCM data.

A limitation common to both methods is that the skill is tied to the simulated precipitation and temperature fields of the GCM, used as “predictors.” While these fields may be depicted less accurately than other potential predictor variables in the GCM (for example atmospheric circulation fields), using GCM precipitation has the advantage of capturing the complexity of physical processes (as represented in the GCM) producing precipitation in a way that using only large-scale circulation indices may not. These considerations are model dependent and should be kept in mind when downscaling data from actual GCM. Regardless of the technique, a final caveat is that of Charles et al. (1999), who noted the validation of a downscaling technique using historic data does not imply it will be equally valid under changed future climate conditions. While both techniques used in this study are shown to provide skill in downscaling, any future changes to the relationships between large scale and fine scale climate cannot be anticipated by them.

Acknowledgements. The authors are grateful to A. Wood at the University of Washington for support and advice related to this study, and to D. Cayan at Scripps Institution of Oceanography for helping to frame this effort. The manuscript benefited from the careful reading and insightful comments of three anonymous reviewers. The work of H. Hidalgo is supported by the California Energy Commission through the California Climate Change Center at Scripps Institution of Oceanography.

Edited by: L. Pfister

References

- Box, G. E. P. and Cox, D. R.: An analysis of transformations, *J. R. Stat. Soc.*, B26, 211–243, 1964.
- Cayan, D. R.: Interannual climate variability and snowpack in the western United States, *J. Climate*, 9, 928–948, 1996.
- Cayan, D. R., Maurer, E. P., Dettinger, M. D., Tyree, M., and Hayhoe, K.: Climate change scenarios for the California region, *Climatic Change*, 87, Suppl. 1, 21–42 doi: 10.1007/s10584-007-9377-6, 2008.
- Charles, S. P., Bates, B. C., Whetton, P. H., and Hughes, J. P.: Validation of downscaling models for changed climate conditions: case study of southwestern Australia, *Clim. Res.*, 12, 1–14, 1999.
- Christensen, J. H., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, I., Jones, R., Kolli, R. K., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C. G., Räisänen, J., Rinke, A., Sarr, A., and Whetton, P.: Regional Climate Projections, in: *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.
- Christensen, N. S., Wood, A. W., Voisin, N., Lettenmaier, D. P., and Palmer, R. N.: The effects of climate change on the hydrology and water resources of the Colorado River basin, *Climatic Change*, 62, 337–363, 2004.
- Daly, C., Neilson, R. P., and Phillips, D. L.: A statistical-topographic model for mapping climatological precipitation over mountainous terrain, *J. Appl. Meteorol.*, 33, 140–158, 1994.
- Dettinger, M. D., Cayan, D. R., Meyer, M., and Jeton, A. E.: Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basins, Sierra Nevada, California, 1900–2099, *Climatic Change*, 62, 283–317, 2004.
- Giorgi, F., Hewitson, B., Christensen, J., Hulme, M., von Storch, H., Whetton, P., Jones, R., Mearns, L., and Fu, C.: Regional Climate Information- Evaluation and Projections, in: *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Houghton, J. T., Cambridge University Press, Cambridge, UK, 583–638, 2001.
- Glahn, H. R. and Lowry, D. A.: The use of model output statistics (MOS) in objective weather forecasting, *J. Appl. Meteorol.*, 11, 1203–1211, 1972.
- Groisman, P. Y., Knight, R. W., Karl, T. R., Easterling, D. R., Sun, B. M., and Lawrimore, J. H.: Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations, *J. Hydrometeorology*, 5, 64–85, 2004.
- Hamlet, A. F. and Lettenmaier, D. P.: Columbia River streamflow forecasting based on ENSO and PDO climate signals, *J. Water Res. Plan. Mgmt.*, 125, 333–341, 1999.
- Harpham, C. and Wilby, R. L.: Multi-site downscaling of heavy daily precipitation occurrence and amounts, *J. Hydrol.*, 312, 235–255, 2005.
- Hayhoe, K., Cayan, D., Field, C. B., Frumhoff, P. C., Maurer, E. P., Miller, N. L., Moser, S. C., Schneider, S. H., Cahill, K. N., Cleland, E. E., Dale, L., Drapek, R., Hanemann, R. M., Kalkstein, L. S., Lenihan, J., Lunch, C. K., Neilson, R. P., Sheridan, S. C., and Verville, J. H.: Emissions pathways, climate change, and impacts on California, *Proc. National Academy Sci.*, 101, 12 422–12 427, 2004.
- Haylock, M. R., Cawley, G. C., Harpham, C., Wilby, R. L., and Goodess, C. M.: Downscaling heavy precipitation over the UK: a comparison of dynamical and statistical methods and their future scenarios, *Int. J. Climatol.*, 26, 1397–1415, 2006.
- Hidalgo, H. G., Dettinger, M. D., and Cayan, D. R.: Downscaling with constructed analogues: daily precipitation and temperature fields over the United States, Report No. CEC-500-2007-123, California Energy Commission, Sacramento, CA, 48 pp., 2008.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., and Reynolds, B.: The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Met. Soc.*, 77, 437–472, 1996.
- Kalnay, E. and Cai, M.: Impact of urbanization and land-use change on climate, *Nature*, 423, 528–531, 2003.
- Kalnay, E., Cai, M., Li, H., and Tobin, J.: Estimation of the impact of land-surface forcings on temperature trends in eastern United States, *J. Geophys. Res.*, 111, D06106, doi:10.1029/2005JD006555, 2006.
- Kharin, V. V., Zwiers, F. W., and Zhang, X.: Intercomparison of near-surface temperature and precipitation extremes in AMIP-2 simulations, reanalyses, and observations, *J. Climate*, 8, 5201–5223, 2005.
- Kharin, V. V., Zwiers, F. W., Zhang, X., and Hegerl, G. C.: Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, *J. Climate*, 20, 1419–1444, 2007.
- Klein, W. H., Lewis, B. M., and Enger, I.: Objective prediction of five-day mean temperature during winter, *J. Meteorol.*, 16, 672–682, 1959.
- Knowles, N., Dettinger, M., and Cayan, D.: Trends in snowfall versus rainfall for the Western United States, *J. Climate*, 19, 4545–4559, 2006.
- Lebassi, B., Gonzalez, J. E., Fabris, D., Maurer, E., Bornstein, R., and Miller, N. L.: Observed long-term California temperature-trends: coastal cooling and inland warming, *Proc. Am. Meteorol. Soc.*, 87th Annual Meeting, 13–18 January, San Antonio TX, 2007.
- Leung, L. R., Qian, Y., Bian, X. D., and Hunt, A.: Hydroclimate of the western United States based on observations and regional climate simulation of 1981–2000. part II: Mesoscale ENSO anomalies, *J. Climate*, 16, 1912–1928, 2003.
- MacCracken, M. C., Barron, E. J., Easterling, D. R., Felzer, B. S., and Karl, T. R.: Climate change scenarios for the U.S. national assessment, *Bull. Am. Met. Soc.*, 84, 1711–1723, doi:10.1175/BAMS-1784-1712-1711, 2003.
- Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R. C.: A Pacific interdecadal climate oscillation with impacts on salmon production, *Bull. Am. Met. Soc.*, 78, 1069–1079, 1997.
- Mantua, N. J. and Hare, S. R.: The Pacific Decadal Oscillation, *J. Oceanogr.*, 58, 35–44, 2002.
- Maurer, E. P., O'Donnell, G. M., Lettenmaier, D. P., and Roads, J. O.: Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalyses using an off-line hydrologic model, *J. Geophys. Res.*, 106, 17 841–17 862, 2001.
- Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P., and Nijssen, B.: A long-term hydrologically-based data set of land

- surface fluxes and states for the conterminous United States, *J. Climate*, 15, 3237–3251, 2002.
- Maurer, E. P., Lettenmaier, D. P., and Mantua, N. J.: Variability and predictability of North American runoff, *Water Resour. Res.*, 40, W09306, doi:10.1029/2003WR002789, 2004.
- Maurer, E. P. and Duffy, P. B.: Uncertainty in projections of streamflow changes due to climate change in California, *Geophys. Res. Lett.*, 32, doi:10.1029/2004GL021462, 2005.
- Maurer, E. P.: Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California under two emissions scenarios, *Climatic Change*, 82, 309–325, doi:10.1007/s10584-10006-19180-10589, 2007.
- Maurer, E. P., Brekke, L., Pruitt, T., and Duffy, P. B.: Fine-resolution climate change projections enhance regional climate change impact studies, *Eos, Transactions, American Geophysical Union*, 88, p. 504, 2007.
- Mote, P. W., Hamlet, A. F., Clark, M. P., and Lettenmaier, D. P.: Declining mountain snowpack in western North America, *Bull. Am. Met. Soc.*, 86, 39–49, 2005.
- Panofsky, H. A. and Brier, G. W.: *Some Applications of Statistics to Meteorology*, The Pennsylvania State University, University Park, PA, USA, 224 pp., 1968.
- Payne, J. T., Wood, A. W., Hamlet, A. F., Palmer, R. N., and Lettenmaier, D. P.: Mitigating the effects of climate change on the water resources of the Columbia River Basin, *Climatic Change*, 62, 233–256, 2004.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: *Numerical Recipes: The Art of Scientific Computing*, 3rd Edition, Cambridge University Press, New York, 1256 pp., 2007.
- Robertson, A. W. and Ghil, M.: Large-scale weather regimes and local climate over the Western United States, *J. Climate*, 12, 1796–1813, 1999.
- Salathé, E. P.: Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin, *Int. J. Climatol.*, 23, 887–901, 2003.
- Salathé, E. P.: Downscaling simulations of future global climate with application to hydrologic modelling, *Int. J. Climatol.*, 25, 419–436, 2005.
- Schmidli, J., Frei, C., and Vidale, P. L.: Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling, *Int. J. Climatol.*, 26, 679–689, 2006.
- Stewart, I. T., Cayan, D. R., and Dettinger, M. D.: Changes toward earlier streamflow timing across western North America, *J. Climate*, 18, 1136–1155, 2005.
- Stocker, T. F., Clarke, G. K. C., Le Treut, H., Lindzen, R. S., Meleshko, V. P., Mugara, R. K., Palmer, T. N., Pierrehumbert, R. T., Sellers, P. J., Trenberth, K. E., and Willebrand, J.: Physical climate processes and feedbacks, in: *Climate Change 2001: The Scientific Basis*, edited by: Houghton, J. T., Ding, Y., Griggs, D. G., et al., Cambridge University Press, Cambridge, 417–470, 2001.
- Tebaldi, C., Hayhoe, K., Arblaster, J. M., and Meehl, G. A.: An intercomparison of model-simulated historical and future changes in extreme events, *Climatic Change*, 79, 185–211, doi:10.1007/s10584-10006-19051-10584, 2006.
- Trenberth, K. E.: Rural land-use change and climate, *Nature*, 427, p. 213, 2004.
- van den Dool, H. M.: Searching for analogues, how long must one wait?, *Tellus*, 2A, 314–324, 1994.
- Van Rheezen, N. T., Wood, A. W., Palmer, R. N., and Lettenmaier, D. P.: Potential implications of PCM climate change scenarios for Sacramento-San Joaquin River Basin hydrology and water resources, *Climatic Change*, 62, 257–281, 2004.
- Vose, R. S., Karl, T. R., Easterling, D. R., Williams, C. N., and Menne, M. J.: Impact of land-use change on climate, *Nature*, 427, 213–214, 2004.
- Widmann, M. and Bretherton, C. S.: Validation of mesoscale precipitation in the NCEP reanalysis using a new grid-cell precipitation dataset for the Northwestern United States, *J. Climate*, 13, 1936–1950, 2000.
- Widmann, M., Bretherton, C. S., and Salathe, E. P.: Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor, *J. Climate*, 16, 799–816, 2003.
- Wilby, R. L. and Wigley, T. M. L.: Downscaling general circulation model output: a review of methods and limitations, *Prog. Phys. Geography*, 21, 530–548, 1997.
- Wilby, R. L., Hay, L. E., Gutowski, W. J., Arritt, R. W., Takle, E. S., Pan, Z., Leavesley, G. H., and Clark, M. P.: Hydrological responses to dynamically and statistically downscaled climate model output, *Geophys. Res. Lett.*, 27, 1199–1202, 2000.
- Wilby, R. L. and Harris, I.: A framework for assessing uncertainties in climate change impacts: low-flow scenarios for the River Thames, UK, *Water Resour. Res.*, 42, W02419, doi:10.1029/2005WR004065, 2006.
- Wilks, D. S.: Multisite downscaling of daily precipitation with a stochastic weather generator, *Clim. Res.*, 11, 125–136, 1999.
- Wilks, D. S.: *Statistical Methods in the Atmospheric Sciences*, 2 ed., Academic Press, New York, NY, USA, 627 pp., 2006.
- Wood, A. W., Maurer, E. P., Kumar, A., and Lettenmaier, D. P.: Long-range experimental hydrologic forecasting for the eastern United States, *J. Geophys. Res.*, 107, 4429, doi:10.1029/2001JD000659, 2002.
- Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Climatic Change*, 62, 189–216, 2004.