

# Principal components-based regionalization of precipitation regimes across the southwest United States and northern Mexico, with an application to monsoon precipitation variability

Andrew C. Comrie<sup>1,\*</sup>, Erik C. Glenn<sup>2</sup>

<sup>1</sup>University of Arizona, Department of Geography and Regional Development, Tucson, Arizona 85721, USA

<sup>2</sup>University of Arizona, Office of Arid Lands Studies, 1955 E Sixth St, Tucson, Arizona 85719, USA

**ABSTRACT:** We determine precipitation regions for the United States-Mexico border region based on seasonality and variability of monthly precipitation at 309 stations for the period 1961 to 1990. Using a correlation matrix of input data to avoid the effect of elevation on precipitation, we apply principal components analysis with oblique rotation to regionalize this large, climatologically complex study area. We examine the applicability of the method, 2 techniques for defining region boundaries, the various defined regions themselves, and the effects of transforming input data and changing obliquity of component rotation. We obtain 9 consistent and largely contiguous regions from each of the analyses, including regions for the North American monsoon, the low deserts, the California Mediterranean regime, and for summer precipitation regimes adjoining the Gulf of Mexico. The derived regions and associated boundaries make physical sense in terms of the driving atmospheric processes, and they are robust to transformed input data and changes in rotation procedures. The central border regions are remarkably consistent across analyses, with small changes to peripheral regions. We also identify 4 monsoon sub-regions, and we illustrate the applicability of the regionalization via an analysis of relationships between monsoon precipitation variability and 500 mb pressure heights. Significantly different 500 mb circulation patterns are associated with wet and dry monsoon seasons in each of the sub-regions, and it appears that shifts in 500 mb circulation relative to the geographic position of each sub-region influence seasonal precipitation variability, directly or indirectly. There are important differences between some sub-regions, but in general wet monsoons are associated with northward meridional bulging of the subtropical anticyclone over the continental monsoon areas, while dry monsoons are associated with zonal stretching of the subtropical anticyclone over adjacent oceans with slightly higher pressure-heights. Overall, the study provides a clear regionalization of the precipitation climatology for the southwest United States and northern Mexico, and shows its utility for studies of climate variability.

**KEY WORDS:** Climate regionalization · Precipitation · Southwest United States · Northern Mexico · Principal components analysis · Monsoon variability

## 1. INTRODUCTION

Precipitation is the critical climate variable in the arid borderlands of the United States and Mexico. Although climate is of course contiguous across the political boundary, historically, broad-scale analyses of precipitation across this border region have most often treated the U.S. or Mexican sides separately, partly for

reasons of data availability. The region is physically and climatically complex, making simple comparisons of precipitation awkward. Yet, there are several well-recognized climatic controls on precipitation across the region that reflect its sub-tropical location. The subtropical anticyclones and associated upper-level ridges centered over the region and adjacent Atlantic and Pacific oceans exert a dominant influence year-round, modulated by the role of the mid-latitude westerly circulation in winter and the North American monsoon in

\*E-mail: comrie@climate.geog.arizona.edu

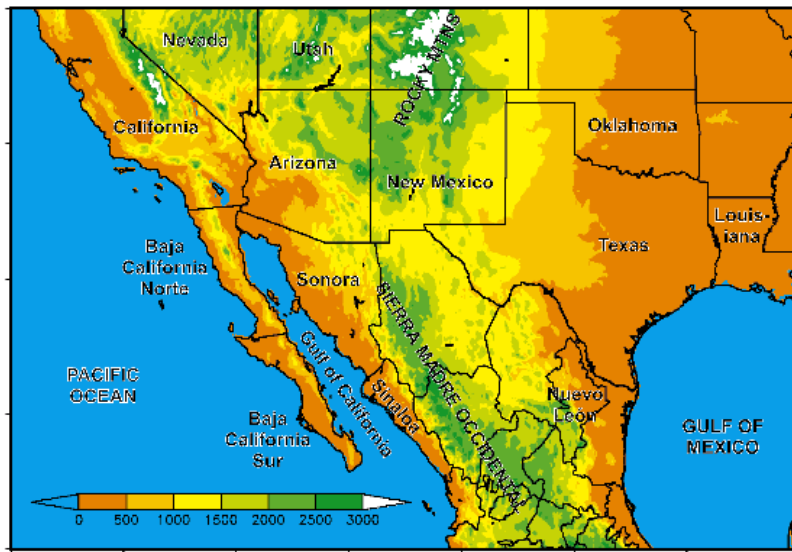


Fig. 1. Terrain elevation (m) of the study area with major place names and physical features mentioned in the text

summer (Carleton 1987, Cavazos & Hastenrath 1990, Burnett 1994, Comrie 1996, Adams & Comrie 1997, Woodhouse 1997). Sub-regionally, physiography (Fig. 1) and atmospheric circulation combine to create a spatially complex precipitation climatology (Fig. 2), with strong seasonal contrasts and areas of considerable interannual variability.

For many analytical purposes (e.g. examination of temporal trends) it is convenient to divide the spatial continuum of a precipitation climatology into a manageable number of quasi-homogeneous areas. Climate regionalization is a useful technique that enables generalization about areas on the basis of a spatially and temporally varying parameter such as precipitation. Eigenvector-based techniques are widely used for delimiting climate regions (Barring 1988, Malants & Feyen 1990, Fernández Mills et al. 1994, Lyons & Bonell 1994, Fernández Mills 1995), and such approaches have been successful in discriminating precipitation regions based on similar seasonality and long-term variability characteristics (Kutzbach 1967, Richman & Lamb 1985, 1987, Ogallo 1989, Eklundh & Pilesjö 1990, White et al. 1991, Bunkers et al. 1996, Cahalan et al. 1996). The work of White et al. (1991) reviews this approach and compares a range of rotation algorithms for eigenvector-based climate regionalization. They

found oblique rotations to be generally the most stable for climate regionalization, while orthogonally rotated and unrotated solutions were less stable.

In this study we use principal components analysis (PCA) to identify precipitation regions for the U.S.-Mexico border region based on the above characteristics. We explicitly avoid relying on the absolute precipitation totals in order to exclude the otherwise overwhelming elevation effect. This is a region of complex terrain in which local precipitation amounts are governed by topographic controls at finer scales than the atmospheric controls of overall seasonality and variability. Consequently, we aim to identify spatially cohesive precipitation regions based on common patterns of seasonality and variability rather than

on simple precipitation totals. There are no studies that explicitly regionalize precipitation in this area, although several authors have examined related topics such as harmonic analysis of precipitation, predominant seasonal air masses, and synoptic climatology (Horn & Bryson 1960, Mitchell 1976, Barry et al. 1981, Comrie 1996).

Specifically, for this investigation we apply the general recommendations of White et al. (1991) in a broad-scale precipitation regionalization over the large domain of the border region. We examine the perfor-

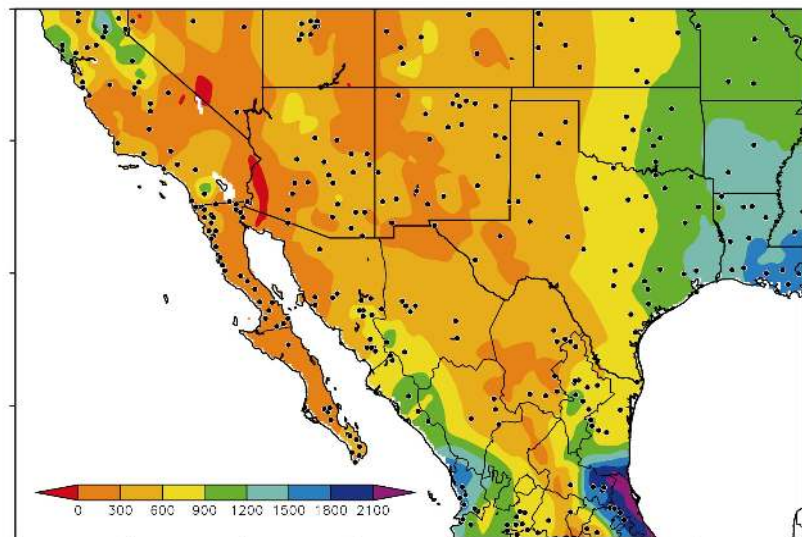


Fig. 2. Station locations used in the study and interpolated mean annual precipitation (mm)

mance of the method itself, we evaluate the various defined regions, and we compare 2 techniques for locating region boundaries. Because the value of any climate classification lies in its usefulness, we illustrate the applicability of the regionalization via an analysis of precipitation variability and 500 mb circulation for the North American monsoon region. Thus, the broader goals behind the study are first to simply identify and describe these regions as they exist across both sides of the political boundary, noting their characteristics and controls, and second to provide appropriate spatial units for analyses of precipitation variability in this and future work.

## 2. DATA

We use precipitation data from a network of 309 stations throughout the study region, which covers land areas between 20° and 40° N and 125° and 90° W (see Figs. 1 & 2). The domain was chosen to be somewhat larger than the perceived 'edges' of the immediate border region so that precipitation regions extending beyond this zone might be resolved with respect to neighboring areas. The data set comprises monthly precipitation totals from 1961 to 1990 ( $n = 360$ ) for stations across the study area, extracted from the Global Historical Climatology Network (GHCN; Vose et al. 1992) archived at the National Climatic Data Center in Asheville, North Carolina. Several other data sources were evaluated, but the quality control of the GHCN data (systematic screening for missing data, outliers, station moves, internal consistency, etc.) made them the optimal choice. A dense pattern of points was chosen, as precipitation is highly variable (White et al. 1991), and a regular spacing of stations was attempted because some studies have shown that a PCA performed on unevenly spaced points can alter the loading patterns significantly (Karl et al. 1982). To do this, we essentially ranked all stations by the proportion of missing data for each. We used 173 stations in the United States that had at least 96.6% data present. For Mexico, we used 136 stations that had at least 75.0% data present. While many of the Mexican stations had better records than this, some of the lower ranking stations were needed for better spatial coverage. Also, much of the missing Mexican data are after the 1985 Mexico City earthquake, following which record keeping was disrupted. No estimates of these missing data were made, as they tended to be missing in chronological blocks for adjacent stations. The lower data percentages are of some concern, as correlation matrices generated from data containing missing values can yield negative eigenvalues (Rummel 1970). We circumvented this problem to some degree by pairwise (rather than listwise) deletion across the matrix. For the application to

precipitation variability in the monsoon region, we used 1961 to 1990, 500 mb geopotential height data from the NCEP/NCAR reanalysis project (Kalnay et al. 1996), available from the National Center for Atmospheric Research.

## 3. METHODS

We applied S-mode PCA (multiple stations over time) using a correlation matrix with pairwise deletion as stated above. Use of the correlation matrix, as opposed to the covariance matrix, allows dry stations in the deserts to be directly compared to relatively wet stations in the mountains. Thus, stations with the same seasonal timing of monthly rainfall were correlated and grouped together, presumably because of the same atmospheric controls, even if absolute precipitation amounts differed simply because of the elevation effect. The selection of the number of principal components (PCs) to retain was relatively clear in this case, and based on scree tests (Cattell 1966), eigenvalues  $> 1.0$ , and the eigenvalue separation test (North et al. 1982), we retained 9 PCs (i.e. regions) explaining 65% of the variance in the data. White et al. (1991), in their comparison of rotation techniques, found that oblique rotations generally produced the best results for climate regionalization. Therefore, our retained components were subjected to direct oblimin (oblique) rotation to enhance interpretability, and to avoid Buell (1975, 1979) patterns. The value of  $\gamma$  controls obliquity. Negative values create less oblique solutions (greater orthogonality), and positive values result in more obliquity (greater collinearity) (Clarkson & Jennrich 1988). Initially, we used  $\gamma = 0$ , and in a later analysis we used  $\gamma = -0.2$ . For this regionalization method, no clustering algorithm is applied after the PCA. Instead, the rotated loading values ( $\gamma$ ) for each PC are plotted on a map. An interpolation algorithm can produce isolines, and regions can be produced from them using the  $\gamma = 0.4$  loading contours as boundaries (Richman & Lamb 1985, 1987, White et al. 1991). We also evaluated an alternative maximum-loading approach in which each station is assigned to the component upon which it loads most highly. Regions can then be drawn around stations assigned to the same PC. Both of these methods are used and compared. The choice of 9 retained components was confirmed via favorable comparison of mapped results between this and several alternative solutions we obtained, as well as to analyses of the United States and Mexican subsets alone.

Thus, in using oblique rotation, we are following the general recommendations of White et al. (1991) in applying this regionalization method to the U.S.-Mexico border region. However, in this case there are several

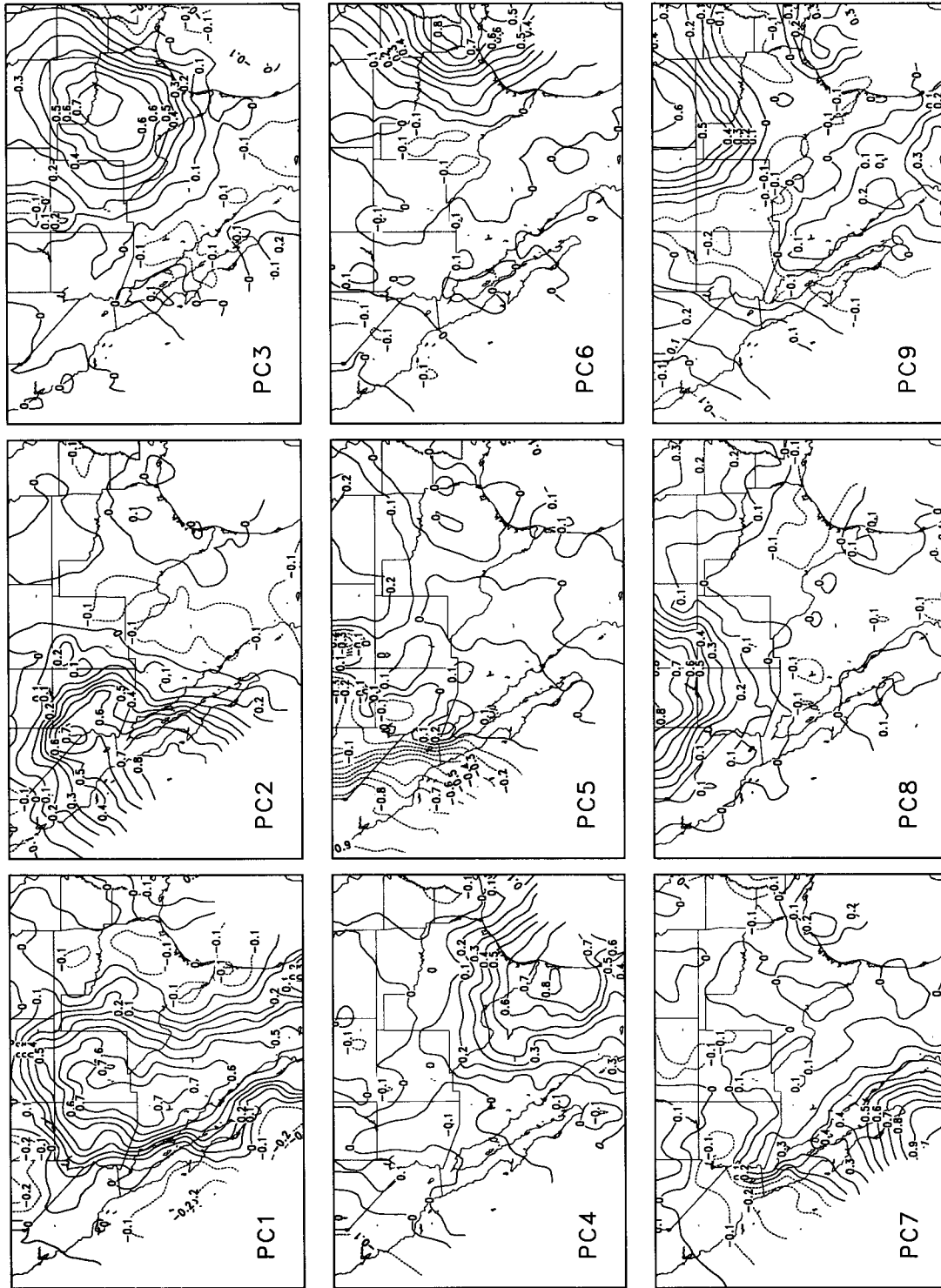


Fig. 3. Rotated loadings for the original 9 principal components (PC1 to PC9)

distinctions: the domain is an order of magnitude larger, comprising more than the 10 border-region states versus the state of Pennsylvania; the precipitation patterns are more complex with strong elevation and rain-shadow effects; the seasonal regimes are also more complex and are not limited to classic frontal mid-latitude systems; and the spatio-temporal data coverage, while the best available, is somewhat less than ideal.

Once the initial results were examined, we performed 4 subsequent analyses. The first and second of these were carried out as checks for robustness of the regions: an analysis of precipitation data transformed closer to normal, and a change in the obliquity of the rotation algorithm. The third analysis was a further sub-regionalization of the monsoon region to provide

an additional level of detail for this large area. This formed the basis for the fourth analysis, an examination of interannual monsoon precipitation variability, in which we performed a compositing analysis of the 500 mb circulation based on anomalous wet and dry summers in each of the monsoon sub-regions.

## 4. RESULTS AND DISCUSSION

### 4.1. Loadings maps

Fig. 3 shows the rotated loadings patterns for the 9 PCs, with PC1 representing the most common variance in the rotated solution, and PCs 2 through 9 representing successively less thereof. The 9 regions are each distinct, physically reasonable, and intuitively satisfying with respect to the climatology of the area. All regions have a clear core, with almost no confusing secondary focus areas. Fig. 4a shows the regions delineated using the  $\gamma = 0.4$  contour rule ( $-0.4$  for the nominally negative loadings in PC5), and Fig. 4b shows the regions delineated using the maximum loading rule described above. The results are essentially the same for both mapping methods; for the contour rule, overlapping lines indicate varying transition gradients across region boundaries, while the maximum loading rule is categorical. Depending on the application, these may be positive or negative attributes. Note that a small number of stations at the northern limit do not fall within any 0.4 contours and might be labeled as unclassified. However, the value of 0.4 is convenient but arbitrary, so all stations could be classified using 0.35, or far fewer stations classified using the 0.5 contour. All stations are classified under the maximum loading rule, and although a minimum loading threshold could be used, we did not do so. Because stations on the periphery of the study area by definition fall within climatic regions underrepresented in the data set, our concern is less with them than with those in the center, which were all well-classified by both rules. Because of applications to some future work, and for consistency in this paper, we show subsequent results using the maximum loading rule.

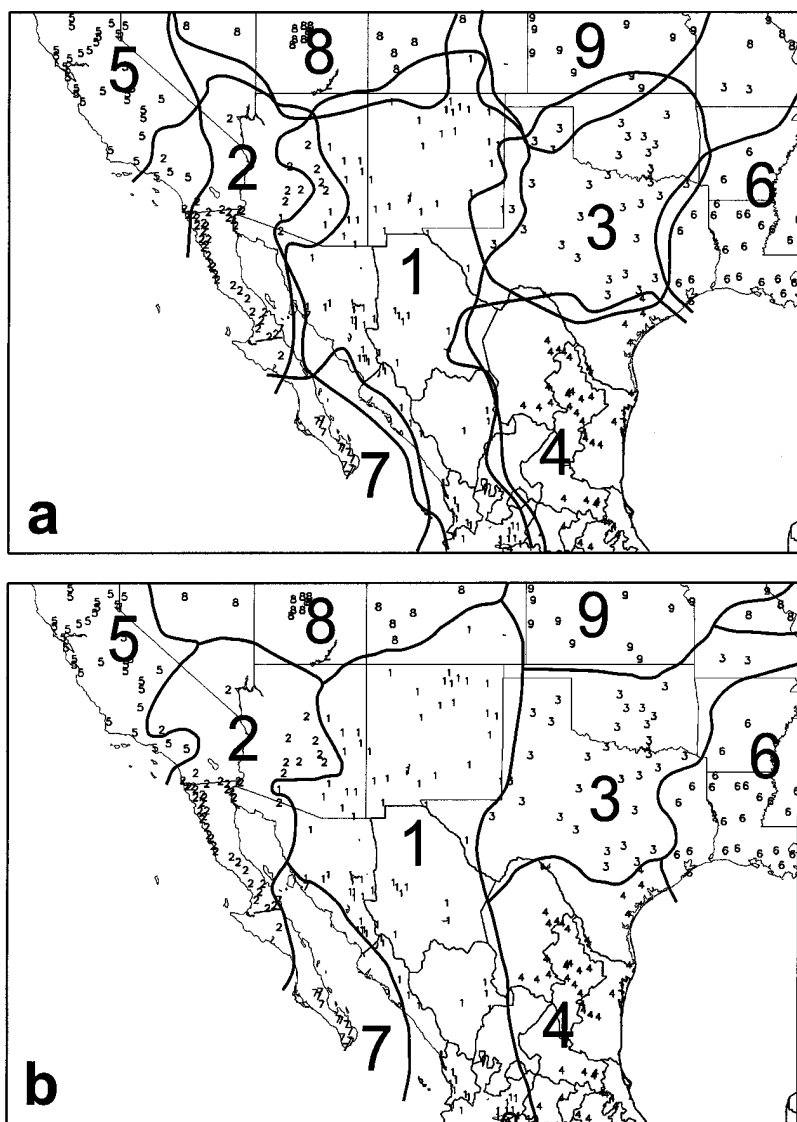


Fig. 4. The 9 regions delineated using (a) the  $\gamma = 0.4$  contour rule and (b) the maximum loading rule

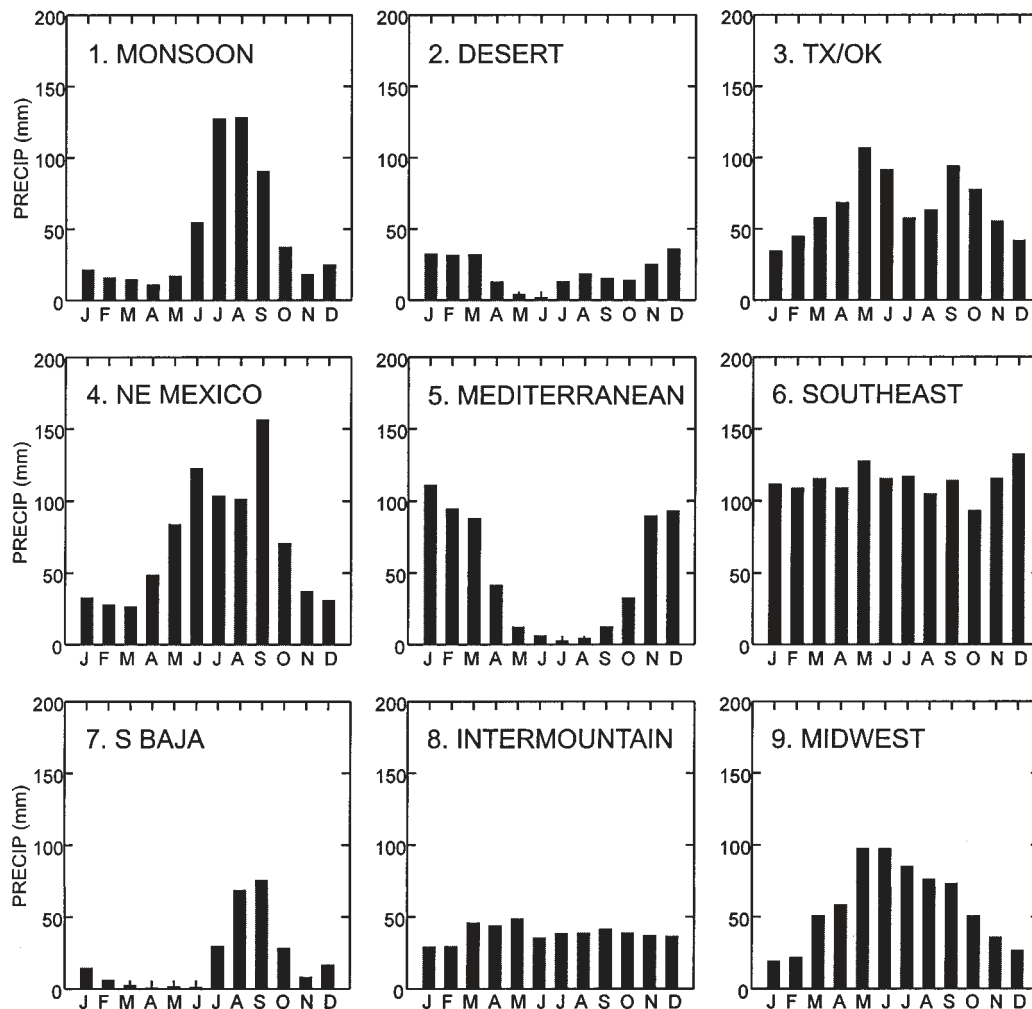


Fig. 5. Mean monthly precipitation of all stations in each of the regions

#### 4.2. Regions

To facilitate discussion of the regions, Fig. 5 presents the mean monthly precipitation across all stations in each of the 9 regions. Determining the spatial form of these 9 seasonality regimes was the primary purpose of the PCA. Recall that use of the correlation matrix focuses on the shape of the seasonal cycles over the study period and not the amount of precipitation. Region 1 is the monsoon region, which stretches along the Sierra Madre Occidental in NW Mexico and follows the Continental Divide through E Arizona and most of New Mexico in the U.S. This region is characterized by the strong mid- to late summer precipitation maximum with considerably less precipitation during the rest of the year. The western boundary appears to skirt the escarpment along the Gulf of California and then snake around the SE highlands and the Mogollon Rim of Arizona, closely mirroring the patterns of thun-

derstorm activity during the monsoon (Adams & Comrie 1997) and the first harmonic of precipitation (Horn & Bryson 1960). The northern and eastern boundaries likewise provide a reasonable match with precipitation-based definitions of the North American monsoon, and they include the extreme southern and western portions of Colorado and Texas respectively, running south along the Mexican *Altiplano* (highlands). The narrow southern boundary of the monsoon region was not resolved within the domain, and it most likely blends into the ITCZ (inter-tropical convergence zone)-related rainfall of central and southern Mexico. This region covers a large latitude range, and it is central to the climate of the entire borderlands, so we performed an additional analysis to identify sub-regions, described later in Section 4.5.

Region 2 is a desert region to the west of the monsoon region, and it encompasses areas surrounding the Mojave desert and the lower Colorado River valley,

including SE California, SE Nevada, W Arizona, far NW Sonora, and Baja California Norte. It is dry essentially year-round (approximately  $300 \text{ mm yr}^{-1}$ ) with marginally more precipitation in winter. The monsoon precipitation peak is notably absent, and the region is at the southernmost limit of mid-latitude frontal systems that supply the limited winter precipitation. The coastal ranges of S California and the Sierra Nevada appear to create the boundary with region 5 over California, which has a distinct wet winter and dry summer Mediterranean regime caused by a classic pattern of dominant westerlies in winter and the dominant north Pacific anticyclone and associated west coast ridge in summer. Region 7 covers Baja California Sur and a narrow coastal strip of the Mexican mainland just across the Gulf of California. The summer rainfall peak and low winter precipitation partly relate this region to the monsoon, but the rains peak in September rather than in July and August, possibly because of increased tropical storm activity (Adams & Comrie 1997, Stensrud et al. 1997).

To the east of the monsoon region, region 3 covers much of central Texas and Oklahoma, and it has dual early/late summer precipitation peaks in May and September. This seasonal regime is similar to that for region 4, covering NE Mexico and adjacent coastal areas in Texas, where the early summer peak is in June and the September peak is relatively a little higher (region 3 likely experiences greater frontal influence early and late in the shoulder seasons than region 4). Region 4 corresponds to the regime examined by Cavazos (1997). Both regimes are strongly controlled by the north Atlantic subtropical anticyclone and its interaction with regional air masses and the continental elevated mixed layer (Cavazos & Hastenrath 1990, Lanicci & Warner 1991, Cavazos 1997), and both experience the mid-summer dry spells known in Mexico as *caniculas*. The 2 maxima in regions 3 and 4 are partly associated with tropical storms and hurricanes over the Caribbean/Atlantic basin. Regions 6, 8, and 9 are peripheral to the border region, but they serve to reinforce the boundaries of the core border regions. Region 6 over E Texas, Louisiana, and the SE United States receives year-round precipitation, while regions 8 and 9 appear related to regimes beyond the study area over the Intermountain West and the Midwest United States.

Overall, the regions correspond well with the respective climatological controls on seasonality. Region 2, the desert region, is perhaps the least definite because of the intrinsically flat precipitation distribution with very little seasonal signal from low precipitation amounts (e.g. the Pacific coast of Baja California Norte might arguably be more Mediterranean in nature). Climatic controls in this region are likely a

marginal mix of those for the neighboring monsoon and Mediterranean regions. Nonetheless, most regions are quite distinct, and, because common seasonality within regions implies common controls, the set of regions provides justifiable, cohesive climatic units for study.

### 4.3. Transformed data

Pearson's correlation coefficient used in the PCA input matrix can be sensitive to non-normality (White et al. 1991). In practice, PCA seems quite robust to moderate departures from normality of input data, but given the unusual precipitation distributions in some parts of the study area (especially stations with many zero precipitation months in the record), it seemed prudent to evaluate the regionalization using data transformed somewhat closer to normal. We examined frequency distributions across stations and experimented with several transformations, including the gamma distribution that is bounded on the left by zero and positively skewed, and therefore used in some precipitation studies. However, a simple square root transformation of monthly precipitation ( $p$ ) appeared most useful when considered across all stations (i.e.  $p^{0.5}$ ). Fig. 6 shows frequency distributions for 4 example stations before and after the  $p^{0.5}$  transform (Tucson, Arizona; Albuquerque, New Mexico; Monterrey, Nuevo León; and Cabo San Lucas at the tip of Baja California Sur); following the transformation all distributions are moderately normal except the most extreme example (Cabo San Lucas). Using the transformed data as input, we performed the analyses exactly as before. Remarkably, perhaps, the appropriate number of PCs to retain was again 9, this time explaining 64.8% of the variance. The rotation (oblique,  $\gamma = 0$ ) and mapping (maximum loading rule) procedures were carried out as before. Fig. 7 shows the resulting 9 regions, in which the PC numbering has changed to reflect the new explained variances of each, but where the overall regions are quite similar to the original analysis.

The monsoon region is largely unchanged, except for the elimination of some stations in NE New Mexico and the appearance of a southern boundary. The latter adjoins a small, new peripheral central Mexico region. The desert, Mediterranean, and southern Baja regions are identical to the initial analysis. The boundaries of the NE Mexico/SE Texas region and the SE United States region are the same except for a few stations where the 2 regions meet. The core of the Texas/Oklahoma region is much the same, but along the northern edge of the study area there are several changes in boundaries, with an additional

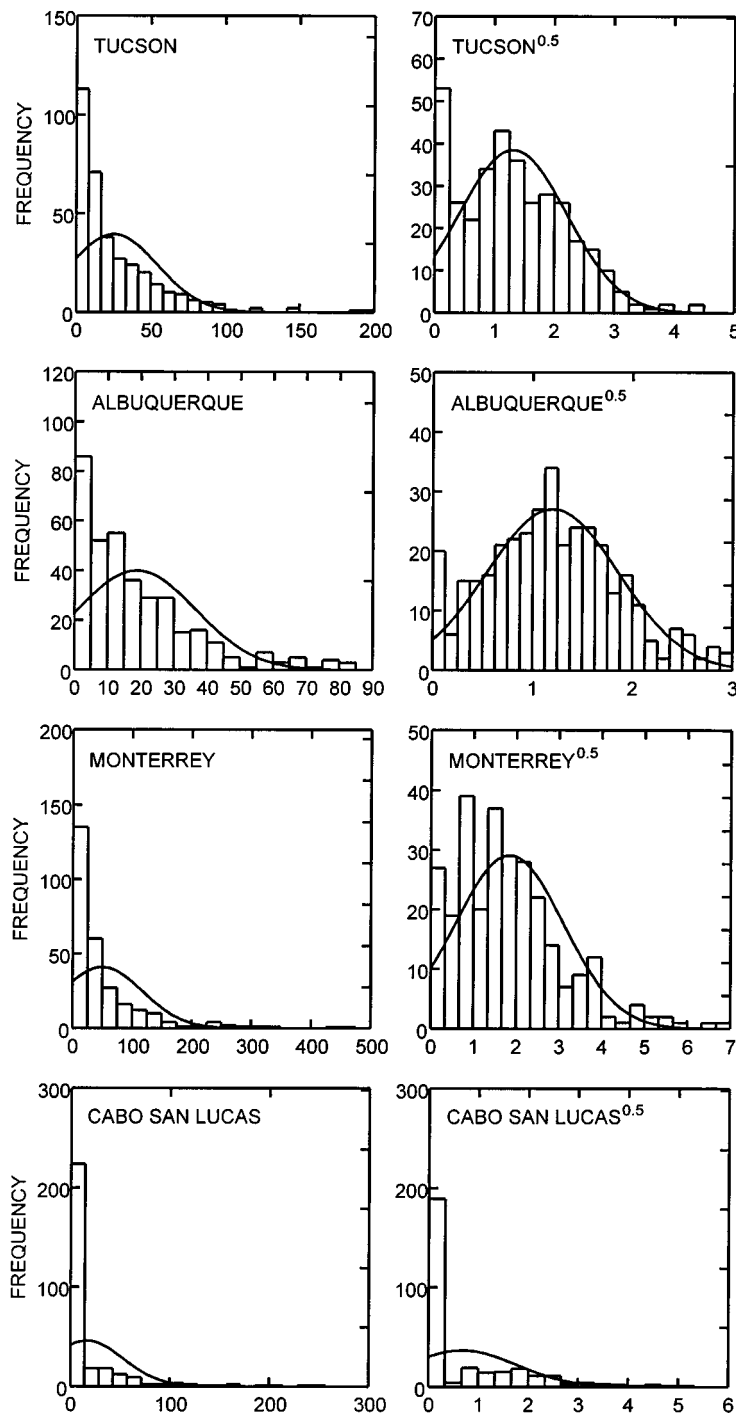


Fig. 6. Frequency distributions for 4 example stations, showing original monthly precipitation ( $p$ ) data (left) and square root transform ( $p^{0.5}$ ) data (right) with normal curves based on sample means and standard deviations (units in mm)

split region. Thus, overall, apart from edge effects, the central border precipitation regions are the same for data transformed using  $p^{0.5}$ .

#### 4.4. Changing obliquity

One of the main points made by White et al. (1991) in their comparison of PCA rotations was the greater stability of obliquely rotated solutions for regionalization (and hence the use of oblique rotation in this study). We used the direct oblimin algorithm because of its availability in our statistical package, but White et al. (1991) mention that it may result in slightly over-oblique solutions compared to some other algorithms. Their application of direct oblimin used  $\gamma = 0$ , the standard case, as we did above.

To evaluate a less oblique regionalization solution (approximating other algorithms), therefore, we performed direct oblimin rotation with a less oblique  $\gamma = -0.2$  on the original (untransformed) PCA results, and mapped the results using the maximum loading rule. Fig. 8 shows the 9 regions for the  $\gamma = -0.2$  solution, which again are consistent with previous patterns and show very similar numbering. Apart from a few individual station changes, the most notable difference in the monsoon region is the movement of the boundary with southern Baja, where the mainland coastal stations along the Gulf of California join the latter region. A new central Mexico region appears, as in the transformed data, with some concomitant minor changes in the boundary with the NE Mexico region. However, the new region is 1 of 2 disconnected parts, the other part being the original Midwest region (which added a few stations in the Texas panhandle and NE New Mexico). Other than these small differences, the SE United States, Texas/Oklahoma, desert, Mediterranean, and Intermountain West regions are basically unchanged, although the latter region has a smaller disconnected area (one station) over Missouri than before. Again, for the core border regions of principal interest, the results for changed obliquity are very similar to the initial analysis.

#### 4.5. Monsoon sub-regions

As mentioned earlier, the monsoon region covers a broad area over a large range of latitude from central Mexico to Colorado, as well as straddling the Continental Divide. To examine the region more closely, we replicated the overall analysis methodology for the 89 stations in region 1



(i.e. stations with maximum loading on PC1 in the initial analysis). Following S-mode PCA of the monthly precipitation data, 4 components were retained, explaining 72% of the variance. Again, the 4 PCs were subjected to direct oblimin rotation ( $\gamma = 0$ ), and the results were mapped.

Fig. 9 shows the 4 sub-regions, and Fig. 10 illustrates the mean monthly precipitation for each sub-region. Sub-region 1 lies within Mexico, sub-regions 2 and 4 straddle the international border and lie to the west and east of the Continental Divide respectively, while sub-region 3 covers NE New Mexico and southern Colorado. The Mexican sub-region (1) has the highest rainfall, as expected, with the summer rains initiating in June, peaking in July and August, and dissipating in October. This rainfall is due not only to the monsoon, but also to other tropical systems in the Caribbean/Atlantic basin mentioned earlier. This region is also affected by late-summer and early-fall tropical storms off the west coast of Mexico. The other sub-regions receive less rainfall each month over a slightly shorter season, being on the northern edge of the monsoon circulation and further from moisture sources. Sub-region 2 has a longer dry period in the early summer, with monsoon rainfall principally in July and August and a small winter precipitation signal. In sub-region 4 to the east, there is

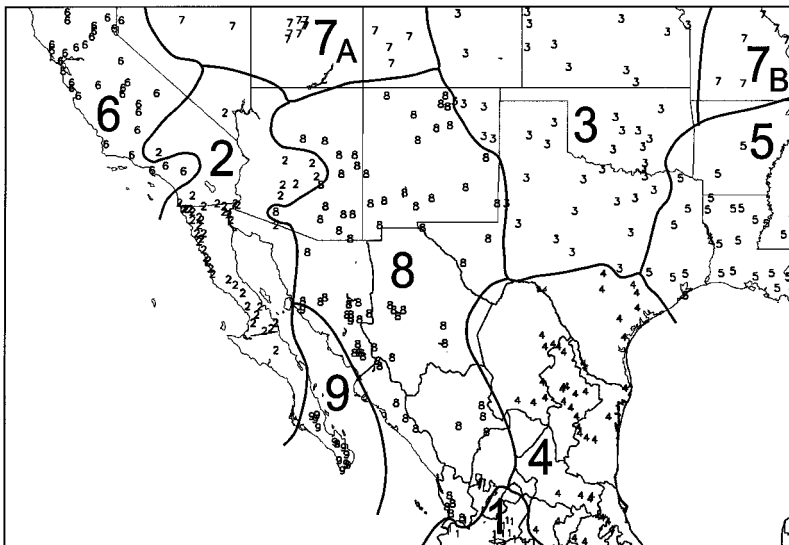


Fig. 7. The 9 regions from the square root transform ( $p^{0.5}$ ) analysis, delineated using the maximum loading rule

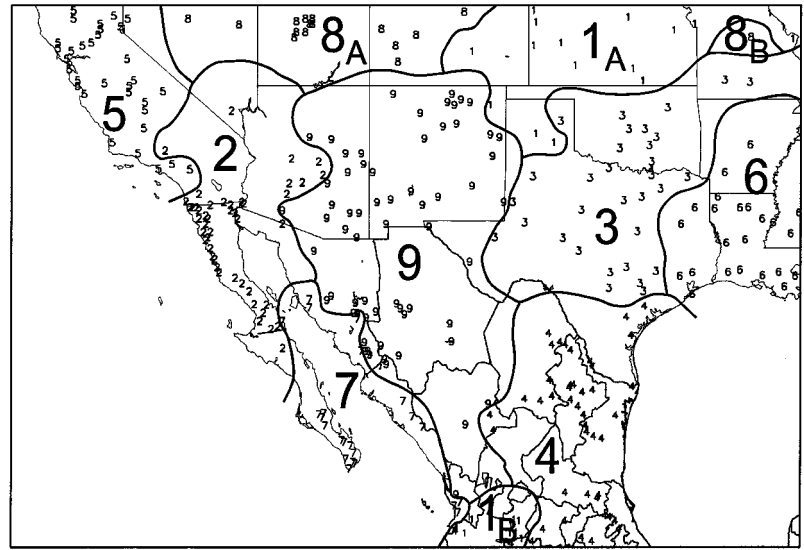


Fig. 8. The 9 regions from the  $\gamma = -0.2$  oblique rotation analysis, delineated using the maximum loading rule

some rainfall in May and June (early summer), but the peak summer rainfall falls between July and September. In sub-region 3 the early summer lacks the typical pre-monsoon dry period (perhaps a result of being the northernmost sub-region and under a more westerly circulation), but the mid- and late summer does have the characteristic July and August rainfall peak. These timing differences across W Texas, S New Mexico, E Arizona and neighboring parts of Mexico match well with our understanding of the onset of the North American monsoon, which begins generally in the east of the greater monsoon region and progresses westward (Adams & Comrie 1997).

#### 4.6. Monsoon sub-regional precipitation variability

The North American monsoon is central, geographically and intellectually, to understanding the climates of the southwestern United States and northern Mexico. It is known that important seasonal and interannual controls on monsoon variability occur at both the synoptic scale and the mesoscale, although the role of each and their interactions are not well explored. However, there is an ongoing paucity of mesoscale climate data, precluding analysis at that scale. At the synoptic scale, most studies have

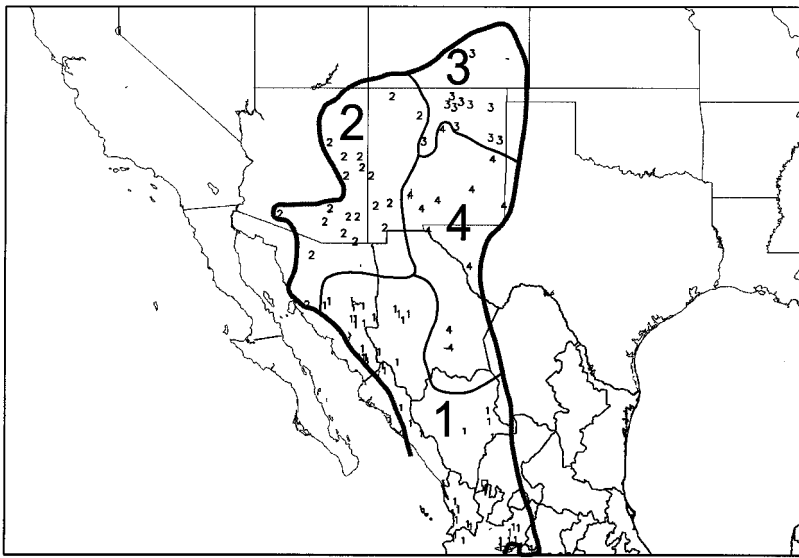


Fig. 9. The 4 monsoon sub-regions delineated using the maximum loading rule

focused on the monsoon margin that falls in the United States (Adams & Comrie 1997), often at the 500 mb level. Therefore, to extend such analyses over the entire binational monsoon region, and to illustrate application of the regionalization results, we examine relationships between precipitation variability in the monsoon sub-regions and the 500 mb circulation.

We calculated anomalies for total summer (July, August, and September) precipitation during the study period for each of the monsoon sub-regions. The standardized anomaly time series are shown in Fig. 11, representing interannual variability that ranges up to plus or minus 100 mm across all sub-regions. Because the mean summer precipitation in sub-region 1 (542 mm) is greater than in the other sub-regions (166 to 183 mm), this represents roughly a 20% variability in sub-region 1 and as much as 50% variability in sub-regions 2 through 4. While there are a number of coincident wet and dry summers between certain sub-regions, precipitation variability across the sub-regions can vary considerably (correlations range between  $-0.3$  and  $0.3$ ).

We selected the 7 driest and 7 wettest summers in each anomaly time series (representing roughly the 25th and 75th percentiles, arbitrarily) and composited the matching daily 500 mb pressure heights ( $n = 644$ ). Fig. 12 shows the mean

500 mb circulation patterns associated with the wet and dry summers for each sub-region. By subtracting the mean of the remaining 'middle' 15 summers ( $n = 1380$ ), we also calculated matching wet and dry anomaly fields and their significance. Fig. 13 illustrates these 500 mb anomalies for wet and dry summers, with shaded areas indicating values significant at  $\alpha \leq 0.05$ .

In monsoon sub-region 1 to the south over Mexico, wet summers exhibit anomalous ridging over northern monsoon regions and over the central United States, while in dry summers the high intensifies, leading to an anomalous west-east trending ridge directly over the sub-region. In sub-region 2, wet summers show an anomalous ridge over the western United States as the high expands over the sub-region and areas to the north-

west, but in dry summers the high expands over large areas to the west-southwest and east, with relatively strong anomalous ridging west of the Baja peninsula and over the eastern United States. Sub-region 3 experiences fairly broad expansion of the anticyclone in wet

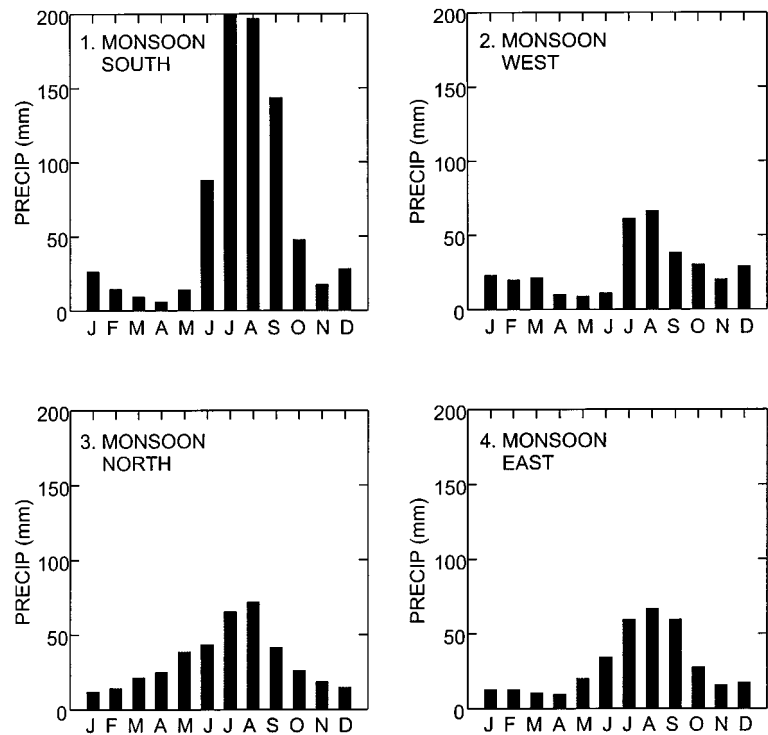


Fig. 10. Mean monthly precipitation of all stations in each of the 4 monsoon sub-regions

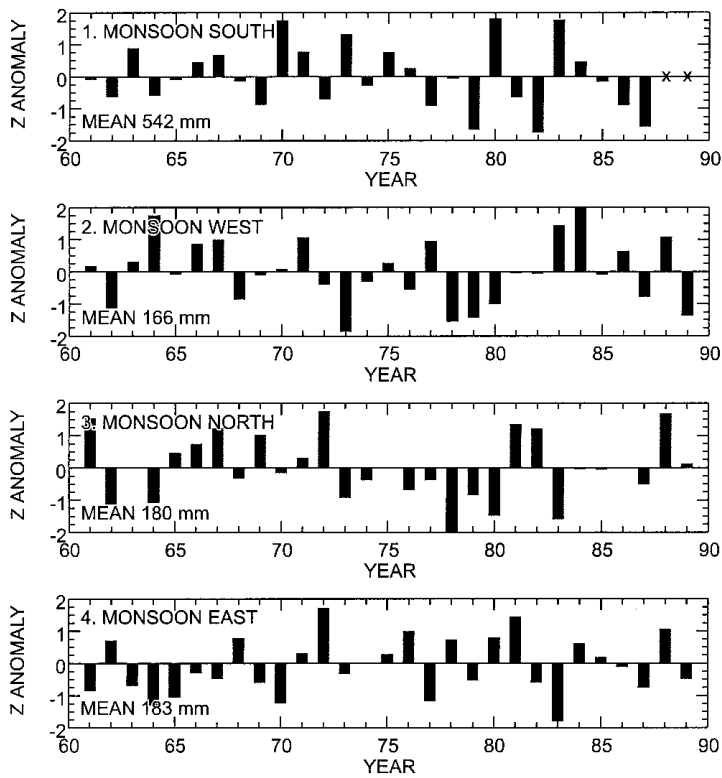


Fig. 11. Standardized anomalies of July-August-September precipitation totals in the 4 monsoon sub-regions (x denotes missing data)

summers, focused on an anomalous ridge to the northwest similar to sub-region 2, while in dry summers the anticyclone intensifies and causes a relatively strong west-east aligned anomalous ridge right over the sub-region. Sub-region 4 along the east of the monsoon region is the most different from the other sub-regions. In wet summers the entire anticyclone shifts east, leading to an anomalous ridge-trough-ridge pattern (with the trough partially over the sub-region), but in dry summers the ridge expands to the south and east of its typical position with an anomalous ridge moving over the sub-region and an anomalous trough to the northwest.

Recognizing that the 500 mb circulation reflects just one of many sources of monsoon variability, these maps nonetheless reveal important changes in the shape and intensity of the subtropical anticyclone related to monsoon precipitation variability. Generally, in wet summers, the anticyclone undergoes some form of expansion to the north with associated continental ridging (except sub-region 4) while maintaining its position over the southern part of the study region, showing meridional bulging of the subtropical circulation into mid-latitude continental areas. In dry summers, the anticyclone expands in various configura-

tions to the south with higher pressure-heights over a broad area, displaying some zonal stretching and possible intensification of the subtropical (and perhaps tropical) circulation. Overall, these results are consistent with, and extend earlier findings based on, Arizona precipitation alone (Carleton et al. 1990, Adams & Comrie 1997). There are, however, noteworthy sub-regional differences from these general patterns, such as those for sub-region 4.

In explaining these results, it is tempting to infer that dry summers have slightly increased subsidence, which decreases monsoonal convection and associated precipitation in a particular sub-region (with perhaps a contrasting analog for wet summers). Although plausible, this inference is probably too simplified, keeping in mind the complex terrain and the absence of climatological information on lower-level mesoscale monsoon dynamics such as 'Gulf surges' of moisture and low-level jets along the Gulf of California. Yet, there are significantly different 500 mb circulation patterns associated with wet and dry monsoons, and it appears that shifts in 500 mb circulation relative to the geographic position of each sub-region influence seasonal precipitation variability, directly or indirectly.

## 5. CONCLUDING DISCUSSION

This study has determined climate regions for the entire United States-Mexico border region based on seasonality and variability of precipitation (rather than precipitation totals). In doing so, we have demonstrated the applicability of the White et al. (1991) regionalization method to a large, diverse area at the sub-continental scale. The method performs well and quite consistently under various analyses and transformations. The resulting contiguous regions, sub-regions, and associated boundaries in the core of the study area make physical sense, and they are robust to transformed input data and changes in rotation procedures. Certainly, the results convey a congruous picture of precipitation regimes across the borderlands.

In some parts of the study area, improving the sparse spatial distribution of stations might offer some greater definition of region boundaries (e.g. Mojave desert region, SW Texas/northern Mexico). In areas where the data are good and behave coherently (including some of the sparser areas), the 2 techniques for delineating region boundaries map closely to one another

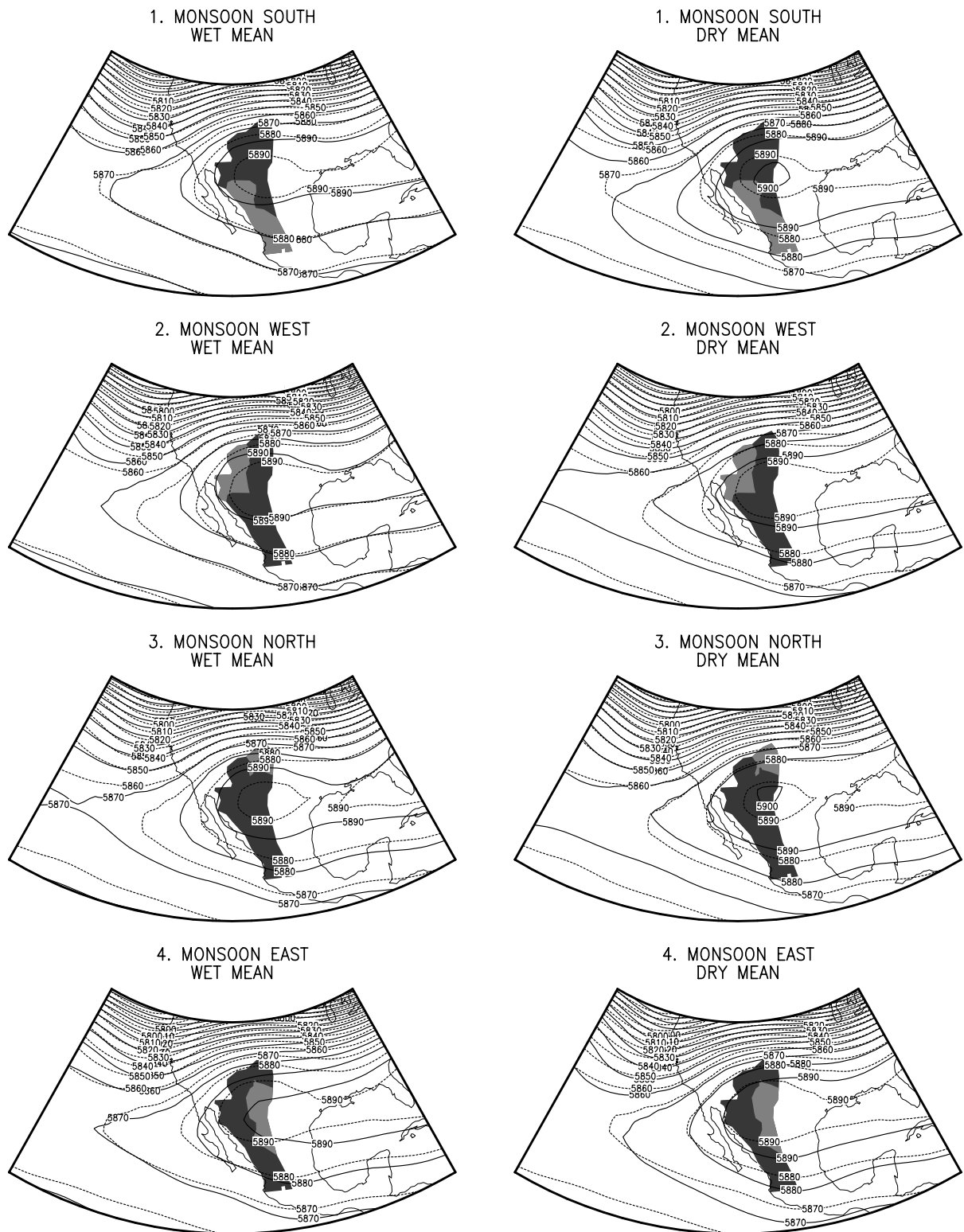


Fig. 12. Composite 500 mb heights (gpm) for the 7 wettest and 7 driest summers (solid lines) in the 4 monsoon sub-regions, as compared to remaining 15 'middle' years (dashed lines), with shaded areas indicating individual sub-regions (light) and the whole monsoon region (dark)



Fig. 13. The 500 mb height anomalies (gpm) for the data in Fig. 12. Shaded areas indicate significant differences for  $\alpha \leq 0.05$ . Monsoon region and appropriate sub-region are outlined on each map

and, as expected, they diverge where the data do not behave coherently. Yet, overall, the regions are remarkably consistent among analyses, and they provide a climatologically satisfying and useful division of border region precipitation climatology that relates well to the driving atmospheric processes.

The contiguous groups of stations identified in this study were created to enable analyses of regional variability, secular trends, and other signals of global change. To illustrate the utility of the regionalization we applied the regionalization results for the monsoon region to an examination of relationships between summer monsoon precipitation variability and 500 mb circulation. The results show significantly different patterns in the overlying pressure-height fields associated with anomalous wet and dry conditions in each of 4 monsoon sub-regions.

Currently, the regionalization results are being used for a follow-on study of summer and winter precipitation variability across the entire border region, and they are also being employed in an integrated assessment project for the southwest United States. Further work on monsoon variability should elucidate in more detail the potential controlling mechanisms, including among others the role of the Hadley circulation and the inter-tropical convergence zone in the shape and strength of the subtropical circulation over the region.

More generally, the results of our study give some indications about the use of this methodology for further applications in other regions. The success of the basic oblimin regionalization over Pennsylvania (White et al. 1991) and over this quite different region suggest that it has wide applicability to different climatological regimes. Yet it is all too easy to blindly apply such a recommendation when in reality there is no 'right' answer. A range of statistics always provides the researcher with greater insight. Therefore, for most studies we recommend also that several data and rotational transformations be examined to supply a measure of reliability to the regionalization results.

*Acknowledgements.* The authors thank David Adams, Connie Woodhouse, and the anonymous reviewers for their useful comments on the manuscript.

#### LITERATURE CITED

- Adams DK, Comrie AC (1997) The North American monsoon. *Bull Am Meteorol Soc* 78(10):2197–2213
- Barring L (1988) Regionalization of daily rainfall in Kenya by means of common factor analysis. *J Climatol* 8:371–389
- Barry RG, Kiladis G, Bradley RS (1981) Synoptic climatology of the western United States in relation to climatic fluctuations during the twentieth century. *J Climatol* 1:97–113
- Buell CE (1975) The topography of empirical orthogonal functions. 4th Conference on Probability and Statistics in Atmospheric Science, Tallahassee, FL. American Meteorological Society, Boston, p 188–193
- Buell CE (1979) On the physical interpretation of empirical orthogonal functions. 6th Conference on Probability and Statistics in Atmospheric Science, Banff, Alberta. American Meteorological Society, Boston, p 112–117
- Bunkers MJ, Miller JR, DeGaetano AT (1996) Definition of climate regions in the northern plains using an objective cluster modification technique. *J Clim* 9:130–146
- Burnett AW (1994) Regional-scale troughing over the southwestern United States: temporal climatology, teleconnections, and climate impact. *Phys Geogr* 15:80–98
- Cahalan RF, Wharton LE, Wu ML (1996) Empirical orthogonal functions of monthly precipitation and temperature over the United States and homogeneous stochastic models. *J Geophys Res* 101(D21):26309–26318
- Carleton AM (1987) Summer circulation climate of the American Southwest: 1945–1984. *Ann Assoc Am Geogr* 77:619–634
- Carleton AM, Carpenter DA, Weber PJ (1990) Mechanisms of interannual variability of the Southwest United States summer rainfall maximum. *J Clim* 3:999–1015
- Cattell RB (1966) The scree test for the number of factors. *Multivar Behav Res* 1:245–276
- Cavazos T (1997) Downscaling large-scale circulation to local winter rainfall in north-eastern Mexico. *Int J Climatol* 17:1069–1082
- Cavazos T, Hastenrath S (1990) Convection and rainfall over Mexico and their modulation by the Southern Oscillation. *Int J Climatol* 10:377–386
- Clarkson DB, Jennrich RI (1988) Quartic rotation criteria and algorithms. *Psychometrika* 53:251–259
- Comrie AC (1996) An all-season synoptic climatology of air pollution in the U.S.-Mexico border region. *Prof Geogr* 48(3):237–251
- Eklundh L, Pilesjö P (1990) Regionalization and spatial estimation of Ethiopian mean annual rainfall. *Int J Climatol* 10:473–494
- Fernández Mills G (1995) Principal component analysis of precipitation and rainfall regionalization in Spain. *Theor Appl Climatol* 50:169–183
- Fernández Mills G, Lana X, Serra C (1994) Catalanian precipitation patterns: principal component analysis and automated regionalization. *Theor Appl Climatol* 49:201–212
- Horn LH, Bryson RA (1960) Harmonic analysis of the annual march of precipitation over the United States. *Ann Assoc Am Geogr* 50:157–171
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds R, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77:437–471
- Karl TR, Koscielny AJ, Diaz HF (1982) Potential errors in the application of principal component (eigenvector) analysis to geophysical data. *J Appl Meteorol* 21:1183–1186
- Kutzbach JE (1967) Empirical eigenvectors of sea-level surface temperature and precipitation complexes over North America. *J Appl Meteorol* 6:791–802
- Lanicci JM, Warner TT (1991) A synoptic climatology of the elevated mixed-layer inversion over the southern Great Plains in spring. Part 1: Structure, dynamics, and seasonal evolution. *Weather Forecasting* 6:181–197
- Lyons WF, Bonell M (1994) Regionalization of daily mesoscale rainfall in the tropical wet/dry climate of the Townsville area of north-east Queensland during the 1988–1989 wet season. *Int J Climatol* 14:135–163

- Mallants D, Feyen J (1990) Defining homogeneous precipitation regions by means of principal components analysis. *J Appl Meteorol* 29:892–901
- Mitchell VL (1976) The regionalization of climate in the western United States. *J Appl Meteorol* 15:920–927
- North GR, Bell TL, Cahalan RF, Moeng FJ (1982) Sampling errors in the estimation of empirical orthogonal functions. *Mon Weather Rev* 110:699–706
- Ogallo LJ (1989) The spatial and temporal patterns of the east African seasonal rainfall derived from principal component analysis. *Int J Climatol* 9:145–167
- Richman M, Lamb PJ (1985) Climatic pattern analysis of 3- and 7-day summer rainfall in the central United States: some methodological considerations and a regionalization. *J Clim Appl Meteorol* 24:1325–1343
- Richman M, Lamb PJ (1987) Pattern analysis of growing season precipitation in southern Canada. *Atmos Ocean* 25(2): 137–158
- Rummel RJ (1970) Applied factor analysis. Northwestern University Press, Evanston, IL
- Stensrud DJ, Gall RL, Nordquist MK (1997) Surges over the Gulf of California during the Mexican monsoon. *Mon Weather Rev* 125:417–437
- Vose RS, Schmoyer RL, Steurer PM, Peterson TC, Heim R, Karl TR, Eischeid J (1992) The Global Historical Climatology Network: long-term monthly temperature, precipitation, sea level pressure, and station pressure data. ORNL/CDIAC-53, NDP-041, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN
- White D, Richman M, Yarnal B (1991) Climate regionalization and rotation of principal components. *Int J Climatol* 11: 1–25
- Woodhouse CA (1997) Winter climate and atmospheric circulation patterns in the Sonoran Desert Region, USA. *Int J Climatol* 17:859–873

*Editorial responsibility: Brent Yarnal,  
University Park, Pennsylvania, USA*

*Submitted: June 24, 1998; Accepted: August 19, 1998  
Proofs received from author(s): October 19, 1998*