Annual and Warm Season Drought Intensity–Duration–Frequency Analysis for Sonora, Mexico

MICHELLE HALLACK-ALEGRIA AND DAVID W. WATKINS JR.

Department of Civil and Environmental Engineering, Michigan Technological University, Houghton, Michigan

ABSTRACT

Located in northwestern Mexico, Sonora is a region affected by the North American monsoon (NAM). The region covers nearly 50% of the North American Sonoran Desert and is characterized by climatic conditions ranging from extremely arid to semiarid. The region has suffered from drought since 1995, and consequently, water supplies are threatened. The objectives of this work are to characterize the spatial and temporal variabilities of precipitation in Sonora and to conduct a meteorological drought intensity–duration–frequency analysis based on annual and warm season precipitation records. Monthly precipitation data are compiled from 76 meteorological stations located in Sonora, along with 19 stations in the neighboring American state of Arizona, for the period 1961–2004. For increased reliability, data are pooled within five plausible climatic regions. Among the results reported herein are summaries of precipitation variability, drought frequency estimates for annual and seasonal durations and return periods of 10–100 yr, and an estimate of the return period of the most recent multiyear drought.

1. Introduction

Water scarcity is a reality in many parts of the world today. While many live in regions affected by endemic drought, others face droughts on an irregular basis and, therefore, may be less prepared for times of water scarcity. Growing population and contamination of surface and groundwater supplies further increase the need for water resources to be managed effectively to avoid severe economic, social, and environmental consequences. In the case of Sonora, Mexico (Fig. 1), approximately 95% of the region is considered arid or semiarid and is characterized by high temperatures and lack of precipitation. Due to low or irregular streamflows, groundwater supplies are very important for Sonora's agricultural and industrial activities. However, the overexploitation of groundwater and the lack of aquifer recharge along the Gulf of California coast have caused a drop in water levels and the intrusion of saltwater (Moreno 1994). Due to a lengthy drought starting in 1995, the growing population has relied almost solely on groundwater withdrawals, which has led to significant drawdown of water levels. As a result, Sonora is potentially facing one of the most severe water conflicts in its history.

A better understanding of climate variability and drought intensity–duration–frequency relationships is warranted and would likely benefit farmers, planners, and water managers in the State of Sonora. Much of the state is located within the Sonoran Desert, which covers approximately 260 000 km^2 (100 000 mi^2) and includes most of the southern half of Arizona, southeastern California, most of the Baja California peninsula, and the islands of the Gulf of California. Low annual precipitation, low cloud cover, and year-round warm temperatures over most of the state are due mainly to a subtropical high pressure ridge that exists for much of the year over the region. However, since the state is located between the midlatitude and subtropical atmospheric circulation regimes, climatic variability can result from shifts in these regimes. Orographic effects, and the state's proximity to the Pacific Ocean and Gulf of California, also play a role in the temporal and spatial variabilities of climate across the region (Sheppard et al. 2002).

Average annual precipitation in Sonora ranges from 15 cm (6 in.) in the coastal zone to 100 cm (40 in) in the

Corresponding author address: David W. Watkins Jr., Dept. of Civil and Environmental Engineering, Michigan Technological University, Houghton, MI 49931.
E-mail: dwatkins@mtu.edu

Fig. 1. Location of Sonora, Mexico.

mountain regions. Most of the area has a biseasonal rainfall pattern, with a significant portion of the annual precipitation attributed to the North American monsoon (NAM), occurring in the summer (July–September). The NAM is the northernmost portion of a more extensive region of heavy precipitation that first develops over southern Mexico during the spring and then spreads northward along the western slopes of the Sierra Madre Occidental (Douglas et al. 1993; Stensrud et al. 1995). The monsoon circulation brings warm, humid air from the south, and short-lived thunderstorms result from convection as the air is lifted by orographic effects. Along with the summer monsoon, Pacific Ocean tropical storms can also influence Sonora warm season rainfall. In contrast, winter (November–April) precipitation is generally associated with relatively long-lived frontal systems that approach Sonora from the Pacific Ocean.

Along with the prevalence of the NAM, Sonora's climate is characterized by a high degree of intraseasonal and interannual variability. The NAM season is noted for having periods of heavy thunderstorm activity with intervening dry periods, associated with atmospheric instability resulting from intense surface heating and high topographic relief. Conversely, a decrease in convective precipitation can result from a northward shift in the subtropical high pressure ridge (Carleton 1986). As discussed by Higgins et al. (1998), the total depth of warm season precipitation is correlated with the onset of the NAM, such that years with early onset typically have above-average total rainfall and vice versa. Winter climate variability is mainly associated

with atmospheric circulation patterns that affect the movement of cyclonic storms over the region. For instance, above-average precipitation can be linked to the Pacific–North American teleconnection pattern, characterized by a meridional (sinuous) flow pattern and a southward shift of westerly storm tracks (Woodhouse 1997).

With respect to interannual variability, there is evidence that winter circulation patterns are strongly affected by Pacific sea surface temperatures (SSTs), with El Niño events resulting in wet winters and La Niña events in dry winters (Kiladis and Diaz 1989). Warm season teleconnections do not appear to be as strong, but Higgins and Shi (2000) suggest that interannual variability in the NAM in the southwestern United States is modulated by long-term fluctuations in Pacific SSTs, and furthermore it is known that the Pacific decadal oscillation has the potential of amplifying ENSO effects. McPhee et al. (2004) further suggest that persistence in Pacific SSTs leads to year-to-year fluctuations in climate that can be embedded within multiyear periods during which the duration and intensity of dry or wet conditions remains above or below average. Thus, droughts (multiyear dry periods) may be a normal and expected phenomenon in Sonora.

Two major droughts, in terms of duration and spatial extent, occurred in North America in the twentieth century. Both the drought of the 1930s, which lasted up to 7 yr in the Great Plains, and the drought of the 1950s, which lasted 5 yr in the southwestern United States, caused significant damages in both human lives and economic losses. These droughts also affected Mexico mainly in the northern and central regions of the country (Tinajero 1986). In northwestern Mexico, extended droughts in the middle of the 1970s and the second half of the 1990s produced significant damages to crops and cattle (Kim and Valdes 2002).

The objectives of the present study are 1) to compile and screen monthly precipitation data from gauges in Sonora and southern Arizona, 2) perform a regional analysis of drought intensity and frequency on annual and seasonal time scales, and 3) estimate the return period of the recent 1995–2004 drought. In applying extreme value theory to the precipitation data, this work adds to previous studies of monthly mean precipitation patterns (e.g., Douglas et al. 1993) and factors responsible for interannual variability of the NAM (e.g., Higgins and Shi 2000). In addition, we present analyses of seasonal drought predictability and discuss ways in which these results may be used by water managers and other decision makers in Sonora.

2. Data and methods

Due to the far-reaching effects of drought, in a wide range of social, environmental, and economic contexts, and due to the range of spatial and temporal scales associated with drought, it is difficult to develop a concise definition of drought or an index to measure it. All definitions seem to agree that drought is a condition of insufficient moisture caused by a deficit in precipitation over some time period (McKee et al. 1993). Operational definitions of drought may be classified in four interrelated categories: meteorological drought, hydrological drought, agricultural drought, and socioeconomic drought. Meteorological drought is usually defined by the degree of dryness and the duration of the dry period. Hydrological drought is associated with the effects of periods of precipitation shortfalls on surface or subsurface water supply (i.e., streamflow, reservoir and lake levels, groundwater). Agricultural drought links various characteristics of meteorological and hydrological drought to agricultural impacts. Socioeconomic drought occurs when the demand for an economic good exceeds supply as a result of a weather-related shortfall in water supply (National Drought Mitigation Center 2005).

The focus of this study is on meteorological drought, expressed on the basis of the degree and duration of precipitation deficits. Meteorological drought is frequently described in terms of drought indices, which are convenient and relatively simple to use. Various drought severity indices based on precipitation data have been introduced in the literature. Gibbs and Maher (1967) used deciles of precipitation for characterizing droughts in Australia. Monthly (or annual) precipitation totals are ranked from highest to lowest, and decile ranges are determined from the cumulative frequency distribution. McKee et al. (1993) introduced the standardized precipitation index (SPI), a probability-based index designed to quantify the precipitation deficit for multiple time scales (Shin and Salas 2000).

This study uses a probability-based approach in which either percentile ranges or SPI values may be readily computed. The methodology is adapted from the regional *L*-moment algorithm, as described in detail by Hosking and Wallis (1997). This algorithm has been applied recently to the analysis of high-intensity rainfall (e.g., Sveinsson et al. 2002; Trefry et al. 2005) and floods (e.g., Kumar and Chatterjee 2005), and an early version of the algorithm was used to develop the U.S. *National Drought Atlas* (U.S. Army Corps of Engineers 2005). The methodology involves pooling data within climatic zones, choosing a frequency distribution for
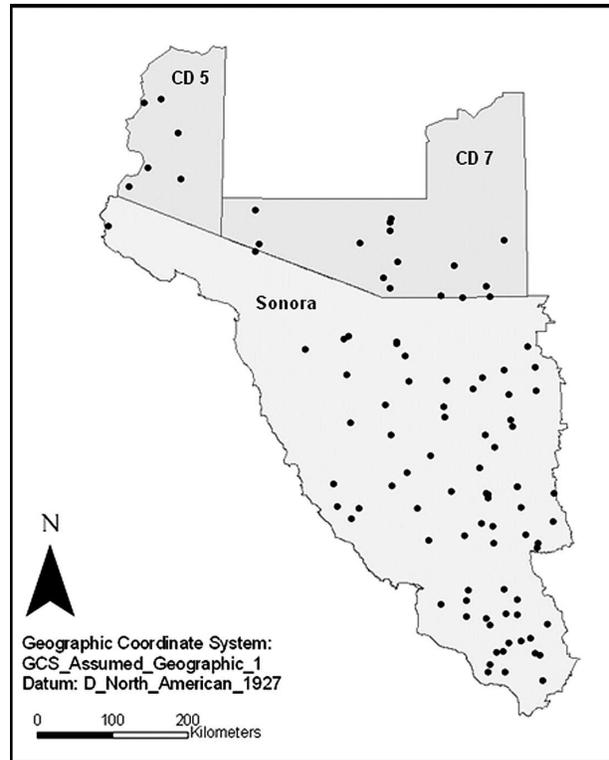

FIG. 2. Location of the precipitation stations used in this study: Sonora, Mexico, and Climatic Divisions 5 and 7 of AZ. In the United States, climatic divisions are considered regions in a state that are reasonably homogeneous with respect to climate and hydrologic characteristics (Sheppard et al. 2002).

each region, and estimating quantiles of the frequency distribution.

a. Data

Although climate is of course continuous across political boundaries, regional analyses of precipitation have most often treated the United States and Mexican sides separately (Comrie and Glenn 1998). In this investigation, the region analyzed comprises the entire State of Sonora and parts of Climatic Divisions 5 and 7 in Arizona (Fig. 2). Most of the monthly precipitation data for Sonora were extracted from the Extractor Rápido de Información Climatológica (ERIC II) database (Quintas 2000) provided by the Mexican Institute of Water Technology (IMTA). Additional data from stations mainly located in the Yaqui River basin were obtained from the National Commission of Water (CNA) in Sonora. Data availability for the stations in Sonora varies, with most of the data being available from 1960 to 1995. In Arizona, data were acquired for 19 stations from the Western Regional Climate Center. These records were selected based on the weather sta-

tion's location, precipitation patterns, and the period of record, and the 19 sites selected were those deemed to have rainfall patterns similar to the Sonora sites. For instance, some gauges in Arizona have a record of zero in annual precipitation, which did not occur for any of the stations in Sonora, so those sites were omitted from the analysis. Further discussion of pooling sites in a regional analysis is included in the next section.

A given site's precipitation records are included in the frequency analysis if the following conditions are met: 1) there is at least 15 yr of data at the site and 2) the records are consistent spatially. A total of 97 meteorological stations meet these criteria, 78 in Sonora and 19 in Arizona, with records covering the period 1961–2004. Thus, the total maximum record length is 44 yr, and the sample size for all of the stations varies from 15 to 44 yr. It is not uncommon for stations to have gaps in their reporting records.

b. Precipitation regions

Sonora is characterized by great variations in topography, diverse landscapes, and arid or semiarid climate. Therefore, it is essential to derive regionally consistent precipitation zones for use in a variety of practical applications and in climate impact research (Giddings et al. 2005). The regional drought frequency analysis procedure presented herein was initiated by identifying quasi-homogeneous climate zones.

The initial formation of precipitation regions was made by identifying clusters based on the following four site characteristics: latitude, longitude, elevation, and mean annual precipitation. Initial clustering did not involve at-site statistics indicating the shape of the frequency distribution of precipitation. This step did, however, involve screening of the data for gross errors and inconsistencies to be eliminated, including checks that the data appear to be homogeneous over time. A statistical discordancy measure was then used to identify sites that appear to be grossly dissimilar to the group as a whole. The discordancy measure is defined as follows (Hosking and Wallis 1997).

With N sites in the group, let $\mathbf{u}_i = [t^{(i)} t_3^{(i)} t_4^{(i)}]^{\mathrm{T}}$ be a vector containing the L-moments t, t_3, and t_4 parameter values of site i, where the superscript T denotes the transposition of a vector or matrix. Then compute the group average as

$$\bar{\mathbf{u}} = N^{-1} \sum_{i=1}^{N} \mathbf{u}_i,$$

and define the matrix of sums of squares and cross products,

TABLE 1. Critical values for the discordancy statistic $\mathbf{D}_i$ (Hosking and Wallis 1997).

No. of sites in region	Critical value	No. of sites in region	Critical value
5	1.333	10	2.491
6	1.648	11	2.632
7	1.917	12	2.757
8	2.140	13	2.869
9	2.329	14	2.971
		≥ 15	3.000

$$\mathbf{A} = \sum_{i=1}^{N} (\mathbf{u}_i - \bar{\mathbf{u}})(\mathbf{u}_i - \bar{\mathbf{u}})^{\mathrm{T}}.$$

The discordancy measure of site i is then

$$\mathbf{D}_i = \frac{1}{3} N(\mathbf{u}_i - \bar{\mathbf{u}})^{\mathrm{T}} \mathbf{A}^{-1}(\mathbf{u}_i - \bar{\mathbf{u}}).$$

Site i is said to be discordant if $\mathbf{D}_i$ is large, where "large" depends on the number of sites in the group. Hosking and Wallis (1997) suggest that a site be regarded as discordant if its $\mathbf{D}_i$ value exceeds the critical value given in Table 1.

In this study, any sites with discordance greater than 3 were closely inspected for errors such as missing months in an annual total. If the data were deemed to be erroneous, either that year was removed from the record, or the site was removed from the zone.

Following deletion of discordant sites, the remaining sites were grouped into regions, preferably but not restricted to being geographically contiguous. In an iterative procedure, the homogeneity of each proposed region was tested by calculating summary statistics of the at-site data and comparing the between-site variability of these statistics with what would be expected of a homogeneous region. Specifically, Ward's algorithm (Ward 1963), with the k-means algorithm (Hartigan and Wong 1979), was used to define regions, and L-moment variability statistics proposed by Hosking and Wallis (1997) were used to test the heterogeneity of each proposed region. A zone was considered to be quasi-homogeneous if the L-moment variability statistics were less than 2.0.

c. Spatial and temporal variability

Topographic differences play a major role in the spatial patterns of precipitation in Sonora, and temporal variations in climate are mainly influenced by hemispheric atmospheric circulation patterns (Sheppard et al. 2002). The southwestern United States and northwest Mexico are subject annually to an inflow of atmospheric moisture in association with the North Ameri-

can monsoon, which is a summertime shift in the atmospheric circulation stretching from the Caribbean Sea, to Mexico, and into the southwestern United States. Across Sonora the seasonal shift from winds with generally a westerly component in winter and spring to winds with more of a southerly component typically establishes itself in early July and persists through mid-September. Convective instability associated with surface heating of the moist air combined with orographic uplift produces frequent convective precipitation events often associated with intense rainfall, lightning, hail, and damaging winds (e.g., McCollum et al. 1995). As much as 50%–70% of the annual rainfall across Sonora results from thunderstorms generated during the summer monsoon season (Carleton et al. 1990; Douglas et al. 1993; Higgins et al. 1997; Mitchell et al. 2002; Sheppard et al. 2002).

Various analyses were performed in order to evaluate the spatial and temporal variabilities in the monthly gauge precipitation data for Sonora. Standardized precipitation index (SPI) values were computed for each gauge (McKee et al. 1993), and time series of annual precipitation totals and SPI values were plotted for each of the identified climate regions. Correlation with the El Niño–Southern Oscillation (ENSO) was also investigated. Kriging methods were used to map average annual precipitation over the entire state, and correlation analysis was performed to investigate the spatial coherence of precipitation across climate regions.

d. Frequency distributions

Following the regional L-moment algorithm (Hosking and Wallis 1997; Trefry et al. 2005), various frequency distributions were screened for an appropriate fit to the data in each climate region. This regional drought frequency analysis differs in several important ways from traditional at-site analysis. Regional frequency analysis aims to overcome sampling variability associated with short record lengths by "trading space for time," that is, using data from nearby or comparable sites to derive drought frequency estimates for any given site in a homogeneous region (Stedinger et al. 1993). As there are more data included in a regional drought analysis, it is more likely that there will be greater accuracy in the estimated drought frequency values, and less chance for data corruption from inaccuracies in measurements. Furthermore, it has been found that if a region is even moderately heterogeneous, regional analysis will still produce more accurate quantile estimates than an at-site analysis (e.g., Lettenmaier et al. 1987; Hosking and Wallis 1988; Hosking and Wallis 1997). The L-moment algorithm is also characterized by the use of L-moment estimators instead of method-of-moment or maximum likelihood estimators. The L-moments are defined as expectations of certain linear combinations of order statistics (Hosking 1990). Similar to standard moments, they measure location (mean), scale (standard deviation), and shape (skewness and kurtosis), but unlike standard moments, they are linear combinations of the ranked observations and are, therefore, less sensitive to outliers.

The frequency distributions considered included the generalized logistic (GLO), generalized extreme value (GEV), Pearson type III (PE3), and generalized normal (lognormal, LN3) distributions. Additionally, the gamma (GAM) distribution was considered, as it is a special case of the PE3 distribution with one parameter value held constant. In general, three-parameter distributions are preferred to two-parameter distributions, as estimates of extreme quantiles can be severely biased if the shape of the tail of the true frequency distribution is not well approximated by the fitted distribution. It has been found by several authors that frequency distributions of hydrologic data tend to be heavy tailed (e.g., Houghton 1978; Landwehr et al. 1979). When the parameters of three-parameter distributions are estimated accurately, these distributions yield less biased estimates of extreme quantiles than their two-parameter counterparts (Hosking and Wallis 1997).

The following goodness-of-fit statistic, the Z statistic, was one criterion used to select a distribution (Hosking and Wallis 1997):

$$Z^{\text{DIST}} = (\tau_4^{\text{DIST}} - t_4^{\text{R}} + B_4)/\sigma_4,$$

where τ_4^{DIST} is the theoretical L kurtosis for the candidate distribution, t_4^{R} is the regional average L kurtosis weighted by record length, B_4 is a simulation-based value to correct for the bias associated with estimating t_4^{R}, and σ_4 is an estimate of the standard deviation of t_4^{R} obtained from repeated simulation of a homogeneous region whose sites have the candidate frequency distribution and the same record lengths as the observed data. This statistic measures how well the theoretical L kurtosis of the fitted distribution matches the regional-average L kurtosis of the observed data. The fit of the distribution is declared satisfactory if $|Z^{\text{DIST}}| \leq 1.64$, which corresponds to the failure to reject the hypothesized distribution at a confidence level of 90% (Hosking and Wallis 1997). It is recognized that use of a single statistic to select a frequency distribution could give misleading results. For this reason a qualitative (visual) assessment of the goodness of fit, particularly in the lower tail of the distribution corresponding to drought events, is also used.

e. Parameter and quantile estimates

In estimating frequency distribution parameters for each climate region, it was convenient to first scale the data from each gauge by the at-site mean. Sample L-moment ratios for each site were then weighted according to record length and combined to give regional average L-moment ratios, and these ratios were used to estimate the regional distribution parameters and quantiles (Hosking and Wallis 1997). After regional quantile estimates were calculated, drought intensity estimates, $\hat{Q}_i(T)$, were computed at site i for the T-year drought using the following equation:

$$\hat{Q}_i(T) = \lambda_1^{(i)}\hat{q}(T).$$

where $\hat{q}(T)$ is the regional quantile estimate and $\lambda_1^{(i)}$ is the site-specific scaling factor, conveniently defined as the at-site mean. It should be noted that the at-site mean itself is subject to sampling error, and thus to smooth sampling variability and determine the drought frequency estimates for any location within the zone, the at-site mean values may be spatially interpolated or smoothed using geostatistical methods (e.g., Watkins et al. 2005).

The above procedure was carried out for both annual and seasonal durations, with a warm season of May–October and a cold season of November–April.

f. Multiyear drought frequency

Multiyear drought analysis may be made based on single-site data (Yevjevich 1967; Dracup et al. 1980) or multisite data (Tase 1976; Santos et al. 1983; Guttman et al. 1992; Soule 1992), depending on the purpose of the study at hand. Shin and Salas (2000) proposed to analyze and quantify the spatial and temporal patterns of multiyear meteorological droughts based on annual precipitation data. As the focus of this study is on regional drought, our analysis was based on annual data for the multiple sites in each climate region. Following the method of Fernàndez and Salas (1999), the return period of the 1995–2004 drought was estimated based on the regionally averaged annual SPI values, along with estimates of the interannual correlation in precipitation. This method assumes a multiyear drought to be a series of annual droughts occurring in a sequence of Bernoulli (binary) trials.

3. Results

a. Precipitation regions

Following the deletion of one discordant site, the remaining sites were assigned to two or more regions, or
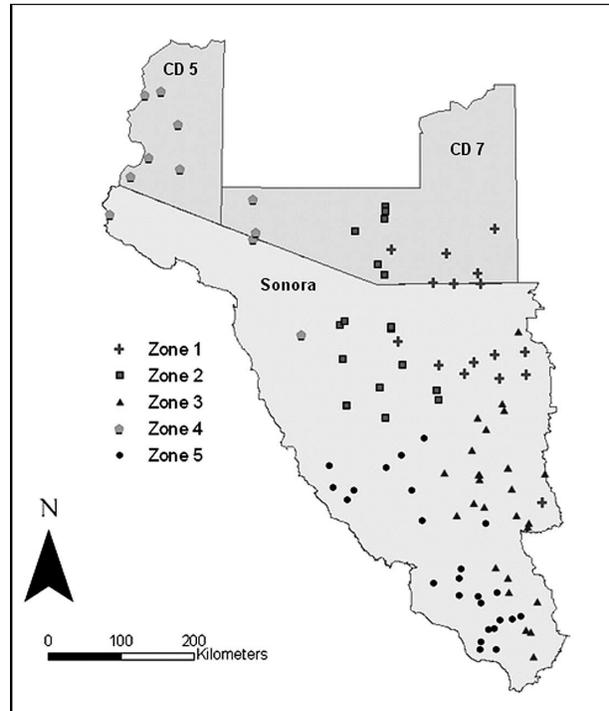

FIG. 3. Proposed precipitation regions.

zones, using cluster analysis algorithms. In each case, some minor adjustments were made to create nearly contiguous zones, which were then tested for statistical homogeneity. After testing different numbers of zones, five zones were deemed sufficient to improve the reliability of drought frequency estimates while adequately representing spatial variability across the state. The five zones are shown in Fig. 3, and the heterogeneity measures for each zone are shown in Table 2. As mentioned, values larger than 2.0 indicate that the zones are not strictly homogeneous, and hence are termed "quasi-homogeneous." While more zones could have been defined in order to ensure statistical homogeneity, a trade-off exists in greater sampling variability and reduced reliability in drought frequency estimates.

The subjectivity of this procedure was minimized by using objective cluster analysis algorithms, but judg-

TABLE 2. Heterogeneity statistics for each region. Bold indicates values corresponding to acceptably homogeneous regions.

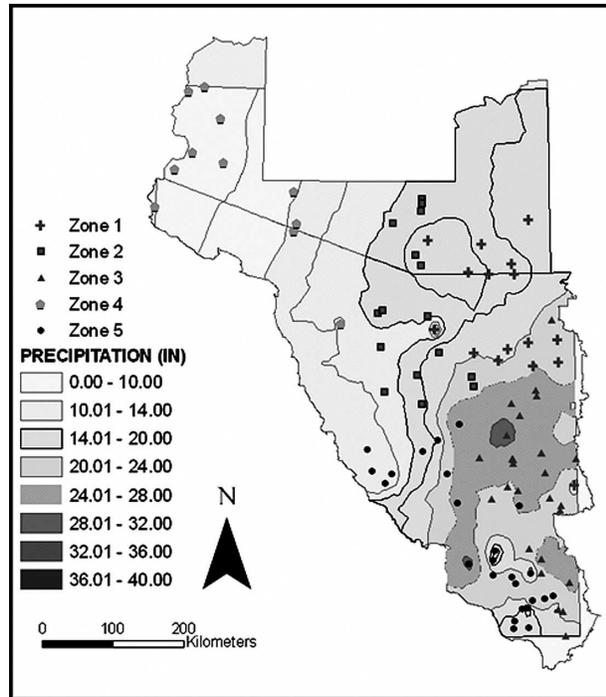| Zone	L-CV	L skew	L kurtosis
1	3.70	2.33	**1.78**
2	2.80	**0.23**	**0.32**
3	2.05	**1.17**	**1.90**
4	2.21	**0.22**	**0.44**
5	3.20	**0.74**	**0.72**

FIG. 4. Average annual precipitation in Sonora and southern AZ, 1961–2004.

ment was required to select five zones as the best number to represent spatial variability while achieving reliable frequency estimates. Also, there is not a unique grouping of sites that have approximately the same frequency distributions (apart from a site-specific factor), and regions can be defined using various objective and

subjective procedures (Hosking and Wallis 1997). We recommend using an objective procedure, such as cluster analysis methods, with an increasing number of regions specified (first a single region, then two regions, etc.), checking at each step for a significant decrease in one or more of the region's heterogeneity statistics, while still maintaining at least 10–15 sites in each region for reliability. Following the cluster analysis procedure, a few sites may be reassigned to neighboring regions to provide better geographical convenience, but the heterogeneity statistics should be recomputed to ensure that they did not increase significantly with the reassignment of sites.

b. Spatial and temporal variability

As shown in Fig. 4, average annual precipitation is highly variable across Sonora, with the lowest precipitation totals found in the coastal areas, and the highest totals found along the Sierra Madre, where orographic effects increase precipitation. Precipitation is also highly variable from year to year, although there is a strong spatial coherency in the annual totals. For instance, dry and wet years correspond quite well between zone 3 and zone 4, despite a difference of 40 cm (15 in.) in average annual precipitation and several hundred miles between both zones, as illustrated in Fig. 5. The annual values for each zone were estimated by averaging the gauge totals within each zone.

Previous studies have demonstrated a relationship between the southwest United States and northwest Mexico climate variability and ENSO, particularly for the cold season (McPhee et al. 2004), and these results
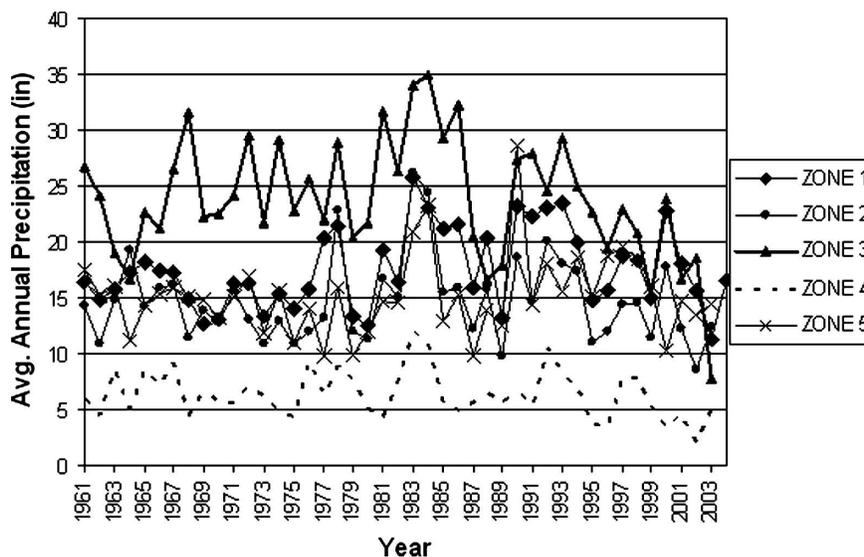

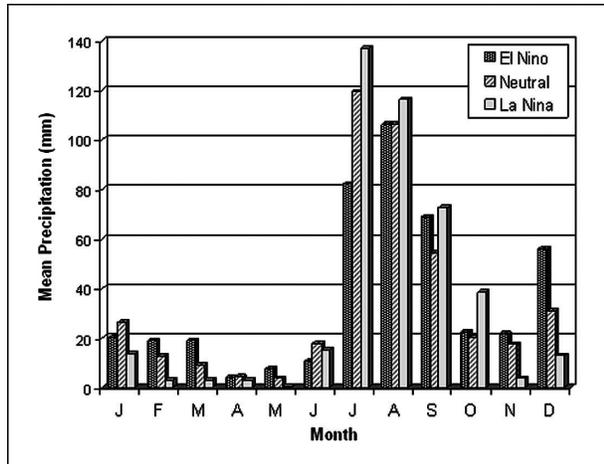FIG. 5. Annual precipitation for Sonora's five climatic zones, 1961–2004.

FIG. 6. Seasonal precipitation pattern for Sonora during ENSO events.

TABLE 3. Classification of SPI values (McKee et al. 1993).

SPI values	
2.0 +	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderate wet
−0.99 to 0.99	Near normal
−1.0 to −1.49	Moderately dry
−1.5 to −1.99	Severely dry
−2.0 and less	Extremely dry

were essentially confirmed in this data. Figure 6 shows the monthly average precipitation (for all zones) during seasons associated with El Niño and La Niña events, compared with neutral (non-ENSO) seasons, for the period 1961–2000. ENSO events were characterized as occurring in winter (November–March) or summer (May–September) based on definitions published by the National Oceanic and Atmospheric Administration–Cooperative Institute for Research in Environmental Sciences (NOAA–CIRES) Climate Diagnostics Center (CDC; information available online at http://www.cdc.noaa.gov/ENSO/Compare). This resulted in the consideration nine El Niño events (1965–66, 1968–69, 1972–73, 1982–83, 1986–87, 1991–92, 1993, 1994, and 1997–98) and six La Niña events (1964–65, 1970–71, 1973–74, 1975–76, 1988–89, and 1998–99). While these are small sample sizes, it appears that there are significant differences in seasonal rainfall distributions associated with ENSO events. In particular, La Niña appears to concur with a stronger summer monsoon but lower winter precipitation than during El Niño or neutral years. In contrast, El Niño events are correlated with weak summer monsoons and higher than average winter precipitation.

c. Annual drought frequencies

For each climate zone, drought intensity–duration–frequency estimates were derived in terms of 1) standardized precipitation index (SPI) values and 2) quantiles of the selected frequency distribution. McKee et al. (1993) recommend computing SPI by fitting a gamma (GAM) distribution to precipitation totals, but since SPI conceptually represents the number of standard deviations above or below the mean of a normal-

transformed variate, any frequency distribution that appropriately fits the data may be used. Guttman (1999) compared various distributions and recommended the PE3 distribution. In this study, we compute SPI using the GAM distribution as suggested by Wu et al. (2005), but recognize that other distributions would be suitable alternatives. Following selection of a distribution and parameter estimation, drought (for any duration) may then be defined objectively using SPI values as classified in Table 3.

A summary of the goodness-of-fit measures for the candidate distributions applied in each zone is shown in Table 4. Values in boldface have met the criterion of $|Z| \leq 1.64$, which corresponds to the failure to reject the hypothesized distribution at a confidence level of 90% (Hosking and Wallis 1997). The fit of the various distributions is also illustrated in Fig. 7, with distribution parameter values computed from sample data from zone 1. As illustrated in Fig. 7, each of the extreme value distributions takes on a similar shape overall, but significant differences may be recognized in the tail behavior of the distributions. Thus, care must be taken in selecting a distribution.

Based on the Z statistic alone, the LN3, GEV, and GLO distributions provided acceptable fits for the majority of the zones, while the PE3 distribution (of which the GAM distribution is a special case) proved acceptable for at least two zones. A Z statistic is not reported for the GAM distribution since, as a two-parameter distribution, it has a fixed value for the L-moment ratio t_4. However, since its relative, the PE3 distribution, has Z statistics that are within or nearly within the recommended range, quantiles of the GAM distribution were computed for comparison with the other candidate distributions.

Results of parameter and quantile estimation for the five candidate distributions are shown in Table 5. The values shown are the quantiles expressed as a fraction of the mean. For instance, for a site in zone 1, all distributions indicate that an annual precipitation total as small as 65% of the mean annual precipitation is estimated to occur with a frequency of 0.1, or on average

TABLE 4. Goodness-of-fit statistics, Z^{DIST}, for candidate distributions fit to annual precipitation data. Boldface values indicate acceptable fits based on the criterion $|Z| \leq 1.64$.

	Z^{DIST}				
Distribution	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5
Generalized logistic	3.56	2.06	**0.13**	**0.60**	**0.60**
Generalized extreme value	**0.22**	**0.25**	3.46	**1.41**	**1.58**
Generalized normal	**0.63**	**0.47**	2.94	**1.45**	1.66
Pearson type III*	**0.39**	**1.17**	3.10	1.87	2.17

* Gamma distribution is a special case.

once in 10 yr. It can be seen that the five distributions provide very consistent estimates for annual drought intensities in the range of 10- to 20-yr return periods, but there is some divergence for the 1% and 2% chance events (100- and 50-yr droughts), especially for zone 4. This zone is the driest of the five zones and experiences the largest variability in annual precipitation, and hence it is expected to have the largest uncertainty in drought frequency estimates.

Given the significant divergence in the 100-yr drought intensity estimates, the quantile estimates from each distribution were compared to empirical estimates (based on a Weibull plotting position), with the closest fit indicated in the table. For example, considering the zone 1 data, approximately 1% of the annual precipitation values are less than 140 mm, which is approximately 31% of the mean annual precipitation for this zone. Thus, the GLO quantile estimate (37% of the mean) appears to be more reliable than those of the other candidate distributions (43%–45% of the mean). The fact that the GLO distribution provides the most reliable estimate of the 0.01 quantile in four of the five zones is an indication that its lower tail behavior is more consistent with the data than is that of the other distributions. This is also evident upon inspection of Fig. 7.

d. Seasonal drought frequencies

The regional frequency analysis was repeated using seasonal data corresponding to two 6-month periods: a warm season (May–October) and a cold season (November–April). With the focus on the NAM, only warm season results are reported here. These results are summarized in Table 6 for the five candidate distributions. As for the annual precipitation frequency analysis, the candidate distributions provide consistent estimates for most zones and most return periods. For the warm season analysis, there is some discrepancy in the estimates of the 50- to 100-yr droughts (1%–2% chance events)

TABLE 5. Annual quantile estimates for four different distributions for frequencies of 0.01, 0.02, 0.05, and 0.1. Values indicate the fraction of the mean annual value corresponding to the given quantile. Boldface value indicates best fit to data at the 0.01 quantile.

	0.1	0.05	0.02	0.01
Zone 1				
Generalized extreme value (GEV)	0.65	0.57	0.49	0.44
Generalized logistic (GLO)	0.66	0.56	0.45	**0.37**
Generalized normal (LN3)	0.65	0.57	0.49	0.44
Pearson type III (PE3)	0.65	0.57	0.49	0.44
Gamma (GAM)	0.65	0.58	0.50	0.45
Zone 2				
GEV	0.61	0.54	0.47	**0.42**
GLO	0.62	0.53	0.43	0.37
LN3	0.61	0.54	0.47	0.43
PE3	0.61	0.55	0.48	0.45
GAM	0.61	0.53	0.46	**0.42**
Zone 3				
GEV	0.64	0.56	0.47	0.42
GLO	0.65	0.55	0.43	**0.34**
LN3	0.64	0.56	0.47	0.41
PE3	0.64	0.56	0.47	0.41
GAM	0.65	0.58	0.50	0.43
Zone 4				
GEV	0.44	0.33	0.21	0.142
GLO	0.45	0.31	0.16	**0.054**
LN3	0.44	0.33	0.22	0.149
PE3	0.44	0.33	0.23	0.172
GAM	0.45	0.35	0.25	0.20
Zone 5				
GEV	0.61	0.53	0.45	0.40
GLO	0.61	0.51	0.41	**0.34**
LN3	0.60	0.53	0.45	0.41
PE3	0.60	0.53	0.46	0.42
GAM	0.60	0.52	0.45	0.39

for all zones, and there is significant discrepancy for zone 4. In this zone, three distributions have 0.01-quantile values that are less than zero, which for practical purposes are interpreted as equaling zero or only a trace (<1 mm). In other words, based on the use of these distributions to fit the observed data, sites in this zone may expect to receive essentially zero warm season precipitation an average of once every 50–100 yr. The other distributions indicate that under the 100-yr seasonal drought, the sites in zone 4 will receive 2%–7% of the average seasonal precipitation. This variability is due to the existence of both very small (<1 mm) and large precipitation values in the seasonal data.

While the focus of this paper is on annual and warm season precipitation, due to the well-defined influence of the NAM on climate in Sonora, winter precipitation
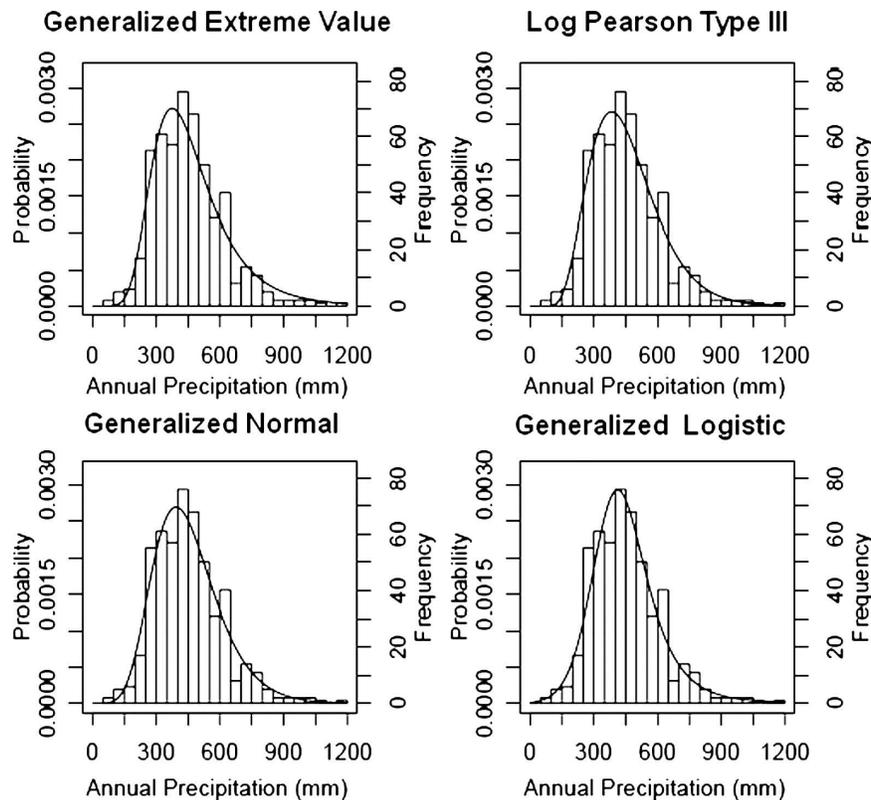
FIG. 7. Extreme value distributions fit to sample data from zone 1.

is crucial to water management in this region. Winter precipitation is usually associated with relatively long-lived frontal systems that approach Sonora from the Pacific Ocean, and these events are typically more effective in terms of recharging soil moisture and groundwater supplies than warm season precipitation events. This is mainly due to two factors: 1) summer precipitation is often very intense, falling at high rates in short periods of time over discontinuous areas, and large amounts of water sometimes run off the surface rather than infiltrating deep into the soil, and 2) high summer temperatures cause high evaporation rates, leaving little or no surplus of surface moisture for storage (McPhee et al. 2004). Also, since the atmospheric phenomena that originate warm and cold season precipitation are very different, seasonal rainfall totals tend to be uncorrelated with each other, and thus separate cold season drought frequency analysis is warranted. Details of drought frequency analysis for the cold season may be found in Hallack-Alegria (2005), who also investigated the potential for cold season precipitation forecasting. Preliminary analysis showed significant correlation between warm season ENSO phenomena and winter precipitation, with La Niña events consistently leading drier-than-average winters.

e. Multiyear drought frequency

The severity of the recent 1995–2004 drought in Sonora is apparent from time series plots of average annual SPI values for zones 1–4, as shown in Fig. 8. (Zone 5 is not included since it has little station data available since 1987.) In each of these zones, there were at least 7 yr of below-average precipitation (i.e., average SPI values ≤ -0.10). Assuming a gamma distribution for the annual precipitation amounts, these SPI values over this 10-yr period correspond to nonexceedance probabilities ranging from 0.10 to 0.80.

To estimate the return period of the 10-yr drought using the method of Fernàndez and Salas (1999), the interannual correlation of precipitation totals must be estimated. Using the normalized average annual precipitation for all sites in each zone, the following correlation coefficients were estimated for the four zones: $\rho_1 = 0.41$, $\rho_2 = 0.23$, $\rho_3 = 0.19$, and $\rho_4 = 0.29$. Estimating the return period of the 10-yr drought as the probability of a 10-yr run in a dependent Markov chain results in the following estimates of return period T for the four zones: $T_1 = 30$ yr, $T_2 = 80$ yr, $T_3 = 75$ yr, and $T_4 = 35$ yr. These return periods are rough estimates for several reasons, including the fol-

TABLE 6. Warm season (May–October) quantile estimates for four different distributions for frequencies of 0.01, 0.02, 0.05, and 0.1. Values indicate the fraction of the mean annual value corresponding to the given quantile. Negative values are to be interpreted as zero precipitation. Boldface values indicate best fit to data at the 0.01 quantile.

	0.1	0.05	0.02	0.01
Zone 1				
Generalized extreme value (GEV)	0.57	0.48	0.37	0.31
Generalized logistic (GLO)	0.58	0.46	0.32	0.22
LN3	0.57	0.47	0.37	**0.30**
PE3	0.57	0.47	0.37	0.31
GAM	0.60	0.52	0.44	0.39
Zone 2				
GEV	0.51	0.42	0.32	0.26
GLO	0.52	0.40	0.27	**0.18**
LN3	0.51	0.42	0.32	0.26
PE3	0.51	0.42	0.33	0.28
GAM	0.51	0.42	0.33	0.28
Zone 3				
GEV	0.64	0.55	0.46	0.40
GLO	0.64	0.54	0.41	**0.32**
LN3	0.64	0.55	0.45	0.39
PE3	0.64	0.55	0.45	0.39
GAM	0.64	0.57	0.49	0.44
Zone 4				
GEV	0.24	0.11	−0.010	−0.084
GLO	0.24	0.09	−0.067	**−0.163**
LN3	0.24	0.12	0.013	−0.048
PE3	0.23	0.14	0.057	0.020
GAM	0.25	0.17	0.100	0.070
Zone 5				
GEV	0.56	0.46	0.36	0.30
GLO	0.56	0.45	0.31	**0.22**
LN3	0.56	0.46	0.36	0.30
PE3	0.56	0.46	0.36	0.30
GAM	0.55	0.46	0.37	0.32

lowing: there is a high degree of variability in the interannual correlation coefficient estimate ρ, there is a relatively small amount of data compiled for recent years (2001–2004), and the assumptions in the method of Fernàndez and Salas (1999) may not be completely valid.

4. Summary and conclusions

The unpredictable characteristics of droughts, including their initiation and termination; frequency; and severity make drought both a hazard and a disaster: a hazard because drought is a natural phenomenon of unpredictable occurrence but of recognizable recurrence and a disaster because it may correspond to the disruption of water supplies for natural and agricultural ecosystems as well as to other human activities (Frick et al. 1990). While accurate forecasting of when a drought will begin or end is extremely difficult, better understanding of the expected intensity, duration, and recurrence frequency of droughts is essential to drought mitigation planning. This is particularly true for the agriculture sector, in which appropriate planning, monitoring, and prediction can greatly reduce economic hardships in the event of drought (Wilhite 2002; Rossi 2003).

In this study we provide an overview of seasonal, annual, and interannual precipitation variability in Sonora, Mexico, where 40%–80% of the annual precipitation falls during the summer monsoon season and there is large year-to-year variability (Stensrud et al. 1997; Comrie and Glenn 1998). Using 15–44 yr of monthly precipitation data from 95 gauges in Sonora and Arizona, the spatial and temporal variabilities of precipitation were characterized for five quasi-homogeneous climate zones, and seasonal effects of El Niño and La Niña (Southern Oscillation) were explored. Using a regional frequency analysis procedure, annual and seasonal drought intensity estimates were derived for each of the zones for return periods of 10, 25, 50, and 100 yr. In addition, simplifying assumptions were made in order to estimate the return period of the recent 10-yr drought. With a return period of 30–80 yr, depending on the climate zone, the recent drought is unusual but not extremely rare. Planning with knowledge of the intensity and probability of this event should help to diminish future impacts of drought on agriculture, industry, and other water uses in Sonora, Mexico.

Future work should continue to investigate teleconnections between large-scale climate anomalies, such as ENSO and the PDO, and precipitation in Sonora. Seasonal forecasts with even small skill may be useful to water managers seeking to prepare, adapt, or respond to drought impacts. Better understanding of long-term climate variability can also be valuable, as known climate trends or cycles may effectively modify drought return period estimates. Care must be taken, however, not to place too much weight on recent records, lest important information about climate variability be lost (e.g., Wu et al. 2005). To improve planning for multiyear droughts, the assumptions made for estimating the frequency of the 1995–2004 drought should be tested, and estimates revised as needed. Finally, on shorter time scales, SPI and other metrics of water scarcity may be applied in operational settings to monitor ongoing droughts and assess drought recovery.
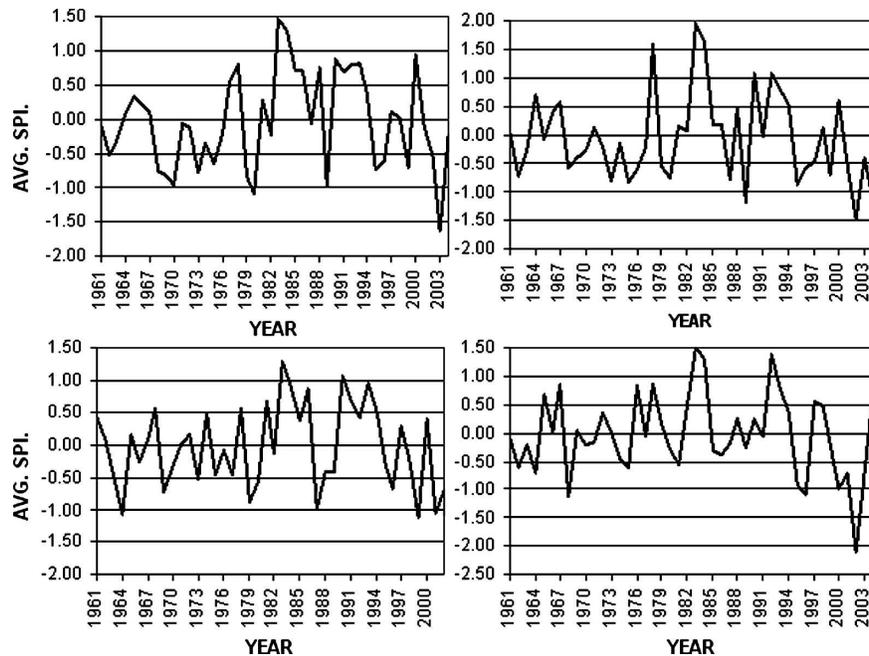
FIG. 8. Time series of annual SPI values for each zone. SPI was computed using the gamma distribution.

made possible through international programs at the University of Sonora. Data were provided by the Mexican Institute of Water Technology (IMTA) and National Commission of Water (CNA). Comments and suggestions made by A. Maclean, A. Mayer, and B. Barkdoll of Michigan Technological University are appreciated, as is the assistance of R. Nikolov at North Carolina State University—Fayetteville. The constructive comments of three anonymous reviewers helped to improve this paper considerably, but any errors or omissions are the sole responsibility of the authors.

REFERENCES

Carleton, A. M., 1986: Synoptic-dynamic character of 'bursts' and 'breaks' in the southwest U.S. summer precipitation singularity. *J. Climatol.,* **6,** 605–623.

——, D. A. Carpenter, and P. J. Webster, 1990: Mechanisms of interannual variability of the southwest United States summer rainfall maximum. *J. Climate,* **3,** 999–1015.

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.

Douglas, M. W., R. A. Maddox, K. Howard, and S. Reyes, 1993: The Mexican monsoon. *J. Climate,* **6,** 1665–1677.

Dracup, J. A., K. S. Lee, and E. G. Paulson, 1980: On the definition of droughts. *Water Resour. Res.,* **16,** 297–302.

Fernàndez, B., and J. D. Salas, 1999: Return period and risk of hydrologic events, I: Mathematical formulation. *J. Hydrol. Eng.,* **4,** 297–307.

Frick, D. M., D. Bode, and J. D. Salas, 1990: Effect of drought on urban water supplies, I: Drought analysis. *J. Hydr. Engrg. ASCE,* **116,** 733–753.

Gibbs, W. J., and J. V. Maher, 1967: Rainfall deciles as drought indicators. Bureau of Meteorology Bull. 48, Melbourne, Australia, 84 pp.

Giddings, L., M. Soto, B. M. Rutherford, and A. Maarouf, 2005: Standardized precipitation index zones for Mexico. *Atmosfera,* **18,** 33–55.

Guttman, N. B., 1999: Accepting the standardized precipitation index: A calculation algorithm. *J. Amer. Water Resour. Assoc.,* **35,** 311–322.

——, J. R. Wallis, and J. R. M. Hosking, 1992: Spatial comparability of the Palmer drought severity index. *Water Resour. Bull.,* **28,** 1111–1119.

Hallack-Alegria, M., 2005: Drought frequency analysis and prediction for Sonora, Mexico. M.S. thesis, Dept. of Civil and Environmental Engineering, Michigan Technological University, Houghton, MI, 65 pp.

Hartigan, J. A., and M. A. Wong, 1979: Algorithm AS 136: A *K*-means clustering algorithm. *Appl. Stat.,* **28,** 100–108.

Higgins, R. W., and W. Shi, 2000: Dominant factors responsible for interannual variability of the summer monsoon in the southwestern United States. *J. Climate,* **13,** 759–776.

——, Y. Yao, and X. L. Wang, 1997: Influence of the North American monsoon system on the U.S. summer precipitation regime. *J. Climate,* **10,** 2600–2622.

——, K. C. Mo, and Y. Yao, 1998: Interannual variability of the U.S. summer precipitation regime with emphasis on the southwestern monsoon. *J. Climate,* **11,** 2582–2606.

Hosking, J. R. M., 1990: L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. Roy. Stat. Soc.,* **52B,** 105-124.

——, and J. R. Wallis, 1988: The effect of intersite dependence on

regional flood frequency analysis. *Water Resour. Res.,* **24,** 588–600.

——, and ——, 1997: *Regional Frequency Analysis.* Cambridge University Press, 224 pp.

Houghton, J. C., 1978: Birth of a parent: The Wakeby distribution for modeling flood flows. *Water Resour. Res.,* **14,** 1105–1110.

Kiladis, G. N., and H. F. Diaz, 1989: Global climate anomalies associated with extremes of the Southern Oscillation. *J. Climate,* **2,** 1069–1090.

Kim, T. W., and J. B. Valdes, 2002: Frequency and spatial characteristics of droughts in the Conchos River basin, Mexico. *Water Int.,* **27,** 420–430.

Kumar, R., and C. Chatterjee, 2005: Regional flood frequency analysis using L-moments for north Brahmaputra region of India. *J. Hydrol. Eng.,* **10,** 1–7.

Landwehr, J. M., N. C. Matalas, and J. R. Wallis, 1979: Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resour. Res.,* **15,** 1055–1064.

Lettenmaier, D. P., J. R. Wallis, and E. F. Wood, 1987: Effect of regional heterogeneity on flood frequency estimation. *Water Resour. Res.,* **23,** 313–323.

McCollum, D. M., R. A. Maddox, and K. W. Howard, 1995: Case-study of a severe mesoscale convective system in central Arizona. *Wea. Forecasting,* **10,** 643–665.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. Preprints, *Eighth Conf. on Applied Climatology,* Anaheim, CA, Amer. Meteor. Soc., 179–184.

McPhee, J., A. Comrie, and G. Garfio, 2004: Drought and climate in Arizona. CLIMAS Final Rep., The University of Arizona, Tucson, AZ, 24 pp.

Mitchell, D. L., D. Ivanova, R. Rabin, T. J. Brown, and K. Redmond, 2002: Gulf of California sea surface temperatures and the North American monsoon: Mechanistic implications from observations. *J. Climate,* **15,** 2261–2281.

Moreno, J. L., 1994: El agua en Sonora (Water in Sonora). University of Sonora Rep., Hermosillo, Mexico.

National Drought Mitigation Center, cited 2005: Drought science: Understanding and defining drought. [Available online at http://drought.unl.edu/whatis/concept.htm.]

Quintas, I., 2000: ERIC II: Documentación de la base de datos climatológica y del programa extractor (ERIC II: Documentation of the climatologic database and data extraction program). Progreso, Morelos, Instituto Mexicano de Tecnología del Agua.

Rossi, G., 2003: An integrated approach to drought mitigation in Mediterranean regions. *Tools for Drought Mitigation in Mediterranean Regions,* G. Rossi et al., Eds., Kluwer, 3–18.

Santos, M. A., 1983: Regional droughts: A stochastic characterization. *J. Hydrol.,* **66,** 183–211.

Sheppard, P. R., A. C. Comrie, G. D. Packin, K. Angersbach, and M. K. Hughes, 2002: The climate of the U.S. southwest. *Climate Res.,* **21,** 219–238.

Shin, H.-S., and J. D. Salas, 2000: Regional drought analysis based on neutral networks. *J. Hydrol. Eng.,* **5,** 145–155.

Soule, P. T., 1992: Spatial patterns of drought frequency and duration in the contiguous USA based on multiple droughts event definitions. *Int. J. Climatol.,* **12,** 11–24.

Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou, 1993: Frequency analysis of extreme events. *Handbook of Hydrology,* D. R. Maidment, Ed., McGraw-Hill, 18.1–18.66.

Stensrud, D. J., R. L. Gall, S. L. Mullen, and K. W. Howard, 1995: Model climatology of the Mexican monsoon. *J. Climate,* **8,** 1775–1794.

——, ——, and M. K. Nordquist, 1997: Surges over the Gulf of California during the Mexico monsoon. *Mon. Wea. Rev.,* **125,** 417–437.

Sveinsson, O. G. B., J. D. Salas, and D. C. Boes, 2002: Regional frequency analysis of extreme precipitation in northeastern Colorado and Fort Collins flood of 1997. *J. Hydrol. Eng.,* **7,** 49–63.

Tase, N., 1976: Area–deficit–intensity characteristics of drought. Hydrology Paper 87, Colorado State University, Fort Collins, CO, 45 pp.

Tinajero, G. J., 1986: Analisis de la sequia en Mexico en el periodo 1976–1980 las variaciones climaticas, la sequia y la disertificacion (Analysis of the drought in Mexico, 1976–1980: Climate variability, drought, and desertification). Secretaria de Agricultura y Recursos Hidraulucos, Mexico.

Trefry, C. M., D. W. Watkins Jr., and D. L. Johnson, 2005: Regional rainfall frequency analysis for the state of Michigan. *J. Hydrol. Eng.,* **10,** 437–449.

U.S. Army Corps of Engineers, cited 2005: *U.S. National Drought Atlas.* [Available online at http://www.iwr.usace.army.mil/iwr/atlas/Atlasintro.htm.]

Ward, J. H., 1963: Hierarchical grouping to optimize an objective function. *J. Amer. Stat. Assoc.,* **58,** 236–244.

Watkins, D. W., Jr., G. A. Link, and D. L. Johnson, 2005: Mapping regional precipitation intensity–duration–frequency estimates. *J. Amer. Water Resour. Assoc.,* **41,** 157–170.

Wilhite, D. A., 2002: Combating drought through preparedness. *Natl. Resour. Forum,* **26,** 275–285.

Woodhouse, C. A., 1997: Winter climate and atmospheric circulation patterns in the Sonoran Desert region, USA. *Int. J. Climatol.,* **17,** 859–873.

Wu, H., M. J. Hayes, D. A. Wilhite, and M. D. Svoboda, 2005: The effect of the length of record on the standardized precipitation index calculation. *Int. J. Climatol.,* **25,** 505–520.

Yevjevich, V., 1967: An objective approach to definitions and investigations of continental hydrologic droughts. Hydrology Paper 23, Colorado State University, 19 pp.