Comparative Evaluation of Four Regional Flood-Frequency Analysis Methods

BABAK NAGHAVI, JAMES F. CRUISE, AND SENARATH EKANAYAKE

Four popular methods for analyzing regional flood frequency were investigated using Louisiana streamflow series. The state was divided into four homogeneous regions and all undistorted, long-term stream gauges were used in the analysis. The generalized extreme value (GEV), two-component extreme value, and regional log Pearson Type III methods were applied to this data base and compared in terms of descriptive capabilities. On the basis of several factors, the GEV method was selected as the overall superior method. The GEV parameters were estimated using the probability-weighted moments (PWMs). Indexing was accomplished using the first PWM (the mean). A procedure to apply this method to ungauged watersheds using regression equations and a regional nondimensional flood distribution was developed. It was found that the procedure performed well when applied to data not used in the calibration of the model. The regional GEV procedure was compared with the method of the U.S. Geological Survey (USGS) and showed significant improvement over the USGS equations in terms of fit to the observed data. This method is easier to apply and more accurate in terms of descriptive and probably predictive ability than other feasible methods for Louisiana data.

Often in hydrologic work, discharges must be estimated for sites at which stream gauge records are unavailable. Several techniques have been developed over the years to do this. Many of these methods are based on some type of regional frequency analysis. The Louisiana Department of Transportation and Development employs the U.S. Geological Survey (USGS) regression technique (*1*) to obtain discharge estimates at ungauged sites in the state. These equations contain a fair degree of error and have not been compared to alternative techniques. The USGS equations are based on regression analysis of at-site frequency estimates, which in turn are based on the regional log Pearson Type III (LP3) distribution. However, this distribution does not lend itself to regionalization techniques because of the variability of the skew coefficient used in LP3 parameter estimation (*2*). Also, LP3 parameters are not easily related to physical watershed characteristics (*3*). Furthermore, the error reported for the USGS equations (typically 40 to 50 percent) represents the standard error of the regression estimates and does not include the error inherent in fitting the LP3 to the samples. This error has been shown to run anywhere from 10 to 30 percent for Louisiana stations (*4*).

Another widely used regional analysis method, recommended by the Interagency Advisory Committee on Water Data (IACWD), is also based on the LP3 distribution but uses a weighted generalized skew coefficient (*5*). The use of a generalized skew coefficient instead of the sample skew coefficient results in a more reliable flood-frequency analysis for streams with short records (*5*).

Alternate regional frequency techniques have been proposed by Dalrymple (*6*) and Stedinger (*7*). Greis and Wood (*8*) recommended an indexing method similar to that of Dalrymple (*6*), but with extreme value Type 1 (EV1) as the base distribution and parameters estimated by probability-weighted moments. This parameter estimation method, first proposed by Greenwood et al. (*9*), has been shown to possess attractive asymptotic characteristics when it is used to estimate the parameters of several distributions, especially when the samples exhibit wide variability (*10*). This characteristic makes the method useful for regional frequency analyses. In support of this, Potter and Lettenmaier tested 10 commonly used frequency methods and found that the GEV index method possessed predictive characteristics superior to the other methods tested (*2*).

Another highly regarded method is the two-component extreme value (TCEV). Rossi et al. (*11*) applied the TCEV with the maximum likelihood method of parameter estimation to regional data series.

The purpose of this study was to formulate two alternative methods of regional frequency analysis using Louisiana annual peak streamflows; compare these methods with the LP3 on the basis of generalized skew coefficients; select the best method based on the basis of statistical comparison indexes of descriptive capabilities and the ease of use (requiring less physical data); and compare the selected regional method to the USGS regression equations. The two regional methods investigated are the TCEV (*11,12*) and the GEV (*13*), indexed by the method of PWM (*9*) outlined by Greis and Wood (*8*).

REGIONALIZATION

The state of Louisiana was divided into four hydrologically homogeneous regions that were determined by soil, geologic, topographic, climatic, and streamflow similarities. The purpose of this analysis was to divide the state into regions such that the hydrologic responses of watersheds in each region are comparable. Thus, the regions should have relatively homogeneous soil and topographic characteristics. In addition, the watersheds in each region should be subjected to similar climatic conditions. Information needed to make the determinations was readily available from previously published sources. The *Atlas of Louisiana* (*14*) and the *General Soil Map of Louisiana* (*15*) were used in forming the regional

---

B. Naghavi, Louisiana Transportation Research Center, 4101 Gourrier Avenue, Baton Rouge, La. 70808. J. F. Cruise, S. Ekanayake, Department of Civil Engineering, Louisiana State University, Baton Rouge, La. 70803.

groupings. The *Geological Map of Louisiana* shows that the state is divided into four general regions by the Mississippi alluvium. The regional groupings were further compared on the basis of climatic and soils information available. A complete description of the methodology used in determining the homogeneous regions is given by Naghavi et al. (*16*).

Once preliminary regions had been identified, the annual peak stream flows of gauged watersheds within each region were analyzed for similarities. This was accomplished by plotting the logarithm of the mean ($\log Q_M$) of the annual flood series (in log space) against the corresponding drainage area (A) for each watershed in the region. A curve through the points was fitted by standard regression techniques. The regression equations for the four regions are as follows:

- *Southeast*:

$$\log Q_M = 2.695\,A^{0.072} \qquad R^2 = .86,\ \mathrm{CV} = 3.1 \tag{1}$$

- *Southwest*:

$$\log Q_M = 2.561\,A^{0.076} \qquad R^2 = .84,\ \mathrm{CV} = 3.22 \tag{2}$$

- *Northwest*:

$$\log Q_M = 2.836\,A^{0.052} \qquad R^2 = .76,\ \mathrm{CV} = 2.509 \tag{3}$$

- *Northeast*:

$$\log Q_M = 2.406\,A^{0.063} \qquad R^2 = .97,\ \mathrm{CV} = 1.36 \tag{4}$$

In analyzing these equations, the coefficient of determination (R^2) represents the percentage of the total variance of the dependent variable ($\log Q_M$) explained by its relationship with the area. The coefficient of variation (CV) represents a dimensionless measure of the error in the regression fit. Thus, the relationship between log mean annual flood values and drainage areas appears to be well confirmed in these cases. Watersheds that fell outside this linear trend (by visual inspection) would not be expected to behave similarly to the other basins within the region. In this way, minor revisions to the regional groupings were determined. These regional boundaries are delineated in Figure 1. The locations of all the stream gauges used in the analysis are also plotted in this figure.

DATA

The data were obtained for all stream gauges in the physiographical regions of the state with a minimum of 20 years of systematic record. A few gauges that fell in the general physiographical regions of Louisiana but that were physically outside state boundaries were included in the analysis. Locations of all gauges are shown in Figure 1. The data set consisted of 110 long-term, continuous stream gauge records. These records were then screened for possible anomalies resulting from flow diversions, interbasin transfers at high discharges, or missing records. The records that passed this screening were further analyzed for consistency within the homogeneous regions previously defined. It was ascertained that gauges with drainage areas of fewer than 10 mi^2 generally did not follow the trend of the rest of the data. Therefore, these records were excluded from the analysis. In the end, 85 gauges passed the screening process and formed the data base for the rest of the analysis. There were 24 gauges in the Southeast region, 32 in the Southwest region, 24 in the Northwest region, and five in the Northeast region. A listing of these gauges, their drainage areas, periods of record, and skews of the log-transformed data are given in Tables 1 through 4.

FLOOD-FREQUENCY ANALYSIS

Regional frequency analyses were performed for each homogeneous region on the basis of all of the screened annual peaks observed in each region. Flood-frequency analysis consists of fitting preselected probability distributions to recorded flood data at individual sites and then estimating the magnitude (quantile) of flood events corresponding to given exceedance probabilities from the distributions. However, using the observed data from only the site under investigation can result in unreliable estimates. This is especially true when the length of record at a single site is relatively short in comparison with the recurrence intervals to be estimated from the data. For instance, it may be necessary to estimate the 100-year flood from only 20 to 30 years of record at an individual site. This is the reason that regional flood-frequency analysis has received much attention in recent engineering literature. Regional frequency analysis consists of using data at other sites considered similar to the site in question to augment the information at an individual site. This reduces the uncertainty inherent in short, systematic records.

Two-Component Extreme Value

TCEV has been derived as a mixture of two exponential marginal distributions from a Poisson counting process (*10*). Thus, its cumulative distribution function can be expressed as the product of two extremal distributions:

$$F(x) = \exp[-\lambda_1 \exp(-x/\theta_1) - \lambda_2 \exp(-x/\theta_2)] \tag{5}$$

where λ and θ are the shape and the scale parameters, respectively, and $F(x)$ is the nonexceedance probability of an event of magnitude x. This distribution attempts to account for the possibility that two distinct subdistributions make up the total annual distribution of flood peaks. In cases in which the marginal distributions can be shown to be exponential or the asymptotic distribution is Gumbel, the TCEV has been shown to give accurate results.

In the original formulation (*11*), TCEV parameter estimation was accomplished by maximum likelihood. However, Arnell and Gabriele (*17*) found that maximum likelihood estimates of TCEV regional parameters sometimes failed to converge and resulted in relatively variable quantile estimates. Therefore, in this study the TCEV was fitted to the regional data series by the method of maximum entropy proposed by Fiorentino et al. (*12*). This method has been shown to require less cumbersome computation and to be more reliable than the maximum likelihood procedure originally proposed by Rossi et al. (*11*).

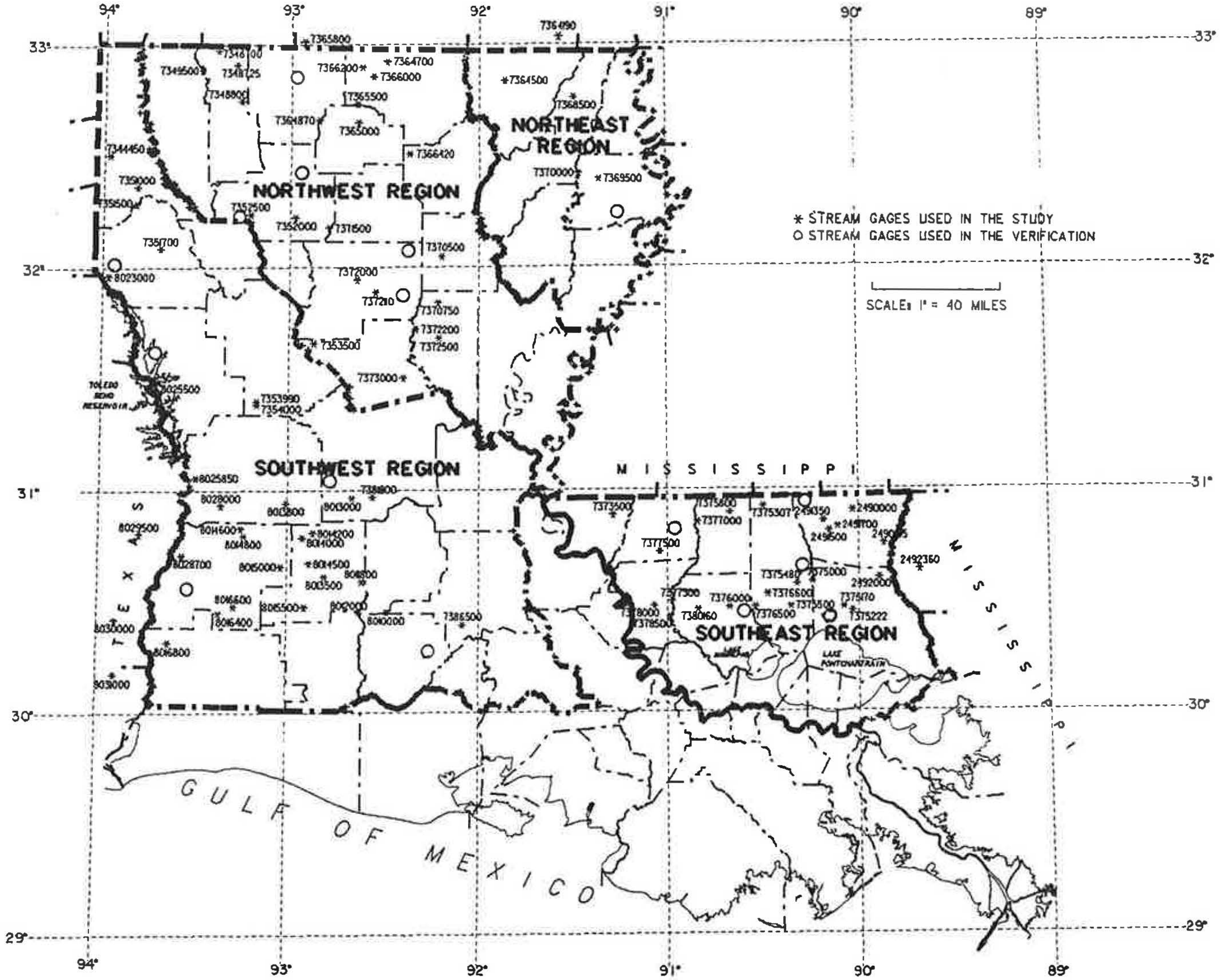

FIGURE 1 Hydrologic regions of Louisiana.

In the regionalization technique, two dimensionless parameters, $\theta = \theta_2/\theta_1$ and $\lambda = \lambda_2/\lambda_1^{1/\theta}$, are assumed to be constant for the homogeneous region; the other two parameters, θ_1 and λ_1, are allowed to vary from site to site. The parameters θ_1 and λ_1 represent the basic component, and θ_2 and λ_2 represent the outlying component of the compound distribution. The parameters θ and λ represent the regional component of the distribution. Conceptually, θ_1 and λ_1 represent the smaller, more frequently occurring events that would be expected to vary from site to site within the region. θ_1 essentially represents the mean flood for this distribution, and λ_1 represents the number of floods per year over the watershed. The parameters θ and λ represent the regional distribution; they are expected to behave similarly within the homogeneous region. As in the previous case, θ represents the mean flood of this distribution and λ represents the number of such events occurring per year. The maximum entropy procedure results in four equations to be solved for the four unknowns described previously.

Generalized Extreme Value Index Method

The index method has been receiving a great deal of attention in recent engineering literature, although its basic premise was outlined by Dalrymple some 30 years ago (*6*). In this procedure, an assumed distribution is fitted to the observed flood series at each site in a hydrologically similar region. The statistics (or parameters) of the distributions at each location are standardized by dividing by the at-site mean in each case. Regional estimates of the parameters are obtained by averaging the parameter estimates for the region. These regional parameters are then used to generate flood quantiles for the site of interest and are subsequently readjusted to account for the differences in scale between watersheds.

The index method has gained popularity since the PWM method of parameter estimation was introduced by Greenwood et al. (*9*). It has recently been used by Greis and Wood (*8*), Landwehr et al. (*10*), and Stedinger (*7*). PWM, which is usually applied only to distributions that can be expressed in

TABLE 1 PERTINENT DATA ON WATERSHEDS IN SOUTHEAST LOUISIANA

STATION No.	AREA IN (sq.mile)	YEARS OF OBS.	SKEW OF LOG TRAN. DATA	SRMSE		
				GEV	TCEV	LP3
02492000	1213	50	-0.08	0.256	0.317	0.327
02492360	175	21	-0.02	0.149	0.107	0.111
02490105	73	22	0.12	0.209	0.222	0.215
02491500	990	66	-0.34	0.171	0.186	0.201
02491700	44	20	-0.69	0.280	0.236	0.188
02491350	42	21	0.70	0.186	0.188	0.179
02490000	12	20	-0.63	0.357	0.319	0.173
07378500	1280	49	-0.12	0.122	0.142	0.130
07375222	46	22	-0.69	0.324	0.227	0.244
07380160	20	33	-0.34	0.298	0.111	0.084
07375170	88	20	0.33	0.144	0.145	0.169
07376000	247	47	-0.20	0.129	0.152	0.108
07376500	80	44	-0.08	0.183	0.097	0.090
07375500	646	49	-0.14	0.157	0.211	0.193
07377300	884	35	0.17	0.159	0.110	0.125
07376600	14	32	-0.89	0.394	0.122	0.081
07375480	91	20	-0.23	0.191	0.200	0.166
07375000	103	44	-0.13	0.266	0.244	0.164
07377000	580	39	-0.44	0.183	0.150	0.198
07375800	90	32	0.24	0.439	0.411	0.379
07375307	52	22	0.20	0.406	0.329	0.262
07378000	284	44	-0.53	0.189	0.069	0.090
07377500	145	45	-0.22	0.215	0.171	0.179
07373500	35	21	-0.32	0.172	0.110	0.104
	REGIONAL AVG.		-0.21	0.232	0.191	0.173

TABLE 2 PERTINENT DATA ON WATERSHEDS IN SOUTHWEST LOUISIANA

STATION No.	AREA IN (sq.mile)	YEARS OF OBS.	SKEW OF LOG TRAN. DATA	SRMSE		
				GEV	TCEV	LP3
07386500	19	28	-1.33	0.346	0.100	0.110
07381800	68	33	-0.22	0.169	0.168	0.105
08012000	527	49	0.95	0.188	0.247	0.321
08010000	131	49	-0.96	0.355	0.155	0.087
08011800	44	24	-0.32	0.153	0.110	0.109
08015500	1700	49	0.46	0.215	0.255	0.351
08013500	753	49	-0.17	0.104	0.098	0.165
08014500	510	48	0.16	0.656	0.642	0.720
08014000	171	27	0.29	0.263	0.314	0.323
08014200	94	37	-0.02	0.370	0.387	0.422
08013000	499	44	-0.46	0.139	0.131	0.113
08016800	177	31	0.08	0.186	0.272	0.328
08016400	148	39	0.21	0.161	0.179	0.168
08016600	82	38	0.36	0.278	0.211	0.161
08015000	238	31	0.02	0.262	0.218	0.181
08014800	120	24	-0.30	0.111	0.129	0.121
08014600	26	20	0.13	0.249	0.284	0.270
08013800	10	21	-0.50	0.116	0.150	0.103
08031000	83	34	-0.78	0.221	0.199	0.147
08030000	69	32	-0.17	0.199	0.156	0.145
08028700	13	26	0.68	0.173	0.253	0.332
08029500	128	36	0.84	0.453	0.445	0.514
08028000	365	36	0.38	0.430	0.352	0.301
08025850	10	20	0.80	0.306	0.371	0.437
08025500	148	31	0.72	0.461	0.419	0.457
08023000	97	28	-0.25	0.140	0.136	0.119
07354000	21	30	-0.71	0.353	0.176	0.118
07353990	37	22	-0.02	0.326	0.285	0.219
07351700	20	26	0.36	0.978	0.981	1.050
07351500	66	49	-1.12	0.121	0.095	0.219
07351000	79	43	-1.12	0.192	0.136	0.270
07344450	81	31	0.05	0.354	0.372	0.352
	REGIONAL AVG.		-0.06	0.282	0.263	0.273

TABLE 3 PERTINENT DATA ON WATERSHEDS IN NORTHWEST LOUISIANA

STATION No.	AREA IN (sq.mile)	YEARS OF OBS.	SKEW OF LOG TRAN. DATA	SRMSE		
				GEV	TCEV	LP3
07373000	51	46	0.03	0.285	0.295	0.164
07372500	92	31	1.15	0.518	0.566	0.769
07372200	1899	30	-0.31	0.124	0.142	0.208
07370750	48	30	0.53	0.138	0.229	0.318
07372110	24	23	0.72	0.443	0.433	0.517
07372000	654	42	-1.10	0.320	0.254	0.275
07370500	271	30	-1.07	0.194	0.195	0.280
07371500	355	49	-0.44	0.074	0.148	0.123
07366420	113	22	0.16	0.462	0.463	0.533
07365000	355	28	-0.34	0.162	0.185	0.140
07364870	47	22	-1.27	0.173	0.130	0.230
07365500	178	30	0.96	0.547	0.561	0.765
07366000	462	43	0.12	0.385	0.424	0.524
07366200	208	32	-0.13	0.357	0.395	0.431
07364700	141	22	1.28	0.737	0.725	0.875
07362100	385	49	0.04	0.176	0.203	0.327
07365800	180	29	0.39	0.969	0.894	1.044
07352000	154	47	-0.12	0.183	0.240	0.097
07352500	423	43	0.17	0.337	0.289	0.147
07348700	605	30	-0.03	0.173	0.237	0.256
07349500	546	49	-0.36	0.285	0.172	0.122
07348725	33	22	-1.71	0.314	0.213	0.377
07348800	67	24	-0.01	0.094	0.165	0.212
07353500	47	26	-0.17	0.311	0.270	0.180
	REGIONAL AVG.		-0.06	0.323	0.328	0.380

TABLE 4 PERTINENT DATA ON WATERSHEDS IN NORTHEAST LOUISIANA

STATION No.	AREA IN (sq.mile)	YEARS OF OBS.	SKEW OF LOG TRAN. DATA	SRMSE		
				GEV	TCEV	LP3
07369500	309	51	-0.58	0.068	0.943	0.038
07370000	782	60	-0.43	0.102	1.270	0.104
07368500	42	28	-0.55	0.048	1.070	0.075
07364500	1645	52	-1.93	0.071	1.103	0.097
07364190	1170	45	-1.92	0.089	1.088	0.101
	REGIONAL AVG.		-1.08	0.076	1.095	0.083

inverse form such as Gumbel and GEV, offers a method of parameter estimation that may be more robust and less biased than the traditional methods. The GEV can be expressed in inverse form as (*13*)

$$x(F) = \xi + \alpha[1 - (-\log F)^k]/k \qquad k \neq 0$$

$$x(F) = \xi - \alpha \log(-\log F) \qquad k = 0 \tag{6}$$

where F is the nonexceedance probability corresponding to the quantile x, and ξ, α, and k are the parameters of the distribution. When $k = 0$, the GEV reduces to the EV1. The index procedure is applied by calculating the PWMs from the observed data at each site in the region. The PWMs are standardized at each site by dividing each PWM by the at-site mean. The standardized PWMs are then averaged over all of the sites in the region. These regional average PWMs are used to obtain the parameters of the regional GEV distribution. Regional indexed quantiles can be generated for any exceedance probability $(1 - F)$ from Equation 6. These quantiles are then rescaled for any site of interest by multiplying by the at-site mean. The at-site mean flood can be determined from the plot of log mean Q versus drainage area for any gauged or ungauged site.

Log Pearson Type III

The regional procedure recommended in the IACWD guidelines (*5*) involves the LP3 distribution. The probability density function of the LP3 is:

$$f(x) = \frac{1}{|a| x \Gamma(b)} \left[\frac{\ln(x) - c}{a} \right]^{b-1} \exp\left[- \frac{\ln(x) - c}{a} \right] \tag{7}$$

where x is the raw (untransformed) flood magnitude, and a, b, and c are the scale, shape, and location parameters, respectively. $\Gamma(b)$ is the gamma function of the parameter b where b is always positive. The LP3 density function is very flexible and can take many forms. Parameters a, b, and c are estimated by the method of logarithmic moments (*4*).

The variability of the skew coefficient of the station record is sensitive to extreme events and sample size, thus making it difficult to obtain accurate skew estimates from small samples. For this reason, the generalized skew values are used in place of at-site skew values, or the at-site skew values are adjusted using the generalized skew when skew estimates are to be obtained from small samples. A generalized skew coefficient for each region was obtained from the arithmetic mean of the station skew values. The generalized skew value was then used to estimate LP3 parameters. To generate regional quantiles at each site of interest, at-site mean and standard deviation of the logarithms of the observed data series, together with the regionalized skew value, are used. In this study, in contrast to Bulletin 17B (*5*), only the generalized skew values were used.

COMPARATIVE ANALYSIS

Each of the three regional frequency methods was fitted to the data by the procedures previously described using the observed annual series at the 85 stream gauges. The purpose of this analysis was to select the most accurate method, on the basis of the comparisons to the observed data, among the three methods. At-site quantiles were generated from the regional distributions for each gauge location in the study. These quantiles were compared to the observed data at each site in terms of standardized root mean square error (SRMSE). The SRMSE between observed and predicted values is given by

$$\text{SRMSE} = \left\{ (1/N) \sum_{i=1}^{N} [(\hat{x}_i - x_i)/\bar{x}]^2 \right\}^{1/2} \tag{8}$$

where

x_i = observed value of standardized variate x,
$\hat{x}_i$ = predicted value of variate at the same probability point as x_i,
N = sample size, and
$\bar{x}$ = sample mean—used to standardize the root mean square error (RMSE).

$\hat{x}_i$ is calculated as $F^{-1}[p(x_i)]$, where $p(x_i)$ is approximated by the Weibull plotting position formula. The RMSE is standardized by dividing by the sample mean to remove the effects of scale and to make the comparison meaningful. This index only measures the descriptive capability of the methods. That is, SRMSE is an index of the ability of each method to interpolate the observed data at each gauged location.

The SRMSE results for the three methods are given in Tables 1 through 4. As can be seen from these results, no one method gave superior fits for all four regions. The TCEV resulted in the lowest SRMSE for the Southwest region, the LP3 method gave superior results in the Southeast region, and the GEV resulted in superior fits to observed data in both the Northwest and Northeast regions. However, the difference between the methods did not appear to be significant in many cases. The TCEV and LP3 methods performed about equally in the Southeast region and both performed significantly better than the GEV for this region. All three methods performed about the same in the Southwest region, where the average SRMSE difference between the methods were less than 10 percent. In the Northwest region, the GEV and TCEV performed evenly and resulted in significantly better fits to observed data than did the LP3, whereas the LP3 and GEV outperformed the TCEV by a considerable margin in the Northeast region. Thus, each method was clearly inferior to its counterparts in one region, clearly superior in one region each, and about equal elsewhere. It would appear difficult to choose between them on a statistical goodness-of-fit basis.

On the basis of the extreme ease with which the GEV can be extended to ungauged sites when compared with the other methods, it was selected as the superior method. The only geomorphological relationship needed is between the indexing factor (mean flood, Q_M) and basin characteristics. Because past studies have shown that the mean flood is highly correlated to the drainage area (as shown by Equations 1 through 4), a simple Q_m–versus–drainage area relationship is all that is required to apply this method to ungauged sites.

Another important factor in the selection of the GEV is that parameter estimation is done by PWM. It has been shown by Greenwood et al. (*9*) and Hosking et al. (*13*) that PWMs are more robust and less biased than conventional methods. Thus, estimates obtained by this method should be better in these respects than those obtained from other methods. This was confirmed in a study by Potter and Lettenmaier (*2*).

REGIONAL COMPARATIVE ANALYSIS

Regional comparative analysis was performed between the USGS equations and the GEV. The combined records of all the gauges within each region composed the data base for that particular region. The GEV regional procedure was applied by using Equations 1 through 4 to approximate the means at each location in the study. Using the mean values, the at-site quantiles corresponding to recurrent intervals of 2, 5, 10, 25, 50, and 100 years were recalculated from the regional values. These quantiles were then compared to the observed data at each site by the SRMSE. The regional average SRMSE results are given in Table 5. The table shows that the error in the procedure averages about 48 percent for the Southeast, Southwest, and Northwest regions and about

TABLE 5 MODEL COMPARISON BASED ON SRMSE FOR EACH REGION

REGION	NUMBER OF STATIONS	REGIONAL AVG. SRMSE		% DIFF.
		GEV/PWM	USGS/REG	
SE	24	0.468	0.536	+ 15
SW	32	0.491	0.695	+ 42
NW	24	0.532	0.872	+ 64
NE	5	0.132	0.563	+ 327
WEIGHTED AVERAGE		0.475	0.692	+ 31

13 percent for the Northeast region. However, the error in the quantile estimates from the distribution itself will be greater for the Northeast region because of the small data base.

Table 5 also shows the average SRMSE values obtained by comparing the USGS equations with the observed data at each site in each region. The USGS equations were derived by fitting the LP3 distribution to the data representing 217 gauging stations with more than 10 years of recorded data. On the basis of the results of this analysis, a regression equation was developed for quantile estimation. The general form of this equation is

$$\log Q_x = \log a + w \log A + y \log(P - 35) + z \log S \qquad (9)$$

where

Q_x = peak discharge for a given recurrence interval (x),
a = regression constant,
A = drainage area (mi^2),
P = average annual precipitation (in.),
S = average stream channel slope (ft/mi), and
w,y,z = regression coefficients.

This equation was calibrated for quantiles corresponding to recurrence intervals of 2, 5, 10, 25, 50, and 100 years using the LP3 results. Thus, the comparison of this method with the regional GEV can be based only on the analyses of these quantiles.

The results show, in every case, that the GEV procedure showed a significant improvement (greater than 10 percent) over the USGS equations in terms of fit to the observed data. The overall weighted average for all regions was 31 percent.

It is assumed that if a method accurately describes the data at gauged sites, it will probably describe the ungauged data within a hydrologic homogeneous region. Of course, not only must a frequency method describe the observed data accurately, but it should be capable of extending the data as well. Many times quantiles, which are beyond the systematic record, must be predicted. The SRMSE index does not directly measure this ability. However, studies by Greis and Wood (*8*), Hosking et al. (*13*), Landwehr et al. (*10*), and Potter and Lettenmaier (*2*) have examined the predictive capabilities of various regional and at-site frequency techniques. From the Monte Carlo or Boot Strap sampling methods, the studies concluded that methods based on PWMs possessed asymptotic characteristics in terms of bias and variability of long-term quantile estimates that were superior to other conventional methods.

VERIFICATION OF RESULTS

To verify the GEV regional procedure, the procedure was evaluated using short-term data not used in the development and calibration of the distribution. Five gauges were selected in each region except the Northeast, where only one gauge was available. Because of the lack of adequate data in the Northeast region, verification of results would not be meaningful for this region. The sites from the other three regions were selected in order to gain maximum coverage of each region. The locations of these gauges are shown by the open circles on the regional map in Figure 1.

In performing this analysis, the sites were treated as ungauged areas. The mean floods were estimated from the appropriate drainage area plots and used to scale the respective regional quantiles for each test site. The regional at-site quantiles were then compared with original data for each gauge record by SRMSE. Each gauge used in this phase of the study had between 15 and 20 years of record. Thus, the SRMSE values are based on the number of events in each case.

The SRMSE values shown in Table 6 result from analysis of each site by the GEV regional method, the at-site LP3 and the USGS equations. The LP3 distribution is used for the comparison, considering that the at-site LP3 would give the best possible distributional fit to the observed data. Analysis of the results in the table shows that the average SRMSE value by the GEV regional method was .278 for the Southeast region, .483 for the Southwest region, and .546 for the Northwest region. Comparison of these values with those given in Table 5 reveals that the method performed as well or better with the new data as it did with the data used in its derivation. Furthermore, the GEV method was generally superior by a wide margin to the USGS equations and even compared fairly well with the at-site LP3 in two regions. These results suggest that the method can be used confidently throughout the regions delineated on Figure 1.

TABLE 6 VERIFICATION OF REGIONAL GEV MODEL

REGION	STATION NO.	SRMSE		
		REGIONAL GEV/PWM	USGS REGRESSION	AT-SITE LP3
SE	07375050	0.220	0.433	0.201
	07376520	0.230	0.623	0.140
	07375463	0.314	0.315	0.339
	07377190	0.449	0.407	0.248
	02491200	0.176	0.307	0.169
	AVG.	0.278	0.417	0.219
SW	08010500	0.435	--	0.147
	08012900	0.578	0.824	0.277
	08016700	0.661	0.158	0.356
	08022765	0.515	0.389	0.102
	08024000	0.225	0.530	0.267
	AVG.	0.483	0.475	0.230
NW	07370700	0.402	0.520	0.339
	07370600	0.145	0.113	0.161
	07365300	0.888	1.140	0.682
	07352700	0.638	1.291	0.367
	07351980	0.658	1.151	0.155
	AVG.	0.546	0.843	0.341

LIMITATIONS

The applications of the results of this study are limited by the range of data available. First, the procedure should not be applied outside the physical bounds of the areas for which gauge data were available. These areas are delineated on Figure 1 and should be adhered to strictly. This eliminates the coastal zones and the Mississippi alluvium (except the Northeast region) from applicability. Second, the range of drainage basin sizes and the corresponding land uses available in each region also limit the application of this procedure. Note that the drainage basins represent undeveloped conditions. The drainage areas of each basin used in the study are given in Tables 1 through 4. The method should not be applied to drainage areas smaller than 10 mi^2, because preliminary work clearly showed that these areas respond differently to a storm event than do the larger areas. Not enough of these small gauges were available to perform a separate study.

CONCLUSIONS

The results of this study indicate that the GEV distribution fitted by the PWM method describes the annual flood series of Louisiana streams better than the other methods examined. The overall weighted average improvement of GEV index method over the USGS regional method was 31 percent. Also, verification results revealed that the GEV procedure describes data better than the USGS method in the vast majority of cases. Past Monte Carlo studies have shown that this procedure also possesses superior predictive capability in the cases for which flood estimates are required that may be out of the range of the recorded data. Therefore, on the basis of the results of this analysis as well as previous studies cited in this report, it is concluded that the GEV-PWM procedure results in overall superior flood estimates from both descriptive and predictive points of view and can be used confidently throughout the regions delineated in Figure 1. GEV-PWM is easily extended to the case of ungauged watersheds by using the relationship between the mean of the observed data (indexing factor) and corresponding drainage area of the watershed (Equations 1 through 4) for each region. However, this procedure should not be applied outside the physical bounds of the areas used in its development and verification. Particularly, the method should not be applied to drainage areas smaller than 10 mi^2, because preliminary work clearly showed that these areas respond differently to a storm event than do the larger areas.

ACKNOWLEDGMENTS

The authors wish to express their thanks to Andrea Populus and Karen Jarreau for their assistance in preparation of this document. Funding for this project was made available through FHWA.

REFERENCES

1. *Floods in Louisiana, Magnitude and Frequency*, 4th ed. Water Resources Technical Report 36. Louisiana Department of Transportation and Development, Baton Rouge, 1985.
2. K. W. Potter and D. P. Lettenmaier. A Comparison of Regional Flood Frequency Estimation Methods Using a Resampling Method. *Water Resources Research*, Vol. 26, No. 3, 1990, pp. 415–424.
3. J. R. Stedinger. Design Events with Specified Flood Risks. *Water Resources Research*, Vol. 19, No. 2, 1983, pp. 511–522.
4. B. Naghavi, J. Cruise, and K. Arora. A Comparative Evaluation of Three Estimators of Log Pearson Type 3 Distribution. In *Transportation Research Record 1279*, TRB, National Research Council, Washington, D.C., 1990.
5. Interagency Advisory Committee on Water Data. *Guidelines for Determining Flood Flow Frequency*. Bulletin 17B. U.S. Government Printing Office, Washington, D.C., 1981.
6. T. Dalrymple. Flood Frequency Analyses. *Manual of Hydrology, Part 3*. U.S. Geological Survey Water Supply Paper 1543-A. U.S. Geological Survey, 1960.
7. J. Stedinger. Estimating a Regional Flood Frequency Distribution. *Water Resources Research*, Vol. 19, No. 2, 1983, pp. 503–510.
8. P. Greis and E. Wood. Regional Flood Frequency Estimation and Network Design. *Water Resources Research*, Vol. 17, No. 4, 1981, pp. 1167–1177.
9. J. Greenwood, J. Landwehr, N. Matalas, and J. Wallis. Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressible in Inverse Form. *Water Resources Research*, Vol. 15, No. 5, 1979, pp. 1049–1054.
10. J Landwehr, N. Matalas, and J. Wallis. Probability Weighted Moments Compared with Some Traditional Techniques in Estimating Gumbel Parameters and Quantiles. *Water Resources Research*, Vol. 15, No. 5, 1979, pp. 1055–1064.
11. F. Rossi, M. Fiorentino, and P. Versace. Two-Component Extreme Value Distributions for Flood Frequency Analysis. *Water Resources Research*, Vol. 20, No. 7, 1984, pp. 847–856.
12. F. Rossi, M. Fiorentino, K. Arora, and V. P. Singh. Two-Component Extreme Value Distributions for Flood Frequency Analysis: Derivation of a New Estimation Method. *Journal of Stochastic Hydrology and Hydraulics*, No. 1, 1987, pp. 199–208.
13. J. R. M. Hosking, J. R. Wallis, and E. F. Wood. Estimation of the Generalized Extreme Value Distribution by the Method of Probability Weighted Moments. *Technometrics*, Vol. 27, No. 3, 1985, pp. 251–261.
14. M. B. Newton, Jr. *Atlas of Louisiana*. Miscellaneous Publication 71-2. School of Geosciences, Louisiana State University, Baton Rouge, 1972.
15. S. A. Lytle and M. B. Sturgis. *General Soil Areas and Associated Soil Series Groups of Louisiana*. Agricultural Experiment Station, Louisiana State University, Baton Rouge, 1962.
16. B. Naghavi, J. Cruise, and S. Ekanayake. *Prediction of Flood Quantiles at Ungauged Watersheds in Louisiana*. Report 229. Louisiana Transportation Research Center, Baton Rouge, 1989.
17. N. W. Arnell and S. Gabriele. The Performance of the Two-Component Extreme Value Distribution in Regional Flood Frequency Analysis. *Water Resources Research*, Vol. 24, No. 6, 1988, p. 879–887.