

Classification of Regional Climate Variability in the State of California

JOHN T. ABATZOGLOU, KELLY T. REDMOND, AND LAURA M. EDWARDS

Division of Atmospheric Sciences, Western Regional Climate Center, Desert Research Institute, Reno, Nevada

(Manuscript received 23 June 2008, in final form 15 January 2009)

ABSTRACT

A novel approach is presented to objectively identify regional patterns of climate variability within the state of California using principal component analysis on monthly precipitation and temperature data from a network of 195 climate stations statewide and an ancillary gridded database. The confluence of large-scale circulation patterns and the complex geography of the state result in 11 regional modes of climate variability within the state. A comparison between the station and gridded analyses reveals that finescale spatial resolution is needed to adequately capture regional modes in complex orographic and coastal settings. Objectively identified regions can be employed not only in tracking regional climate signatures, but also in improving the understanding of mechanisms behind regional climate variability and climate change. The analysis has been incorporated into an operational tool called the California Climate Tracker.

1. Introduction and motivation

A current goal of applied climate science is to improve knowledge at regional and local levels. The smaller the scale at which such information can be provided, the greater the relevance to users for most applications. The state of California provides a distinct challenge against efforts to describe, monitor, and explain its temporal and spatial climatic characteristics. Large-scale features of the general circulation, including the semipermanent subtropical high and the Aleutian low, dominate weather and climate patterns across California over the course of the seasonal cycle. In addition, the diverse geography of the state—complex topography, maritime influence, and time-varying land use patterns—contributes to an additional set of spatial–temporal physical controls on regional climate variability (Fig. 1). The climatic complexity within the state provides a suitable test bed for characterizing regional- and state-scale climate variability. California has made a substantial commitment to address climate change; however, objectively defined, easily interpretable, and widely used climate monitoring products to track change and variability have

not been heretofore available in the public arena. This study develops a novel approach to objectively define regional modes of climate variability as a means to not only create improved regional-scale climate monitoring products but also to improve our understanding of the processes that drive such patterns.

Although there exists a clear indication of change in the recent global surface temperature record, the regional manifestation of climate change is presently not well quantified. For the globe as a whole, the National Climatic Data Center (NCDC) has developed a land–air–sea surface reconstruction that provides a monthly temperature time series for the earth's surface (Smith and Reynolds 2005). At the much finer scale of a single station, climate variability and trends can also be analyzed, often with a number of caveats including climate inhomogeneities (e.g., Peterson et al. 1998) and the ability of station subsets with unique microclimates to misrepresent regional climate signals. The latter issue is likely to be highly important in the complex topography that defines the western United States.

Historically, seven climate divisions defined by NCDC (Guttman and Quayle 1996) have been heavily used to characterize regional climate in California. The delineation of climate division boundaries was a partly subjective process that involved alignment with major watersheds, agricultural administrative districts, geographic convenience, and expert judgment, with little guidance from

Corresponding author address: Dr. John T. Abatzoglou, Division of Atmospheric Sciences, 2215 Raggio Parkway, Reno, NV 89512-1095.

E-mail: john.abatzoglou@dri.edu

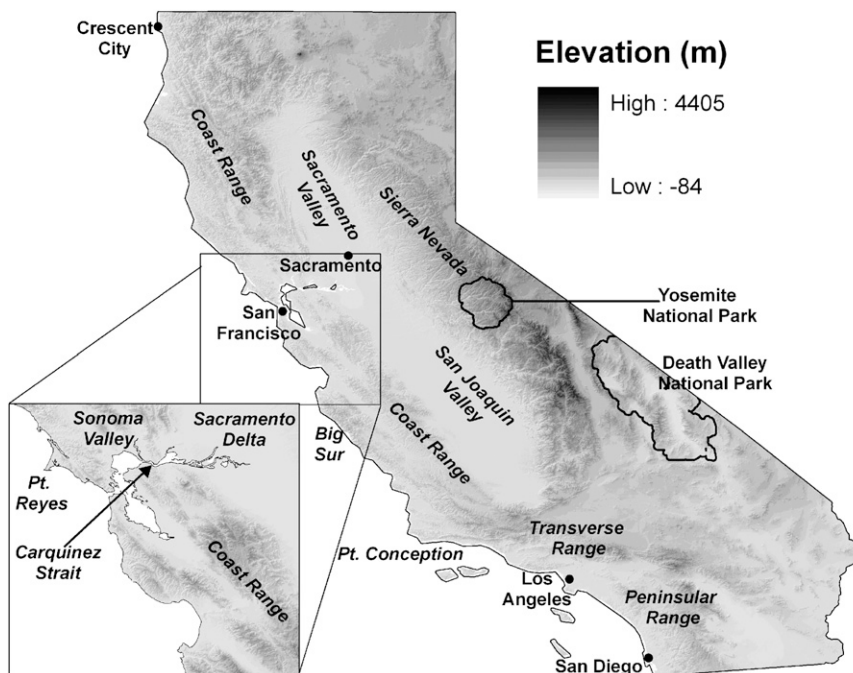


FIG. 1. Major terrain and geographic features of the state of California, and the San Francisco–Sacramento Delta (inset).

objective analyses reflecting physical processes that lead to covariability of clusters of stations. Some California climate divisions span an elevation range from sea level to over 4400 m (division outlines shown in Fig. 2), and others contain major internal differences in land use and maritime influence. Objective methods to characterize spatial–temporal patterns of climate variability have been demonstrated with promising results (Willmott 1977; Comrie and Glenn 1998; Wolter and Allured 2007). Fovell and Fovell (1993) used a clustering algorithm to divide the conterminous United States into climate regions using the NCDC climate division dataset. They noted that their clustering algorithm had the poorest performance within the state of California “due to the poor resolution of the NCDC (divisional) dataset in this region.”

The applications for aggregated regionalized data are extremely diverse, spanning scientific and utilitarian needs, and experience over many years has shown that some amount of compromise among competing considerations is unavoidable. This can be recognized by considering that the basis for practical regionalized datasets suitable for hydrologic planning may differ from those suitable for electrical demand planning, and that objective regionalized datasets suitable for characterizing temperature patterns may differ from those suitable for characterizing precipitation patterns. While the final choice of regions depends on the intended

application, the major purpose of this analysis is to maximize the objective contribution to this process by putting forth a methodology to objectively identify regional modes of variability in a multivariate dataset. The goal of the regionalization process is not to group areas with similar climatic means, but rather areas that covary together in time. The larger goal of this effort is to develop a tool to track spatial and temporal variability in climate in a manner that 1) is understandable to the average person, 2) is logistically and practically computable, 3) reflects physical drivers and correlation structures of the regional climate, including ties to continental and hemispheric scales, and 4) offers potential insights into climate process explanations of the observed temporal behavior and spatial variability.

Section 2 discusses datasets used in the climate regionalization process, and issues of preprocessing monthly datasets, including identification and adjustment for climate inhomogeneities. The techniques used to obtain variability patterns are addressed in section 3. Results of the regionalization process are discussed in section 4, and the robustness of these results is tested using a variety of methods. Section 5 presents the regional datasets for California and highlights an example that illustrates its utility in understanding the mechanisms that control regional climate variability. The final section provides concluding thoughts.

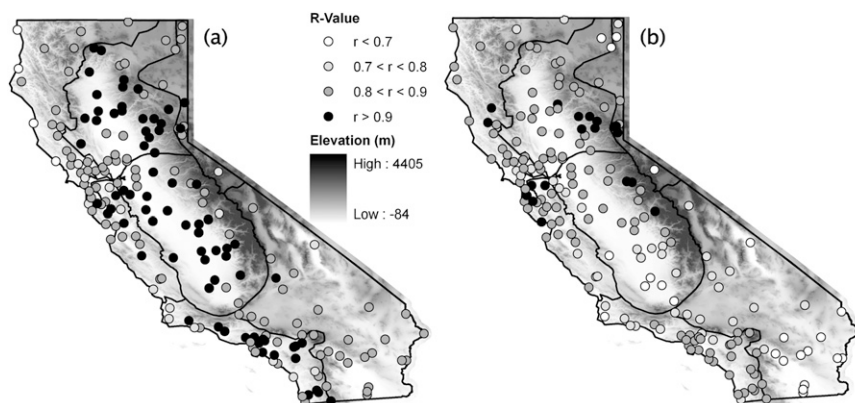


FIG. 2. Average monthly correlation of (a) standardized mean temperature, and (b) standardized precipitation between COOP stations and their respective NCDC climate divisions (outlined) for the time period 1948–2005.

2. Data

Station data for this study are from the NCDC Summary of the Day (SOD) database of daily observations from the National Weather Service (NWS) cooperative observer network [COOP; for a description see, e.g., NRC (1998)]. The analysis is restricted to temperature and precipitation. These values are aggregated for each individual month in the record to form time series of mean temperatures and total precipitation for each station. Mean monthly maximum and minimum temperatures are treated as separate elements, and mean monthly temperature is created from their arithmetic average. Months with more than five missing days are excluded and the data is set to missing. Records must be sufficiently complete (see below), and stations must be in operation as of January 2006 to meet operational needs. Records from more than 600 candidate COOP stations in California were examined. A total of 195 stations statewide that met the following criteria were employed in the subsequent analysis:

- (i) station reports daily maximum and minimum temperature,
- (ii) station reports daily precipitation accumulation,
- (iii) complete original data exist in more than 75% of all months from January 1949 to December 2005, and
- (iv) station was in operation and reported data during January 2006.

Changes in location, observer, equipment, observational methodology, and other factors can cause temporal climate inhomogeneities in station records. Although archived SOD data have been subjected to quality control (e.g., Hubbard et al. 2007), the techniques historically employed inherently cannot detect inhomogeneities. Three preprocessing procedures, described next, were

applied to the data first to identify potential erroneous data, second to infill all missing values, and finally to identify and adjust for inhomogeneities. The latter process can be fraught with traps and a conservative approach is adopted.

a. Initial quality control

The goal of initial quality control is to identify outliers in the dataset that may bias the results of the regionalization process. The following steps were employed to screen COOP data separately for monthly mean maximum temperature, monthly mean minimum temperature, and monthly precipitation accumulation for erroneous data:

- 1) Monthly data are transformed into nondimensional standardized anomalies. Precipitation data can be highly skewed and in all cases are subjected to a cube-root transformation. All anomalies are computed with reference to the same period of record (1949–2005).
- 2) Outliers that fall more than two standard deviations from the distribution of the remaining stations within the state are marked as erroneous and set as missing. This step is iterated until the dataset for that month is void of outliers.

Prior experience with complex western topography has shown many cases of large but legitimate differences between closely spaced stations, associated with elevational and coastal gradients. Step 2 makes an intercomparison of all stations within the state, therein assuming that each station belongs to a group of stations that exhibit similar variability on a regional scale, a condition that should hold even for the sparsely sampled regions of the state. The initial quality control identified approximately 160 monthly values (0.04%) as outliers.

b. Infilling missing data

The statistical methods employed require a temporally complete dataset devoid of missing values at every station. Multiple linear regression is applied for infilling missing data (less than 4% of all data). This process is conducted separately for each station, element, and month. For each month and climate element, a correlation matrix is computed by taking the mean correlation using overlapping 11-yr moving windows. Prior experience has shown that interstation relationships can vary through time through both climatic and non-climatic means; the 11-yr window is intended to isolate abrupt changes (i.e., inhomogeneities) in such relations to no more than about one-decade duration. A set of four comparison stations is chosen for each candidate station based on the highest four correlations for that element–month combination. With real data, both comparison and candidate stations can have flaws; however, four stations provide ample opportunities to detect flawed comparison stations. Comparison stations for each climate element are reassessed on a monthly basis to account for seasonally varying relationships among stations. This is crucial, as such interstation relations can vary greatly (sometimes changing sign), depending on season, elevation, coastal proximity, land use, large-scale flow regime, and climate element, therein reflecting a wide variety in physical coupling mechanisms.

The top four monthly correlation values across all elements exceed $r = 0.65$ (median value), with correlations generally higher in regions of high station density and during the cool season (November–April). Each candidate station is examined for its distance from and elevational difference between each of their four comparison stations. Precipitation shows the highest sensitivity to distance, with a majority of the highest correlations found with interstation distances of less than 20 km. Results are less clear for both maximum and minimum temperature whereby there is sensitivity to both elevation difference and interstation distance.

Climate inhomogeneities (see section 2c) can produce spurious features in datasets that may otherwise be interpreted as trends and/or variability. To minimize the effects of inhomogeneities from contaminating the data infill of candidate stations, three measures are taken. First, comparison stations are selected based on monthly correlations calculated using overlapping 11-yr moving windows over the period of record. Second, the regression coefficients (m_{vi} , b_{vi}) are recomputed each time using only the contemporaneous monthly values from within the moving correlation window. Finally, the use of four comparison stations mitigates the contribution from inhomogeneities from any single station.

Multiple linear regression infilling is performed as follows:

$$\phi_v(t) = \frac{\sum_{i=1}^n \{r_{vi}^2 [m_{vi}x_i(t) + b_{vi}]\}}{\sum_{i=1}^n r_{vi}^2}, \quad (1)$$

where $\phi_v(t)$ is the estimated value, r_{vi} is the interstation correlation, m_{vi} and b_{vi} are the linear regression coefficients, $x_i(t)$ is the value of the i th comparison station at time t , and n is the number of valid comparison stations (defined here, $n = 4$). Estimates from each of the selected comparison stations are taken from the standard anomaly weighted by the square of the correlation coefficient. Finally, standardized anomalies are transformed back into data with measured units. Monthly data infilled by this process are not used to estimate other data.

c. Climate inhomogeneities

By definition, a homogeneous climate time series is one that fluctuates only in response to weather and climate (Conrad and Pollak 1950). Methodological influences (hereinafter “climate inhomogeneities”) are present in most long-term observational records, and as much as possible these should be identified and removed. Inhomogeneities can include abrupt, step change–like behavior or more gradual, trendlike behavior, and can be both chronic and intermittent (e.g., DeGaetano 2006). Although inhomogeneities may often appear minor compared to the inherent high degree of daily–interannual variability, nonclimatic effects can significantly bias a time series and lead to faulty interpretations of results. This is quite apparent when realizing that assessed surface air temperature trends ($+0.5^{\circ}$ – 1.0°C over the last century; e.g., IPCC 2007) are very close to the uncertainty in a single temperature measurement and within the range of influence associated with the detection (or nondetection) of a climate inhomogeneity.

Station history information (metadata) is maintained for each COOP station. Metadata include information on station location, instrumentation, and observational practices. However, documentation is often itself incorrect, incomplete, or missing, can vary in quality through time, or is not accessible in digital format. For these reasons and for this analysis, it is highly desirable to develop an automated objective method to diagnose potential inhomogeneities without prior knowledge of the metadata.

1) IDENTIFICATION OF INHOMOGENEITIES

A delicate balance exists between detection–adjustment of errors and preservation of the original data. Loose

criteria may lead to excessive correction of suspected errors in a given dataset (type I errors), and strict criteria may lead to insufficient correction (type II errors). Numerous methodologies have been developed and implemented to detect and correct for climate inhomogeneities (e.g., Peterson et al. 1998), with each technique having strengths and weaknesses (DeGaetano 2006), and no technique a clear favorite.

Every method to identify inhomogeneities will occasionally do so erroneously (e.g., Menne and Williams 2005). To reduce this likelihood, two complementary but independent techniques are used in this study to screen the data. This increases confidence in the detection of changepoints (time of inhomogeneity) and reduces type I errors. The methods used in this paper are variations on the multiple linear regression (MLR) method (Vincent 1998) and double-mass method (Kohler 1949). Each method considers the relationship between a given (“candidate”) station and a group of comparison stations. The methods are applied to the standardized monthly anomalies obtained from transformations of the original data.

(i) Multiple linear regression

The multiple linear regression method of Vincent (1998) compares a time series of observations (T_O) from a candidate station with time series of estimates (T_E) derived from comparison stations selected on a month-by-month basis. Unique to this method is the consideration of the autocorrelation of the residual ($T_E - T_O$). If the time series is free of inhomogeneities, the residual should take the form of random noise, as opposed to abrupt step changes or trends. Following Vincent (1998) a series of models are applied to the data in an iterative procedure until each segment of the residual has negligible serial correlation assessed using the Durbin–Watson D statistic.

The MLR method is illustrated using again the monthly mean minimum temperatures for the Yosemite Valley Headquarters COOP. Figure 3a shows the residuals of the estimates minus the actual data. An iterative procedure used to detect changepoints in the data by fitting the residuals to the given models identifies two potential inhomogeneities denoted by vertical lines in 1965 and 1978. Data lying between these changepoints are determined to be free of inhomogeneities given the condition that the residuals exhibit insignificant serial correlation.

(ii) Double mass

The double-mass curve (Kohler 1949) shows the deviation of a candidate station’s cumulative sum (e.g., summation of temperature anomalies) versus the cumulative sum of a comparison station. Changes in the re-

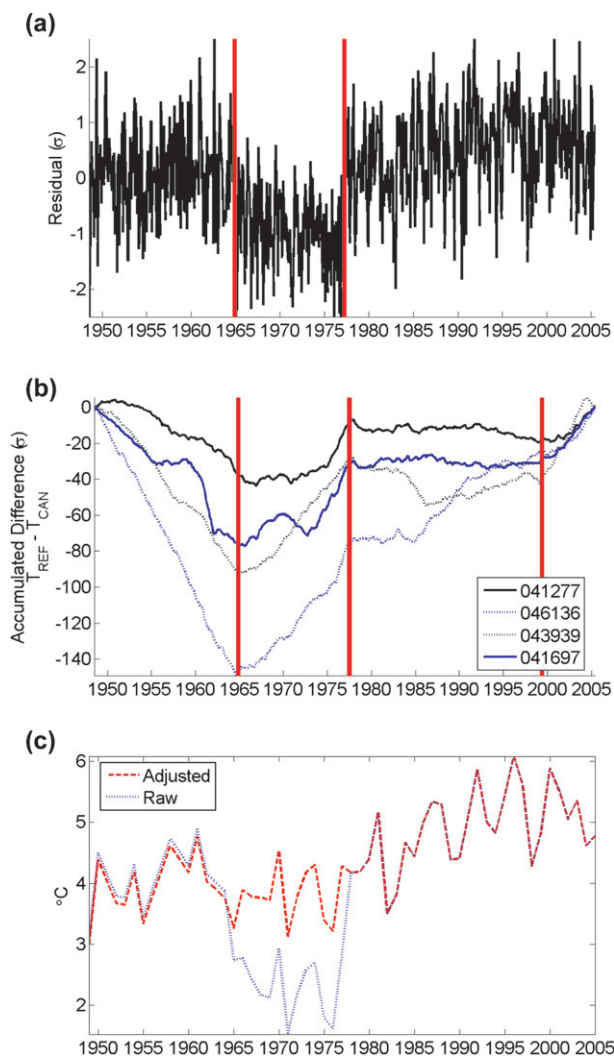


FIG. 3. Inhomogeneity tests for monthly mean minimum temperatures at Yosemite Park Headquarters COOP. (a) MLR analysis: $T_E - T_O$, residual in standard anomalies. Potential changepoints detected using the iterative MLR procedure are shown by the vertical line in May 1965 and June 1978. (b) Double-mass analysis: slope residual from comparison stations: Calaveras Big Trees, Nevada City, Hetch Hetchy, and Cherry Valley Dam. Vertical red lines in (a) and (b) indicate potential changepoints identified by the MLR and double-mass analysis, respectively. (c) Annual mean minimum temperatures for the original data (dotted blue line) and dataset after adjustment for inhomogeneities (dashed red line).

lationship between two sites manifest themselves as a change in slope of the curve. A residual can be computed by subtracting the linear least squares fit from the double-mass curve. Examination of the residual provides an effective means to identify nonlinearities in the relationship between stations. A perfect record, free of nonlinearities in observational record and free of climate inhomogeneities, would reveal a double-mass

plot that is approximately linear and a residual void of abrupt changes.

Arndt and Redmond (2004) used the double-mass method to subjectively identify inhomogeneities in temperature and precipitation observations for COOP stations. Though inhomogeneities assessed from double-mass residual curves may be apparent to the trained eye, manual screening of large datasets is quite laborious. An automated technique is needed that can identify an uncharacteristically large change in the slope of the residual. Monte Carlo resampling is used to objectively determine what constitutes a “large” (statistically significant) change in the residual slope. This procedure, described in the appendix, uses matched resampling and variance adjustment on the paired dataset to create a probability distribution of slope change, stratified by time scale. The double mass method uses four comparison stations, selected as described above but based on correlations taken over all months of the year, and identifies potential changepoints when the slope of the residual slope exceeds the 99% confidence bounds.

An example of the double mass analysis performed on monthly mean minimum temperature for the Yosemite Park Headquarters COOP is shown in Fig. 3b. The key features that stand out are time intervals when the residual accumulates approximately linearly with time, separated by abrupt changes in the slope of the residual. Finer detail reveals cyclic patterns with distinct annual periodicities indicative of physically explainable interstation seasonal differences that pertain to topographic effects in regions of complex terrain (e.g., Lundquist and Cayan 2007). The set of three vertical lines denote periods when at least three of the four comparison stations agree upon a changepoint. One can visually note other nonlinearities in the residual for individual stations in Fig. 3b; hence, several comparison stations indicating a particular changepoint in a candidate station increase the confidence in attributing the inhomogeneity to the candidate station as opposed to the comparison station. A test is run to ensure that changepoints detected with either the MLR or double-mass method do not coincide with data that have been infilled; however, due to the reliance of interstation relationships in the infilling process and the paired inhomogeneity methods, changepoints did not coincide with infilled data.

2) CHANGEPPOINT IDENTIFICATION AND DATA ADJUSTMENTS

Potential inhomogeneities identified concurrently by both the double-mass and MLR methods are marked as changepoints. For such near-simultaneous inhomogeneities (within three months of one another), a subsequent analysis is employed to identify the exact month

of the changepoint by examining a time series 25 months in length centered about the period of interest. The MLR method is then applied on this short time series to identify the changepoint. In the example for Yosemite Park Headquarters, both methods identify inhomogeneities in June 1965 and June 1978.

Once inhomogeneities are identified, it is necessary to decide if and how to adjust the data. First, an MLR estimate of the time series is generated using comparison stations. Second, adjustments of each segment (data between inhomogeneities) are made on the condition that the estimate deviates from the reported data at the 95% confidence level using a Student's *t* test. Under this condition, the segment is shifted accordingly to conform to the present relation between the estimate and the reported data (e.g., no adjustment is permitted for the latest values) using standard anomalies, rather than fixed values (e.g., degrees Celsius). Changepoints that meet the latter criteria are marked as adjusted changepoints.

Figure 3c shows the raw time series of annual average minimum temperatures at Yosemite Park Headquarters (blue dotted line), along with the adjusted time series (red dashed line). Station metadata reveal seven station observation changes during the 1949–2005 period, with automated methods detecting adjusted changepoints coinciding with the two significant station moves (station moved 250 m south and downhill in June 1965, and 460 m NNW in June 1978). To test the ability of the automated changepoint detection method in identifying documented changepoints, a subset of 46 stations composing the U.S. Historical Climatology Network (USHCN) is considered. Documented changepoints from station metadata are compared to objectively adjusted changepoints. Approximately three-quarters (77%) of all objectively determined changepoints coincided with documented changepoints, not only illustrating the ability of objective methods to detect known changepoints, but also reinforcing the importance of implementing such methods to reveal changepoints that lack documentation. When metadata are available, a more conservative approach may be to take the union of documented metadata changepoints and objectively determined changepoints. Although precipitation and temperature are both vulnerable to inhomogeneities, efforts described herein refer to the detection of changepoints (date of an apparent inhomogeneity) in the temperature time series.

d. Evaluation of NCDC climate divisions

Correlations of COOP stations with their respective NCDC climate division are computed to highlight intra-divisional structure, and thereby, deficiencies in using divisional data to characterize regional climate. For each month of the year, station-division correlations are calculated

using standardized anomalies of mean monthly temperature and total monthly precipitation. Temperature fields correlate well overall within the state, with a mean station-to-division correlation of $r = 0.87$ averaged over all stations (Fig. 2a); however, correlation values tend to show weaker correlations along a narrow region of the Pacific coast, as well as in regions where divisional boundaries span diverse topographic ranges (i.e., division 5 includes the San Joaquin Valley and the southern Sierra Nevada). Precipitation fields perform modestly across the state (mean $r = 0.81$; Fig. 2b). Relatively high correlations are found along the coast and western slopes of the Sierra Nevada, whereas correlations are consistently lower ($r < 0.7$) on the leeward side of north-south-oriented topographic features and exhibit increasing deficiencies for divisions that span large distances. Not shown is the fact that station-to-division correlations have a strong seasonal signal, suggesting that while NCDC divisions may be appropriate during certain seasons, they are less so during others. Overall, the paucity of high correlations over regions the size of traditionally used climate divisions suggests that improvements could be made to better characterize regional climate variability.

3. Regionalization process

The immediate goal of this analysis is to capture regional-scale variability in multivariate meteorological data to reduce a state-sized region with numerous climate stations into a manageable and physically realistic set of smaller regions. The use of monthly data is superior to the use of annual data because regional-scale climate variations may be present only during certain months of the year, and otherwise obscured by the utilization of annual mean data.

Previous studies have used objective means to define climate regions using principal component analysis (PCA) (e.g., White et al. 1991) and clustering analysis (e.g., Fovell and Fovell 1993). Both approaches were explored in this study, but hereinafter PCA results are emphasized. PCA seeks structures that explain the maximum amount of variance in a two-dimensional dataset (time versus space) by isolating a set of structures in the spatial dimension [empirical orthogonal functions (EOFs)] that concisely capture the variability within a given dataset. Although PCA is a powerful tool used to investigate variability, the resulting patterns are affected by choice of domain shape, domain size, and sampling. These choices often result in eigenvectors, or loading patterns, resembling a series of predictable geometric patterns called Buell patterns (Buell 1979). Likewise, standard PCA analysis often has limitations in isolating individual modes

of variability due to the orthogonality constraint, as two distinctly separate processes governed by different physical processes (which tend to be nonorthogonal) may be superposed into a single extracted component.

To counter these well-known problems, rotated EOFs (rEOFs) are employed. Two main classes of rEOFs were considered in the analysis: orthogonal and oblique rotation. Both methods offer a compromise between the strict mathematical constraints of an unrotated EOF and the ability to identify physically realistic patterns. Rotated EOFs tend to have more stable, physically realistic patterns as they typically minimize sampling errors compared to unrotated EOFs (Richman 1986). Rotation is also useful when loading patterns are distributed among many modes, as is the case when performing a regionalization process. Orthogonal rotation fixes the principal components to be orthogonal, while the loading patterns are allowed to correlate. Oblique rotation removes the constraint that the principal components are orthogonal. White et al. (1991) found that oblique rotation was most appropriate for generalizing stable regionalizations. Comrie and Glenn (1998) used obliquely rotated EOFs to isolate quasi-homogeneous regions of variability in precipitation to distinguish among the dynamical and physical processes associated with the North American monsoon region.

Most prior studies have defined regions based on a single climate element (i.e., precipitation or temperature); however, this study seeks to maximize the variance of multiple elements with the goal of justifying climate regions that work for both temperature and precipitation. To combine elements, standardized anomalies are employed and the analysis herein is constrained to monthly precipitation and monthly mean temperature (the numerical mean of monthly mean maximum and minimum temperature), so as to equally weight precipitation and temperature variables. A spatial mode EOF analysis is performed, therein effectively giving each station equal weighting in determining the patterns of variability. Both orthogonal varimax rotation and oblique rotation (using the "Direct Oblimin" method; see, e.g., Richman 1986) were applied to the EOF loading patterns. Though orthogonal rotation is able to capture significant modes of variability, oblique rotation was found to be preferable in achieving higher communalities.

4. Results

A fundamental question when assessing regional climate variability across a region as complex as California is, How many distinct regions are needed to adequately characterize climate variability? The answer depends on the intended applications, which as mentioned previously

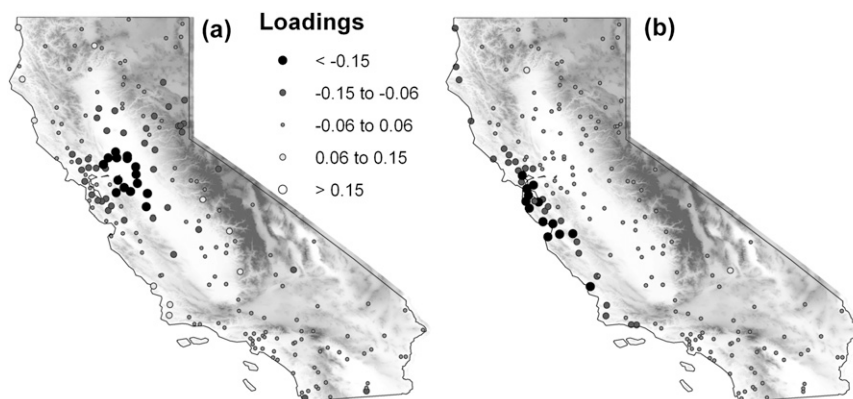


FIG. 4. Rotated loadings of monthly multivariate standardized temperature and precipitation for (a) component one (rEOF1) and (b) component three (rEOF3) from the station-based dataset. Dark large circles denote strong positive loadings (nondimensional), small gray circles are neutral loadings, and white circles are negative loadings.

can be quite diverse. Objectively, the number can be determined by noting that the scree plot (not shown) reveals a distinct drop-off in variance explained between mode 11 and 12. A total of 11 EOFs (e.g., 11 regions) is retained, accounting for 83.5% of the cumulative variance in temperature and precipitation fields.

The loading patterns associated with the first and third retained (rotated) modes are shown in Figs. 4a and 4b, respectively. The first loading pattern (rEOF1) isolates a mode of variability encompassing the lower Sacramento Valley, the Sacramento–San Joaquin Delta, and the interior valleys of the San Francisco Bay area. In contrast, the third loading pattern (rEOF3) isolates a narrow-coast strip that extends from Point Conception north to Point Reyes, with a strong gradient in loadings perpendicular to the coastline. The separation of rEOF1 and rEOF3 is physically consistent with controls on the intrusion of maritime air mass and associated variability in temperature fields. During the summer, variations in coastal SST and upwelling modify low-level stratus and temperature for stations dominated by rEOF3, while a different set of physical controls dictates temperature variability farther inland (e.g., Alfaro et al. 2004).

The “maximum loading rule” is used to delineate regions. This simple procedure classifies stations by the component on which they have the highest loading (e.g., Comrie and Glenn 1998). At most stations, loadings were heavily weighted toward a single component. However, at a few stations the loading on one component is marginally more than on another component. These so-called transitional stations were all found to lie in a region of overlapping influences, or in a buffer zone. Although it is desirable that a well-defined transition exists from one climate region to another, in reality the distinction between the two is not always sharp.

Figure 5 shows the 195 stations used in this study, coded with their assigned climate region according to the maximum loading rule. Regions broadly follow physical boundaries, most notably those dictated by coastal and topographic features, as well as latitudinal bands. The regionalization process suggests that coherent regions of precipitation variability are stratified primarily by latitude, except on the lee side of the Sierra Nevada and in the eastern desert and high plateau regions. These results are largely consistent with the analysis by Willmott (1977), who employed monthly precipitation (1961–70) from 90 stations across California to identify four distinct precipitation regions stratified by latitude and topography. By contrast, coherent regions of temperature variability tend to be stratified according to coastal and topographic features. In California, these geographic features are largely oriented north–south, leading to sharp longitudinal gradients in temperature regions.

Analogous to the station-to-NCDC division correlations computed previously (Fig. 2), station-to-objectively determined climate region correlations are computed. The regionalization process results in an improvement in station–climate region correlations across the state (Fig. 6). Boundaries defined in Fig. 6 are determined using the gridded database (section 4b) by identifying the largest pixel-to-region correlation with the stipulation that regions must be contiguous (in multiple-mountain areas, discontinuous regions can be physically justifiable even though impractical). A comparison between NCDC climate divisions and objectively determined climate regions (Figs. 2, 6) reveals that the newly defined regions not only more than double the number of stations that show very high correlations ($r > 0.9$), but also eliminate groups of stations with lower ($r < 0.7$) correlations, particularly with

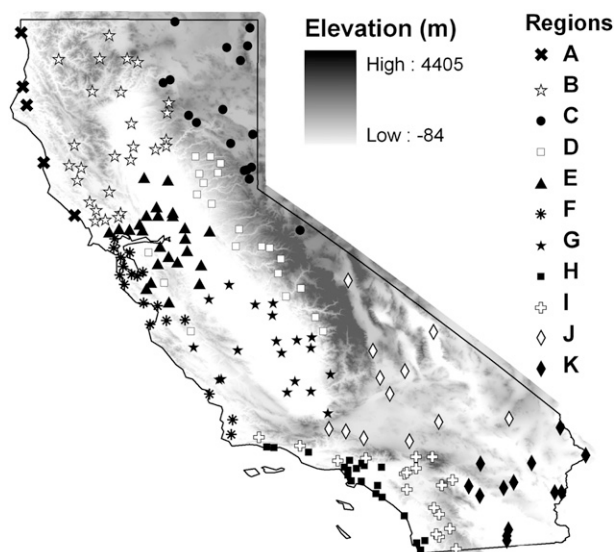


FIG. 5. Regionalization of station-based data using monthly mean temperature and precipitation. Eleven climate regions are isolated: the north coast (A), north-central (B), northeast (C), sierra (D), Sacramento-Delta (E), central coast (F), San Joaquin Valley (G), south coast (H), southern interior (I), Mojave (J), and Sonoran (K) regions.

respect to seasonality (e.g., some stations may show strong station-to-division correlations for 9 months of the year but poor correlations during the other 3). These 11 regions effectively reduce the complexity inherent in temperature and precipitation data fields across an array of COOP stations into a lower-dimensional dataset that effectively characterizes regional climate variability across the state.

a. The climate regions

Three narrow coastal climate regions exist along the north coast (A), central coast (F), and south coast (H),

separated by transitions at Point Reyes and Point Conception. Their existence arises from the moderating influence of maritime air on temperature; coastal regions do not stand out in an analysis of precipitation variability alone. Remarkably, summer [June–August (JJA)] maximum temperatures along the entire coastal strip from San Diego to Crescent City are isolated as a single mode. In other words, interannual variations in summer daytime temperature are coherent along the California coastline that stretches nearly 10° (1100 km) in latitude. These findings appear consistent with Alfaro et al. (2006), who stated that the coast–inland gradient in maximum temperatures likely pertains to the immediate coastal expression of the Pacific decadal oscillation.

Latitudinal variations in precipitation divide these three coastal regions. The north coast region often has a bimodal structure whereby precipitation maxima occur in early and late winter during those winters when the storm track migrates abnormally far south. On the other hand, the south coast region has a distinct tie to ENSO (Redmond and Koch 1991). This appears most robust for cold ENSO events (e.g., La Niña winters have been unambiguously dry in the southern part of the state, whereas El Niño winters have been ambiguously wet). In the central and northern part of California the association between ENSO and precipitation is less well defined.

A distinctly separate set of regions occur in the transition inland from the immediate coastal environment. The north region (B) extends from the Oregon border to the northern parts of the Sonoma Valley, covering both the coastal range and the northern Sacramento Valley. The lack of a topographic barrier in the Sacramento-Delta region (E) allows for sporadic penetration of maritime flow through the Carquinez Strait (e.g., Zaremba and Carroll 1999). This inland penetration of a sea-breeze circulation (e.g., east Pacific subtropical high dominating offshore while a thermal low

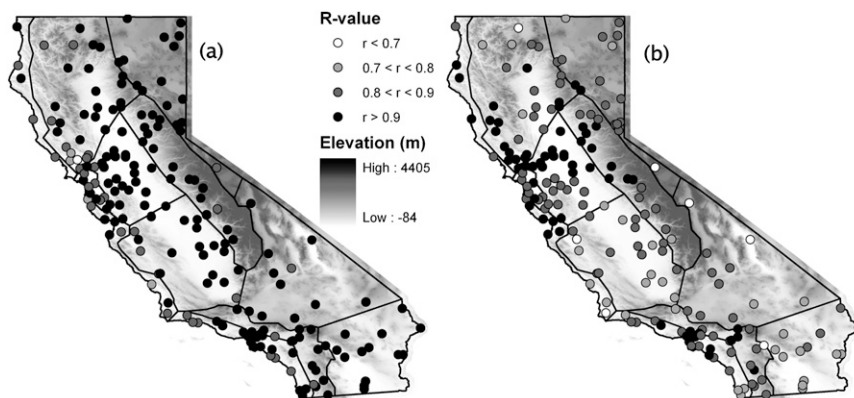


FIG. 6. Monthly correlation of (a) mean temperature, and (b) precipitation between stations and their respective climate regions as identified in this study (region boundaries outlined).

forms over the heated interior) allows for temperatures to moderate, in comparison with locations in the northern and southern ends of the Central Valley that have no direct ventilation route. The lack of an abrupt topographic barrier also results in a wider transition zone between the coastal and more interior climate regions. The San Joaquin Valley region (G) is found to be distinctly separate from the Sacramento-Delta region in both its temperature and precipitation records. Without a direct conduit to maritime air or ventilation, both the San Joaquin Valley and north regions are more vulnerable to persistent stagnation during periods without synoptic disturbances.

The sierra region (D) covers the foothills and higher elevations along the west slope to the crest of the Sierra Nevada south of 40°N. A strong difference is noted in temperature fields between the sierra region and the adjoining San Joaquin Valley and Sacramento-Delta regions. This is most apparent during midwinter during months when pronounced ridging sets up over the region. Consequently, persistent inversions in winter enable radiation fog to suppress maximum temperatures in the valley, while warmer than normal conditions exist in the higher altitudes above the inversion. For similar reasons, temperature variations of moderate elevation stations (>800 m) located on the coast range in central California align more with stations in the sierra region than closer stations in the valley or coast.

The northeast region (C) lies in the rain shadow of the northern Sierra Nevada and southern Cascade ranges. Although progressive synoptic systems impact the north-central, sierra, and northeast regions, the orographic precipitation ratios can vary widely from system to system because of the trajectory of the storm track relative to that of the topography. On longer time scales, modulations associated with large-scale modes of variability may play a role in the bias of the orientation of the storm track as repeated systems are integrated over time to form realized monthly or seasonal values. This is noted by observing that DJF precipitation correlation between the northeast and neighboring north-central regions during the winter months is less than 0.75. In addition, the leeward position of the northeast region makes the area more vulnerable to cold air outbreaks via retrograding polar continental air masses that dive southward into the Great Basin, but are often unable to penetrate west of the Sierra Nevada. During the transitional seasons of spring and fall the monthly temperature correlation between the northeast and Sierra region exceeds $r = 0.95$; this drops substantially ($r < 0.85$) during December and January.

Directly inland of the south coast is the southern interior (I) region, covering the inland valleys and Peninsular

and Transverse Ranges of southwestern California. Precipitation and temperature of the southern interior and the south coast are tightly coupled for most of the year; however, these regions decouple during late summer through early fall. Competing influences from the west due to interannual variability in the inland penetration of the marine layer, and from the east due to interannual variability along the westernmost periphery of the monsoonal flow, appear to result in the division of these regional modes of variability.

Topographic barriers separate the Mojave (J) and Sonoran (K) regions from progressive eastward-moving synoptic systems, largely decoupling wintertime precipitation in these regions from that of the rest of the state. These regions are also unique within the state, as they exist on the northwestern extent of the North American monsoon region and contain a relatively strong summertime precipitation signal, in contrast to the prevailing statewide winter precipitation domination. Comrie and Glenn (1998) isolated a region covering both the Mojave and Sonoran regions along with the lower Colorado River basin and the northern half of the Baja Peninsula from the remaining portion of California into a mode of variability using monthly precipitation data. The division between the Mojave and Sonoran regions is evident during the summer months, as the correlation between the temperature records is significantly reduced ($r < 0.75$, as opposed to $r = 0.95$ for remaining months). These two regions also represent distinctly different ecological provinces, arising from climate differences.

b. Robustness to subdomains and subsampling

Two analysis limitations that may affect regionalization results are the nonuniform station density (biased toward the more heavily populated parts of the state) and the chosen domain (restricted to within the political border of the state). Karl et al. (1982) showed that irregularly spaced points (here, stations) can significantly alter PCA loading patterns. Because the regionalization procedure is blind to station location, the analysis will tend to give added weight to a mode of variability tied to a dense cluster of correlated stations that essentially provide samples of the same climate, and relatively less weight to true physical modes that have limited sampling. Similar arguments can be made for truncated domains. To test whether the identified climate regions reflect true physical modes of variability, an examination is performed to discern whether coherent centers of action retain their shape with respect to (i) varying the density of stations, (ii) redefining the domain, and (iii) redefining the period of record.

Previous studies have addressed the issue of station density by removing stations to achieve a regular spacing

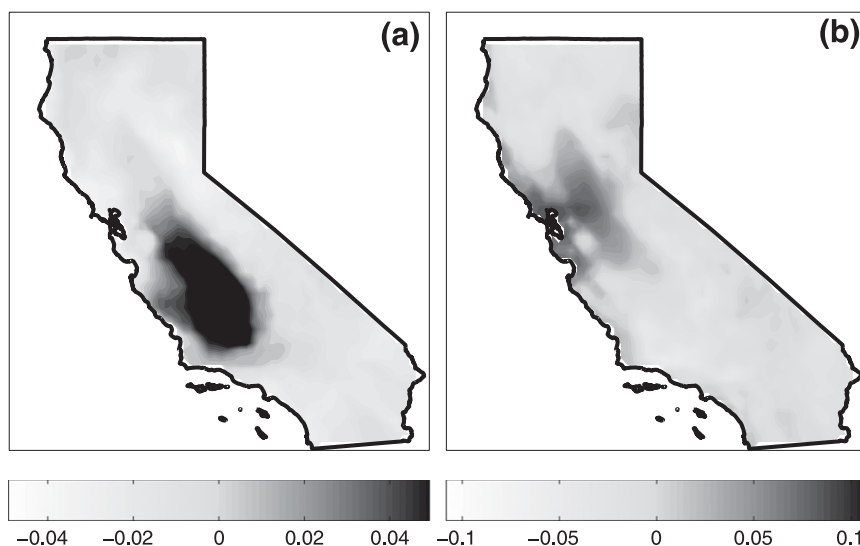


FIG. 7. Rotated loadings monthly multivariate standardized temperature and precipitation for rEOF2 and rEOF4 using the PRISM dataset. Shading denotes nondimensional loadings.

of stations. Here, the deletion of stations was contradictory to the goal of identifying regional-scale variability using the most complete dataset available. To circumvent this issue, and to add robustness to the results, a different version of the monthly dataset was employed—the gridded Parameter-Elevation Regressions on Independent Slopes Model (PRISM) dataset (Daly et al. 2002). PRISM incorporates climate data from up to 8000 stations across the United States along with a digital elevation model to provide a high-resolution (4 km) gridded analysis of climate variables on monthly time scales over the period of record (1895–present). For computational tractability, the 4-km PRISM grid points for California are here aggregated to 16-km resolution within the state of California, with the time series again covering the same period as the station analysis, 1949–2005. The PCA-based regionalization analysis of monthly temperature and precipitation described in section 3 is repeated by essentially considering the grid of points (approximately 1500 points) to be a set of “stations” uniformly distributed across the entire state.

A total of 10 components are retained from the regionalization process using PRISM data and scree tests. The resulting loading patterns bear a striking resemblance to those obtained using station-based data. Figure 7 shows loading patterns for rEOF2 and rEOF4, clearly isolating the San Joaquin Valley and Sacramento-Delta and regions as identified with the station dataset. The only significant difference between the PRISM-based regionalization and the station-based regionalization process was that the PRISM-based regionalization failed to iden-

tify a southern interior region, and instead grouped the valleys and mountain ranges of the southwestern portion of the state together with the coast. Another revealing finding using the gridded dataset is the narrowness of parts of the coastal regions of the state. The coastal strip encompassing the north coast and central coast regions was found to be extremely narrow, typically 1–2 grid points wide (e.g., 16–32 km wide), and undetectable along the central coast near Big Sur where the steepest coastal gradients in North America are found. Along these lines, it is reasoned that the aggregation from 4 to 16 km inhibits the ability of the gridded data to realize the south coast region, therein revealing the importance of finescale spatial resolution, whether it be through observations or high-resolution gridded datasets, in capturing regional modes of variability in complex topography.

Another method used to examine the robustness of the regionalization process is to restrict the spatial domain of stations in the analysis. PCA-based regionalization is performed on a subdomain of the state using only stations located south of 35°N. Three components are isolated as a result of the regionalization procedure (Fig. 8). The most evident mode of variability is a narrow coastal strip (triangles) that encompasses all the stations in the south coast region along with a couple of stations assigned to the central coast region. Stations in the lower desert (crosses) cover all stations assigned to the Sonoran region. The third region encompasses the south interior region (squares) along with several stations along the southern border of the Mojave region. Overall, the regionalization process performed on this

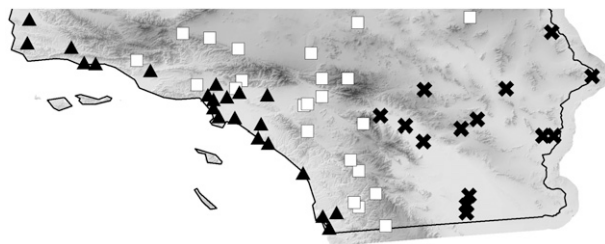


FIG. 8. Regionalization of all California-based COOP stations located south of 35°N. Three modes of variability highlight a coastal region (black triangles), a low desert region (crosses), and higher-elevation inland locales (white squares).

subdomain matches remarkably well with that performed for the state as a whole. A similar level of corroboration was found by regionalizing stations located in the northern half of the state (not shown).

For reasons of operational practicality and computability, the analyses described thus far were necessarily restricted to the use of stations and PRISM grid points located exclusively within the state border. Since political boundaries do not physically correspond to climatic boundaries, it is instructive to examine possible differences in results from analyses based on a larger domain. A PCA performed on a domain incorporating the adjacent states of Oregon, Nevada, and Arizona introduces a few additional regions along state borders (not shown). For example, the Mojave region is divided into one region that encompasses Death Valley National Park, much of the southern tip of Nevada, and the southwestern Colorado River basin, while another region spans the higher-elevation desert on the lee side of the Sierra Nevada covering sections of eastern California and western Nevada.

Another means of verifying the stability and robustness of the identified patterns is to vary the temporal domain of the dataset. PCA analysis was performed on the first half of the record (1949–77) and then the second half (1978–2005). The loading patterns across these two time periods are found to be quite similar to one another and to the loading patterns obtained by considering the entire period of record. Finally, one may question whether regional trends may influence the results of this study. The regionalization process was not sensitive to trends in the dataset; removal of the linear trend at each station did not significantly alter results.

5. Regional datasets and their application

Regional datasets were created for each of the 11 climate regions identified within the state by this study. Regional values are formed in two ways: 1) by taking the simple arithmetic mean of monthly COOP-based

data within a given region, and 2) by taking the areal mean from the gridded PRISM database for each climate element of interest. This provides a basis for calculating updates in an operational setting using available stations just after the end of a given month. The regionalization process was performed using monthly data from 1949 to 2005, a time period that included a more complete dataset; however, inhomogeneity tests described herein were applied to reporting stations back to 1895. Despite a smaller number of reporting stations in the earlier record, the availability of numerous USHCN stations allows for the creation of the monthly time series for each region from 1895 to present. The climate regions and their time series are available via a real-time product routinely updated on a monthly basis (see section 6).

To emphasize the utility of the regionalization processes we examine regional differences between the San Joaquin Valley region and the sierra region, both of which are encompassed by NCDC's climate division 5 (San Joaquin drainage basin). For most of the year, interregional correlation of maximum temperatures exceeds $r = 0.95$; however, during winter [December–February (DJF)] the interregional correlation drops below $r = 0.60$, in agreement with Christy et al. (2006, their Table 4). To better understand the origins of this winter decoupling, all months where the difference between valley and mountain maximum temperatures (in standardized anomalies) exceeds one standard deviation are analyzed. A total of 30 (32) such months are found when San Joaquin Valley standard temperature anomalies are less than (more than) one standard deviation from sierra region standard temperature anomalies.

A composite of 500-hPa geopotential height anomalies for months characterized by relatively warmer maximum temperatures in the sierra region in comparison with the San Joaquin Valley region shows anomalous ridging located upstream of the West Coast (Fig. 9a). This flow pattern is indicative of a blocking pattern off the west coast of North America and a poleward deflection in the storm track over the western United States leading to a significant reduction in precipitation for the central and southern part of the state. Midwinter blocking episodes and associated subsidence inversions over the western United States lead to conditions favorable for the occurrence of long-lived, spatially extensive radiation fog (e.g., Holets and Swanson 1981). These conditions during midwinter can persist for more than two weeks, effectively decoupling climate on a monthly time scale between the San Joaquin Valley and neighboring high-altitude sierra during the winter months. By contrast, nearly symmetric anomalies of opposing sign associated with a deepening of the trough

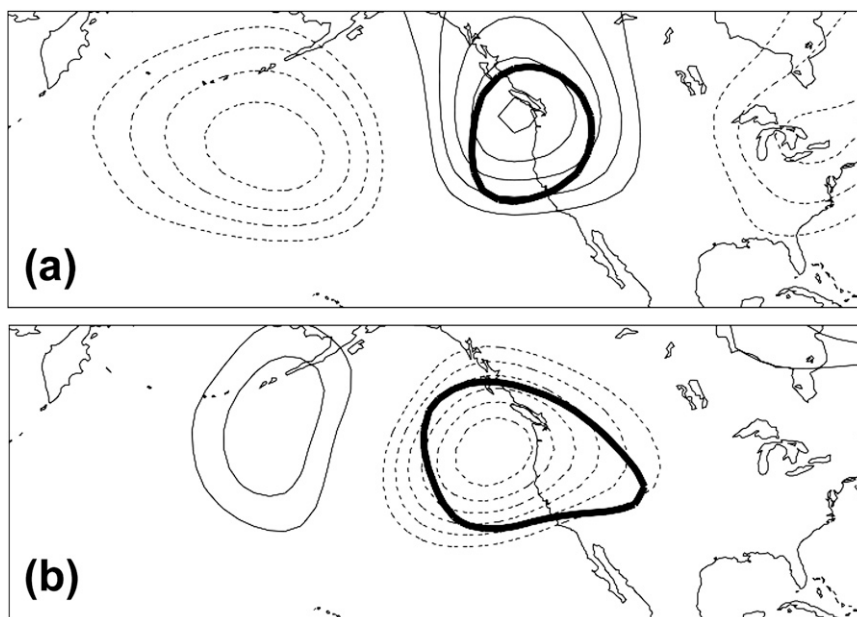


FIG. 9. Composite 500-hPa geopotential height anomaly from 1958 to 2005 reference period for (a) months when maximum temperature anomalies in the sierra region exceeded maximum temperature anomalies in the San Joaquin Valley region by at least one standard deviation and (b) months when maximum temperature anomalies in the San Joaquin Valley region exceeded maximum temperature anomalies in the sierra region by one standard deviation. Contours start at 20 m and are shown every 10 m. The 95% significance level is shown by the thick contour.

off the west coast of North America, a southward shift in the jet, and positive precipitation anomalies across the central portion of the state are present in months characterized by relatively cool temperatures in the sierra (Fig. 9b). Given the elevational gradient between locales, such conditions are consistent with an enhanced lapse rate and the increased prevalence of synoptic disturbances. The ability for objectively defined climate regions to discern significant regional differences in variability provides an impetus for utilizing objective climate regions rather than NCDC climate division datasets in climate assessment studies.

6. Conclusions

The ability to both monitor and understand mechanisms behind regional-scale variations in climate is becoming increasingly relevant in the face of global climate change. Climate variability is a focal point of the interface of the physical–environmental–societal system. By better understanding the drivers of regional-scale climate variations, we may become more adept at projecting the true impacts of climate change on civilization as well as environmental systems such as hydrologic processes and ecosystem dynamics. Although regional climate variability is apparent across the globe, coherent patterns are most pronounced, and with the finest spatial scales,

in the presence of complex physiographic controls. As demonstrated for the state of California, the confluence of large-scale modes of variability with both the maritime influence and topography illuminates distinct regional climate variability signatures. It is likely that a similar set of features that give rise to these regional modes of variability also factors strongly into determining ecological regimes unique to the landscape of each region.

Application of these regions can be used to illuminate physical mechanisms that are responsible for regional climate variability (e.g., section 5). Monthly temporal resolution (rather than annual) is crucial in uncovering true regional-scale modes of variability present only during portions of the annual cycle. This is not surprising in view of the seasonality of the mean large-scale climate drivers, and of the seasonality in their variability properties. Other drivers such as land use patterns and irrigation (e.g., Lobell and Bonfils 2008) are more localized and are sharply segregated by month. The utility of employing objectively defined regions may be particularly useful in better understanding variability and trend behavior in regional climate records, which may otherwise be absent through the use of traditionally employed NCDC divisional datasets (e.g., Christy et al. 2006).

The results of this study have been incorporated into an operational delivery system called the California

Climate Tracker, with monthly updates available on the first day of each month. This can be accessed at the National Oceanic and Atmospheric Administration (NOAA) Western Regional Climate Center (<http://www.wrcc.dri.edu/monitor/cal-mon/>). The purpose is to make the resulting time series accessible to the public, media, and policy sectors, and is particularly vital in assessing the impact of climate variability for energy, agricultural, and hydrological sectors within California. An example of the application of objective climate regionalization for the energy sector is the tracking of JJA maximum temperatures along the narrow coastal zones of the state, home to a majority of the state's population and energy demand. Given the decoupling of coastal locations from the interior, and recent asymmetric trends (Lebassi et al. 2009), the practical monitoring of regional monthly temperature variations provides value-added climate information for energy planning purposes.

Acknowledgments. The authors are appreciative of discussion, critiques, and support provided by the NOAA Western Regional Climate Center, Scripps California Climate Change Center, California Energy Commission Public Interest Energy Research Program, Desert Research Institute Division of Atmospheric Sciences, and the Oregon State University PRISM Group. We also thank the anonymous reviewers for their constructive comments in helping to improve the manuscript. We also thank Crystal Kolden for assistance with figures.

APPENDIX

Objective Double-Mass Analysis

Much previous usage of the double-mass analysis in identifying climate inhomogeneities has relied on subjective “eyeball” methods. An objective means to diagnose inhomogeneities using Monte Carlo methods was developed to improve replicability. Double-mass analysis is performed for each candidate station, using a set of four comparison stations chosen based on peak monthly mean correlation across all months of the year. The algorithm is as follows:

- 1) Data are transformed into standardized anomalies.
- 2) The double-mass method between candidates and each of the comparison stations is applied with “arms” extending forward and backward in time along the curve from the present test date, the site of a hinge point (e.g., Arndt and Redmond 2004). The length of these arms can be any time frame of interest. Dates of discontinuities can be better resolved with shorter-length arms, but at the expense of

greater noise and uncertainty. This algorithm begins with arm lengths of 6 months.

- 3) The difference between the forward-looking and backward-looking slope of the residual from the double-mass analysis between the candidate station and each comparison station is computed at each time point in the period of record.
- 4) A Monte Carlo matched resampling procedure is applied to assess the statistical significance of changes in the slope difference. The double-mass method as described in steps 2–3 is applied for each set of resampled data. Since climate datasets have built in serial correlation, one cannot directly compare Monte Carlo methods, which effectively randomize the data, to the actual dataset. To account for this discrepancy between the resampled and actual datasets variance adjustment is applied on the differential slopes calculated for each dataset. This enables one to define a 99% confidence interval on difference in slope.
- 5) Slope differences exceeding the 99th percentile are marked as potential inhomogeneities.
- 6) If three of the four comparison stations identify a potential inhomogeneity at any instance, it is marked as an actual inhomogeneity. Steps 2–7 are repeated with arm lengths of 12 and 24 months. If a detected inhomogeneity coincides with a previously identified inhomogeneity (or falls within one-half of the arm length), the original changepoint identified using the smaller arm length is retained, since shorter arm lengths identify changepoints with greater precision.

REFERENCES

- Alfaro, E. J., A. Gershunov, D. Cayan, A. Steinemann, D. Pierce, and T. Barnett, 2004: A method for prediction of California summer air surface temperature. *Eos, Trans. Amer. Geophys. Union*, **85**, 553–558.
- , —, and —, 2006: Prediction of summer maximum and minimum temperature over the central and western United States: The roles of soil moisture and sea surface temperature. *J. Climate*, **19**, 1407–1421.
- Arndt, D. S., and K. T. Redmond, 2004: Toward an automated tool for detecting relationship changes within series of observations. Preprints, *14th Conf. on Applied Climatology*, Seattle, WA, Amer. Meteor. Soc., 7.1. [Available online at <http://ams.confex.com/ams/pdfpapers/72826.pdf>.]
- Buell, C. E., 1979: On the physical interpretation of empirical orthogonal functions. Preprints, *Sixth Conf. on Probability and Statistics in Atmospheric Sciences*, Banff, AB, Canada, Amer. Meteor. Soc., 112–117.
- Christy, J. R., W. B. Norris, K. Redmond, and K. P. Gallo, 2006: Methodology and results of calculating central California surface temperature trends: Evidence of human-induced climate change? *J. Climate*, **19**, 548–563.
- Comrie, A. C., and E. C. Glenn, 1998: Principal components-based regionalization of precipitation regimes across the Southwest

- United States and Northern Mexico, with an application to monsoon precipitation variability. *Climate Res.*, **10**, 201–215.
- Conrad, V., and L. W. Pollak, 1950: *Methods in Climatology*. 2nd ed. Harvard University Press, 459 pp.
- Daly, C., 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.
- DeGaetano, A. T., 2006: Attributes of several methods for detecting discontinuities in mean temperature series. *J. Climate*, **19**, 838–853.
- Fovell, R. G., and M.-Y. C. Fovell, 1993: Climate zones of the conterminous United States defined using cluster analysis. *J. Climate*, **6**, 2103–2135.
- Guttman, N. B., and R. G. Quayle, 1996: A historical perspective of U.S. climate divisions. *Bull. Amer. Meteor. Soc.*, **77**, 293–303.
- Holets, S., and R. N. Swanson, 1981: High-inversion fog episodes in central California. *J. Appl. Meteor.*, **20**, 890–899.
- Hubbard, K. G., N. B. Guttman, J. You, and Z. Chen, 2007: An improved QC process for temperature in the daily cooperative weather observations. *J. Atmos. Oceanic Technol.*, **24**, 206–213.
- Karl, T. R., A. J. Koscielny, and H. F. Diaz, 1982: Potential errors in the application of principal component (eigenvector) analysis to geophysical data. *J. Appl. Meteor.*, **21**, 1183–1186.
- Kohler, M. A., 1949: Double-mass analysis for testing the consistency of records for making adjustments. *Bull. Amer. Meteor. Soc.*, **30**, 188–189.
- Lebassi, B., J. Gonzalez, D. Fabris, E. Maurer, N. Miller, C. Milesi, P. Switzer, and R. Bornstein, 2009: Observed 1970–2005 cooling of summer daytime temperatures in coastal California. *J. Climate*, **22**, 3558–3573.
- Lobell, D. B., and C. Bonfils, 2008: The effect of irrigation on regional temperatures: A spatial and temporal analysis of trends in California, 1934–2002. *J. Climate*, **21**, 2063–2071.
- Lundquist, J. D., and D. R. Cayan, 2007: Surface temperature patterns in complex terrain: Daily variations and long-term change in the central Sierra Nevada, California. *J. Geophys. Res.*, **112**, D11124, doi:10.1029/2006JD007561.
- Menne, M. J., and C. N. Williams Jr., 2005: Detection of undocumented changepoints using multiple test statistics and composite reference series. *J. Climate*, **18**, 4271–4286.
- NRC, 1998: *Future of the National Weather Service Cooperative Network*. National Academies Press, 65 pp.
- Peterson, T. C., and Coauthors, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.*, **18**, 1493–1517.
- Redmond, K. T., and R. W. Koch, 1991: Surface climate and streamflow variability in the western United States and their relationship to large scale circulation indices. *Water Resour. Res.*, **27**, 2381–2399.
- Richman, M. B., 1986: Rotation of principal components. *J. Climatol.*, **6**, 293–335.
- Smith, T. M., and R. W. Reynolds, 2005: A global merged land–air–sea surface temperature reconstruction based on historical observations (1880–1997). *J. Climate*, **18**, 2021–2036.
- Solomon, S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen, Eds., 2007: *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 996 pp.
- Vincent, L. A., 1998: A technique for the identification of inhomogeneities in Canadian temperature series. *J. Climate*, **11**, 1094–1104.
- White, D., M. Richman, and B. Yarnal, 1991: Climate regionalisation and rotation of principal components. *Int. J. Climatol.*, **11**, 1–25.
- Willmott, C. J., 1977: A component analytic approach to precipitation regionalization in California. *Arch. Meteor. Geophys. Bioklimatol.*, **24**, 269–281.
- Wolter, K. E., and D. Allured, 2007: New climate divisions for monitoring and predicting climate in the U.S. Intermountain West Climate Summary. [Available online at http://www.climas.arizona.edu/forecasts/articles/climatedivisions_July07.pdf.]
- Zaremba, L. L., and J. J. Carroll, 1999: Summer wind flow regimes over the Sacramento Valley. *J. Appl. Meteor.*, **38**, 1463–1473.

Copyright of *Journal of Applied Meteorology & Climatology* is the property of American Meteorological Society and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.