IMPROVED REGIONALIZED FLOOD FREQUENCY ESTIMATES FOR PUERTO RICO

Rafael Segarra, Ph.D.
Department of Civil Engineering

Project No. 5
Grant Agreement No. 14-08-0001-G1611

FINAL TECHNICAL REPORT
TO
U.S. DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

The work on which this report is based were financed in part by the Department of the Interior, U.S. Geological Survey, through the Puerto Rico Water Resources Research Institute.

July 1991

Improved Regionalized Flood Frequency Estimates for Puerto Rico

by

Rafael I. Segarra García

(ABSTRACT)

The project regionalizes flood frequency information for catchments in Puerto Rico. Regionalized flood frequency curves are obtained by applying the Generalized Extreme Value-Probability Weighted Moments distribution procedure to 30 catchments in Puerto Rico. Discriminant analysis is utilized to assign ungauged watersheds to homogeneous clusters obtained from gauged basins. T-year quantiles estimates for gauged and ungauged basins are obtained. This procedure yields smaller standard errors than those obtained from previous efforts in Puerto Rico utilizing regression techniques.

TABLE OF CONTENTS

Table of Contents (cont.)

List of Tables

List of Tables (Cont.)

List of Tables (Cont.)

List of Tables (Cont.)

List of Tables (Cont.)

List of Figures

List of Figures (Cont.)

List of Figures (Cont.)

CHAPTER I

INTRODUCTION

Estimation of flood frequency is one of the most important aspects of any basin hydrologic study. Studies that require the estimation of the discharge for a particular frequency are numerous. These include flood studies for facilities construction on floodplains, such as bridges and drainage channels, hydrologic studies for flood depth determination, and hazard mitigation studies. These efforts require that accurate information on flood frequency be available.

The only major study addressing flood frequency estimation in Puerto Rico is a 1979 flood frequency report from the U.S.Geological Survey (López et al., 1979), based on regional regression techniques. These techniques are subject to considerable error when available records are short, as the case of Puerto Rico. Furthermore, little consideration has been given to the possible grouping of basins into homogeneous regions so as to improve regional flood frequency quantile estimation. Recent developments in regional flood frequency estimation, particularly the Probability Weighted Moments technique and the statistical classification of homogeneous regions provide significant improvements over the use of regression techniques. The application of these techniques to the updated peak flow data of Puerto Rico, through the present study, seeks to provide more reliable estimates of flood flow frequency quantiles - the T-year flood - for use in all basin hydrologic studies.

The objectives of this study were:

1) To apply the method of the Generalized Extreme Value Probability Weighted Moments to catchments in Puerto Rico in order to obtain regional flood frequency curves.
2) To apply statistical homogeneity tests to identify possible basin clusters in Puerto Rico that share similar attributes. This results in improved regionalized flood frequency parameter estimation.
3) To estimate regionalized flood frequency curves applicable to ungauged catchments.
4) To reduce estimated quantiles standard errors.

CHAPTER II

LITERATURE REVIEW

2.1 Flood Frequency Analysis

Forecasting of flood events can generally reduce damages caused by floods. There are two ways in which this beneficial effect may be achieved. First, warning of an impending event enables people to evacuate a danger area. If sufficient lead time is provided precautions can be taken to minimize property damage and loss of life. The second method of achieving benefits through forecasting is to use flood frequency analysis in the design of structures within the flood plain. The latter is the subject of the present research. Flood frequency regionalization ensures that housing and industrial developments are not located in high-risk zones.

Flood frequency analysis is not only use as an aid in averting flooding damages, but can also be used develop efficient designs. When a hydraulic structure, through inadequate flow estimation or inaccurate data is underdesigned, it may fail. This does not happen very often and so the designer, equating nonfailure with success, is satisfied with his design techniques. Nonfailure, however, does not necessarily means an efficient design. Frequently, structures are overdesigned and hence very safe, but also very expensive. A truly efficient design will be achieved only as a result of studies relating cost to risk and flood frequency analysis.

Flood frequency analysis works with extreme events, such as the annual maximum discharge of a stream, or minimum daily flow. Hydrologic variables, like streamflow, can be treated as random variables. The annual maximum discharge, being the extreme value of a set of random variables is also a random variable. The probability distribution of this extreme value random variable will depend on the statistics of the sample size and the parent distribution from which the sample is obtained.

Consider a random sample of size n consisting of $x_1, x_2, \ldots, x_n$. Let Y be the largest of the sample values. Let $P_y(y)$ be the probability of $Y \leq y$ and $P_{xi}(x)$ be the probability of $X_i \leq x$. Let $p_y(y)$ and $p_{xi}(x)$ be the corresponding probability density functions of y and x.

If the $X_{i's}$ are independently and identically distributed we have

$$P_y(y) = P_{x_1}(y) P_{x_2}(y) \ldots P_{x_n}(y) = [P_x(y)]^n \tag{2-1}$$

$$p_y(y) = \frac{dP_y(y)}{dy} = n\,[P_x(y)]^{n-1}\, p_x(y) \tag{2-2}$$

Therefore, the probability distribution of the maximum of n independently and identically distributed random variables depends on the sample size n and the parent distribution $P_x(x)$ of the sample.

Frequently, the parent distribution, of the extreme value is unknown. If the sample size is large, certain general asymptotic results can be used. This applications depend on limited assumptions concerning the parent distribution to obtain the distribution of extreme values. Much of the original work on extreme value distributions was performed by Gumbel (1954,1958).

Three types of asymptotic distributions exist, (Haan, 1977).

These are:

a. Type I - Parent distribution unbounded in direction of the desired extreme, and all the moments of the distribution exist (exponential type distributions).

b. Type II - Parent distribution unbounded in direction of the desired extreme, and all the moments of the distribution do not exist (Cauchy type distributions).

c. Type III - Parent distribution bounded in the direction of the desired extreme (limited distributions).

The Type II, or Cauchy type, extreme value distributions have found little application in hydrology. Most of the

Therefore, the probability distribution of the maximum of n independently and identically distributed random variables depends on the sample size n and the parent distribution $P_x(x)$ of the sample.

Frequently, the parent distribution, of the extreme value is unknown. If the sample size is large, certain general asymptotic results can be used. This applications depend on limited assumptions concerning the parent distribution to obtain the distribution of extreme values. Much of the original work on extreme value distributions was performed by Gumbel (1954,1958).

Three types of asymptotic distributions exist, (Haan, 1977).

These are:

a. Type I - Parent distribution unbounded in direction of the desired extreme, and all the moments of the distribution exist (exponential type distributions).

b. Type II - Parent distribution unbounded in direction of the desired extreme, and all the moments of the distribution do not exist (Cauchy type distributions).

c. Type III - Parent distribution bounded in the direction of the desired extreme (limited distributions).

The Type II, or Cauchy type, extreme value distributions have found little application in hydrology. Most of the applications are from the Type I and III distributions, and they are discussed further.

2.1.1 Extreme Value Type I.

The Type I extreme value has been been referred in the literature as Gumbel's Extreme Value Distribution, the Extreme Value Distribution, the Fisher-Tippett Type I Distribution, and the Double Exponential Distribution.

The type I asymptotic distribution for maximum values is the limiting model as n approaches infinity, for the distribution of the maximum of n independent values from an initial distribution whose right tail is unbounded, and which is an exponential type. That is, the initial cumulative distribution approaches unity with increasing values of the random variable at least as fast as the exponential distribution approaches unity. The Normal, Lognormal, Exponential, and Gamma distributions exhibit this property.

The Type I extreme value distribution has been used for rainfall depth-duration-frequency studies, and as the distribution of the yearly maximum of daily river flows. Gumbel (1958) states that this latter application assumes: (1) the distribution of daily discharges (the parent distribution) is of the exponential type, (2) n = 365 is a sufficiently large sample, and (3) the daily discharges are independent. Gumbel states that the first and second assumptions cannot be checked since the analytical form of the distribution of discharges is unknown and the third assumption is clearly not true, so that the number of independent observations is less than 365. In spite of violating the latter assumption, experience with the Type I for the maximum of daily discharges has been reasonably good. Maximum annual flood peaks would more nearly fulfill assumption (3), although the sample size could be much less than 365.

The probability density function for type I extreme value distribution is

$$p_x(x) = \exp \frac{\frac{-(x-\beta)}{\alpha} - \exp \frac{-(x-\beta)}{\alpha}}{\alpha} \quad (2\text{-}3)$$

for $-\infty < x < \infty$; $-\infty < \beta < \infty$; $\alpha > 0$. The parameters α and β are scale and location parameters respectively, with β being the mode of the distribution.

The Type I extreme value distribution for maxima has been used to define the "mean annual flood". The "mean annual flood" refers to a flood with a return period of 2.33 years.

2.1.2 Extreme Value Type III.

The Extreme Value Type III distribution arises when the extreme is from a parent distribution that is limited in the direction of interest. This distribution has found its greatest use in hydrology as the distribution of low flows. However it can be used for the estimation of maximum flows. The U.S. Water Resourses Council recommended in 1967 that the

Log-Pearson type III distribution be adopted as the standard flood frequency distribution in all U.S. Government agencies. In describing the investigations behind this recommendation, Benson (1968) claimed that no rigorous statistical criteria exist on which comparison of distributions can be based and therefore the choice of the Log-Pearson Type III was, to some extent, arbitrary.

If the logarithms of a variable X are distributed as a Pearson Type III variate, then the variable will be distributed as a Log-Pearson Type III with probability density function

$$p(x) = \frac{1}{\alpha\, x\, \Gamma(\beta)} \left(\frac{\ln x - \delta}{\alpha}\right)^{\beta-1} \exp \frac{-(\ln x - \delta)}{\alpha} \qquad (2\text{-}4)$$

where α, β, and δ are the scale, shape and location parameters, respectively.

Being a three-parameter distribution operating upon the logarithms of the variable, the Log-Pearson Type III would appear to be an extremely versatile distribution. However, its application in hydrology is strictly limited. For flood frequency analysis the only shape of interest is that which is the unimodal one, continuous from 0 to + ∞, has either an infinitely high order at the lower limit, and is unbounded at the upper limit. The Log-Pearson Type III falls within these criteria only when $\beta > 1$ and $1/\alpha > 0$. If the skew coefficient, δ, is negative, this corresponds to a negative value of α, and is not suitable. Reich (1972) has given examples of the application of this distribution to samples with negative skew in which the computed upper bounds have been lower than the maximum observed events.

A distribution that combines the three types of extreme value distributions, the Generalized Extreme Value distribution, using the Probability Weighted Moments method to estimate its parameters, is discussed next.

2.2 Generalized Extreme-Value Probability-Weighted Moments Method

Various methods have been proposed for estimating the parameters, and consequently the quantiles, of a distribution from a finite length sample. The method of moments and the maximum likelihood have been widely used to estimate the parameters, but need large samples to work efficiently.

A new class of moments, called probability-weighted moments (PWM), was introduced by Greenwood et al. (1979). It can be applied to distributions expressable in inverse form; that is, if X is a random variable and F is the value of the cumulative distribution function of x, i.e. the non-exceedance probability, the value of x may be expressed as a function of F, x = x(F). Greenwood et al. (1979) show that rather simple expressions for the parameters could be obtained for a number of distributions. Landwehr et al. (1979)

obtained inverse forms for the Gumbel distribution.

As demonstrated by Greis and Wood (1981) the PWM methodology allows the regional estimation of upper quantiles of flood frequency by simple averaging. The method is particularly robust for short sample records and is suitable for application to Puerto Rico basins.

The Probability Weighted Moments procedure is suitable for use with the so-called Generalized-Extreme Value distribution, as demonstrated by Hosking et al. (1985a).

Jenkinson (1955) introduced the Generalized-Extreme Value (GEV) distribution, combining into a single form the possible types of limiting distributions (Types I, II, and III) for extreme values, as derived by Fisher and Tippet (1928). This distribution can be written in the inverse form, thus the PWM method can be used for estimating its parameters.

The fundamental procedure for applying the Generalized Extreme-Value Probability-Weighted Moments (GEV/PWM) regional estimation is introduced in Chapter 3, based on the work of Hosking et al. (1985a) and Greis and Wood (1981).

Hosking et al. (1985b) have appraised the flood frequency procedure used in an earlier study of catchments in Great Britain (Flood Studies Report, 1975) utilizing the GEV/PWM procedure as one of the methodologies. Their results suggest the adoption of the GEV/PWM procedure as a substitute for the original methodology employed in the Flood Studies

Report. The GEV/PWM procedure gave consistently less variable estimates than other procedures commonly utilized in practice. Hosking et al. (1985b) explored the comparability of the PWM method with the maximum likelihood and the conventional moments method by means of Monte Carlo experiments, and found that the maximum likelihood was generally the most efficient method of the three, except for the case of small samples, where the PWM was the most efficient.

Hosking et al. (1985a) shows the advantages of using the PWM method for the estimation of the parameters of the GEV distribution. The estimators are straightfoward to compute. Hosking et al. (1985a) uses computer simulation to compare the PWM estimators with maximun likelihood estimators. It was found that the standard deviations of the PWM estimators were less than those of the maximum likelihood estimators for small samples (n = 15, 25). This was because the theoretical justification of the maximum likelihood approach was based on large-sample theory.

2.3 Regionalization Techniques

With regional analysis a combination of records from many stations is possible. This provides an increased data base from which better parameter estimates can be obtained. Specifically , it is possible to obtain regionalized frequency curves. Regionalization in this regard consists of grouping stations into homogeneous groups. Therefore, a test of homogeneity should be a prerequisite of any regional flood frequency analysis.

Regions for flood frequency analysis are often arranged to be coincident with recognized geographical, political or administrative boundaries. Such regions are likely to contain drainage basins with geomorphological diversity, whose flood-related geophysical and climatic characteristics may not be comparable. In this case a region-averaged frequency curve will be poorly defined. A more efficient classification scheme should consider the following approaches:

a. Classification by statistics of basin flood frequency distribution.

b. Classification by basin characteristics.

Classification methods include (Gordon, 1981) the partitioning, hierarchical, clumping, and geometrical methods. The first three methods are recognized as clustering, or cluster analysis methods. They share the property that a certain type of structure is imposed on the data. The last classification type does not explicity impose a structure on the set of objects.

An example of the first approach is the work of Mosley (1981), in New Zealand, using hierarchical cluster analysis to form groups of basins characterized by specific mean annual flood and coefficient of variation. Another example is the work of Wiltshire (1986a, 1986b) using the PWM algorithm to classify Great Britain basins.

As an illustration of the second approach, White (1975) used factor analysis of basin characteristics data to identify collections of physically similar basins in Pennsylvania.

Traditionally, regional flood frequency analysis has been undertaken with two fundamental techniques (Kite, 1977):

a) Index-Flood method.

The basic idea behind the index-flood method is to increase the reliability of the frequency characteristics within a region. If, within a hydrologically homogeneous area, a number of hydrometric stations have been operating and recording the effects of the same meteorologic factors, then a combination of these records will provide, not a longer record, but a more reliable record.

b) Multiple Regression method.

This method utilizes the relationship between discharges at specified return periods and basin characteristics. In general, the relationships take the form:

$$Q_T = f\ (A^a,\ B^b,\ C^c, \ldots,\ Z^z) \qquad (2\text{-}5)$$

where A, B, C,..., Z are independent variables and a, b, c,..., z are constants. Many different procedures are available for determining the relevant parameters in Equation (2-5), including simple linear regression, multiple linear regression, and foward, backward, and stepwise procedures.

A new regional estimation technique using the GEV/PWM procedure has been introduced for the generation of regionalized estimates of flood frequency (Greenwood et al., 1979; Hosking et al., 1985b; Landwehr et al., 1979). Greis and Wood (1981) showed that this technique gives less standard error when compared to regression methods.

The testing and application of the GEV/PWM procedure has shown that it is a robust procedure and highly recommended for use as a regional flood frequency quantile estimator. Therefore, it is highly desirable to apply the regional GEV/PWM methodology to catchments in Puerto Rico for the scientific and practical rewards this entails.

2.4 Previous work in Puerto Rico

The only relevant major scientific study on regional flood frequency in Puerto Rico that has been identified is the U.S.G.S. flood study report by López et al. (1979). In this report, regionalized estimates of the T-year flood are estimated through regression analysis of available flood data. Here, all basin data were used for only one regression. The regression obtained was of the following type:

$$Q_T = kA^x (Ann\ P)^y \qquad (2\text{-}6)$$

where Q_T is the T-year flood; k, x, and y are regression parameters; A is the catchment area, and (Ann P) is the average annual precipitation. The standard errors of prediction of these equations range from -32 to +48 percent.

For example, for computing the 100-year return period peak discharge for the Rio de la Plata basin, López et al. (1979) give k = 286, x = 0.832, y = 0.531, A = 208 mi^2, and (Ann P) = 68 in. Substitution in Equation (2-6) yields:

$$Q_{100} = 286\ (208)^{0.832}\ (68)^{0.531} = 228064\ cfs$$

The use of Equation (2-6) has, more or less, become standarized for flood frequency estimation on the Island. A major limitation of the regression approach in this case is the large standard errors obtained, due in part to the short records available, and to the fact that all stations were grouped together without regard for regional homogeneity. Also, regressions forcefully correlate available data with parameters that may possess considerable internal estimation uncertainty that manifests itself in noisy regional estimates and low correlation. The question also arises whether this particular model is the best for regional flood estimation.

The other methodology that has been extensively applied on the Island for individual basin studies is the Log-Pearson Type III flood frequency curve prescribed by the Water Resources Council (1977). There is no particular reason why this distribution should be the " best" distribution for Island basins. In fact, its use is questionable since the short records available on the Island introduce large uncertainty in the estimation of the higher quantiles obtained from this distribution.

CHAPTER III

RESEARCH PROCEDURE AND METHODOLOGY

The methods and procedures employed in the project are presented in this chapter.

The first task of the project was the acquisition of the relevant information. Available flood records were upgraded by collecting the latest available data. The necessary physical watershed parameters were identified during this task. Federal and Local goverment agencies, such as the U.S. Geological Survey and the Department of Natural Resources of the Commonwealth of Puerto Rico, were the mayor sources of the required information.

The mayor basins for which detailed analyses were undertaken were identified. The criteria for selecting these basins were related to the model data needs. Rivers with adequate and homogeneous records were selected. The major streamflow gaging stations in Puerto Rico were identified. These are shown in Figure 3-1, together with the major hydrologic subdivisions of the Island.

Following the data acquisition effort, the Generalized Extreme Value/Probability Weighted Moments (GEV/PWM) procedure was applied to the available annual flow records within each region. This task was concurrently undertaken with the clustering of basins. The formulation of the fundamental

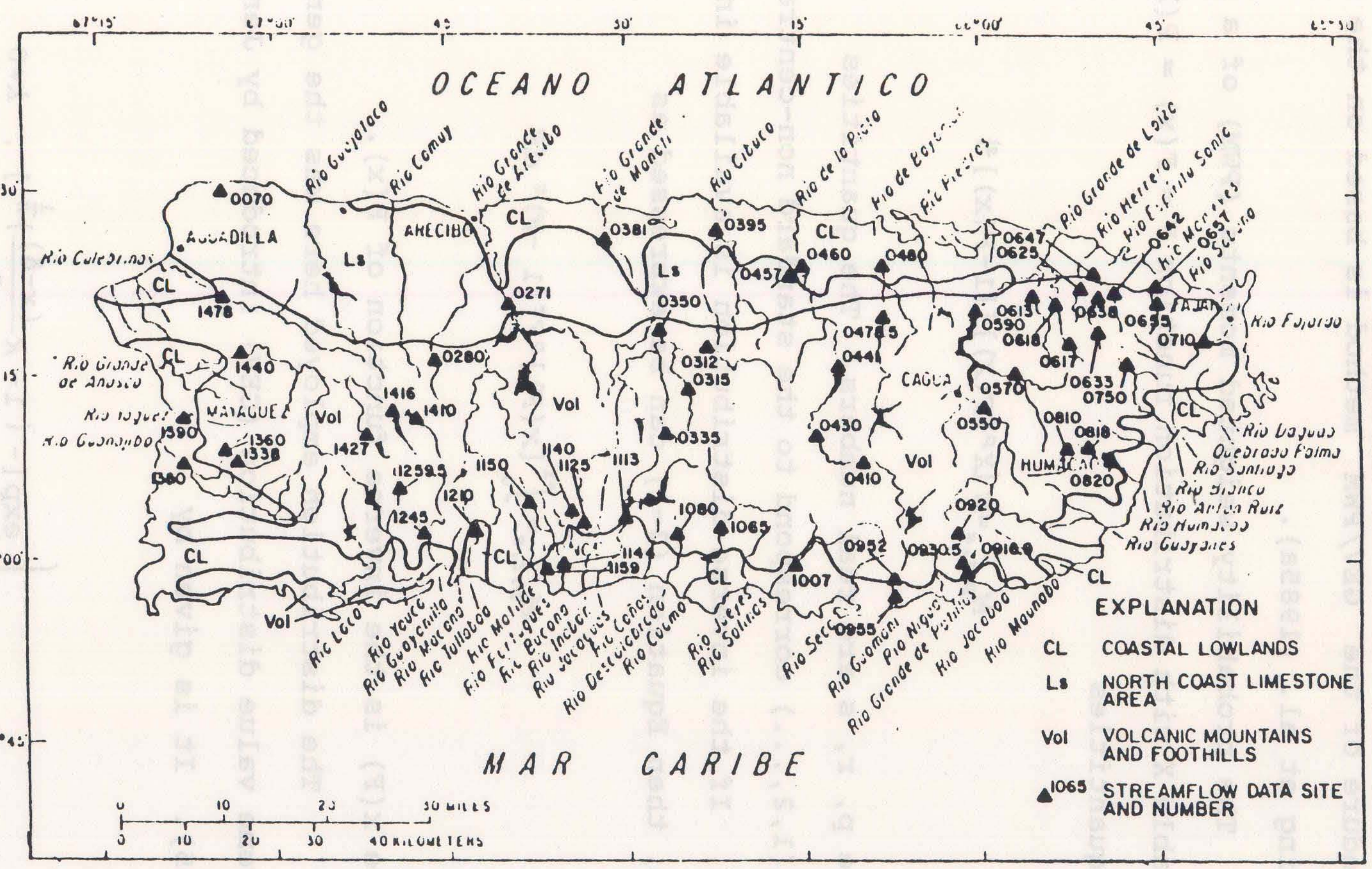

Figure 3-1
Hydrologic Subdivisions, Major River Systems, and Location of Streamflow Stations In P.R. (López et al. 1979)

procedure of the GEV/PWM method is based on the work of Hosking et al. (1985a).

The probability weighted moments (PWM) of a random variable X with distribution function $F = F(x) = P(X \le x)$ are the quantities

$$M_{p,r,s}=E\{X^p[F(X)]^r[1-F(x)]^s\} \qquad (3\text{-}1)$$

where p, r, s are real numbers. The quantities $M_{p,0,0}(1,2,...)$ correspond to the standard non-central moments of X. If the inverse distribution is available in closed form, then Equation (3-1) can be expressed as

$$M_{p,r,s}=\int_0^1 [x(F)]^p F^r (1-F)^s \, dF \qquad (3\text{-}2)$$

where x(F) is the inverse function of F(x).

The distribution employed here was the generalized extreme value distribution (GEV) introduced by Jenkinson (1955). It is given by

$$F(x) = \begin{cases} \exp[-\{1-k\frac{(x-\delta)}{\alpha}\}^{\frac{1}{k}}] , & k\neq 0 \\ \exp[-\exp\{-\frac{(x-\delta)}{\alpha}\}] , & k=0 \end{cases} \qquad (3\text{-}3)$$

with x bounded by $\delta + \alpha/k$ from above if k>0, and from below if k<0. Here, δ and α are location and scale parameters, respectively. The shape parameter k determines which distribution is represented. Fisher - Tippet types I, II, and III correspond to k=0, k<0, and k>0, respectively. When k=0 the distribution reduces to the Gumbel extreme value distribution. Usually k values falls between $-\frac{1}{2}$ and $\frac{1}{2}$.

The inverse distribution of Equation (3-3) is given by

$$x(F) = \begin{cases} \delta + \alpha\,[\,1-(\,-\log F\,)^{k}]/k & , k \neq 0 \\ \delta - \alpha \log(\,-\log F) & , k=0 \end{cases} \qquad (3\text{-}4)$$

Among the available moments, the following was employed, as suggested by Hosking et al. (1985a), for the evaluation of the extreme value distributions:

$$\beta_r = M_{1,r,0} = E\,\{X[F(x)]^r\} \qquad , r=0,1,2,\ldots. \qquad (3\text{-}5)$$

From Equation (3-4) the PWM of the GEV distribution for $k \neq 0$ are given by

$$\beta_r = (r+1)^{-1}\left\{\delta + \frac{\alpha\,[1-(r+1)^{-k}\,\Gamma(1+k)]}{k}\right\} \qquad (3\text{-}6)$$

For $k \leq -1$, β_r does not exist.

Equation (3-6) was used to obtain the PWM estimators δ', α', k' of the parameters δ, α, and k. The solution procedure was to estimate parameters from the first three moments of Equation (3-6). A suitable estimate of the solution (Hosking et al., 1985b) is given by

$$k' = 7.8590\, c + 2.9554\, c^2 \tag{3-7}$$

$$\alpha' = \frac{(2\,\beta_1 - \beta_0)\, k'}{\Gamma(1 + k')\,(1 - 2^{-k'})} \tag{3-8}$$

$$\delta' = \beta_0 + \frac{\alpha'[\,\Gamma(1 + k') - 1]}{k'} \tag{3-9}$$

where,

$$c = \frac{2\,\beta_1 - \beta_0}{3\,\beta_2 - \beta_0} - \frac{\log 2}{\log 3} \tag{3-10}$$

Given the estimated parameters the quantiles of the flood frequency distribution were obtained from Equation (3-4). The quantile corresponding to each specified exceedance probability corresponds to the T-year flood, the inverse of the probability corresponding to the T years between the observation of similar or larger events.

In order to obtain adequate regional flood frequency distributions and avoid large estimation errors it was necessary to define the parameters within hydrologically homogeneous groups. Here, a test of homogeneity of the available basins was conducted and, if necessary, the clustering of basins into new homogeneous groups was done to obtain better regionalized flood flow frequency relationships.

The first undertaking was the application of a homogeneity test to the selected basins. The procedure used was the R test proposed by Wiltshire (1986a). The basic statistic is given by

$$R = \left(\frac{1}{V}\right) \sum_{j=1}^{n} \frac{(G_j' - G)^2}{n_j} \qquad (3\text{-}11)$$

where G_j'= mean of folded non-exceedance probabilities at site j, G is the regional mean of the G_j' values, V is the variance of a uniform distribution (1/12), and n_j is the record length at the j^{th} site. The statistical test was applied to test the homogeneity hypothesis. The R statistic is expected to be Chi-square distributed with (N-1) degrees of freedom. For each cluster, the resultant R-value was compared with the $R_{critical}$ value given by the Chi-square distribution with (N-1) degrees of freedom. If the resultant R-value is less than the $R_{critical}$ value the cluster passes the test, and the null hypothesis accepted.

Ungauged basins were assigned to clusters by utilizing a procedure devised by Wiltshire (1986b), based on discriminant analysis. Discriminant analysis is used to statistically distinguish between two or more groups. To distinguish between the groups a collection of discriminating variables, which measure characteristics in which the groups are expected to differ, are needed. The mathematical device is to weight and linearly combine the discriminating variables so that the groups are forced to be as statistically distinct as possible. The use of discriminant analysis as a classification technique comes after the initial computation. Once a set of variables is found which provides satisfactory discrimination for the cases with known groups, a set of classification functions can be derived which will permit the classification of new cases with unknown memberships.

The method consists in finding a transformation, or combination of basin characteristics, which minimizes the ratio of the difference between the multivariate cluster means of the basin and the multivariate variance of the basin characteristics within clusters. The general form of the discriminant function is given by

$$S = L_0 + L_1U_1 + L_2U_2 + \ldots\ldots + L_AU_A \qquad (3-12)$$

where L_i's are coefficients reflecting the loss obtained when assigning a basin to the wrong cluster, L_0 is a constant, U_i's are vector of basin characteristic values, and S is the discriminant score.

Rao (1973) gives the discriminant score for each cluster as

$$S_i = (B_i^T C^{-1}) X - B_i^T C^{-1} \frac{B_i}{2} + \log W_i \qquad (3\text{-}13)$$

where B^T_i is a vector containing the means of each of the basin characteristics for each cluster i, C is the covariance matrix for the basin characteristics, and W_i is the prior probability of basin X being in cluster i.

The basin geomorphological and climatologic characteristics used were the following: basin area, average annual rainfall, stream frequency, baisn slope, mean annual potential evapotranspiration, 5 & 25 year 24-hour effective rainfall (U.S. Weather Bureau, 1961), and percentage of lakes in the basin area. A 1:250,000 topographic map was used to compute the basin slope and the stream frequency.

After the adequate hydrologic regions were defined, the regionalization process of flood frequency parameters was undertaken. The information available was regionalized within each particular subdivision. The GEV/PWM regionalization procedure follows that suggested by Greis and Wood (1981), and

Hosking et al. (1985b), and is based on the PWM procedure illustrated earlier. The regionalization procedure was conducted by performing the following steps:

a) The probability weighted moments β_r, r=0,1,2, were estimated for each site using the plotting position estimator $\beta_r = \beta_r'[p_j]$, where

$$\beta_r' [p_j] = \frac{1}{n} \sum_{j=1}^{n} p_j^r x_j \qquad (3\text{-}14)$$

and p_j is the plotting position estimate of $F(x_j)$ for each j_{th} smallest value of the ordered samples $x_1 \leq x_2 \leq \ldots \leq x_n$. The value of p_j was chosen as $p_j = (j - 0.35)/ n$. Hosking et al. (1985a) found that the biased estimator p_j outperformed the unbiased estimators for small samples sizes.

b) The estimated probability weighted moments were scaled by β_0, the sample mean, obtaining for each site the quantities

$$t_{j,i} = \frac{\beta_j}{\beta_0} \qquad (3\text{-}15)$$

for j = 1, 2 and i = 1, 2, ..., N.

c) The regional estimators were obtained from

$$T_j = \frac{\sum_i t_{j,i}\, n_i}{\sum_i n_i} \quad , \; j = 1,2 \qquad (3\text{-}16)$$

d) The regional GEV distribution parameters were estimated using the regional probability weighted moments $\beta_0 = 1$, $\beta_1 = T_1$, and $\beta_2 = T_2$ with the estimators described by Equation (3-16).

e) For each site, the at-site quantile x(F) is estimated from $\beta'_0 x'_R$, where x'_R is the corresponding quantile of the regional GEV distribution.

For the case of ungauged watersheds, the procedure is the one proposed by Wiltshire (1986b). In this procedure the estimate of the T-year flood X_T for a basin is given by

$$X_T = \frac{\sum_{i=1}^{M} P_i T_i}{\sum_{i=1}^{M} P_i} \qquad (3\text{-}17)$$

where $\Sigma P_i = 1$, T_i is the i^{th} cluster, or basin grouping, estimate of the T-year flood, P_i is the posterior probability of the new ungauged basin being in each of the M clusters or groupings. The probabilities are obtained from

$$P_i = \frac{\exp\,(S_i)}{\sum_{i=1}^{M} \exp\,(S_i)} \qquad (3\text{-}18)$$

where S_i has been defined previously as the cluster discriminant score.

Equation (3-17) is used to construct a dimensionless frequency curve for a new ungauged basin considering several return periods.

Several basins were set aside in order to estimate their quantiles from the regionalized parameter estimates, for the purpose of comparison with actual data and error estimation. Applications of the methodology to ungauged sites will be illustrated.

CHAPTER IV

REGIONALIZATION APPLICATIONS

4.1 Frequency Regionalization of Selected Basins

From the available flood records (U.S. Geological Survey, 1989), 30 major streamflow gauging stations were chosen. The selection criteria were based on the size of the historical flood records, that is, the ones with adequate data and geographical location. The selected stations are shown in Table 4-1. In Figure 4-1, the geographical location of the selected basins is shown.

In order to apply the GEV/PWM method to the selected basins, a computer program of the procedure was developed. The program computes the estimators of the probability weighted moments ($\beta_0, \beta_1, \beta_2$) using Equation (3-14), and the estimators of the GEV parameters k, α, and δ, by means of Equations (3-7) through (3-10). The program listing and sample output is presented in Appendix A.

The GEV-PWM procedure was applied to the selected basins, and the results are shown in Tables B-1 through B-30, and Figures B-1 through B-30 of Appendix B. Thereon, the quantiles from the GEV/PWM procedure are compared with the historical maximum flows (U.S.G.S., 1989). The return period (years) for the historical data was obtained from

$$Return\ Period = \frac{1}{(1 - p_j)} \qquad (4\text{-}1)$$

TABLE 4-1
SELECTED BASINS FOR REGIONALIZATION ANALYSIS

STATION	BASIN
50038100	Río Grande de Manatí
50039500	Río Cibuco
50046000	Río De La Plata
50048000	Río Bayamón
50059000	Río Grande de Loiza
50062500	Río Herrera
50064200	Río Grande
50065700	Río Mameyes
50071000	Río Fajardo
50075000	Río Icacos
50082000	Río Humacao
50106500	Río Coamo
50111300	Río Jacaguas
50112500	Río Inabón
50114400	Río Bucaná
50115900	Río Portugües
50121000	Río Tallaboa
50124500	Río Guayanilla
50136000	Río Rosario
50144000	Río Grande de Añasco
50147800	Río Culebrinas
50028000	Río Tanamá
50029000	Río Grande de Arecibo
50056400	Río Valenciano
50057000	Río Gurabo
50061800	Río Canóvanas
50063800	Río Espíritu Santo
50114000	Río Cerrillos
50141000	Río Yahuecas
50092000	Río Grande de Patillas

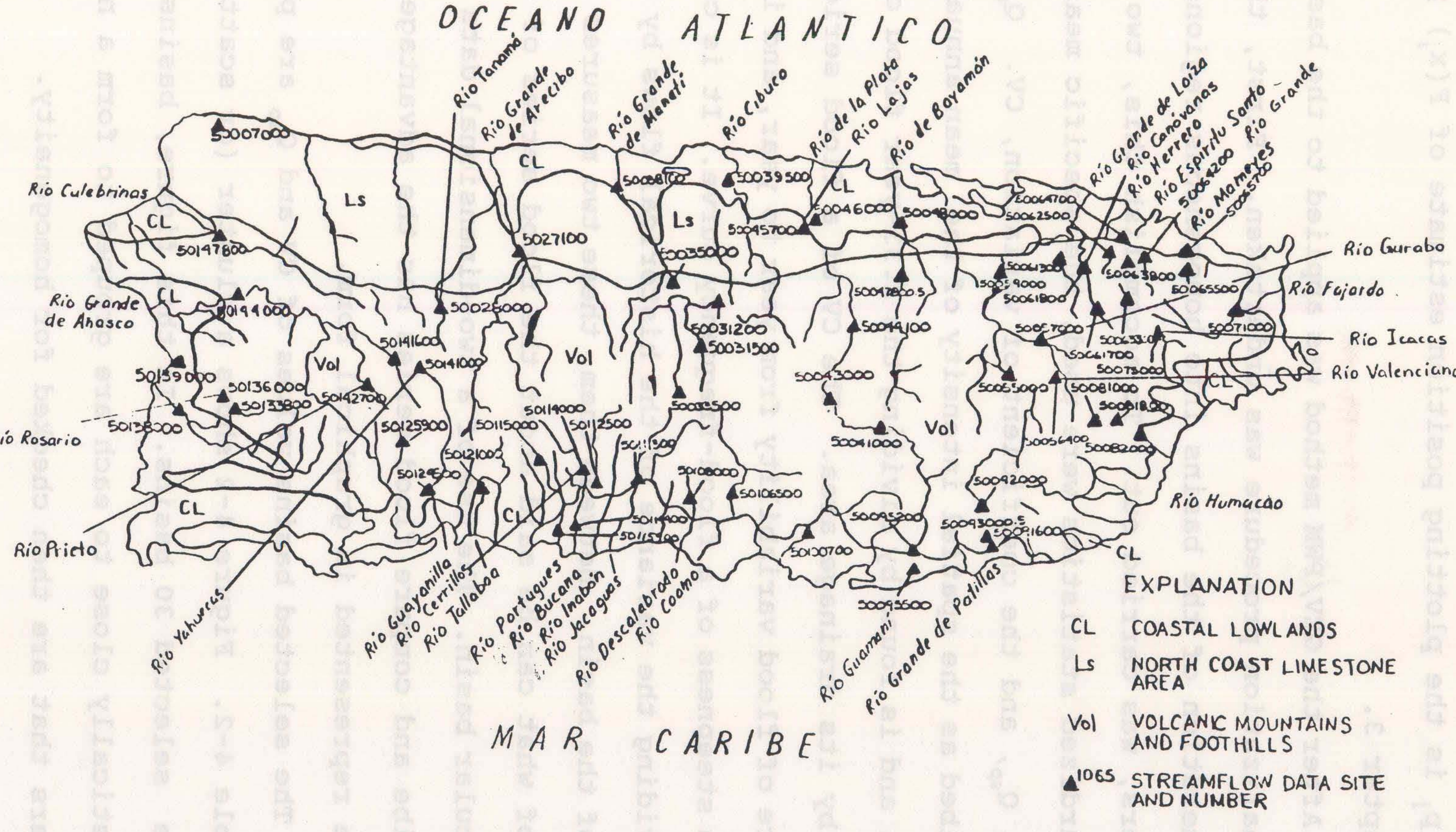

Figure 4-1
Geographical Locations of Selected Basins
(Adapted from López et al. 1979)

where p_j is the plotting position estimate of $F(x_j)$ presented in Chapter 3.

After the GEV/PWM method was applied to the basins, the regionalization procedure was undertaken. First, the agglomeration of the basins into homogeneous regions, or clusters, was carried out. To accomplish this, two standardized statistics were used, the specific mean annual flood, Q_{sp}, and the coefficient of variation, CV. Q_{sp} can be described as the spatial intensity of the mean annual maximum flood, and is found by dividing the 2.33-year flood of a basin (Q_{bar}) by its drainage area. The CV of a flood series is a measure of flood variability from year to year, and is related to the steepness of a flood-frequency curve. It is calculated by dividing the variance of the historical flows by the mean flow of the basin. Between them, these two measures summarize much of what can be said about the flood series of a particular basin. The use of a two-dimensional data space to describe and compare flood series has the advantage that it can be represented in graphical form.

The selected basins values of CV and Q_{sp} are presented in Table 4-2. Figure 4-2 shows a cluster (or scatter) plot of the selected 30 basins. In this figure, basins statistically close to each are grouped to form a number of clusters that are then checked for homogeneity.

A number of trials were conducted to obtain a cluster arrangement that would pass the homogeneity test and thus

TABLE 4-2
SELECTED BASINS VALUES OF CV AND Q_{sp}

Station	CV	Q_{sp}
50038100	0.852	1.32
50039500	0.926	0.82
50046000	1.077	1.62
50048000	1.028	1.78
50059000	0.811	3.28
50062500	0.455	9.50
50064200	0.618	8.93
50065700	0.568	11.22
50071000	0.610	6.48
50075000	0.419	11.55
50082000	1.090	3.83
50092000	0.573	3.07
50106500	1.042	1.92
50111300	1.085	2.85
50112500	1.390	1.67
50011440	1.029	2.97
50115900	1.229	2.25
50121000	0.818	3.22
50124500	1.155	2.55
50136000	0.801	3.97
50144000	1.209	1.74
50147800	0.481	4.11
50028000	0.373	3.75
50029000	1.400	0.55
50056400	0.803	8.42
50057000	0.785	5.30
50061800	0.602	7.69
50063800	0.548	10.07
50114000	0.940	2.28
50141000	0.473	5.40

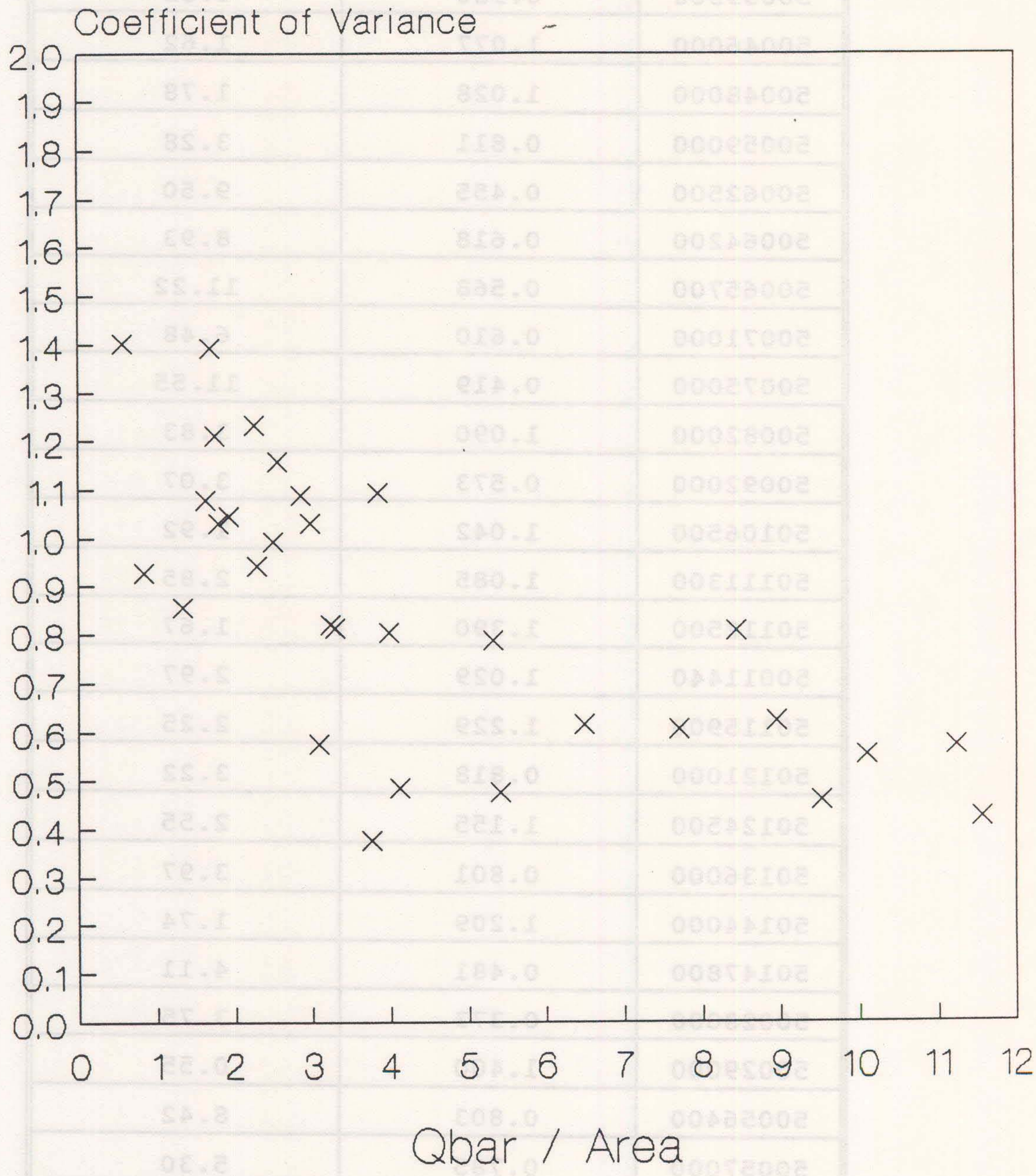

Figure 4-2
Scatter Plot for Cluster Analysis

accept the null hypothesis. This arrangement was obtained with the configuration of the four clusters shown in Figure 4-3. Table 4-3 shows the basin distribution for each cluster.

As indicated in Chapter 3, the R-test was performed to test for homogeneity within the basin clusters. To perform the homogeneity test the probability weighted moments were scaled by β_0 (the sample mean), for each site of the four clusters selected, to obtain the quantities $t_{1,i}$, and $t_{2,i}$, for i=1 to N, by means of Equation (3-15), as described in Chapter 3. With these values the regional estimators T_1 and T_2 for the four regions were computed from Equation (3-16).

These results appear in Appendix C. Tables C-1 and C-2 show the probability weighted moments β_0, β_1, and β_2, the quantities $t_{1,i}$, and $t_{2,i}$, and the regional estimators T_1, and T_2 for each of the four clusters. As an illustration, the regional estimator T_1 and T_2 for Cluster 1 are computed from Equation (3-16) as:

$$T_1 = \frac{\sum t_{1,1}\, n_1}{\sum n_1} = \frac{207.2235}{274} = 0.7563$$

$$T_2 = \frac{\sum t_{2,1}\, n_1}{\sum n_1} = \frac{173.4874}{274} = 0.6332$$

Then, the regional GEV distribution parameters k, α, δ, for each cluster are computed by means of Equations (3-7)

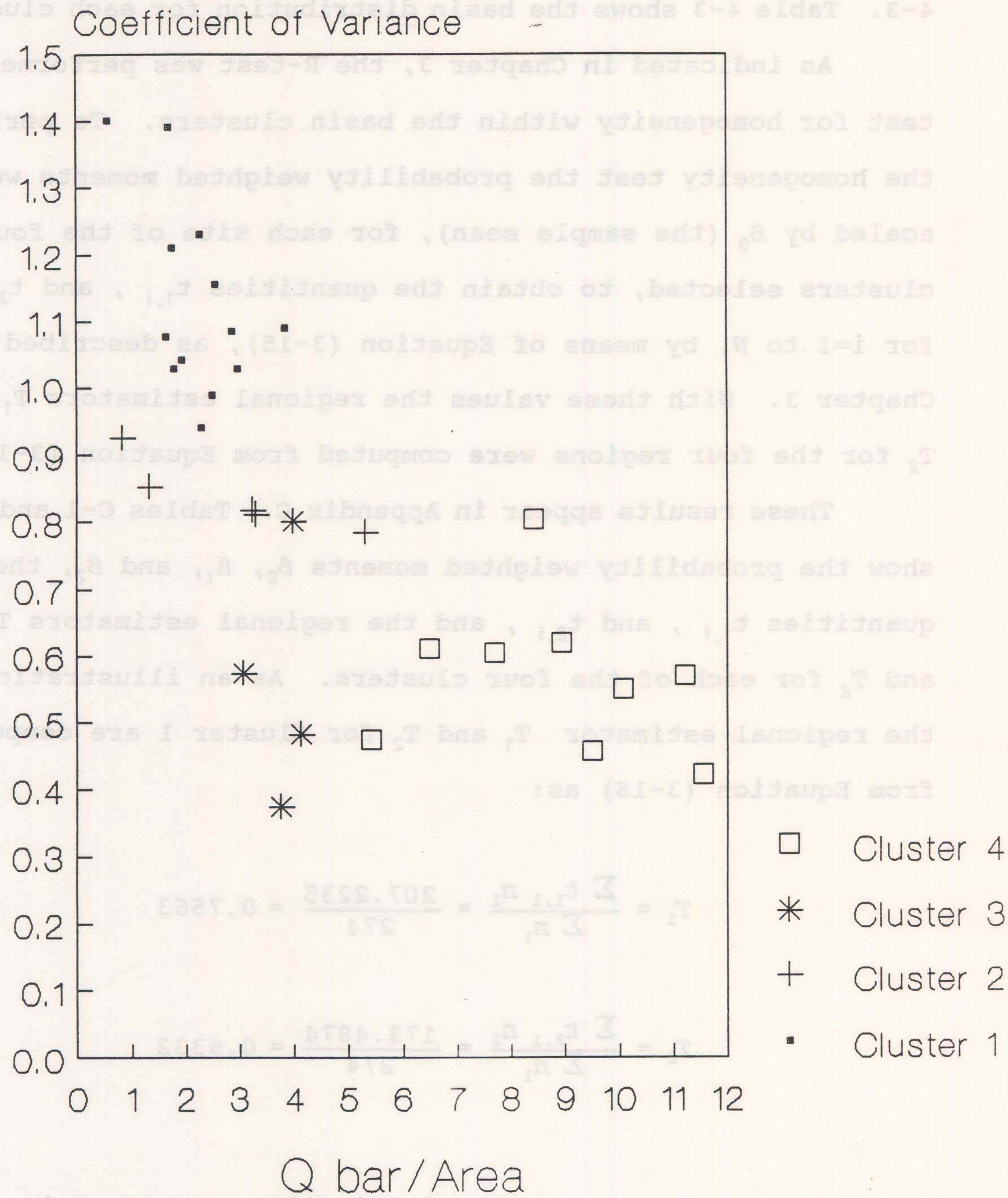

Figure 4-3
Homogeneous Cluster Configuration

TABLE 4-3
SELECTED BASIN ARRANGEMENT INTO CLUSTERS FOR REGIONALIZATION

CLUSTER 1		CLUSTER 2	
STATION	BASIN	STATION	BASIN
50029000	Río Grande Arecibo	50057000	Río Gurabo
50112500	Río Inabón	50121000	Río Tallaboa
50115900	Río Portugües	50059000	Río Grande Loiza
50144000	Río Grande Añasco	50039500	Río Cibuco
50124500	Río Guayanilla	50038100	Río Grande Manatí
50082000	Río Humacao		
50046000	Río De La Plata		
50106500	Río Coamo		
50048000	Río Bayamón		
50011440	Río Bucaná		
50114000	Río Cerrillos		
50111300	Río Jacaguas		
CLUSTER 3		**CLUSTER 4**	
STATION	BASIN	STATION	BASIN
50092000	Río Grande Patillas	50062500	Río Herrera
50147800	Río Culebrinas	50064200	Río Grande
50028000	Río Tanamá	50065700	Río Mameyes
50136000	Río Rosario	50071000	Río Fajardo
		50075000	Río Icacos
		50056400	Río Valenciano
		50061800	Río Canóvanas
		50063800	Río Espíritu Santo
		50141000	Río Yahuecas

to (3-10), by fixing the probability weighted moments as $\beta_0=1$, $\beta_1=T_1$, $\beta_2=T_2$. Table 4-4 shows the resulting regional GEV distribution parameters k', α', and δ', for each of the four clusters.

With these regional GEV distribution parameters the quantities $G_{jk}(x)$ and G'_{jk} are computed for each flood k at each station j of the four clusters. The quantities $G_{jk}(x)$, or the non-exceedance probability, were computed as F(x) from Equation (3-3). These values (G_{jk}) should have a uniform distribution, so their expected value is 0.5 and their variance 1/12. The folded, non-exceedance probabilities, G'_{jk}, transform the G_{jk} values by the expression

$$G'_{jk} = 2 \mid G_{jk} - 0.5 \mid \quad (4-2)$$

The set of G'_{jk} values are the transformed G_{jk} values obtained by "folding" the distribution to the left of the 0.5 value into the right-hand part of the distribution. The results for sites whose distribution is the same as the parent distribution is another identical uniform distribution. At sites whose data do not derive from the parent, systematic departures of G_{jk} points from the expected horizontal uniform distribution will be effectively amplified by the folding process. The values of G_{jk} and G'_{jk} obtained for all clusters are shown in Tables C-3 to C-13 of Appendix C.

The values of G_j' and G necessary to compute the R-value

TABLE 4-4

REGIONAL GEV DISTRIBUTION PARAMETER ESTIMATES

CLUSTER	k	α	δ
1	- 0.46898320	0.3743860	0.4639132
2	- 0.18124230	0.5521330	0.5619375
3	- 0.08701749	0.3586664	0.7593919
4	- 0.04394704	0.4388174	0.7268367

of Equation (3-11) were obtained. The mean of folded non-exceedance probabilities at site j, G'_j , is obtained from

$$G'_j = \frac{1}{n_j} \sum_{k=1}^{n_j} G'_{jk} \tag{4-3}$$

and G, the regional mean of the G_j' values is obtained from

$$G = \frac{\sum_{j=1}^{N} n_j G'_j}{\sum_{j=1}^{N} n_j} \tag{4-4}$$

where N is the number of stations in the cluster. For example the value of G_j' for station 50029000 of Cluster 1 is given by:

$$G'_j = \frac{1}{18} (12.393) = 0.688$$

and the value of G for Cluster 1 is

$$G = \frac{141.774}{274} = 0.517$$

With these, the R-value for the clusters was computed with Equation (3-11). Values of G_j', G, and R for each cluster are presented in Tables 4-5 to 4-8.

TABLE 4-5

VALUES OF G'_{jk}, G'_j, AND R FOR CLUSTER 1

STATION	n	$\Sigma G'_{jk}$	G'_j	R_i
50029000	18	12.393	0.689	6.324
50112500	26	12.159	0.468	0.773
50115900	24	14.050	0.585	1.331
50144000	27	12.329	0.457	1.198
50124500	23	14.036	0.610	2.379
50082000	15	7.101	0.473	0.349
50111300	29	14.934	0.515	0.002
50046000	33	19.967	0.605	3.042
50106500	20	8.332	0.417	2.440
50048000	16	8.579	0.536	0.0682
50011440	22	9.361	0.425	2.232
50114000	21	8.533	0.406	3.110
Σ =	274	141.774		
G = 0.517			R =	20.135

TABLE 4-6

VALUES OF G'_{jk}, G'_j, AND R FOR CLUSTER 2

STATION	n	$\Sigma G'_{jk}$	G'_j	R_i
50121000	21	11.113	0.529	0.280
50059000	13	6.899	0.531	0.158
50057000	30	17.844	0.595	0.375
50039500	31	17.950	0.579	0.101
50038100	31	17.074	0.551	0.052
Σ =	126	70.881		
G =	0.5638		R =	0.966

TABLE 4-7

VALUES OF G'_{jk} , G'_j , AND R FOR CLUSTER 3

STATION	n	$\Sigma G'_{jk}$	G'_j	R_i
50136000	21	8.729	0.416	0.959
50092000	24	14.319	0.597	4.099
50147800	22	10.110	0.460	0.0846
50028000	30	13.143	0.438	0.554
Σ =	97	46.301		
G =	0.477		R =	5.698

TABLE 4-8

VALUES OF G'_{jk} , G'_j , AND R FOR CLUSTER 4

STATION	n	$\Sigma G'_{jk}$	G'_j	R_i
50056400	19	11.982	0.631	3.443
50064200	15	8.280	0.552	0.352
50061800	22	13.155	0.598	2.148
50063800	23	10.919	0.475	0.301
50065700	18	9.718	0.540	0.223
50062500	18	8.394	0.466	0.371
50075000	27	10.319	0.382	5.110
50071000	30	16.372	0.546	0.519
50141000	30	13.428	0.448	1.303
Σ =	202	102.567		
G = 0.508				R = 13.770

After the R-value of each cluster was computed, the remaining task was to check if these values were smaller than $R_{critical}$ at the 1% significance level. The comparisons between the obtained R-value and $R_{critical}$ are shown in Table 4-9. The results show that the selected configuration of four clusters, or regions, are satisfactory.

The geographical cluster distribution is illustrated in Figure 4-4. The distribution shows that basins are, to a great extent, distributed within specific areas, corresponding roughly to climatological regions. Cluster 1 is defined predominantely within the southern region of the Island, Cluster 2 within the northern region, Cluster 3 within the western part, and Cluster 4 within the eastern region. However, not all basins within a given region are to be found within the cluster corresponding to that region. Notable exceptions are Río Grande de Arecibo, found within Cluster 1 while located within the region where Cluster 2 basins abound, and Río Grande de Añasco, found within Cluster 1 while located within the region where Cluster 3 basins predominate. A probable explanation for these results is deferred until results are discussed in Chapter 5.

It must be pointed out that it was attempted to treat all basins as a single region (only one cluster). For this arrangement the R-value was higher than the acceptable $R_{critical}$ value from the Chi-square distribution. So, this indicates that Puerto Rico basins cannot be described by a

TABLE 4-9

HOMOGENEITY STATISTICS FOR THE FOUR REGIONS IN P.R.

REGION NUMBER	NUMBER OF SITES	CALCULATED R	CRITICAL R AT 1%
1	12	20.10	24.70
2	5	0.97	13.30
3	4	5.70	11.30
4	9	13.77	21.70

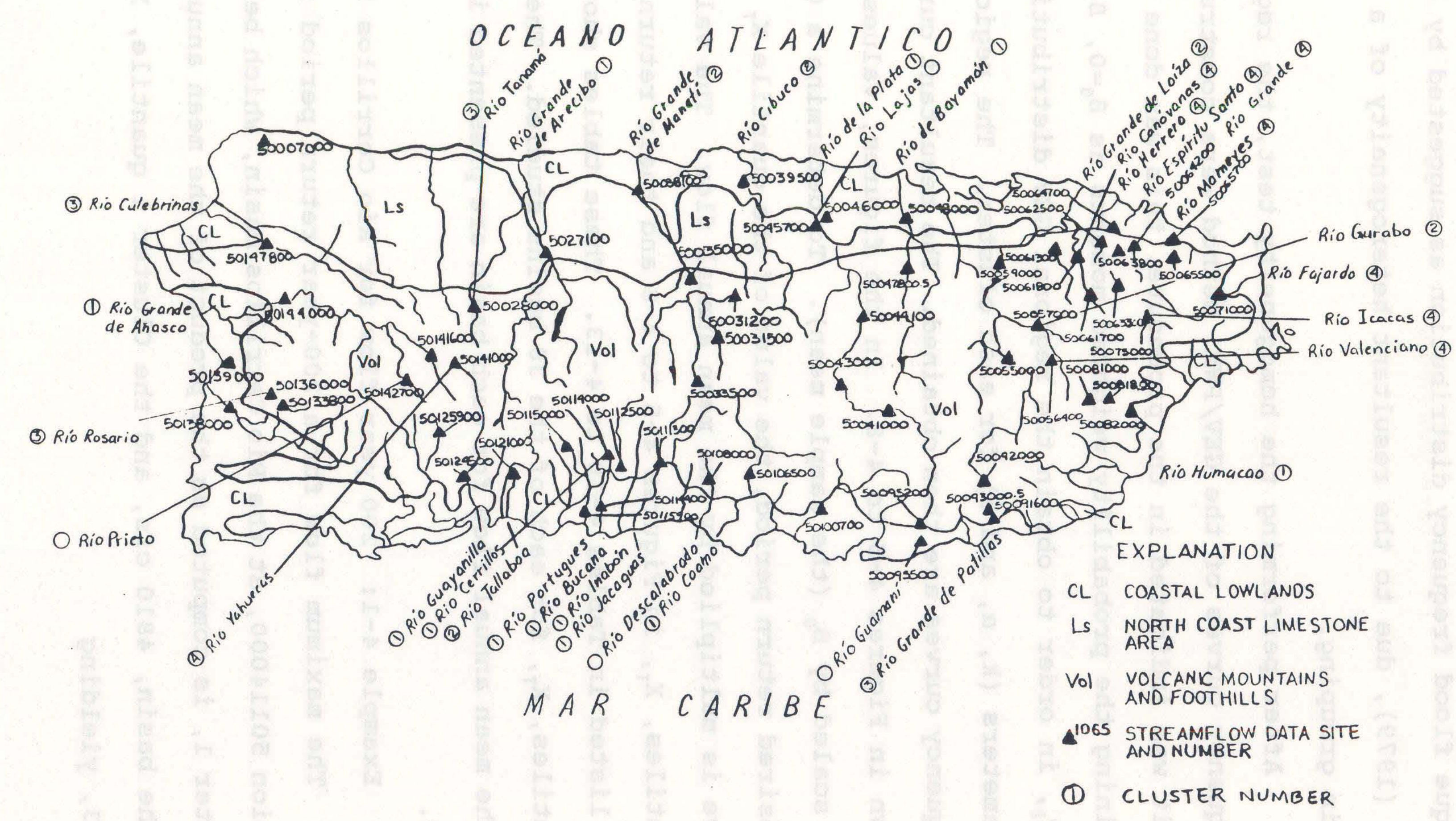

Figure 4-4
Geographical Cluster Distribution
(Adapted from López et al. 1979)

unique flood frequency distribution as suggested by López et al. (1979), due to the resultant heterogeneity of a single basin grouping.

After performing the homogeneity test, the regional frequency curves of the GEV/PWM method were constructed. As it was indicated in the procedure, this was done by defining the probability weighted moments as $\beta_0=0$, $\beta_1=T_1$, and $\beta_2=T_2$, in order to obtain the regional GEV distribution parameters (k, α, and δ) for each cluster. The regionalized frequency curves are then obtained. The resultant curves are shown in Figures 4-5 to 4-8. In the figures, values of Q_{max} are scaled by β_0 (the sample mean). To determine a Q_{max} for a desired return period, the value of the quantile X_T from the curve is multiplied by the mean annual flow. The values the quantiles, X_T, in Figures 4-5 to 4-8 and their return periods are listed in Tables 4-10 to 4-13. These tables show the quantiles, X_T, for each of the 30 basins studied. The values of the mean annual flow for each basin are presented in Table 4-14.

Example 4-1: 200 year flow for Rio Cerrillos Basin.

The maximum flow for a 200-year return period for station 50114000, at the Rio Cerrillos basin, which belongs to Cluster 1, is computed as the product of the mean annual flow of the basin, 4810 cfs, and the Cluster 1 quantile, X_{200} = 9.233, yielding

TABLE 4-10
QUANTILES FOR CLUSTER 1

Return Period (years)	Quantile X_T
2	0.615
3	0.885
4	1.100
5	1.279
10	1.959
15	2.463
20	2.880
25	3.244
30	3.569
35	3.867
40	4.142
45	4.399
50	4.642
55	4.872
60	5.090
65	5.300
70	5.500
75	5.694
80	5.880
85	6.060
90	6.235
95	6.405
100	6.570
200	9.233
300	11.241
400	12.916
500	14.380

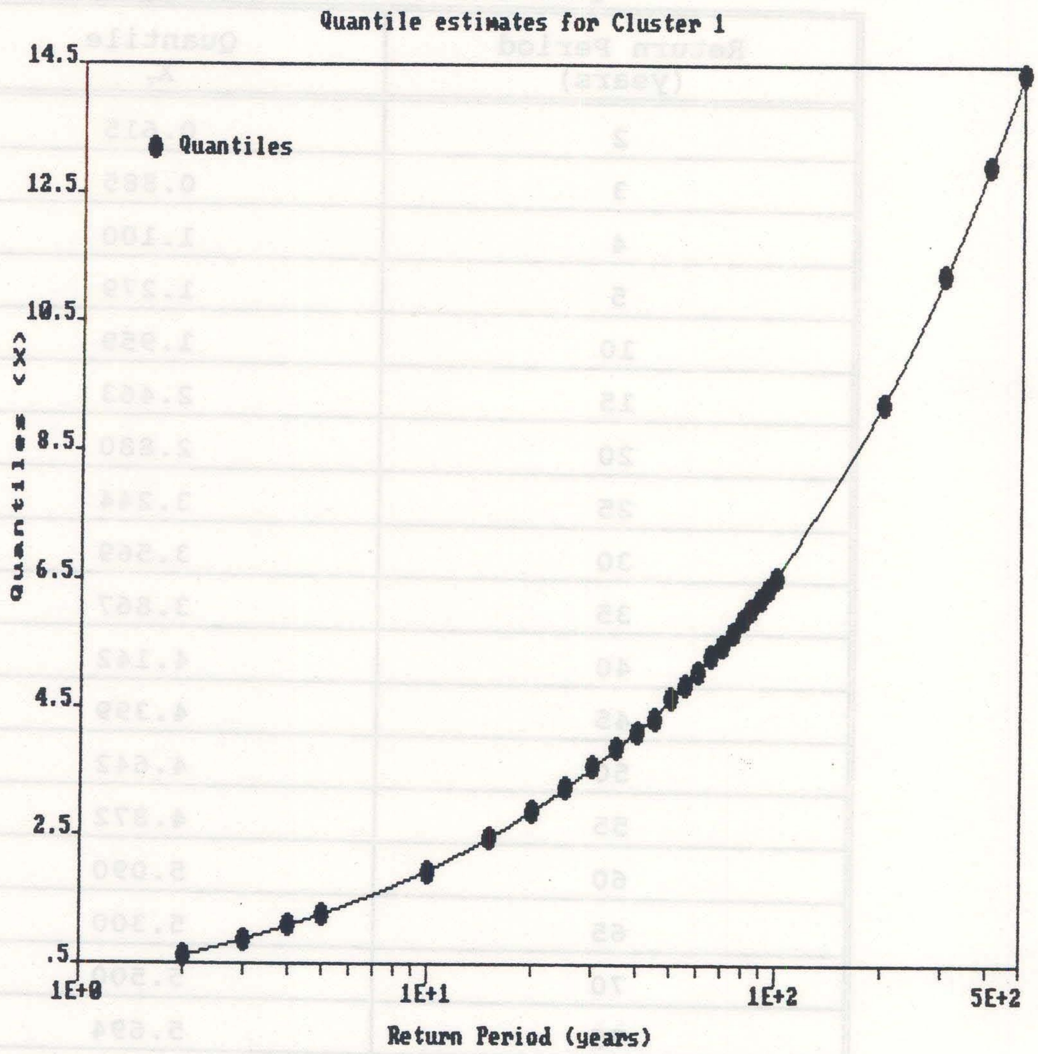

Figure 4-5
Quantile estimates for Cluster 1

TABLE 4-11

QUANTILES FOR CLUSTER 2

Return Period (years)	Quantile X_T
2	0.771
3	1.103
4	1.334
5	1.514
10	2.096
15	2.461
20	2.734
25	2.955
30	3.141
35	3.303
40	3.447
45	3.576
50	3.695
55	3.803
60	3.904
65	3.998
70	4.087
75	4.170
80	4.249
85	4.323
90	4.395
95	4.463
100	4.528
200	5.470
300	6.078
400	6.537
500	6.910

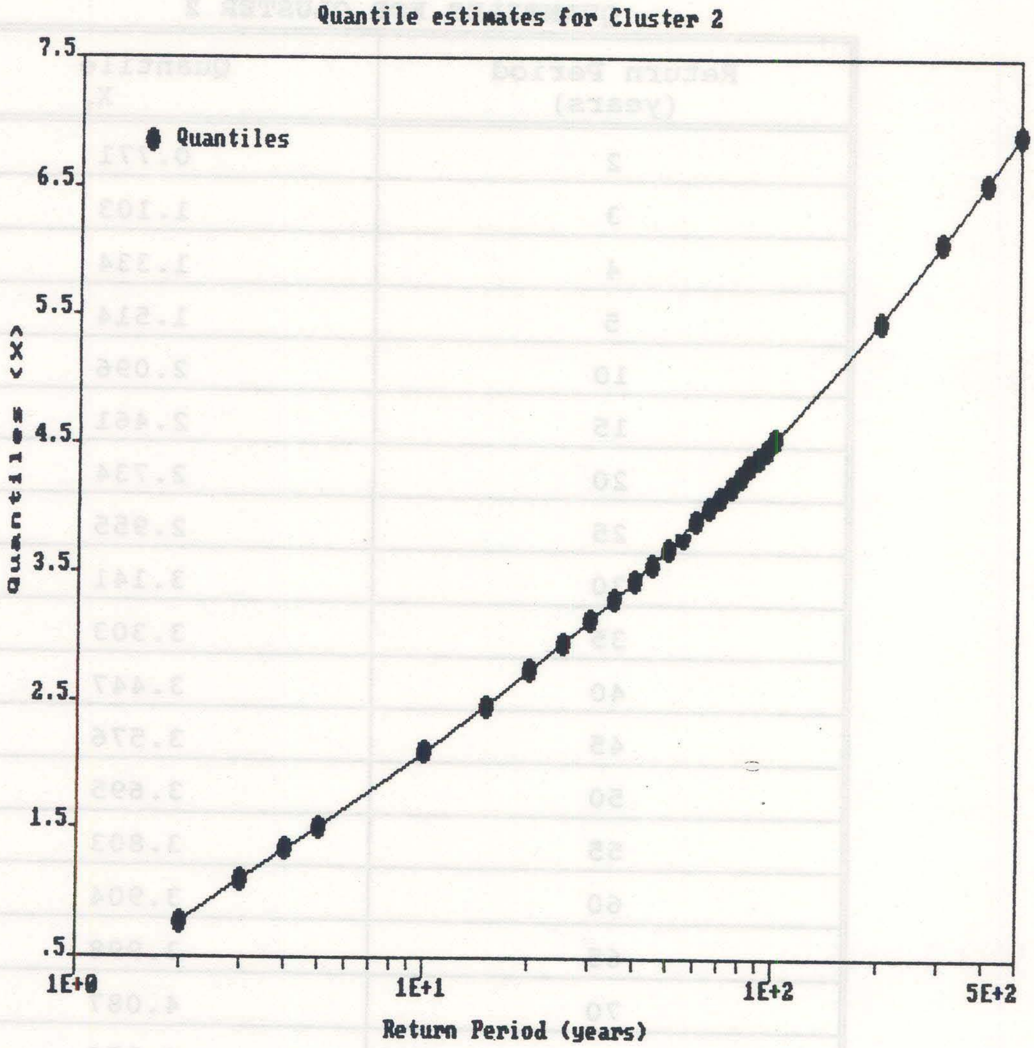

Figure 4-6
Quantile estimates for Cluster 2

TABLE 4-12

QUANTILES FOR CLUSTER 3

Return Period (years)	Quantile X_T
2	0.893
3	1.096
4	1.231
5	1.334
10	1.651
15	1.839
20	1.975
25	2.082
30	2.171
35	2.247
40	2.313
45	2.372
50	2.426
55	2.474
60	2.519
65	2.561
70	2.599
75	2.635
80	2.669
85	2.702
90	2.732
95	2.761
100	2.788
200	3.172
300	3.407
400	3.579
500	3.716

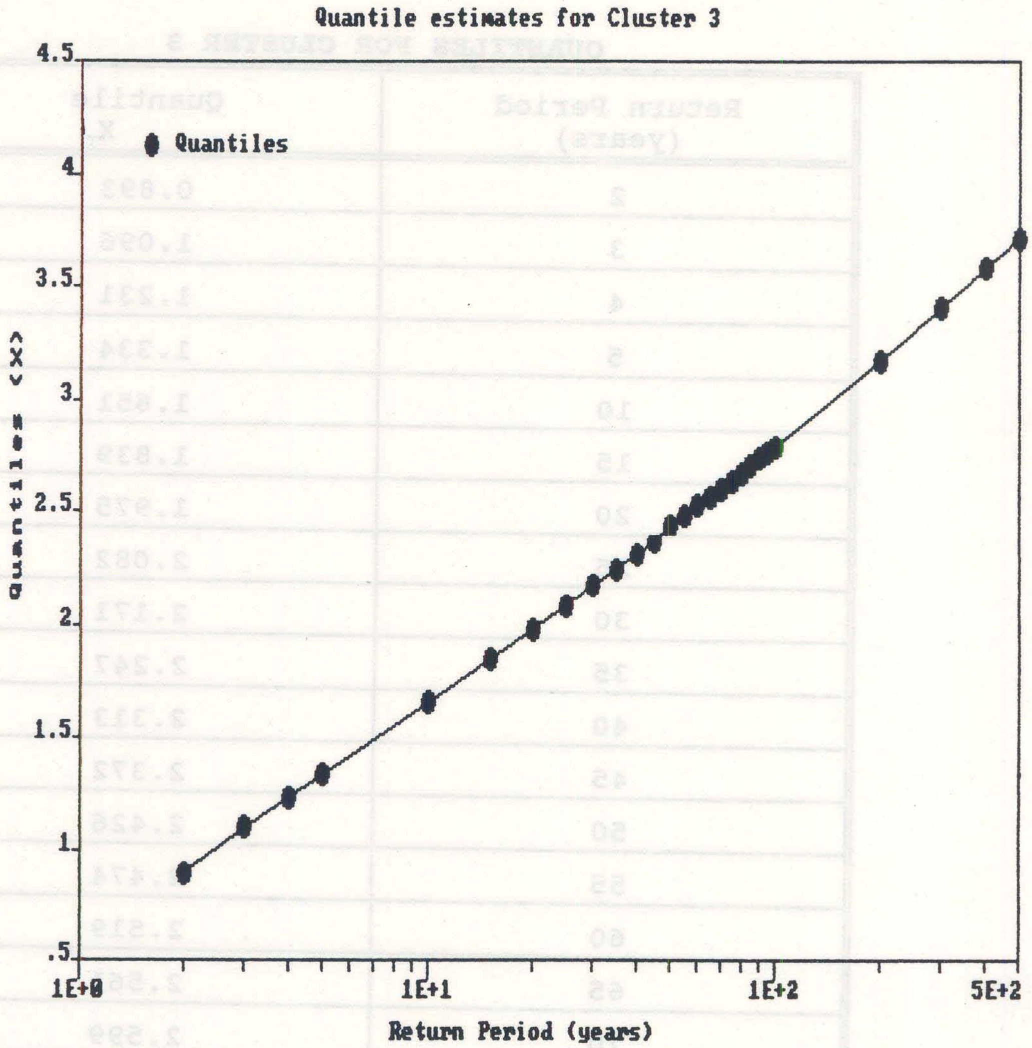

Figure 4-7
Quantile estimates for Cluster 3

TABLE 4-13

QUANTILES FOR CLUSTER 4

Return Period (years)	Quantile X_T
2	0.889
3	1.131
4	1.289
5	1.407
10	1.765
15	1.972
20	2.119
25	2.234
30	2.328
35	2.408
40	2.478
45	2.539
50	2.595
55	2.645
60	2.691
65	2.733
70	2.773
75	2.810
80	2.844
85	2.877
90	2.907
95	2.936
100	2.964
200	3.343
300	3.570
400	3.734
500	3.862

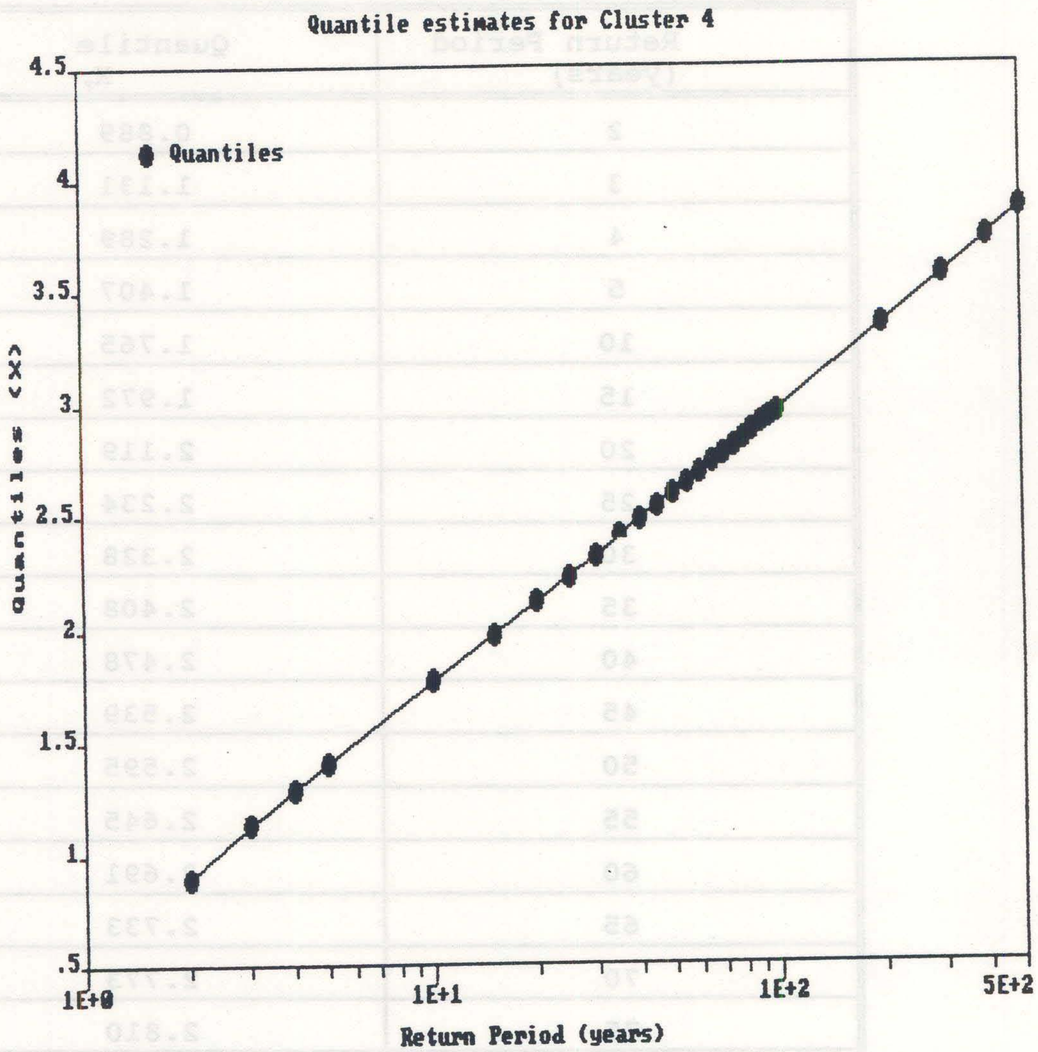

Figure 4-8
Quantile estimates for Cluster 4

TABLE 4-14
MEAN ANNUAL FLOWS FOR THE SELECTED BASINS

STATION	BASIN	Q_{mean} (cfs)
50038100	Río Grande de Manatí	32,542
50039500	Río Cibuco	10,198
50046000	Río De La Plata	34,882
50048000	Río Bayamón	19,199
50059000	Río Grande de Loiza	66,054
50062500	Río Herrera	2,301
50064200	Río Grande	7,437
50065700	Río Mameyes	13,583
50071000	Río Fajardo	9,062
50075000	Río Icacos	1,426
50082000	Río Humacao	8,605
50106500	Río Coamo	10,820
50111300	Río Jacaguas	14,782
50112500	Río Inabón	2,545
50011440	Río Bucaná	9,278
50115900	Río Portugües	6,041
50121000	Río Tallaboa	7,548
50124500	Río Guayanilla	8,844
50136000	Río Rosario	6,529
50144000	Río Grande de Añasco	24,119
50147800	Río Culebrinas	26,467
50028000	Río Tanamá	5,835
50029000	Río Grande de Arecibo	29,201
50056400	Río Valenciano	13,355
50057000	Río Gurabo	25,359
50061800	Río Canóvanas	6,351
50063800	Río Espíritu Santo	7,917
50114000	Río Cerrillos	4,810
50141000	Río Yahuecas	6,769
50092000	Río Grande de Patillas	5,319

$$Q_{200} = 9.233(\ 4810\ cfs\ ) = 44{,}411\ cfs$$

So, the maximim flow for a 200-year return period for the Río Cerrillos basin is 44,441 cfs. A similar procedure is employed for estimating the maximum annual flow at any of the other gauged stations. ■

The procedure is further generalizated to encompass ungauged catchments.

4.2 Application to Ungauged Basins

One of the objectives of this research was to allocate ungauged basins to the clusters formed in the preceding section. A multivariate discriminant analysis was performrd considering basin geomorphological and climatological characteristics as the alocation mechanism. Eight basin characteristics were used to perform the discriminant analysis. Here , a different set of characteristics than those employed in previous studies (Wiltshire, 1986b) were used. These characteristics were found to describe better the geomorphologic and climatologic variables in Puerto Rico. These characteristics are listed and defined in Table 4-15. Table 4-16 shows selected characteristics for the 30 basins previously grouped into four clusters.

To illustrate the application to ungauged basins, four cahtments were selected. These have short historical flow records (n=10), but, for the purpose of evaluating the

TABLE 4-15

BASIN CHARACTERISTICS MEASURED FOR DISCRIMINANT ANALYSIS

CHARACTERISTICS	DEFINITION	UNITS
AREA	Basin drainage area	mi^2
AAR	Average annual rainfall	inches
RSMD1	5 year return period 24 hour rainfall	inches
RSMD2	25 year return period 24 hour rainfall	inches
MAPE	Mean annual potential evapotranspiration	inches
STMFRQ	Stream frequency	segments/area
SLOPE	Stream slope	m / km
LAKE	Fraction of basin area draining through a lake	%

TABLE 4-16
DISCRIMINANT CHARACTERISTICS OF 30 GAUGED BASINS

STATION	AREA	AAR	RSMD1	RSMD2
50038100	197.00	70.0	6.3	9.0
50039500	99.10	65.0	6.1	9.0
50046000	208.00	68.0	6.0	8.5
50048000	71.90	70.0	6.5	8.6
50059000	209.00	89.0	7.5	9.9
50062500	2.75	85.0	7.7	10.5
50064200	7.31	93.0	8.0	11.0
50065700	11.80	80.0	8.0	11.0
50071000	14.90	78.0	8.0	10.9
50075000	1.26	120.0	8.8	11.8
50082000	17.30	85.0	8.2	12.0
50106500	46.00	35.0	6.4	9.5
50111300	43.50	40.0	6.1	9.0
50112500	9.70	37.0	6.3	9.2
50011440	25.60	35.0	6.0	8.9
50115900	18.60	35.0	6.2	9.1
50121000	24.20	45.0	7.5	11.0
50124500	20.80	40.0	7.6	11.5
50136000	16.40	75.0	7.0	9.8
50144000	94.30	100.0	6.0	7.7
50147800	16.70	90.0	6.0	7.5
50028000	18.40	90.0	6.5	8.5
50029000	200.00	75.0	6.0	7.5
50056400	16.40	70.0	7.9	11.0
50057000	60.20	75.0	7.0	10.0
50061800	9.84	100.0	7.5	10.0
50063800	8.62	95.0	8.0	10.9
50114000	17.80	40.0	6.5	9.5
50141000	15.40	85.0	9.0	13.0
50092000	18.30	80.0	7.0	10.3

TABLE 4-16
(Cont.)
DISCRIMINANT CHARACTERISTICS OF 30 GAUGED BASINS

STATION	MAPE	STMFRQ	SLOPE	LAKE
50038100	55.0	0.4670	12.458	1.0
50039500	55.0	0.3330	20.833	0.0
50046000	50.0	0.7596	9.968	3.0
50048000	55.0	0.6537	14.388	0.5
50059000	55.0	0.4737	8.772	3.0
50062500	50.0	0.3636	38.580	0.0
50064200	45.0	0.5472	85.979	0.0
50065700	55.0	0.9322	40.541	0.0
50071000	50.0	1.0067	49.020	0.0
50075000	45.0	0.7937	5.787	0.0
50082000	65.0	0.6936	13.123	0.0
50106500	55.0	0.7609	31.172	0.0
50111300	55.0	0.5287	35.328	1.0
50112500	55.0	0.4124	54.945	0.0
50011440	55.0	0.5859	49.500	0.0
50115900	55.0	0.2688	40.527	0.0
50121000	55.0	0.6198	47.529	0.0
50124500	55.0	0.8654	62.500	0.0
50136000	50.0	0.7927	37.634	0.0
50144000	45.0	0.5938	20.375	1.0
50147800	45.0	2.2754	13.274	0.0
50028000	50.0	0.3261	44.910	0.0
50029000	55.0	0.5450	14.368	2.0
50056400	45.0	0.5488	22.059	0.0
50057000	45.0	0.6645	6.061	0.0
50061800	45.0	0.9146	51.282	0.0
50063800	40.0	0.9281	59.524	0.0
50114000	55.0	0.5618	111.434	0.0
50141000	45.0	0.1948	61.728	0.0
50092000	60.0	0.5464	50.891	0.0

methodology, they were treated as ungauged basins. This allows the computation of the standard errors of the estimated quantiles. The "ungauged basins", and their discriminant characteristics are presented in Table 4-17.

Discriminant analysis of the four ungauged basins, described in Table 4-17, was performed using subroutine DISCRIMINANT of the Statistical Package for Social Sciences, SPSS release 4.0, for the VAX/VMS computer (Nie et al. 1975).

The program computes the covariance matrix, the vectors of the means of each basin characteristics for each cluster, and the prior probability, to supply as output the discriminant scores for each cluster by means of Equation (3-13). The program also gives a "performance matrix" in which basins are notionally allocated to the cluster which yields the highest discriminant score. This notional allocation, based solely on basin characteristic data, is then compared with the previous allocation of basins to the clusters derived from flood statistics in the preceding section.

The performance matrix is presented in Table 4-18, showing the original cluster number (i) and the notional relocation (j) based on basin characteristics. The entry (i,j) in this matrix is the fraction of the basins in cluster i that have been notionally reclassified into cluster j. From the entries in the diagonal of the matrix, which gives the fraction of cluster i that have been notionally incorporated

TABLE 4-17

DISCRIMINANT CHARACTERISTICS OF FOUR UNGAUGED BASINS

STATION	AREA	AAR	RSMD1	RSMD2
50045700	8.40	70.0	6.1	9.0
50108000	12.90	35.0	6.0	8.8
50142700	14.40	85.0	9.0	11.0
50095500	12.30	65.0	6.8	9.5

STATION	MAPE	STMFRQ	SLOPE	LAKE
50045700	50.0	0.7143	7.021	0.0
50108000	55.0	0.6977	51.563	0.0
50142700	40.0	0.4167	57.115	0.3
50095500	55.0	0.8130	44.595	0.0

STATION	BASIN
50045700	Río Lajas
50108000	Río Descalabrado
50142700	Río Prieto
50095500	Río Guamaní

TABLE 4-18
PERFORMANCE MATRIX

ORIGINAL CLUSTER NUMBER (i)	NOTIONAL RELOCATION PROBABILITY TO CLUSTER NUMBER (j)			
	1	2	3	4
1	0.750	0.083	0.167	0.000
2	0.200	0.800	0.000	0.000
3	0.000	0.000	1.000	0.000
4	0.000	0.000	0.000	1.000

into the same cluster, it can be noticed that all the original clusters were well defined, obtaining 100% relocation for Clusters 3 and 4.

Utilizing the discriminant scores, the subroutine computes the probability of a new ungauged basin being in each of the four clusters by means of Equation (3-18). The use of Equation (3-18) implies fractional membership, and provides an attractive alternative to unique allocation to a single cluster since the consequences of allocating the ungauged basin to the wrong cluster are alleviated. These probabilities are presented in Table 4-19.

Quantiles for the ungauged basins are estimated by substituting the probabilities of Table 4-19 into Equation (3-17) according to the procedure. In Equation (3-17), T_i is the i^{th} cluster quantile estimate of the T-year flood. The calculation of a weighted quantile estimate is appealing because it accounts for the uncertainty in assigning a basin to a specific cluster. Therefore, a better use of the available information is obtained.

Example 4-2: 50 year quantile for Rio Lajas Basin.

The quantile of the 50-year flood for Rio Lajas basin from Equation (3-17) is given by:

$$X_{50} = 0.65(\ 2.426\ ) + 0.35(\ 4.642\ ) = 3.202$$

Here, Rio Lajas basin has 65.0% of being in Cluster 3, whose

TABLE 4-19

UNGAUGED BASIN PROBABILITIES OF BELONGING TO EXISTING CLUSTERS

Basin	Cluster number			
	1	2	3	4
Río Lajas	0.35	0.00	0.65	0.00
Río Descalabrado	0.97	0.00	0.03	0.00
Río Prieto	0.00	0.00	0.00	1.00
Río Guamaní	0.40	0.00	0.60	0.00

50-year quantile is 2.426, and 35.0% of being in Cluster 1, whose quantile is 4.642.■

The quantile estimates for the four ungauged basins studied are shown in Table 4-20. Here, quantiles up to of 500-year return period are presented. With the data of Table 4-19, dimensionless frequency curves are constructed for each basin.

These are given in Figures 4-9 to 4-12.

To estimate the T-year flood, the ungauged basin mean flow is needed. To estimate this, a regression analysis between the mean flow and the drainage area of the 30 gauged studied basins was employed. This approach to estimate the ungauged basin mean flow is not a standard procedure, but no mean flow regionalization study has been undertaken for Puerto Rico.

The regression is shown in Figure 4-13. The figure plots, in log-log scale, the mean annual flows of the 30 gauged basins as related to their drainage area. The resulting linear regression equation is given by

$$Q_{mean} = 1495\ (Area)^{0.592} \qquad (4\text{-}5)$$

where the Area is in mi^2 and Q_{mean} in cfs. This regression yields a correlation coefficient of 0.87.

TABLE 4-20
QUANTILES FOR UNGAUGED BASINS

Return Period (years)	Ungauged Basin Quantile, X_T			
	Río Prieto	Rio Lajas	Río Guamaní	Río Descalabrado
2	0.889	0.796	0.782	0.623
3	1.131	1.022	1.012	0.891
4	1.289	1.185	1.179	1.104
5	1.407	1.315	1.312	1.281
10	1.765	1.759	1.774	1.950
15	1.972	2.057	2.089	2.444
20	2.119	2.292	2.337	2.853
25	2.234	2.489	2.547	3.209
30	2.328	2.660	2.730	3.527
35	2.408	2.814	2.895	3.818
40	2.478	2.953	3.045	4.087
45	2.539	3.081	3.183	4.338
50	2.595	3.202	3.312	4.576
55	2.645	3.313	3.433	4.800
60	2.691	3.419	3.547	5.013
65	2.733	3.520	3.657	5.218
70	2.773	3.614	3.759	5.413
75	2.810	3.706	3.859	5.602
80	2.844	3.793	3.953	5.784
85	2.877	3.877	4.045	5.959
90	2.907	3.958	4.133	6.130
95	2.936	4.036	4.219	6.296
100	2.964	4.112	4.301	6.457
200	3.343	5.293	5.596	9.051
300	3.570	6.149	6.541	11.006
400	3.734	6.848	7.315	12.639
500	3.862	7.448	7.982	14.060

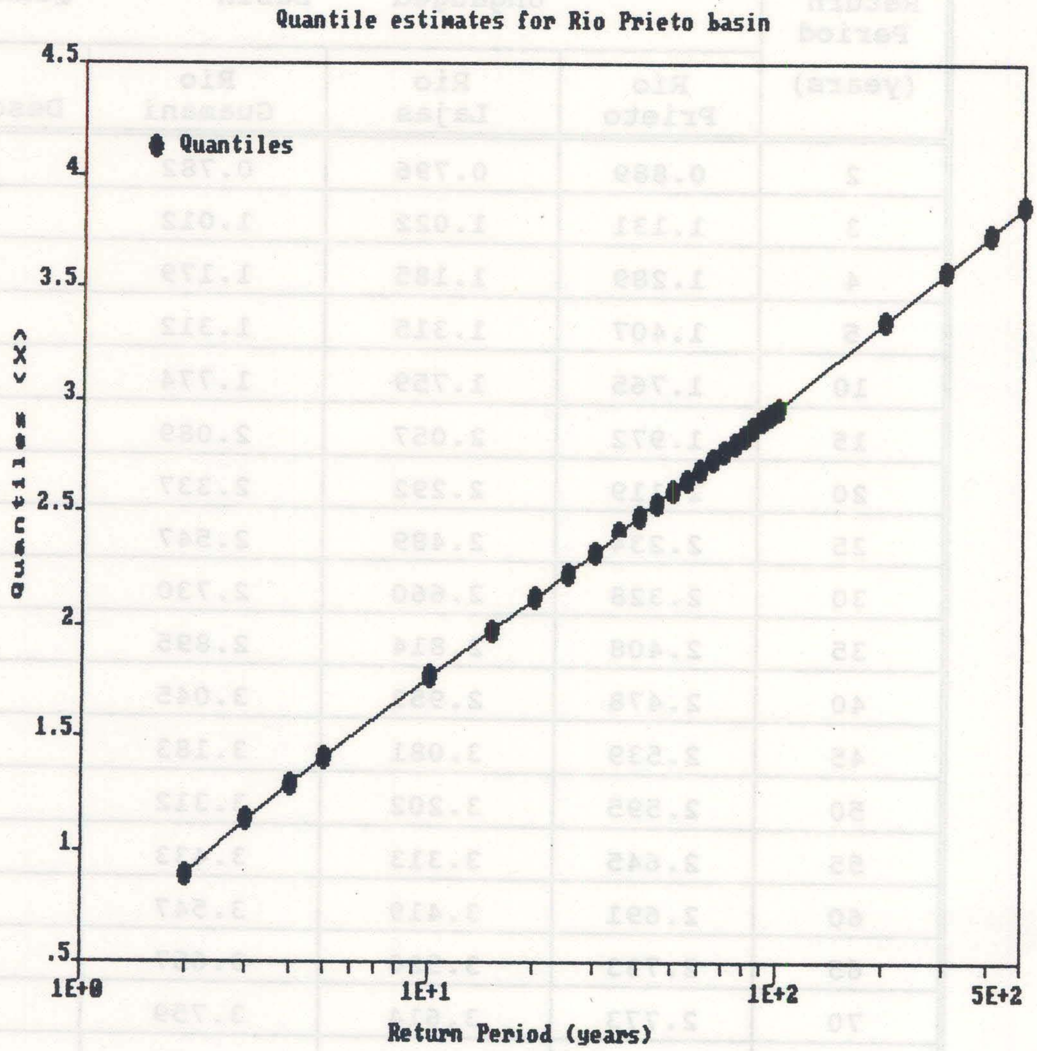

Figure 4-9
Quantile estimates for Río Prieto basin

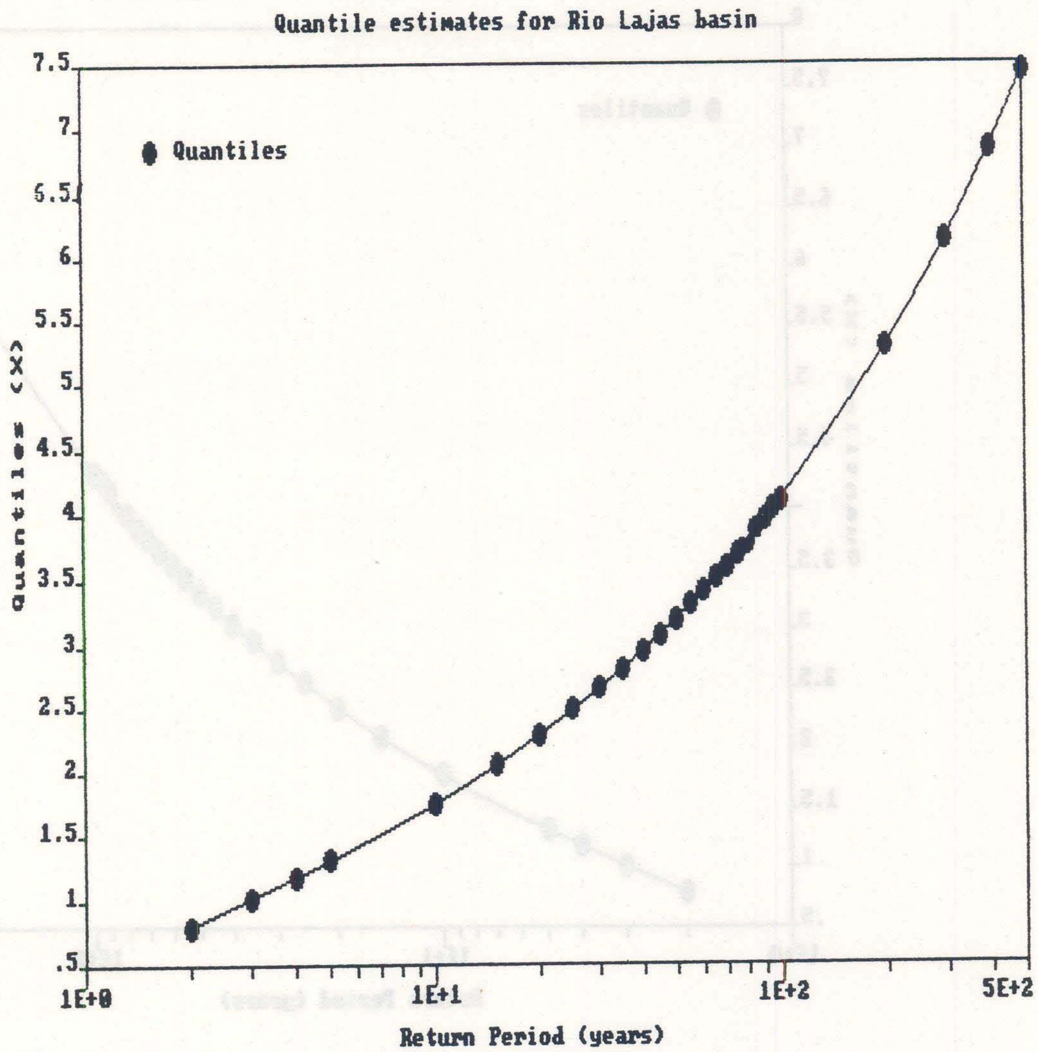

Figure 4-10
Quantile estimates for Río Lajas basin

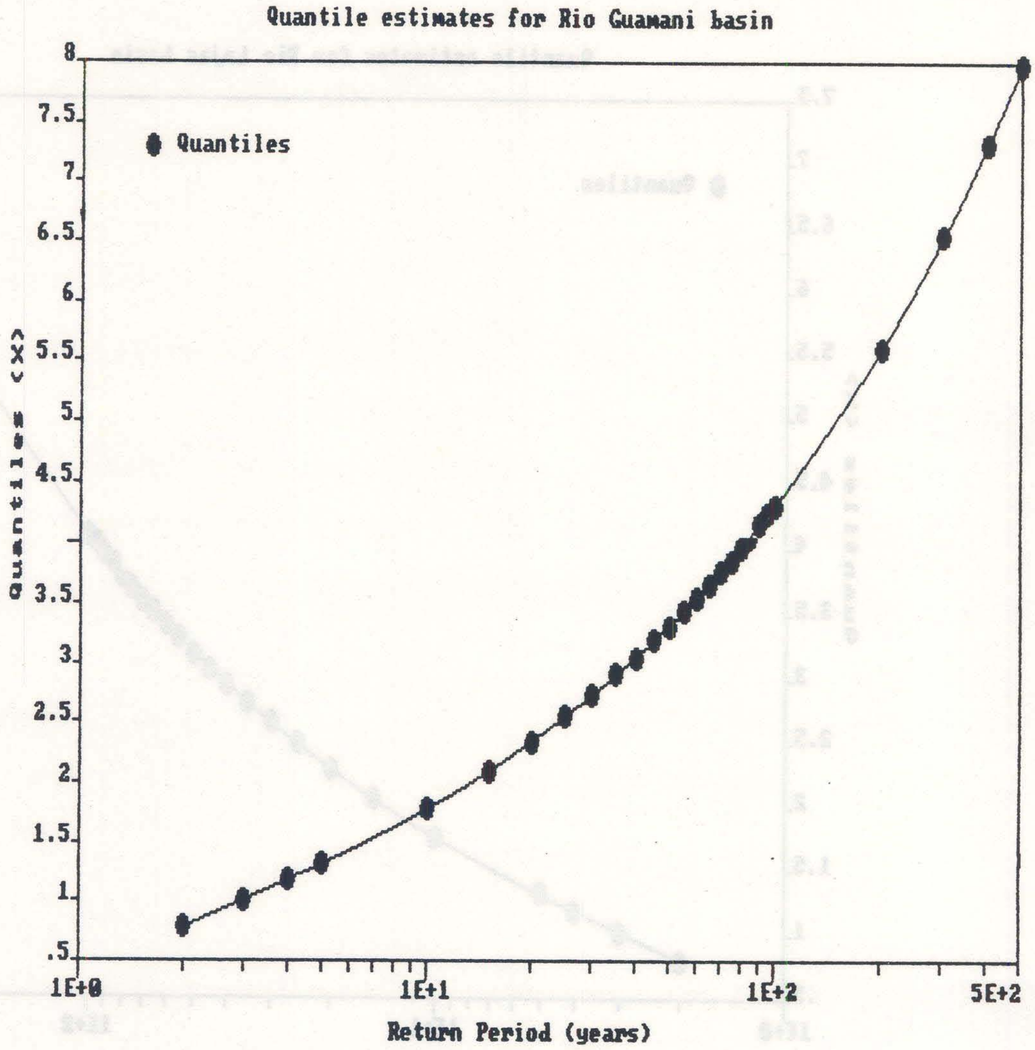

Figure 4-11
Quantile estimates for Río Guamaní basin

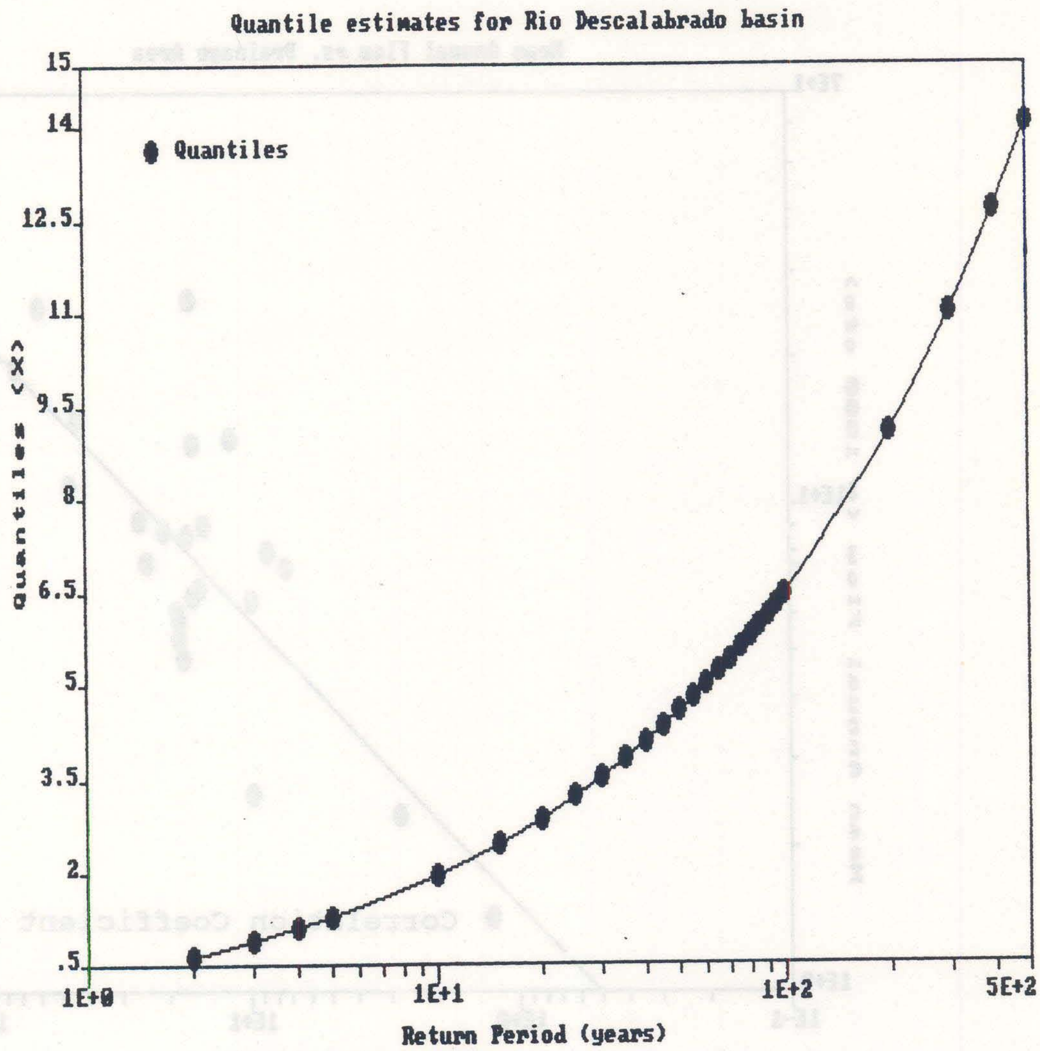

Figure 4-12
Quantile estimates for Río Descalabrado basin

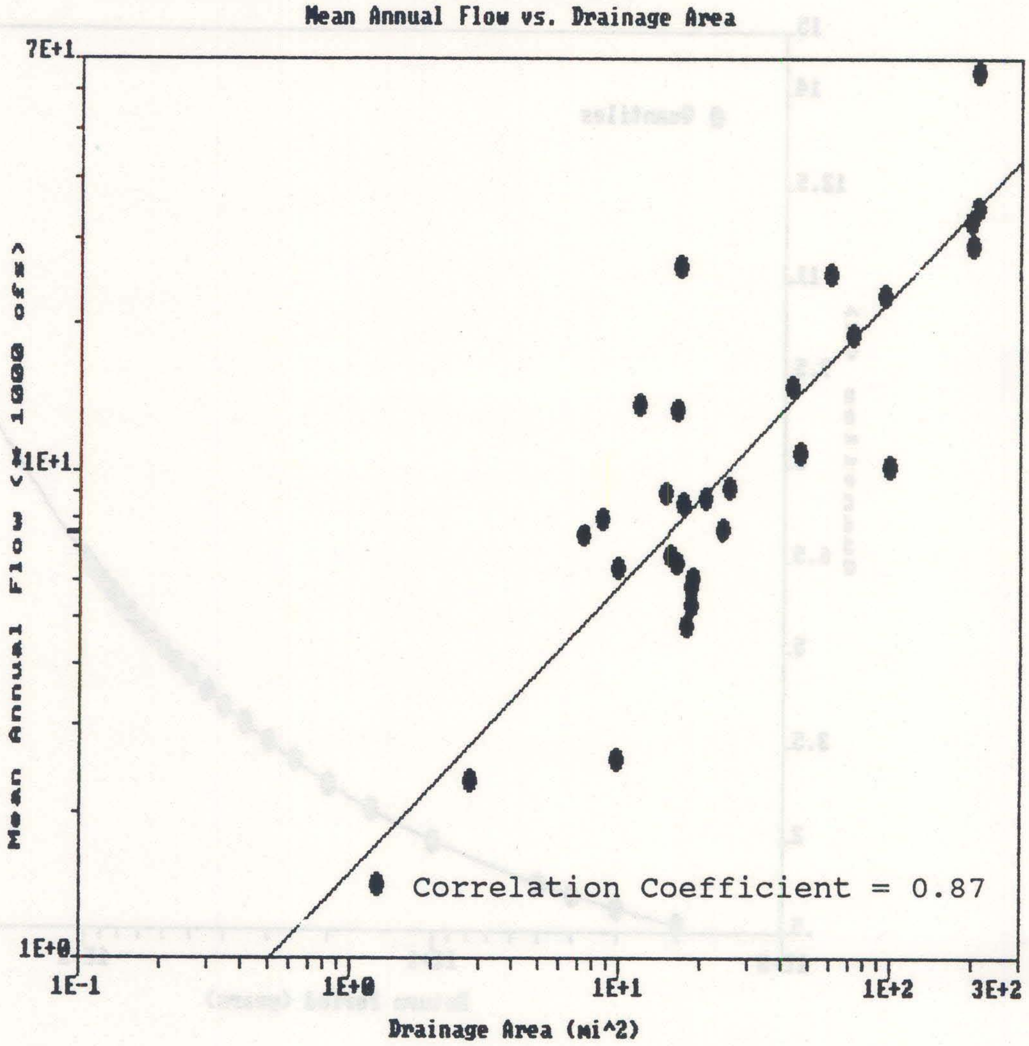

Figure 4-13
Mean Annual Flood, Q_{mean} , vs. Drainage Area

Example 4-3: Mean annual flow of Rio Guamaní basin.

The mean annual flow of the 12.3 mi^2 Rio Guamaní basin by means of Equation (4-5) is:

$$Q_{mean} = 1495\ (12.3)^{0.592} = 6605\ cfs. \blacksquare$$

To obtain the T-year maximum flow for the basin, the T-year quantile, Q_T, is multiplied by the mean annual flow, following the procedure utilized for the gauged basins.

A discussion of the results obtained through the research is presented in the next chapter.

CHAPTER 5

DISCUSSION OF RESULTS

5.1 Evaluation of the methodology

To measure the improvement in the regionalization of flood frequency estimates in P.R., the standard errors are estimated and compared with the errors of the U.S.G.S. approach (López et al., 1979).

The mean standard error of the estimated maximum flow at Cluster j is given by

$$SE_j\ (\%) = (100)\left[\frac{\sum_{i=1}^{N}\left(\frac{Q_{R_i} - Q_{E_i}}{Q_{R_i}}\right)^2}{N}\right]^{1/2} \qquad (5\text{-}1)$$

Where Q_R is the real maximum flow for the basin i, Q_E is the regionalized estimate for the maximum flow at basin j, and N is the number of basins in Cluster j. For the real maximum flows, Q_R, the parent distribution of the historical sample is the at-site GEV distribution.

Results of standard errors from the GEV/PWM and the regression method are compared in Table 5-1 for the 10, 25, 50, and 100 year quantiles. The superiority of the regionalized GEV/PWM procedure is evident when its standard errors for each cluster are compared to the higher ones of the regression technique, SE_R.

TABLE 5-1

STANDARD ERRORS FOR GEV/PWM AND REGRESSION ESTIMATED QUANTILES

Q_T	Standard Error (%) for the GEV/PWM Method				SE_R* (%)
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	
Q_{10}	6.3	1.1	6.0	7.1	+45, -34
Q_{25}	7.8	4.9	13.8	12.6	+50, -33
Q_{50}	10.4	8.3	21.0	17.0	+55, -36
Q_{100}	13.8	12.0	28.7	21.5	+61, -38

* From López et. al. (1979).

Standard errors of estimated flows for the four ungauged basins are obtained from:

$$SE\ (\%) = (100) \left[\left(\frac{Q_R - Q_E}{Q_R} \right)^2 \right]^{1/2} \qquad (5\text{-}2)$$

Here Q_R is the real T-year maximum flow from the parent distribution (GEV distribution) of the available data, and Q_E is the product of the T-year quantile from Table 4-20 and the basin mean annual flow estimate from Equation 4-5. Table 5-2 shows the standard errors of the T-year maximum flow for the ungauged basins under study. The "ungauged basins" do have short records, which are employed here for the evaluation purposes.

The largest source of error in the ungauged basins estimates of the T-year maximum flow is the regression-based estimate of the basin mean annual flood. Greis and Wood (1981) had an average standard error of 59.1% in their study of 16 basins in Arizona. They used short record basins to apply the ungauged basin methodology, as in this research, and then used the sample mean to estimate the maximum flows and compare the standard errors of this approach to the regression based estimates. The same approach was taken here and the results are presented in Table 5-3. Here, the average

TABLE 5-2

STANDARD ERROR (%) OF ESTIMATED QUANTILE USING REGRESSION MEAN

Return Period	Station Number				Average
	5010800	50095500	50142700	50045700	
Q_{10}	68.3	136.7	25.3	53.0	70.8
Q_{25}	72.3	60.4	23.0	24.7	45.1
Q_{50}	74.4	22.6	22.4	7.3	31.7
Q_{100}	76.0	5.0	22.2	7.3	27.9
Station Average	72.8	56.2	23.2	23.1	43.8

TABLE 5-3

STANDARD ERROR (%) OF ESTIMATED QUANTILE USING SAMPLE MEAN

Return Period	Station Number				Average
	5010800	50095500	50142700	50045700	
Q_{10}	5.4	2.6	12.1	4.1	6.1
Q_{25}	8.0	30.4	13.7	21.9	18.5
Q_{50}	9.3	46.8	14.1	32.8	25.8
Q_{100}	10.2	58.8	14.2	41.9	31.3
Station Average	8.2	34.7	13.5	25.2	20.4

standard error is 20.4%, so the magnitude of the difference in standard error between maximum flows obtained using regression-derived mean and the sample mean is indicative of the degree to which regional estimates could be improved with better regionalized estimates of the ungauged basin mean annual flow.

5.2 Discussion

The application of the GEV/PWM procedure has been shown to produce a significant reduction of the uncertainty associated with flood frequency estimation. Several comments are in order regarding the results obtained.

The clustering process is shown to be an effective procedure for grouping basins into homogeneous groups. It has yielded groupings that, not surprisingly, correspond to climatological regions found on Puerto Rico, specifically the Eastern, Southern, Western, and Northern regions. However, some basins have been grouped in clusters that do not correspond to their particular regions. This is somewhat unexpected, but understandable, in view of the uncertainty associated with estimating the parameters of the flood record from small samples. The second moment, the variance, is particularly sensitive to sample size, and may account for the clustering of some basins outside the predominant cluster in their region. Large basins with several high flood values tended to be clustered together, independently of the region they belonged to. This is consistent, since basins should be grouped according to observed hydrologic response, and not according to geographical subdivisions. As observed in Table 4-18, the performance matrix has indicated, with high probability, that the clusters were well-defined.

An analysis of the regionalized frequency curves for the four clusters shows steep curves for Clusters 1 and 2, and milder curves for Clusters 3 and 4. The steepness in the first two curves arise from the relatively large shape parameter obtained from the regionalization. This reflects the fact that large floods are recorded within the short sample records available, and thus a certain, though unknown, degree of skewness is present. For the other clusters the shape parameter is small, actually near zero on the negative side, thus exhibiting a milder slope. Frequency curves with small, positive shape parameters will be asymptotic to some upper bound of flow. The results demonstrate the regional variations that can be obtained within a single geographic area. The implication is that regions may exhibit considerable heterogeneity in their flood frequency characteristics that cannot be assessed with regression methods, but would require regionalization within statistically homogeneous groups for proper flood frequency estimation.

Regionalization within homogeneous groups has been shown to be an effective technique for grouping basins for flood frequency analysis. The procedure for assessing flood quantiles for ungauged basins has also been shown to be effective.

For ungauged basins, discriminant analysis has been successfully employed to estimate flood quantiles. A particular contribution obtained through this research is an expanded set of basin characteristics that has yielded sharper cluster membership probabilities when compared with results obtained elsewhere.

However, it was discovered that estimation of the mean flow from regressions with catchment area may not be generally effective for estimating flows from ungauged sites. While it is known that the mean annual flow correlates well with catchment area, this does not imply that flood quantiles will be estimated with low standard errors. It was not part of the scope of the project to provide estimates of the mean annual flow for the user. This information must be obtained from suitable regionalization of the available mean flow data. However, the research has established that adequate estimates of the mean annual flow for ungauged sites will significantly reduce the standard error associated with flood quantile estimates.

Overall, the regionalized GEV/PWM procedure, coupled to a discriminant classification of basins, provides a novel methodology that has effectively classified Island catchments, yielding flood quantile estimates with significantly lower standard errors than those obtained with traditional regression techniques.

CHAPTER VI

Conclusions and Recommendations

The Generalized Extreme Value Probability-Weighted Moments procedure has been applied to catchments in Puerto Rico, with the purpose of regionalizing the available flood data to obtain better estimates of flood flow frequency for water resources projects. The following conclusions are obtained from the project:

1. The GEV/PWM procedure has been shown to be an effective technique for assessing flood frequency distributions for catchments in Puerto Rico.

2. Cluster analysis has shown that catchments in Puerto Rico can be grouped into four homogeneous clusters for the purpose of regionalizing flood frequency parameters.

3. It has been established that agglomerating all the flood frequency data into one group for regression analysis is not a valid procedure in view of the heterogeneity exhibited by such a grouping. The standard error of the regression approach, proposed by López et al., 1979, far exceeds the error from the regionalized GEV/PWM procedure.

4. Discriminant analysis is a useful tool for assigning ungauged catchments to clusters for estimating flood quantiles. But this approach will yield good results only when adequate regionalized mean flow information is available.

5. The standard errors obtained with the regionalized GEV/PWM method were relatively small, and comparable to literature values for other sites.

The application of the GEV/PWM model yielded good results with the available data. However, some refinements and extentions of the methodology are possible as part of the dynamic of the theoretical evolution of these procedures. With this objectives in mind, the following recommendations are proposed:

1. The discriminant procedure needs of further research along the lines of establishing the adequate, or statistically optimum, number and kind of basin measures that are required for discriminating among basin types.

2. The ungauged basin flood estimation procedure requires an adequate estimate of the mean annual flow. This low standard error estimate can only be achieved by undertaking an adequate regionalization procedure of the available mean flow records on the Island.

3. Further research is required to define the proper distribution that would allow estimation of the confidence intervals for the upper quantiles obtained from the GEV/PWM procedure.

4. Further research into statistical cluster analysis should provide more efficient procedures to form homogeneous groups so as to minimize non-membership probabilities of

selected basins. The application of these techniques to water resources problems has only recently been undertaken.

BIBLIOGRAPHY

Benson, M. A., and N. C. Matalas, Synthetic hydrology based on regional statistical parameters, Water Resources Research, 3(4), 935-935, 1967.

Benson, M. A., Uniform Flood Frequency Estimating Methods for Federal Agencies, Water Resources Research, 4(5), 891-908, 1968.

Fisher, R. A., and Tippett, L. H. C., Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample, Proceedings of the Cambridge Philosophical Society, 24, 180-190, 1928.

Flood Studies Report, Vols. I-V, Natural Environmental Research Council, London, 1975.

U. S. Weather Bureau, Generalized Estimates of Probable Maximum Precipitation and Rainfall-Frequency Data for Puerto Rico and Virgin Islands, Technical Paper no. 42, Washington D. C., 1961.

Gordon, A. D., Classification, Chapman and Hall, New York, N.Y., 1981.

Greenwood, J. A., J. M. Landwehr, M. C. Matalas, and J. R. Wallis, Probability weighted moments: Definition and relation to parameters of several distributions expressed in inverse form, Water Resources Research, 15(6), 1049-1054, 1979.

Greis, N. P., and E. F. Wood, Regional flood frequency estimation and network design, Water Resources Research, 17(4),1167-1177, 1981.

Gumbel, E. J., Statistical Theory of Extreme Values and Some Practical Applications, U.S. National Bureau of Standards Applied Mathematics Series, No. 33, Washington, D.C., 1954.

Gumbel, E. J., Statistics of Extreme Values, Columbia University Press, New York, N.Y., 1958.

Haan, C. T., Statistical Methods in Hydrology, The Iowa State University Press, Ames, Iowa, 1977.

Hosking, J. R. M., J. R. Wallis, and E. F. Wood, An appraisal of the generalized extreme-value distribution by the method of probability-weighted moments, Technometrics, 27(3), 251-261, 1985a.

Hosking, J. R. M., J. R. Wallis, and E. F. Wood, An appraisal of the regional flood frequency procedure in the UK Flood Studies Report, Hydrological Sciences - Journal, 30, 1, 3, 85-109, 1985b.

Jenkinson, A. F., The frequency distribution of the annual maximum (or minimum) of meteorological elements, Quarterly Journal ofthe Royal Meteorological Society, 81, 158-171, 1955.

Kite, G.W., Frequency and Risk Analysis in Hydrology, Water Resources Publications, Fort Collins, Colorado, 1977.

Landwehr, J. M., N. C. Matalas, and J. R. Wallis, Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, Water Resources Research, 15 (5), 1055-1064, 1979.

López, M. A., E. Colón-Dieppa, and E. D. Cobb, Floods in Puerto Rico, magnitude and frequency, U.S.G.S. Water Resources Investigations WRI 78-141, June 1979.

Mosley, M. P., Delimitation of New Zealand hydrological regions, Journal of Hydrology, 49, 173-192, 1981.

Nie, N. H., Hadlai, C., Jenkings, J. G., Steinbrenner, K., Bent, D. H., Statistical Package for Social Sciences, Mc Graw-hill Inc., New York, N. Y., 1975.

Rao, C. R., Linear Statistical Inference and Its Applications, John Wiley & Sons, London, U. K., 1973.

Reich, B. M., Log-Pearson Type III Distribution and Gumbel Analysis of Floods, Second International Symposium in Hydrology, Fort Collins, Colorado, 290-303, 1972.

Solomon, S. I., Parameter regionalization and network design, in Stochastic Approaches to Water Resources, Vol. 1, edited by H. W. Shen, Water Resources Publications, Fort Collins, Colo., 1976.

Water Resources Council, A Uniform Technique to Determine Flood Flow Frequency, Bull. 15, Hydrol. Comm. Washington, D.C., 1967.

Water Resources Council, Guidelines for determining flood flow frequency, Bull. 17-A, Hydrol. Comm. Washington, D. C., 1977.

White, E. L., Factor analysis of drainage basins properties: classification of flood behavior in terms of basin geomorphology, Water Resources Bulletin, 11(4), 676-686, 1975.

Wiltshire, S. E., Regional flood frequency analysis I: Homogeneity statistics, Hydrological Sciences - Journal, 31, 9, 3, 321-333, 1986a.

Wiltshire, S. E., Regional flood frequency analysis II: Multivariate classification of drainage basins in Britain, Hydrological Sciences - Journal, 31, 3, 9, 335- 346, 1986b.

U.S. Geological Survey, Peak flows records for Puerto Rico, U.S.G.S. Water Resources Division, Caribbean District, 1989.

APPENDIX A

GEV/PWM METHOD APPLICATIONS PROGRAM

```
REM  *******   GEV/PWM PROCEDURE    **********


DECLARE SUB PedirDatos ()
DECLARE SUB LeerNombre ()
DECLARE SUB LeerDatos ()
DECLARE SUB GrabarDatos ()


COMMON SHARED N, Q(), Archivo$

CLS
DO

   clave$ = "si"

   LOCATE 12, 20
   INPUT "Archivo Antiguo (A), Archivo Nuevo (N)"; v$
   v$ = UCASE$(v$)

   IF v$ = "A" THEN
         CALL LeerNombre
         CALL LeerDatos

   ELSEIF v$ = "N" THEN
          CALL LeerNombre
          CALL PedirDatos
          CALL GrabarDatos

   ELSE
         clave$ = "NO"

   END IF

LOOP WHILE clave$ = "NO"


CLS : COLOR 9


      REM   CALCULO DE ESTIMADORES B(0),B(1),B(2)

  DIM P(N), B(3)
  A = .35

  FOR R = 0 TO 2
     xx = 0
```

```
      FOR J = 1 TO N
         P(J) = (J - A) / N
         xx = xx + (P(J)) ^ R * Q(J)
      NEXT J

      B(R + 1) = (1 / N) * xx

   NEXT R


   LOCATE 23, 25: PRINT "PRESIONE ENTER PARA CONTINUAR"
   LINE INPUT C$


   CLS
   LOCATE 5, 15: PRINT " VALORES DE ESTIMADORES B(0), B(1), B(2) "
   LOCATE 7, 20: PRINT "B(0)=   ", B(1)
   LOCATE 9, 20: PRINT "B(1)=   ", B(2)
   LOCATE 11, 20: PRINT "B(2)=   ", B(3)


         REM   CALCULO DE ESTIMADOR  k

   AB = 0
   AC = 0
   C = 0
   k = 0

   AB = (2 * B(2)) - B(1)
   AC = (3 * B(3)) - B(1)


   C = (AB / AC) - (LOG(2) / LOG(3))
   k = (7.859 * C) + (2.9554 * C ^ 2)

   LOCATE 14, 15: PRINT " VALOR DE ESTIMADORES  k, ALFA , PSI  "
   LOCATE 16, 20: PRINT " k =   ", k



      REM   CALCULO DE FUNCION GAMMA

  IF k > 0 THEN

     G1 = -.577171652#
     G2 = .988205891#
     G3 = -.897056937#
     G4 = .918206857#
     G5 = -.7567040779#
     G6 = .482199394#
     G7 = -.193527818#
     G8 = .035868343#
```

```
    GA = 0

    GA = 1 + (G1 * k) + (G2 * k ^ 2) + (G3 * k ^ 3) + (G4 * k ^ 4)
         + (G5 * k ^ 5) + (G6 * k ^ 6) + (G7 * k ^ 7) + (G8 * k ^ 8)

ELSE

   k1 = -k
   G1 = -.577171652#
   G2 = .988205891#
   G3 = -.897056937#
   G4 = .918206857#
   G5 = -.7567040779#
   G6 = .482199394#
   G7 = -.193527818#
   G8 = .035868343#
   GA1 = 0
   RA = 0
   RA1 = 0
   GA = 0

   GA1 = 1 + (G1 * k1) + (G2 * k1 ^ 2) + (G3 * k1 ^ 3) +
         (G4 * k1 ^ 4) + (G5 * k1 ^ 5) + (G6 * k1 ^ 6) +
         (G7 * k1 ^ 7) + (G8 * k1 ^ 8)

   RA1 = 3.141592654# * k1

   GA = RA1 / (GA1 * SIN(RA1))

END IF

      REM    CALCULO DE ESTIMADORES ALFA Y PSI

   A1 = 0
   A2 = 0
   ALFA = 0
   PSI = 0

   A1 = AB * k
   A2 = GA * (1 - 2 ^ -k)

   ALFA = A1 / A2
   PSI = B(1) + (ALFA / k) * (GA - 1)

   LOCATE 18, 20: PRINT "ALFA =  ", ALFA
   LOCATE 20, 20: PRINT "PSI =   ", PSI

   LOCATE 23, 25: PRINT "PRESIONE ENTER PARA CONTINUAR"
   LINE INPUT C$
   CLS

   LOCATE 7, 25: PRINT "VALOR DE FUNCION GAMMA"
   LOCATE 9, 25: PRINT " GAMMA = ", GA
```

```
LOCATE 24, 25: PRINT "PRESIONE ENTER PARA CONTINUAR"
LINE INPUT C$
CLS



DIM F(20), Tr(N), Q1(20, 2), TT(8), Q2(8), FF(8)


PP = 0
Z1 = 0
Z2 = 0
Z3 = 0
Z4 = O
ZZ = 1

FOR Z = 5 TO 100 STEP 5

   PP = 1 / Z
   F(ZZ) = 1 - PP
   Z1 = -LOG(F(ZZ))
   Z4 = Z1 ^ k
   Q1(ZZ, 2) = PSI + (ALFA / k) * (1 - Z4)
   Q1(ZZ, 1) = Z
   ZZ = ZZ + 1

NEXT Z

v = 1

FOR H = N TO 1 STEP -1

   Tr(H) = (N + 1) / v
   v = v + 1

NEXT H

RRR = 1

FOR RR = 2 TO 5
   PP = 1 / RR
   FF(RRR) = 1 - PP
   Z1 = -LOG(FF(RRR))
   Z4 = Z1 ^ k
   Q2(RRR) = PSI + (ALFA / k) * (1 - Z4)
   TT(RRR) = RR
   RRR = RRR + 1
NEXT RR

SSS = 5

FOR SS = 200 TO 500 STEP 100
```

```
      PP = 1 / SS
      FF(SSS) = 1 - PP
      Z1 = -LOG(FF(SSS))
      Z4 = Z1 ^ k
      Q2(SSS) = PSI + (ALFA / k) * (1 - Z4)
      TT(SSS) = SS
      SSS = SSS + 1
   NEXT SS

   LOCATE 3, 5: PRINT "RESULTADOS DE  Tr(x)"
   LOCATE 5, 15: PRINT " Q(x)              Tr(x)      "
   LOCATE 6, 15: PRINT "-------          ----------   "

   L1 = 1

   FOR L = 1 TO N

      IF L1 < 16 THEN
         LOCATE 6 + L1, 14: PRINT Q(L), Tr(L)
         L1 = L1 + 1
      ELSE
         LOCATE 23, 25: PRINT " PRESIONE ENTER PARA CONTINUAR"
         LINE INPUT C$
         CLS
         L1 = 1
         LOCATE 3, 5: PRINT "RESULTADOS DE F(x)"
         LOCATE 5, 15: PRINT " Q(x)           Tr(x) "
         LOCATE 6, 15: PRINT "-------         ----------"
         LOCATE 6 + L1, 14: PRINT Q(L), Tr(L)
      L1 = L1 + 1

      END IF

   NEXT L

   LOCATE 24, 25: PRINT "PRESIONE ENTER PARA CONTINUAR"
   LINE INPUT C$
   CLS

   LOCATE 2, 5: PRINT "CAUDALES MAXIMOS Y TIEMPO DE RECURRENCIA"
   LOCATE 3, 15: PRINT "  Tr            Qmax  "
   LOCATE 4, 15: PRINT "-------        -----------"

   FOR NN = 1 TO 20

      LOCATE 4 + NN, 14: PRINT Q1(NN, 1), Q1(NN, 2)

   NEXT NN

   LOCATE 24, 25: PRINT "PRESIONE ENTER PARA CONTINUAR"
   LINE INPUT C$
   CLS
```

```
LOCATE 2, 5: PRINT "CAUDALES MAXIMOS Y TIEMPO DE RECURRENCIA"
LOCATE 3, 15: PRINT "  Tr              Qmax   "
LOCATE 4, 15: PRINT "-------        -----------"

FOR MMM = 1 TO 8

    LOCATE 4 + MMM, 14: PRINT TT(MMM), Q2(MMM)

NEXT MMM

LOCATE 24, 25: PRINT "PRESIONE ENTER PARA CONTINUAR"
LINE INPUT C$
CLS


SUB GrabarDatos

   CLOSE #1

   OPEN "o", 1, Archivo$
   PRINT #1, N

   FOR x = 1 TO N
       PRINT #1, Q(x)
   NEXT x

   CLOSE #1

   END SUB

SUB LeerDatos
CLS

CLOSE #1
OPEN "i", 1, Archivo$
INPUT #1, N
DIM Q(N)
FOR x = 1 TO N

   INPUT #1, Q(x)

NEXT x

CLOSE #1

END SUB

SUB LeerNombre
```

```
   LOCATE 14, 20
   INPUT "Numero de Estacion ######## "; Ne
   Ne$ = STR$(Ne)
   Ne$ = MID$(Ne$, 2, LEN(Ne$))

   Archivo$ = Ne$ + ".est"
   CLS

   END SUB

SUB PedirDatos


   LOCATE 4, 30: INPUT "NUMERO DE DATOS =   "; N
   DIM B(3), Q(N), P(N)
   LOCATE 7, 20: PRINT "CAUDALES EN ORDEN ASCENDENTE"
   LOCATE 9, 23: PRINT "       CAUDAL "

   T1 = 1

   FOR T = 1 TO N
      IF T1 < 10 THEN

         LOCATE 9 + T1, 27: INPUT Q(T)
         T1 = T1 + 1
      ELSE

      CLS
      T1 = 1
      LOCATE 5, 20: PRINT "CAUDALES EN ORDEN ASCENDENTE"
      LOCATE 7, 23: PRINT "          CAUDAL"
      LOCATE 8 + T1, 27: INPUT Q(T)
      END IF

   NEXT T


   END SUB
```

APPENDIX B

RESULTS FROM THE APPLICATION OF

THE GEV/PWM PROCEDURE TO 30 GAUGED BASINS IN P.R.

Table B-1

Return Period for Historical data and GEV-PWM Distribution for Station 50029000 at Río Grande de Arecibo

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.04	2000	2	11930
1.10	3410	3	19875
1.17	3430	4	26529
1.25	3680	5	32448
1.35	3780	10	56485
1.46	5200	15	75856
1.59	6790	20	92773
1.74	7770	25	108089
1.93	8070	30	122245
2.16	8680	35	135508
2.45	10200	40	148053
2.84	15800	45	160004
3.36	21000	50	171451
4.14	23700	55	182464
5.37	26100	60	193097
7.66	76000	65	203393
13.33	105000	70	213388
51.43	195000	75	223113
		80	232591
		85	241843
		90	250888
		95	259743
		100	268420
		200	417033
		300	538419
		400	644949
		500	741657

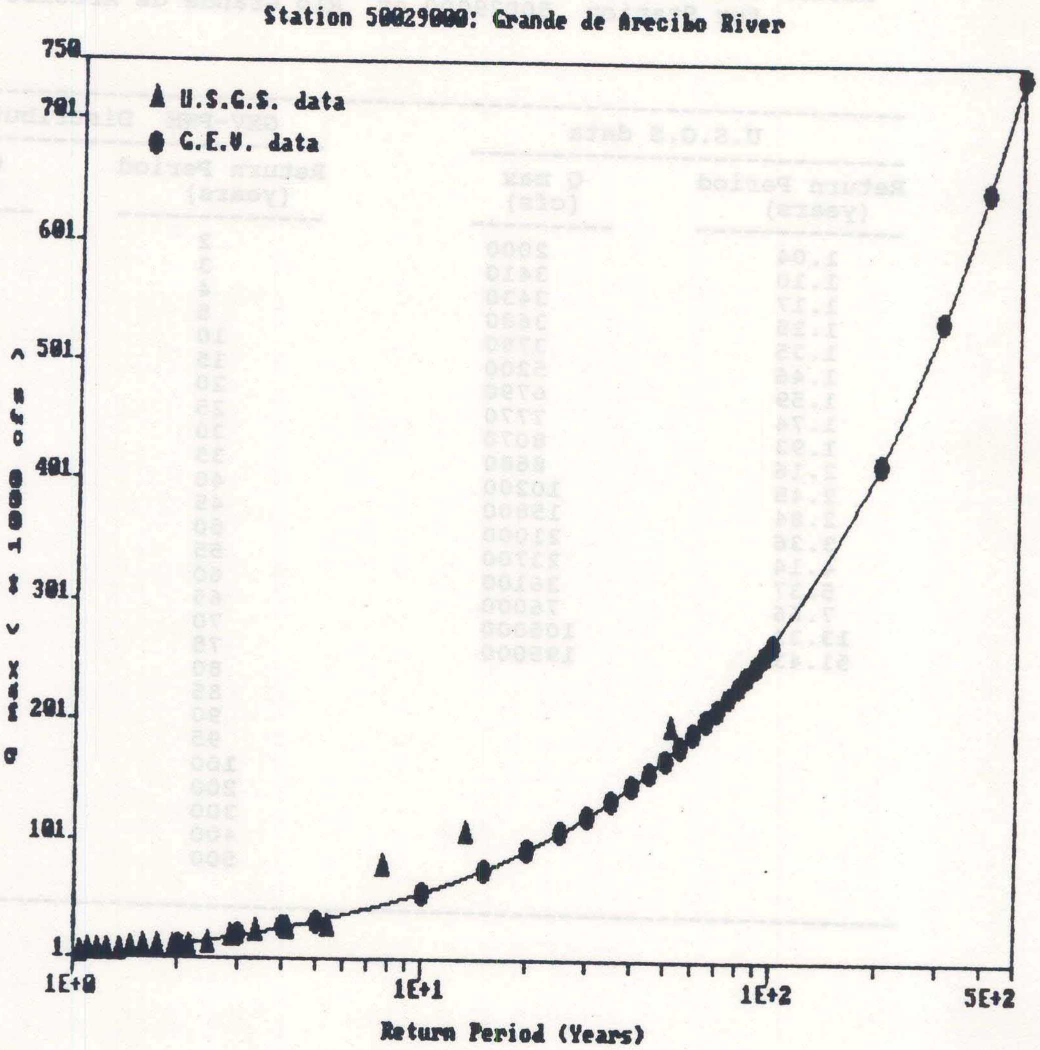

Figure B-1
Flood Frequency Curve for Rio Grande de Arecibo basin

Table B-2

Return Period for Historical data and GEV-PWM Distribution for Station 50144000 at Río Grande de Añasco

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	5950	2	13638
1.07	6980	3	18861
1.11	7450	4	23124
1.15	7830	5	26848
1.21	8000	10	41503
1.27	8020	15	52917
1.33	8090	20	62675
1.39	8800	25	71371
1.47	8960	30	79310
1.56	9300	35	86673
1.65	9320	40	93577
1.77	10700	45	100103
1.88	11600	50	106312
2.02	12500	55	112250
2.19	14400	60	117950
2.38	15000	65	123443
2.61	15600	70	128750
2.89	16000	75	133892
3.23	19300	80	138882
3.67	20300	85	143736
4.25	28600	90	148464
5.05	29700	95	153077
6.21	32300	100	157583
8.06	38700	200	232938
11.49	53600	300	292541
20.00	77200	400	343782
77.14	140000	500	389591

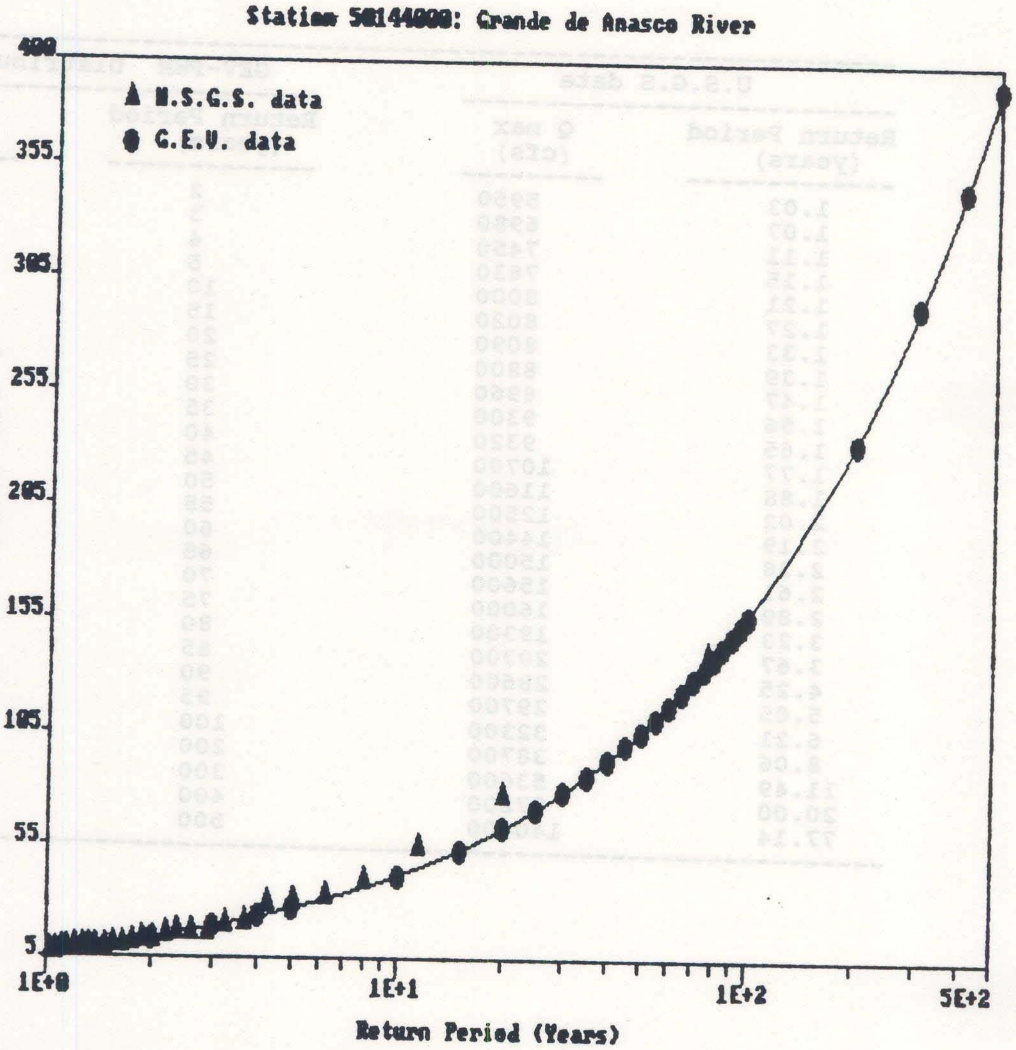

Figure B-2
Flood Frequency Curve for Rio Grande de Añasco basin

Table B-3

Return Period for Historical data and GEV-PWM Distribution for Station 50048000 at Río de Bayamón

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.04	3680	2	12130
1.11	3680	3	17817
1.19	3680	4	22182
1.29	6010	5	25837
1.41	7300	10	39190
1.55	7700	15	48776
1.71	8450	20	56566
1.92	9110	25	63255
2.18	11100	30	69184
2.52	12000	35	74551
2.99	13300	40	79480
3.68	27000	45	84056
4.78	29000	50	88340
6.81	35000	55	92378
11.85	58000	60	96204
45.71	72000	65	99847
		70	103328
		75	106666
		80	109875
		85	112968
		90	115956
		95	118848
		100	121651
		200	165984
		300	198460
		400	225045
		500	247965

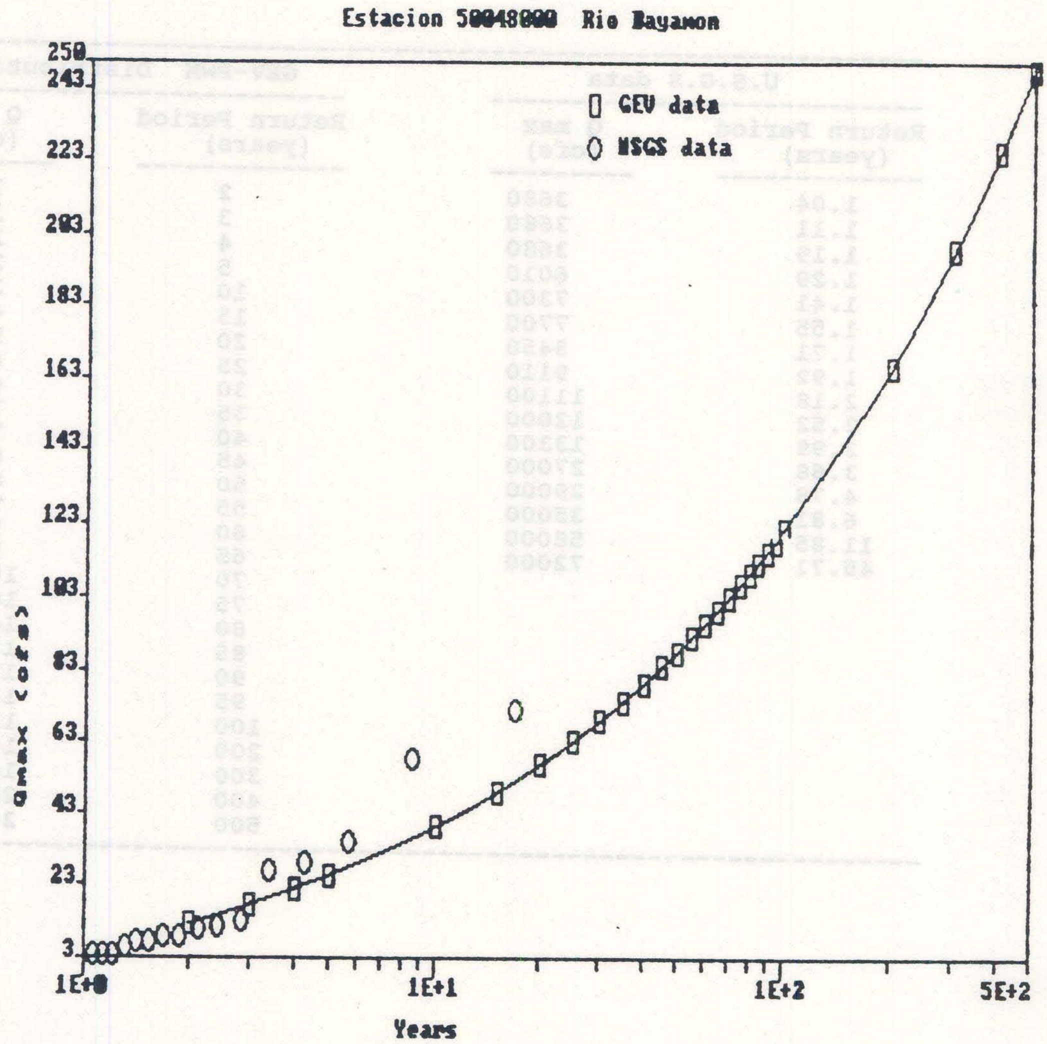

Figure B-3
Flood Frequency Curve for Río Bayamón basin

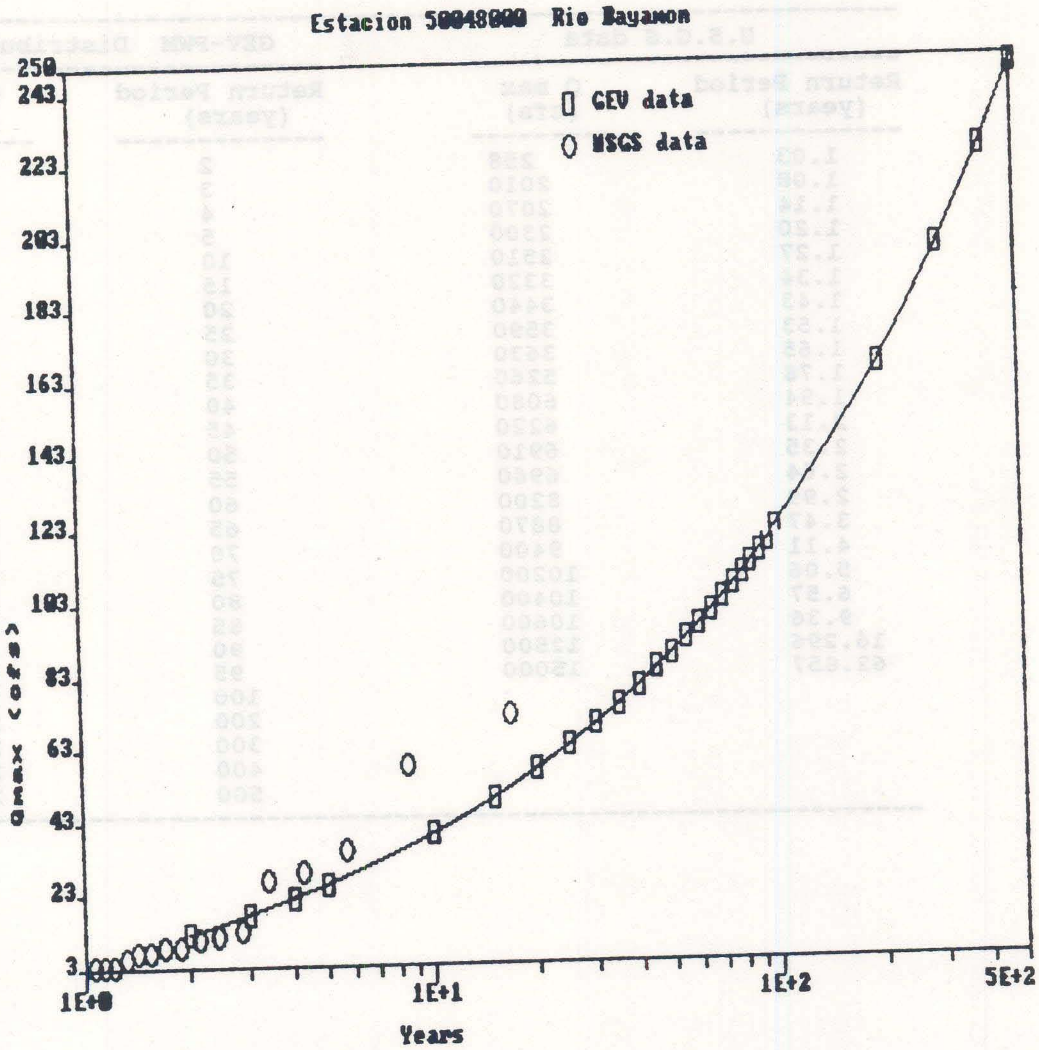

Figure B-3
Flood Frequency Curve for Río Bayamón basin

Table B-4

Return Period for Historical data and GEV-PWM Distribution
for Station 50061800 at Río Canóvanas

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	258	2	5781
1.08	2010	3	7539
1.14	2070	4	8641
1.20	2300	5	9445
1.27	2510	10	11765
1.34	3320	15	13038
1.43	3440	20	13915
1.53	3590	25	14581
1.65	3630	30	15118
1.78	5260	35	15567
1.94	6080	40	15953
2.13	6220	45	16290
2.35	6910	50	16590
2.64	6960	55	16859
2.99	8200	60	17104
3.47	8870	65	17328
4.11	9400	70	17535
5.06	10200	75	17727
6.57	10400	80	17905
9.36	10600	85	18073
16.296	12500	90	18230
62.857	15000	95	18378
		100	18518
		200	20377
		300	21435
		400	21173
		500	22738

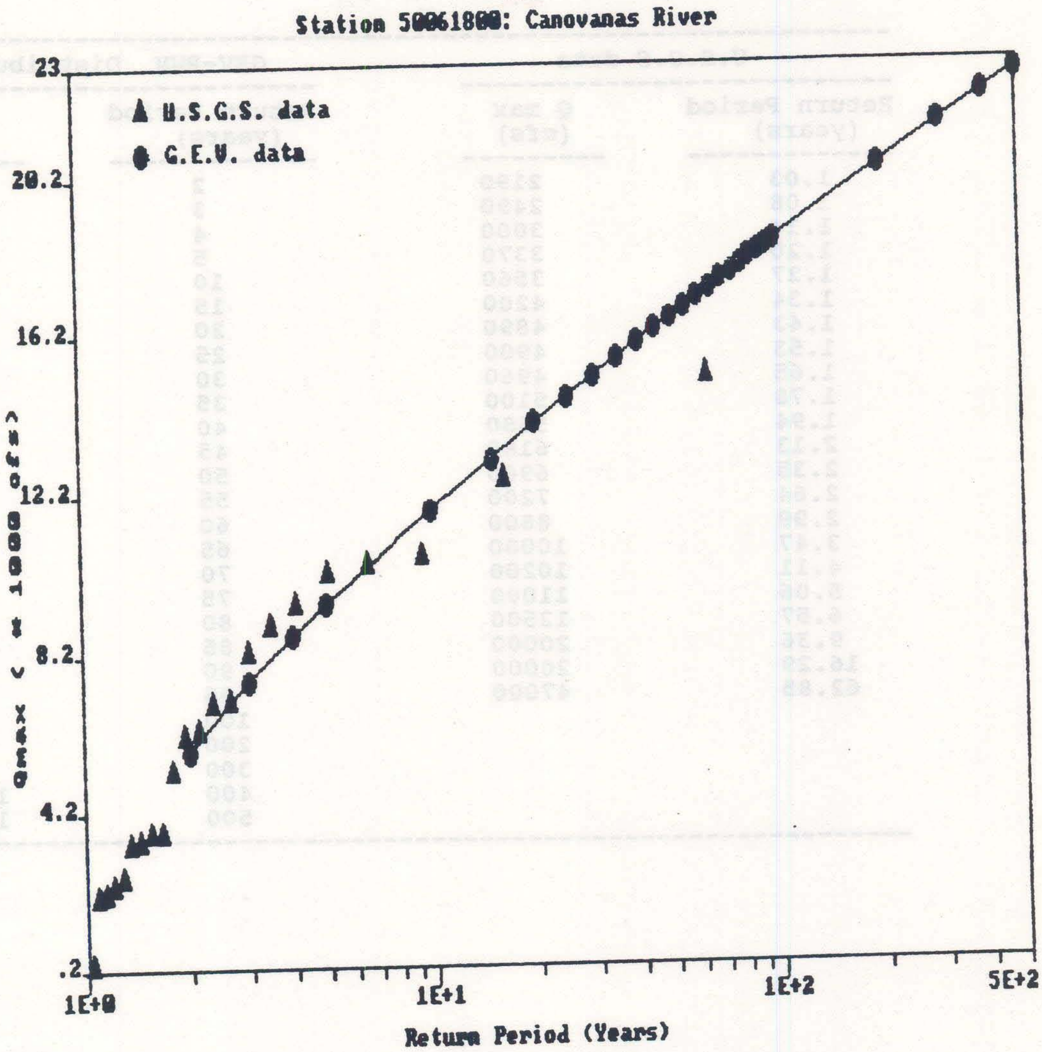

Figure B-4
Flood Frequency Curve for Río Canóvanas basin

Table B-5

Return Period for Historical data and GEV-PWM Distribution for Station 50011440 at Río Bucana

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	2190	2	6178
1.08	2490	3	8415
1.14	3000	4	10169
1.20	3370	5	11647
1.27	3560	10	17187
1.34	4200	15	21263
1.43	4890	20	24625
1.53	4900	25	27545
1.65	4960	30	30157
1.78	5100	35	32538
1.94	5680	40	34738
2.13	6180	45	36792
2.35	6900	50	38724
2.64	7200	55	40553
2.99	8800	60	42293
3.47	10000	65	43956
4.11	10200	70	45550
5.06	11000	75	47083
6.57	12500	80	48561
9.36	20000	85	49990
16.29	20000	90	51373
62.85	47000	95	52715
		100	54020
		200	74991
		300	90713
		400	103773
		500	115156

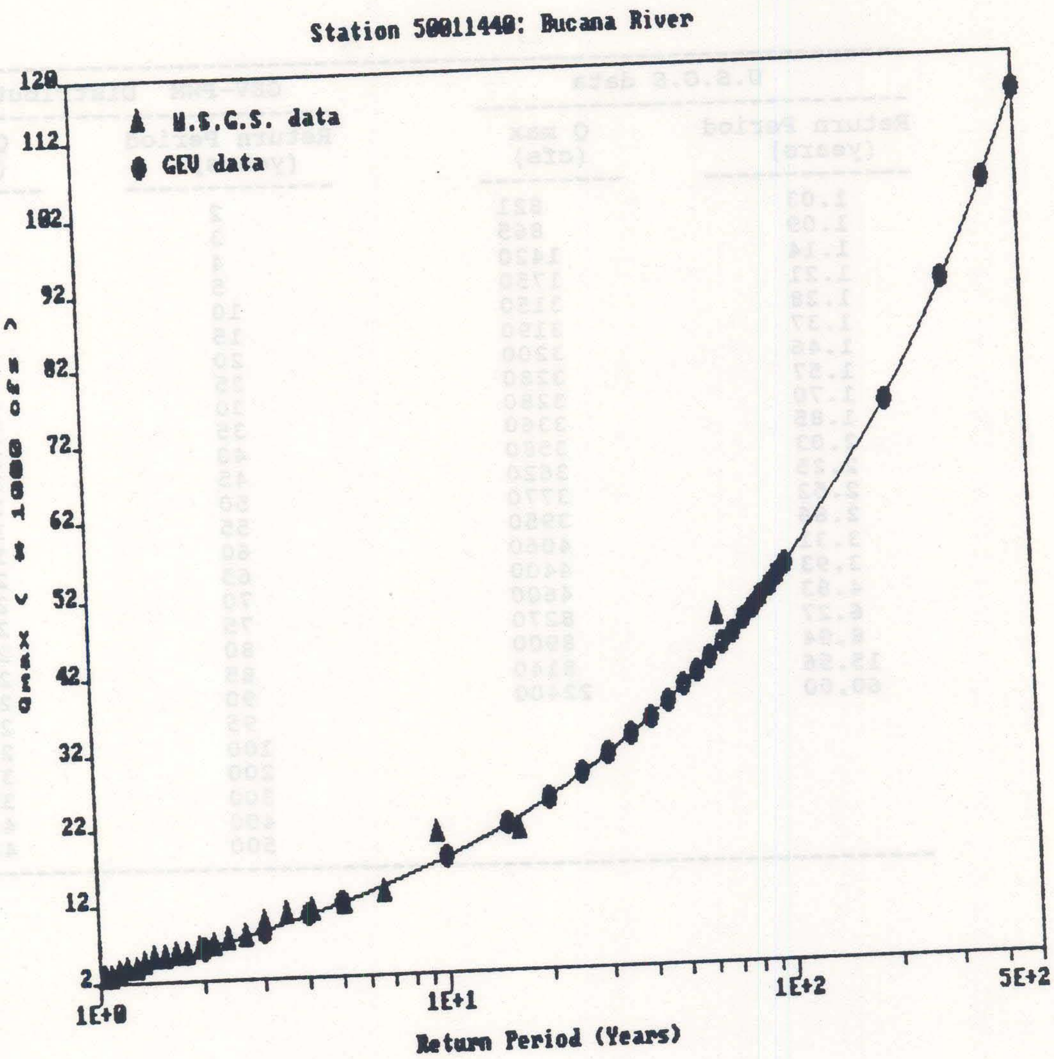

Figure B-5
Flood Frequency Curve for Río Bucaná basin

Table B-6

Return Period for Historical data and GEV-PWM Distribution for Station 50114000 at Río Cerrillos

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	821	2	3437
1.09	865	3	4573
1.14	1420	4	5440
1.21	1750	5	6163
1.28	3150	10	8792
1.37	3190	15	10667
1.46	3200	20	12185
1.57	3280	25	13484
1.70	3280	30	14633
1.85	3360	35	15672
2.03	3580	40	16624
2.25	3620	45	17506
2.52	3770	50	18332
2.86	3950	55	19108
3.31	4060	60	19844
3.93	4400	65	20544
4.83	4600	70	21212
6.27	8270	75	21851
8.94	8900	80	22466
15.56	8140	85	23058
60.00	22400	90	23630
		95	24183
		100	24719
		200	33154
		300	39296
		400	44305
		500	48611

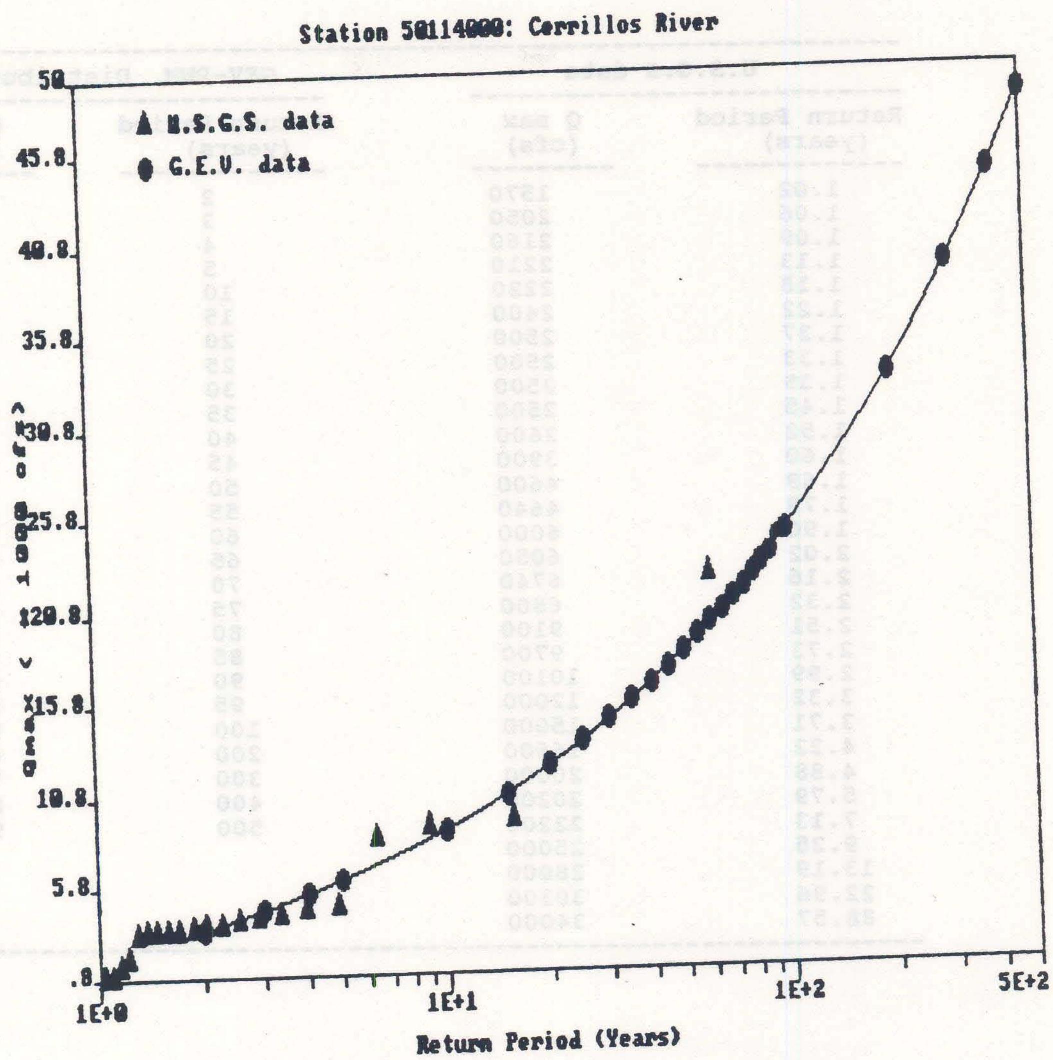

Figure B-6
Flood Frequency Curve for Río Cerrillos basin

Table B-7

Return Period for Historical data and GEV-PWM Distribution for Station 50039500 at Río Cibuco de Vega Baja

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.02	1570	2	7224
1.06	2050	3	10566
1.09	2160	4	12984
1.13	2210	5	14931
1.18	2220	10	21564
1.22	2400	15	25972
1.27	2500	20	29385
1.33	2500	25	32215
1.39	2500	30	34655
1.45	2500	35	36813
1.52	2600	40	38757
1.60	3900	45	40531
1.69	4600	50	42167
1.79	4640	55	43687
1.90	6000	60	45111
2.02	6050	65	46451
2.16	6740	70	47718
2.32	6800	75	48922
2.51	9100	80	50068
2.73	9700	85	51164
2.99	10100	90	52214
3.32	12000	95	53223
3.71	15000	100	54194
4.22	16600	200	68768
4.88	20000	300	78695
5.79	20200	400	86453
7.13	22200	500	92917
9.25	25000		
13.19	28000		
22.96	30300		
88.57	34000		

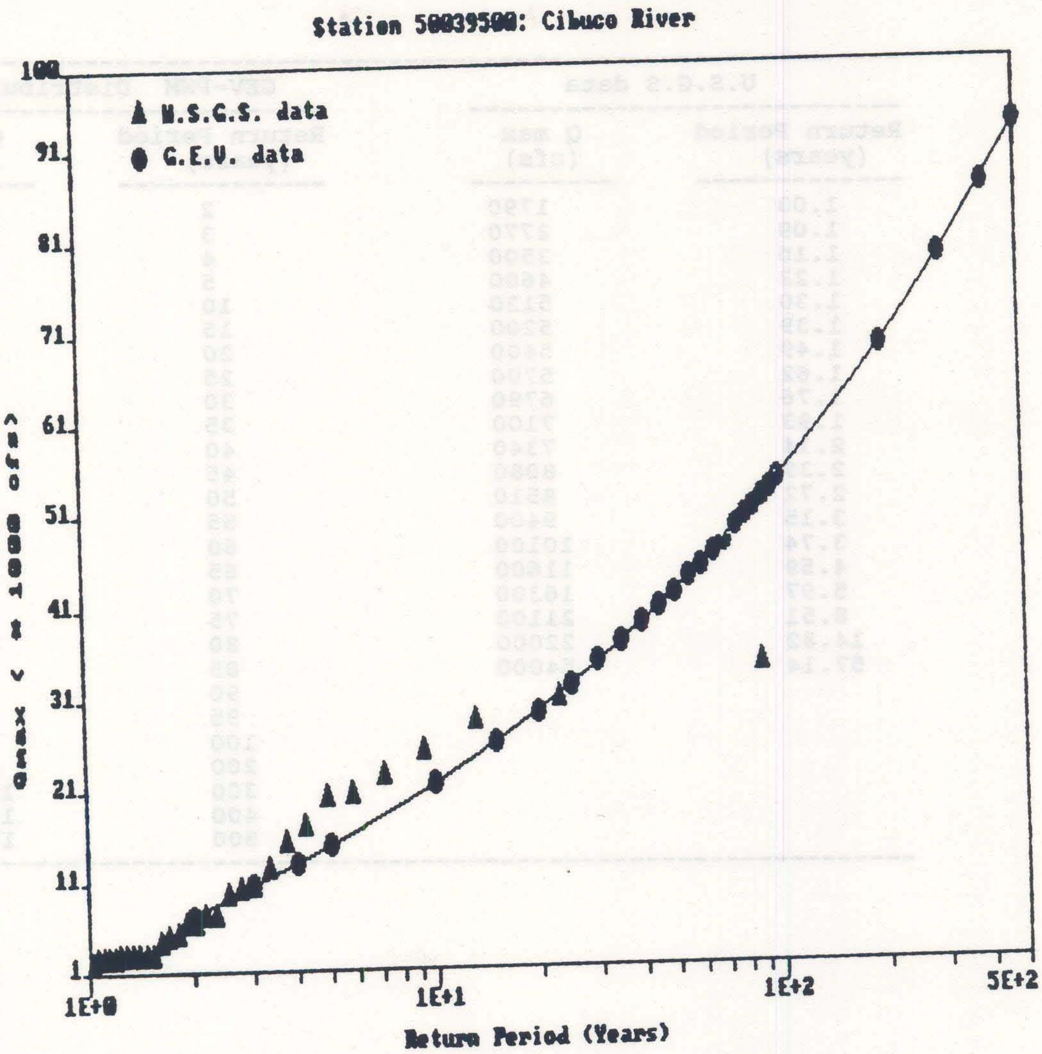

Figure B-7
Flood Frequency Curve for Río Cibuco basin

Table B-8

Return Period for Historical data and GEV-PWM Distribution
for Station 50106500 at Río Coamo

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	1790	2	7155
1.09	2770	3	9779
1.15	3500	4	11833
1.22	4600	5	13577
1.30	5120	10	20100
1.39	5200	15	24908
1.49	5400	20	28879
1.62	5700	25	32330
1.76	6790	30	35418
1.93	7100	35	38235
2.14	7340	40	40839
2.39	8080	45	43271
2.72	8510	50	45560
3.15	9400	55	47727
3.74	10100	60	49790
4.59	11600	65	51760
5.97	16300	70	53651
8.51	21100	75	55469
14.82	22000	80	57222
57.14	54000	85	58917
		90	60559
		95	62152
		100	63701
		200	88625
		300	107340
		400	122903
		500	136478

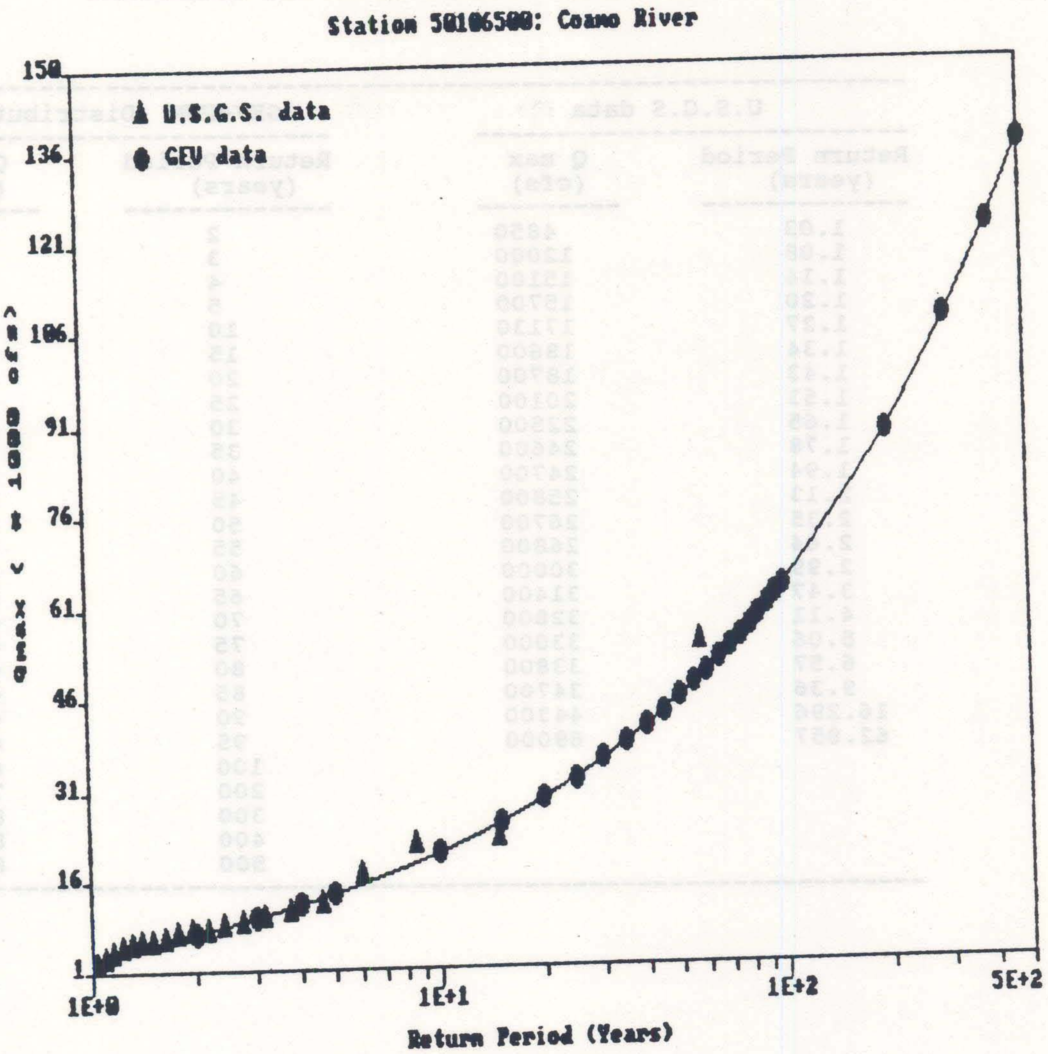

Figure B-8
Flood Frequency Curve for Río Coamo basin

Table B-9

Return Period for Historical data and GEV-PWM Distribution
for Station 501478700 at Río Culebrinas

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	4850	2	24133
1.08	12000	3	29373
1.14	15100	4	32778
1.20	15700	5	35326
1.27	17130	10	42986
1.34	18600	15	47398
1.43	18700	20	50527
1.53	20100	25	52959
1.65	22500	30	54952
1.78	24600	35	56643
1.94	24700	40	58111
2.13	25800	45	59411
2.35	26700	50	60576
2.64	26800	55	61633
2.99	30000	60	62600
3.47	31400	65	63491
4.11	32800	70	64319
5.06	33000	75	65090
6.57	33800	80	65814
9.36	34700	85	66494
16.296	44300	90	67137
62.857	69000	95	67746
		100	68325
		200	76238
		300	80949
		400	84332
		500	86978

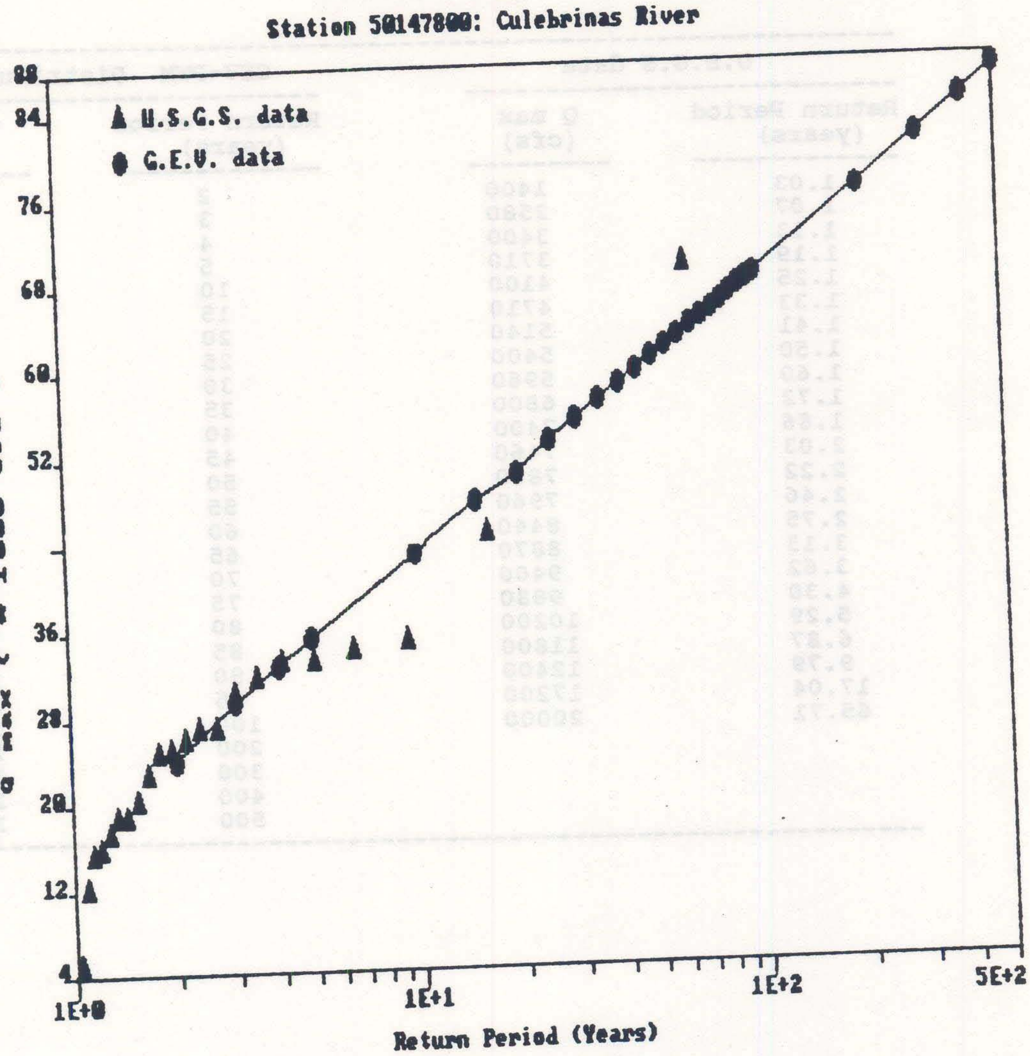

Figure B-9
Flood Frequency Curve for Río Culebrinas basin

Table B-10

Return Period for Historical data and GEV-PWM Distribution for Station 50063800 at Río Espiritu Santo

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	1400	2	7031
1.07	2580	3	8875
1.13	3400	4	10085
1.19	3710	5	10997
1.25	4100	10	13770
1.33	4710	15	15389
1.41	5140	20	16546
1.50	5400	25	17451
1.60	5960	30	18197
1.72	6800	35	18831
1.86	7400	40	19384
2.03	7460	45	19875
2.22	7880	50	20316
2.46	7960	55	20717
2.75	8440	60	21085
3.13	8870	65	21425
3.62	9400	70	21741
4.30	9880	75	22036
5.29	10200	80	22313
6.87	11800	85	22574
9.79	12400	90	22821
17.04	17200	95	23055
65.71	20000	100	23278
		200	26351
		300	28203
		400	29544
		500	30599

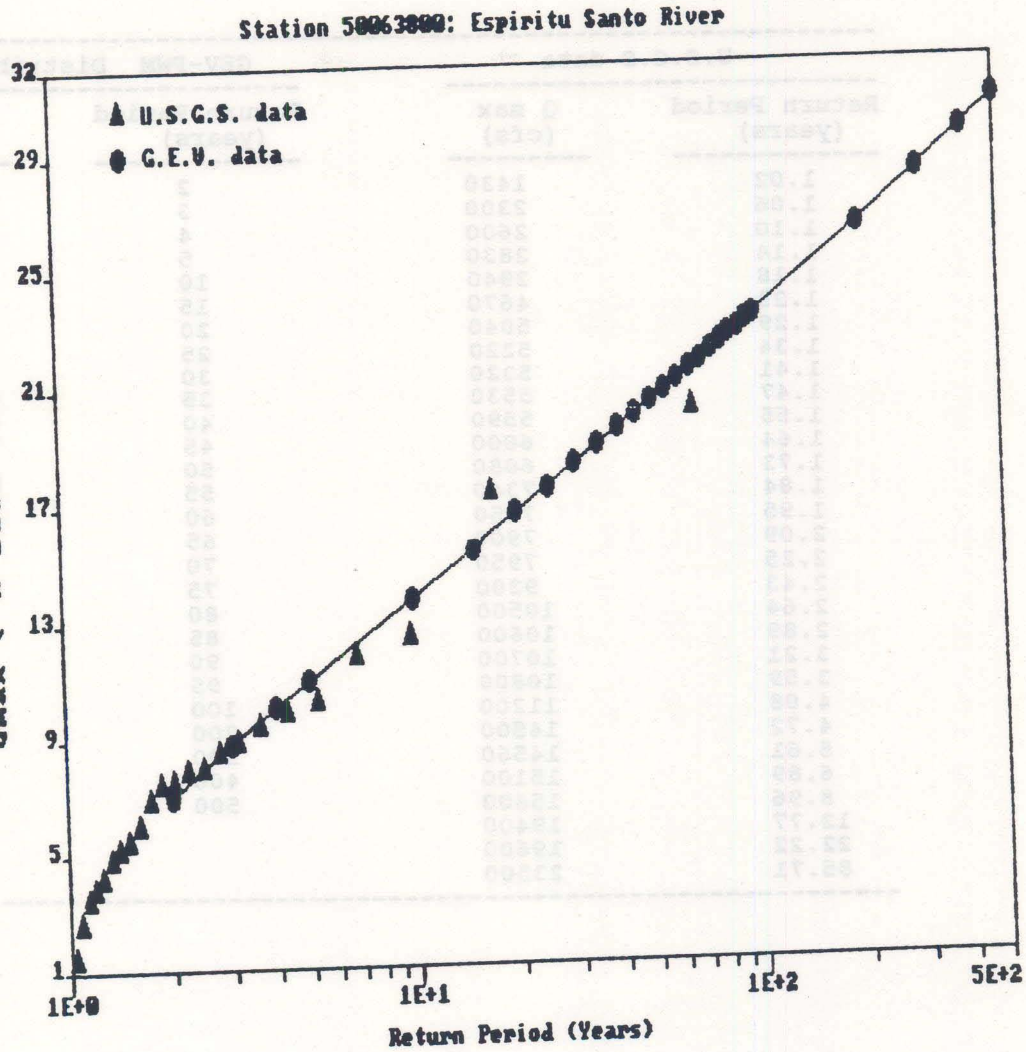

Figure B-10
Flood Frequency Curve for Río Espíritu Santo basin

Table B-11

Return Period for Historical data and GEV-PWM Distribution for Station 50071000 at Río Fajardo

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.02	1430	2	7883
1.06	2300	3	10247
1.10	2600	4	11808
1.14	2830	5	12987
1.18	2940	10	16594
1.23	4670	15	18713
1.29	5040	20	20235
1.34	5220	25	21428
1.41	5320	30	22414
1.47	5530	35	23254
1.55	5590	40	23988
1.64	6000	45	24640
1.73	6080	50	25227
1.84	7340	55	25761
1.95	7850	60	26251
2.09	7900	65	26705
2.25	7950	70	27126
2.43	9200	75	27521
2.64	10500	80	27891
2.89	10600	85	28241
3.21	10700	90	28572
3.59	10800	95	28886
4.08	11200	100	29185
4.72	14500	200	33324
5.61	14560	300	35835
6.89	15100	400	37660
8.96	15600	500	39100
12.77	19400		
22.22	19600		
85.71	23500		

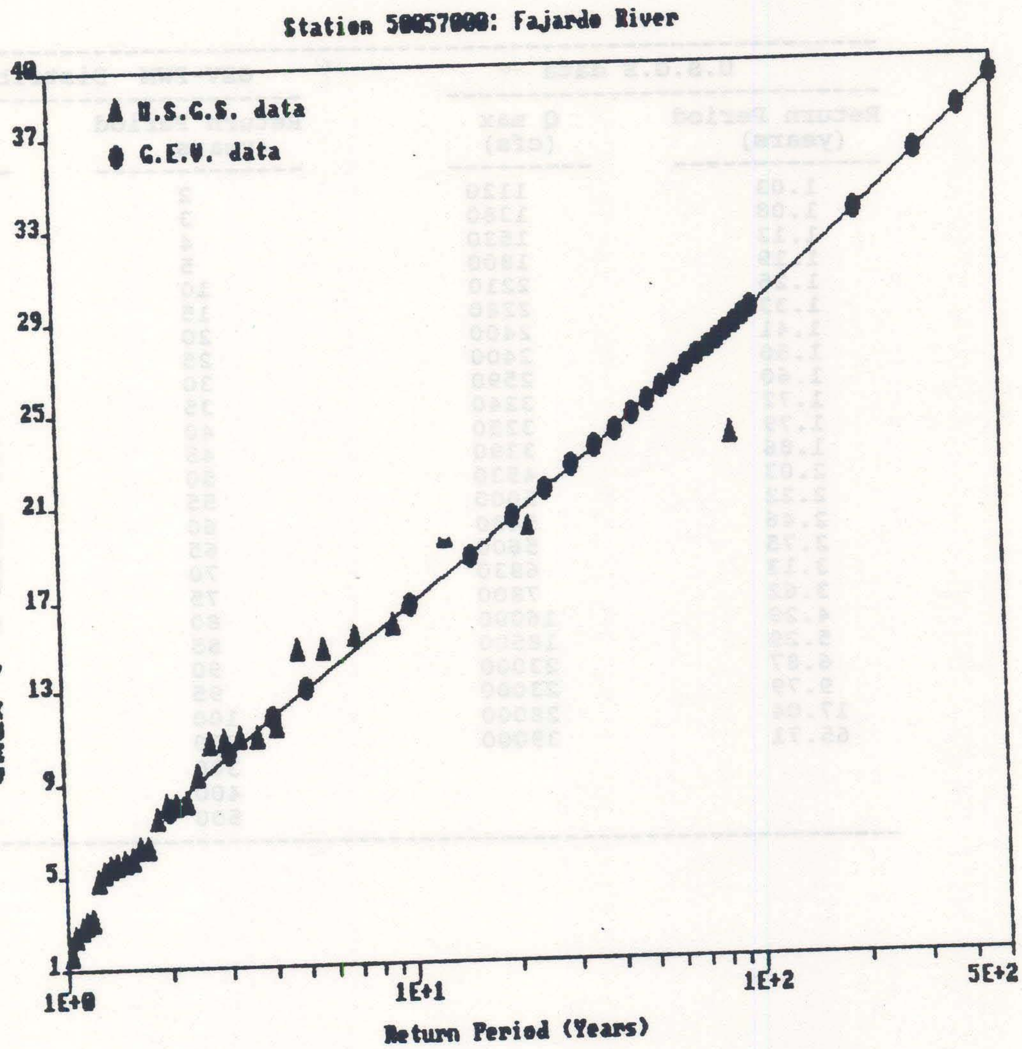

Figure B-11
Flood Frequency Curve for Río Fajardo basin

Table B-12

Return Period for Historical data and GEV-PWM Distribution for Station 50124500 at Río Guayanilla

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	1120	2	5175
1.08	1380	3	7931
1.13	1530	4	10070
1.19	1800	5	11877
1.25	2210	10	18571
1.33	2280	15	23452
1.41	2400	20	27456
1.50	2400	25	30920
1.60	2590	30	34008
1.72	3240	35	36815
1.79	3250	40	39404
1.86	3390	45	41815
2.03	4530	50	44080
2.22	5000	55	46220
2.46	5400	60	48254
2.75	5600	65	50194
3.13	6830	70	52052
3.62	7800	75	53837
4.29	16000	80	55556
5.29	18500	85	57216
6.87	23000	90	58822
9.79	23000	95	60379
17.04	28000	100	61891
65.71	39000	200	86048
		300	104006
		400	118844
		500	131726

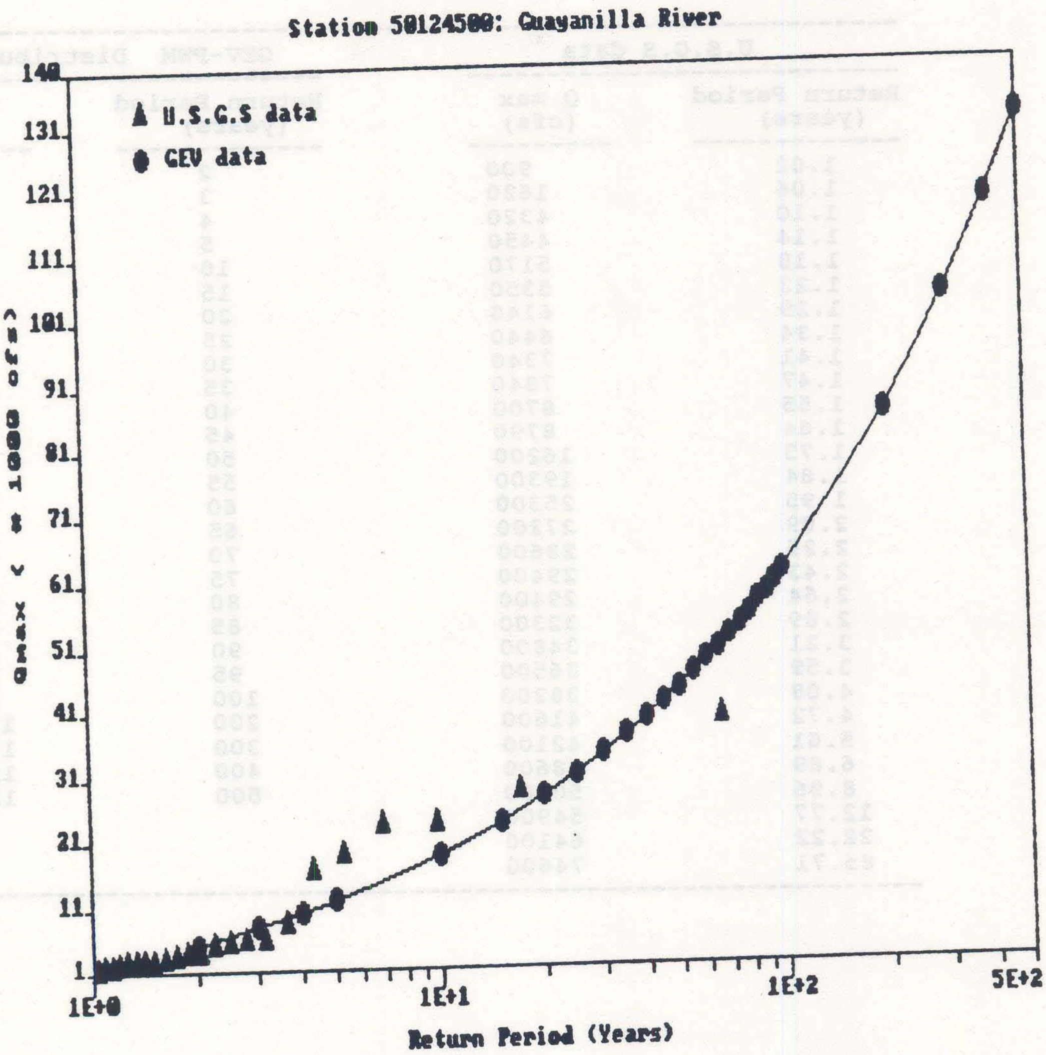

Figure B-12
Flood Frequency Curve for Río Guayanilla basin

Table B-13

Return Period for Historical data and GEV-PWM Distribution
for Station 50057000 at Río Gurabo

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.02	900	2	21485
1.06	1620	3	30147
1.10	4320	4	35781
1.14	4450	5	39996
1.18	5170	10	52679
1.23	5350	15	59988
1.29	6140	20	65173
1.34	6440	25	69205
1.41	7340	30	72510
1.47	7840	35	75313
1.55	8700	40	77749
1.64	8790	45	79904
1.73	16200	50	81838
1.84	19300	55	83591
1.95	25300	60	85196
2.09	27200	65	86676
2.25	28600	70	88049
2.43	29400	75	89330
2.64	29400	80	90531
2.89	32300	85	91661
3.21	34800	90	92728
3.59	36500	95	93740
4.08	38200	100	94701
4.72	41600	200	107845
5.61	42100	300	115677
6.89	48600	400	121301
8.96	50600	500	125302
12.77	54900		
22.22	64100		
85.71	74600		

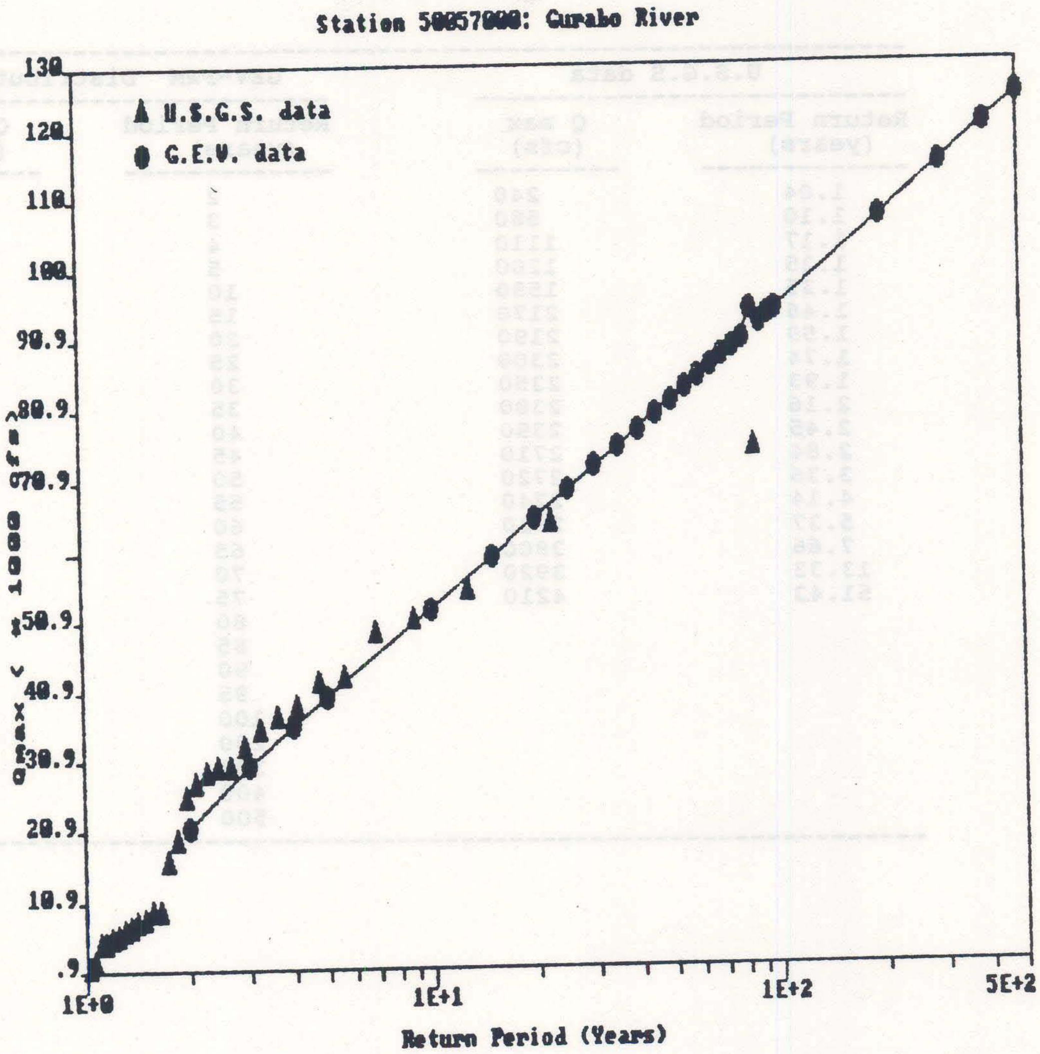

Figure B-13
Flood Frequency Curve for Río Gurabo basin

Table B-14

Return Period for Historical data and GEV-PWM Distribution for Station 50062500 at Río Herrera

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.04	240	2	2288
1.10	580	3	2778
1.17	1110	4	3055
1.25	1200	5	3245
1.35	1580	10	3734
1.46	2170	15	3969
1.59	2190	20	4118
1.74	2300	25	4225
1.93	2350	30	4308
2.16	2380	35	4374
2.45	2390	40	4429
2.84	2710	45	4476
3.36	2720	50	4516
4.14	2740	55	4552
5.37	2820	60	4584
7.66	3800	65	4612
13.33	3920	70	4638
51.43	4210	75	4662
		80	4683
		85	4703
		90	4722
		95	4739
		100	4755
		200	4952
		300	5051
		400	5115
		500	5161

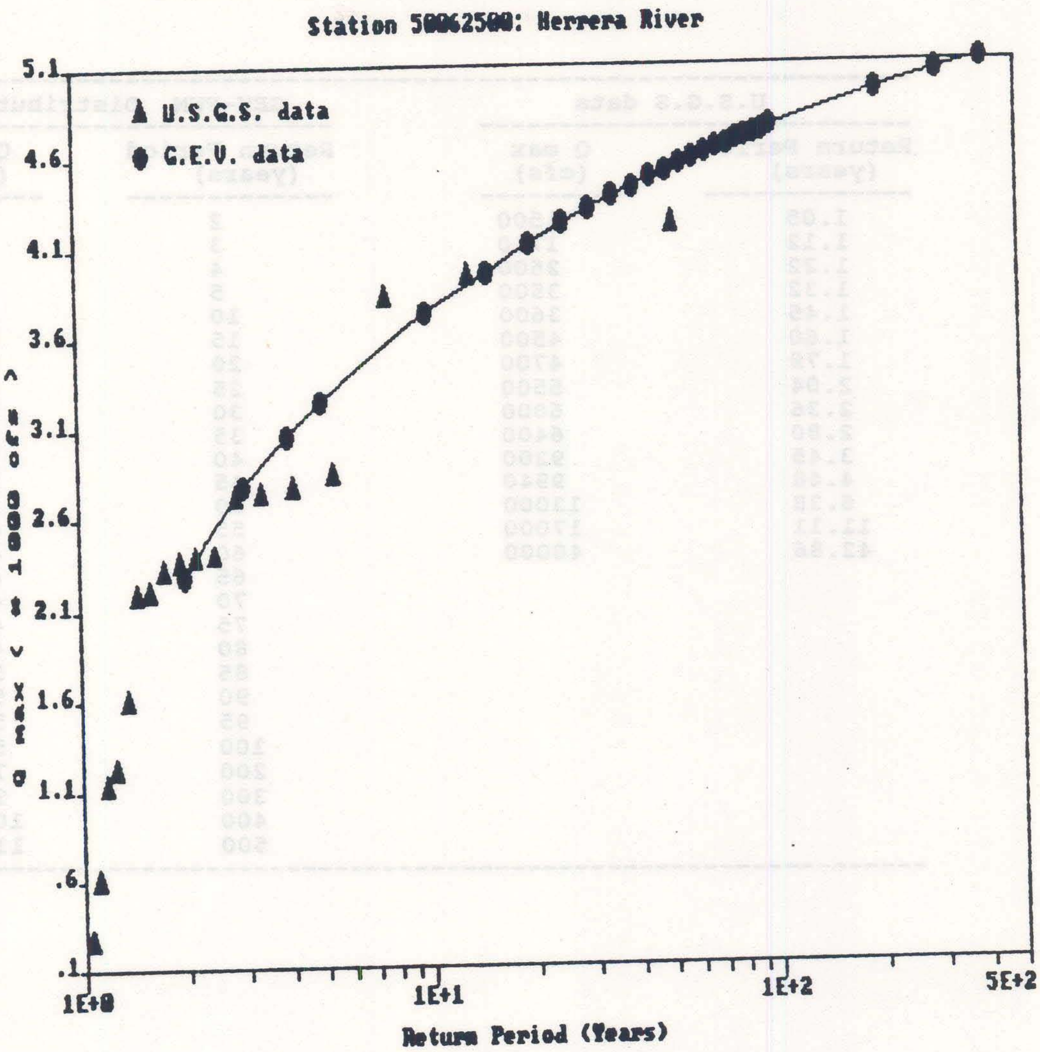

Figure B-14
Flood Frequency Curve for Río Herrera basin

Table B-15

Return Period for Historical data and GEV-PWM Distribution
for Station 50082000 at Río Humacao

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.05	1500	2	5427
1.12	1740	3	7727
1.22	2500	4	9524
1.32	3500	5	11048
1.45	3600	10	16736
1.60	4500	15	20919
1.79	4700	20	24369
2.04	5500	25	27363
2.36	6000	30	30041
2.80	6400	35	32482
3.45	9200	40	34737
4.48	9940	45	36842
6.38	13000	50	38822
11.11	17000	55	40695
42.86	40000	60	42478
		65	44181
		70	45814
		75	47834
		80	48898
		85	50361
		90	51778
		95	53152
		100	54488
		200	75951
		300	92032
		400	105385
		500	117021

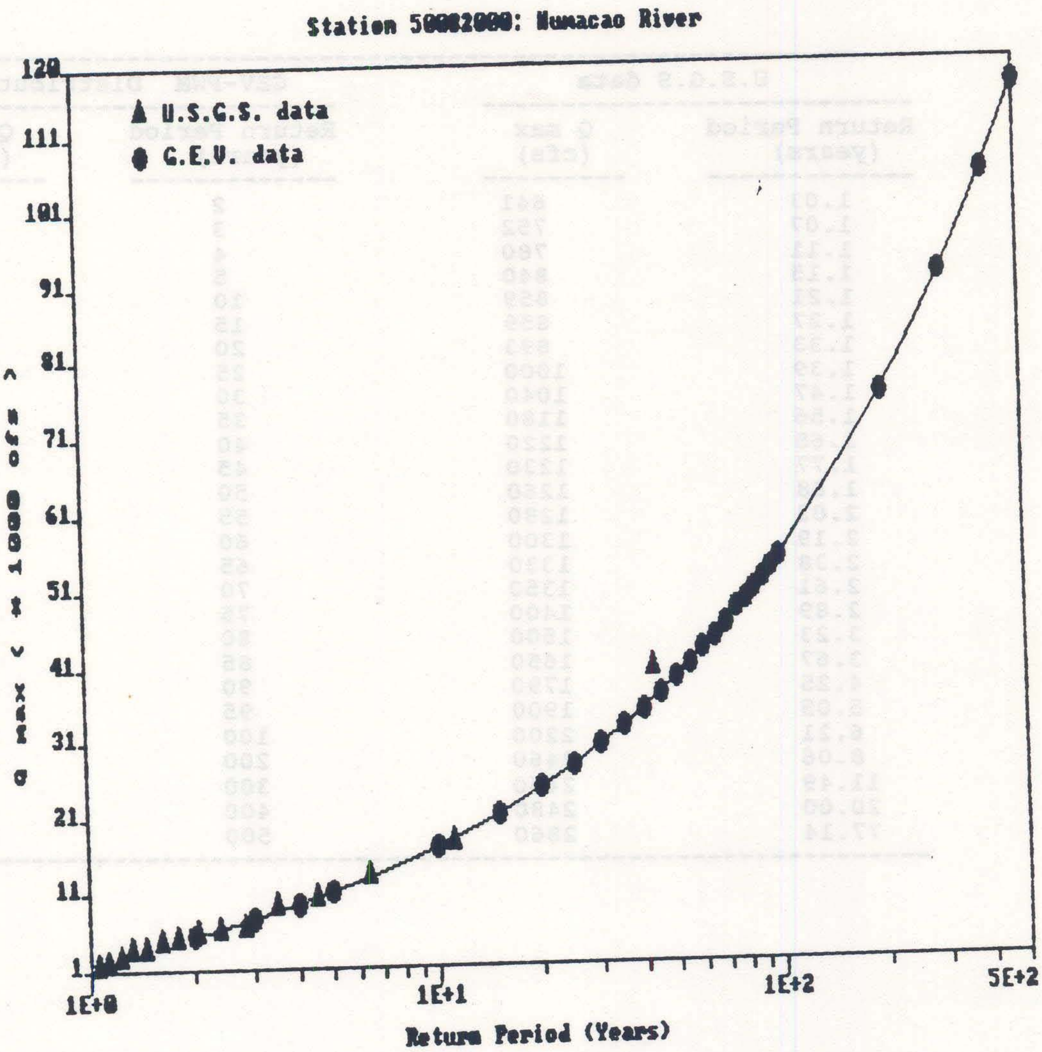

Figure B-15
Flood Frequency Curve for Río Humacao basin

Table B-16

Return Period for Historical data and GEV-PWM Distribution for Station 50075000 at Río Icacos

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	641	2	1285
1.07	752	3	1539
1.11	780	4	1709
1.15	840	5	1839
1.21	859	10	2243
1.27	859	15	2486
1.33	893	20	2663
1.39	1000	25	2803
1.47	1040	30	2919
1.56	1180	35	3019
1.65	1220	40	3106
1.77	1230	45	3185
1.88	1260	50	3255
2.02	1280	55	3320
2.19	1300	60	3379
2.38	1330	65	3434
2.61	1350	70	3486
2.89	1400	75	3534
3.23	1500	80	3580
3.67	1650	85	3622
4.25	1790	90	3663
5.05	1900	95	3702
6.21	2200	100	3739
8.06	2460	200	4257
11.49	2460	300	4578
20.00	2480	400	4814
77.14	2860	500	5002

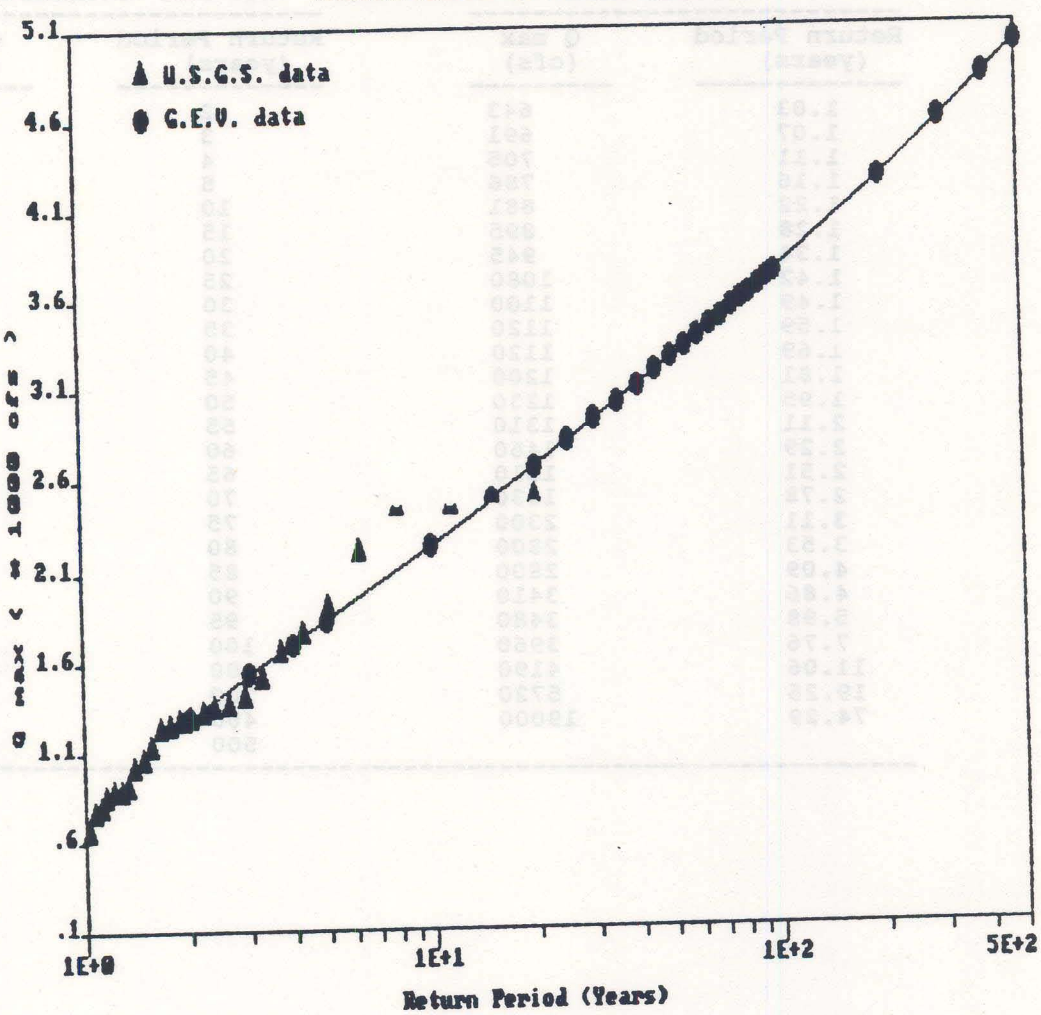

Figure B-16
Flood Frequency Curve for Río Icacos basin

Table B-17

Return Period for Historical data and GEV-PWM Distribution for Station 50112500 at Río Inabón

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	643	2	1475
1.07	691	3	2044
1.11	705	4	2511
1.16	786	5	2920
1.22	881	10	4540
1.28	895	15	5811
1.34	945	20	6903
1.42	1080	25	7878
1.49	1100	30	8771
1.59	1120	35	9601
1.69	1120	40	10381
1.81	1200	45	11119
1.95	1230	50	11822
2.11	1310	55	12495
2.29	1460	60	13142
2.51	1510	65	13766
2.78	1830	70	14370
3.11	2300	75	14955
3.53	2800	80	15524
4.09	2800	85	16077
4.86	3410	90	16617
5.98	3480	95	17144
7.76	3960	100	17658
11.06	4190	200	26310
19.26	5720	300	33198
74.29	19000	400	39145
		500	44478

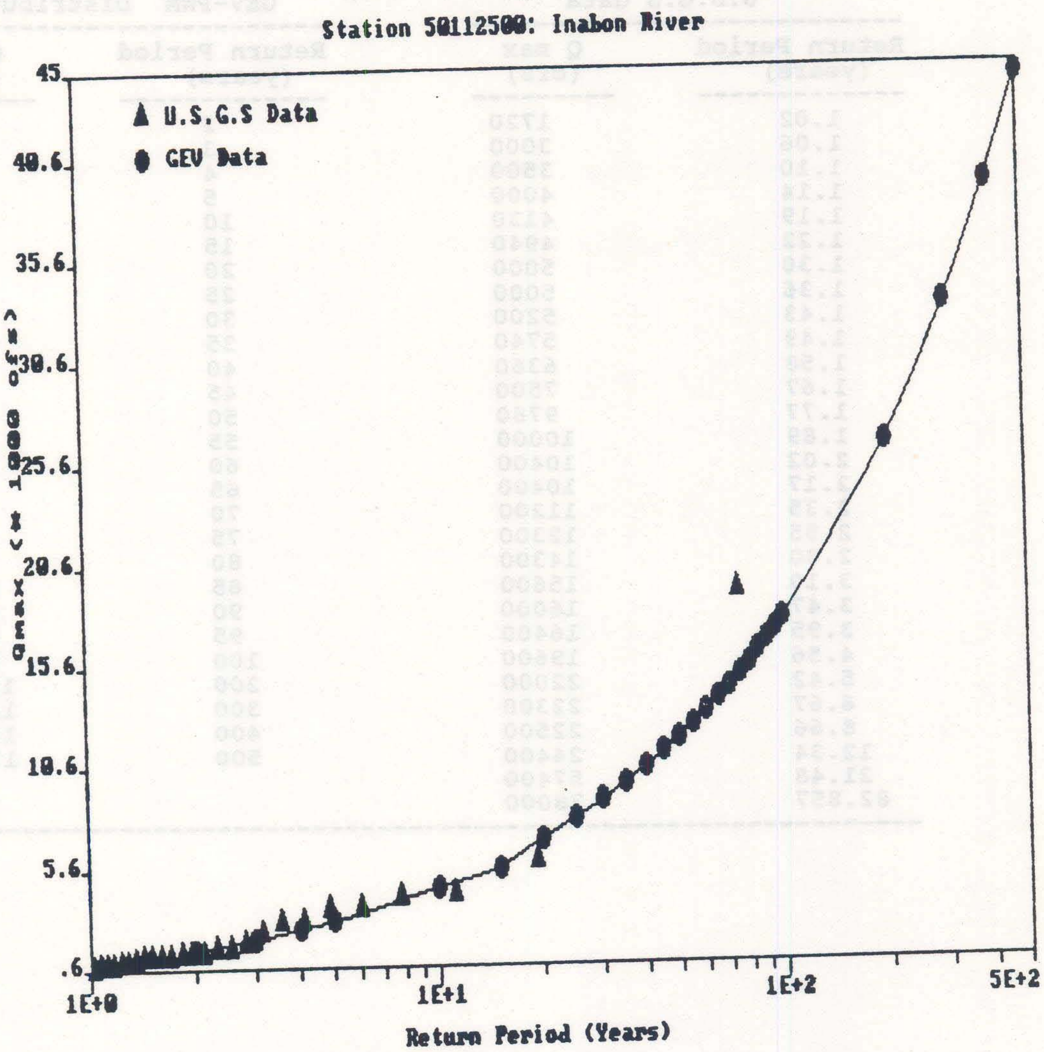

Figure B-17
Flood Frequency Curve for Río Inabón basin

Table B-18

Return Period for Historical data and GEV-PWM Distribution
for Station 50111300 at Río Jacaguas

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.02	1720	2	9796
1.06	3000	3	13902
1.10	3500	4	17041
1.14	4000	5	19662
1.19	4130	10	29191
1.22	4940	15	35996
1.30	5000	20	41506
1.36	5000	25	46227
1.43	5200	30	50404
1.49	5740	35	54178
1.58	6360	40	57639
1.67	7500	45	60849
1.77	9780	50	63851
1.89	10000	55	66678
2.02	10400	60	69354
2.17	10400	65	71900
2.35	11200	70	74331
2.55	12300	75	76661
2.80	14300	80	78899
3.10	15600	85	81055
3.47	16000	90	83137
3.95	16400	95	85151
4.56	19600	100	87102
5.42	22000	200	117844
6.67	22300	300	140252
8.66	22500	400	158535
12.34	24400	500	174261
21.48	57400		
82.857	78000		

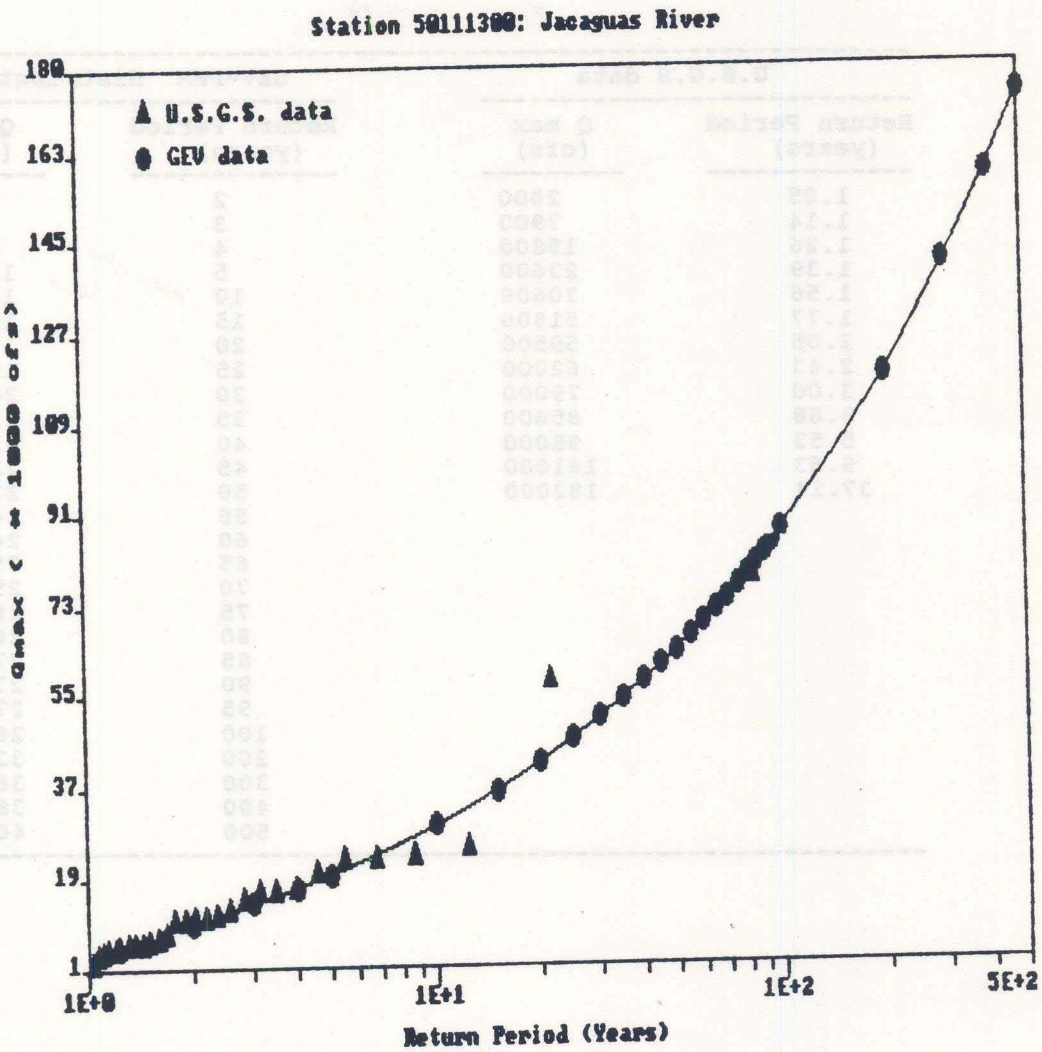

Figure B-18
Flood Frequency Curve for Rio Jacaguas basin

Table B-19

Return Period for Historical data and GEV-PWM Distribution for Station 50059000 at Río Grande de Loíza

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.05	2000	2	52616
1.14	7900	3	75178
1.26	15800	4	90457
1.39	23600	5	102210
1.56	30600	10	139259
1.77	51800	15	161795
2.05	59500	20	178321
2.43	62000	25	191487
3.00	79000	30	202488
3.88	85600	35	211967
5.53	98000	40	220317
9.63	161000	45	227791
37.14	182000	50	234566
		55	240768
		60	246493
		65	251813
		70	256786
		75	261455
		80	265860
		85	270029
		90	273988
		95	277759
		100	281360
		200	332348
		300	364305
		400	387998
		500	406984

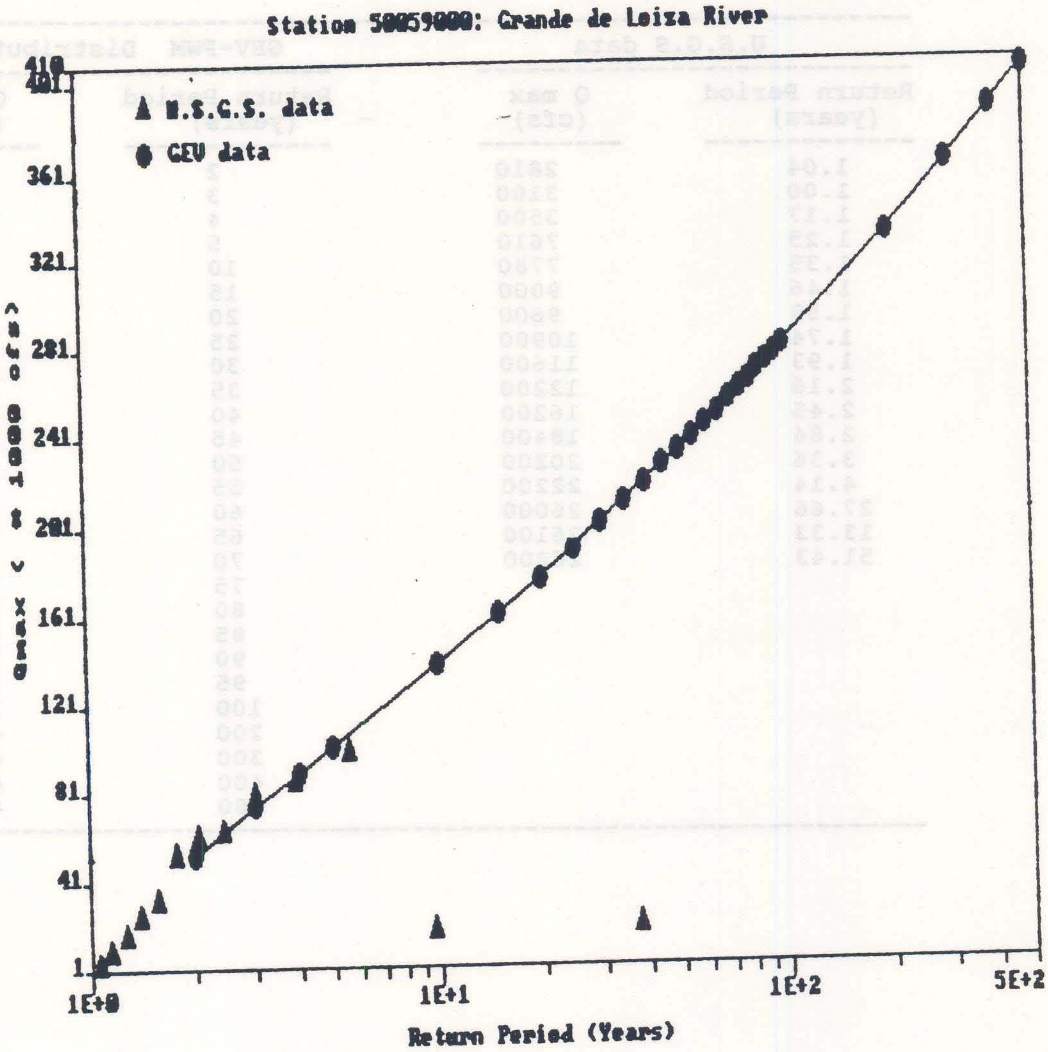

Figure B-19
Flood Frequency Curve for Río Grande de Loiza basin

Table B-20

Return Period for Historical data and GEV-PWM Distribution
for Station 50065700 at Río Mameyes

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.04	2810	2	12419
1.00	3100	3	16005
1.17	3500	4	18252
1.25	7610	5	19892
1.35	7780	10	24624
1.46	9000	15	27220
1.59	9600	20	29007
1.74	10900	25	30367
1.93	11600	30	31462
2.16	12200	35	32378
2.45	16200	40	33164
2.84	18400	45	33853
3.36	20200	50	34464
4.14	22200	55	35014
27.66	26000	60	35514
13.33	26100	65	35971
51.43	26200	70	36393
		75	36784
		80	37148
		85	37489
		90	37810
		95	38112
		100	38398
		200	42191
		300	44349
		400	45854
		500	47007

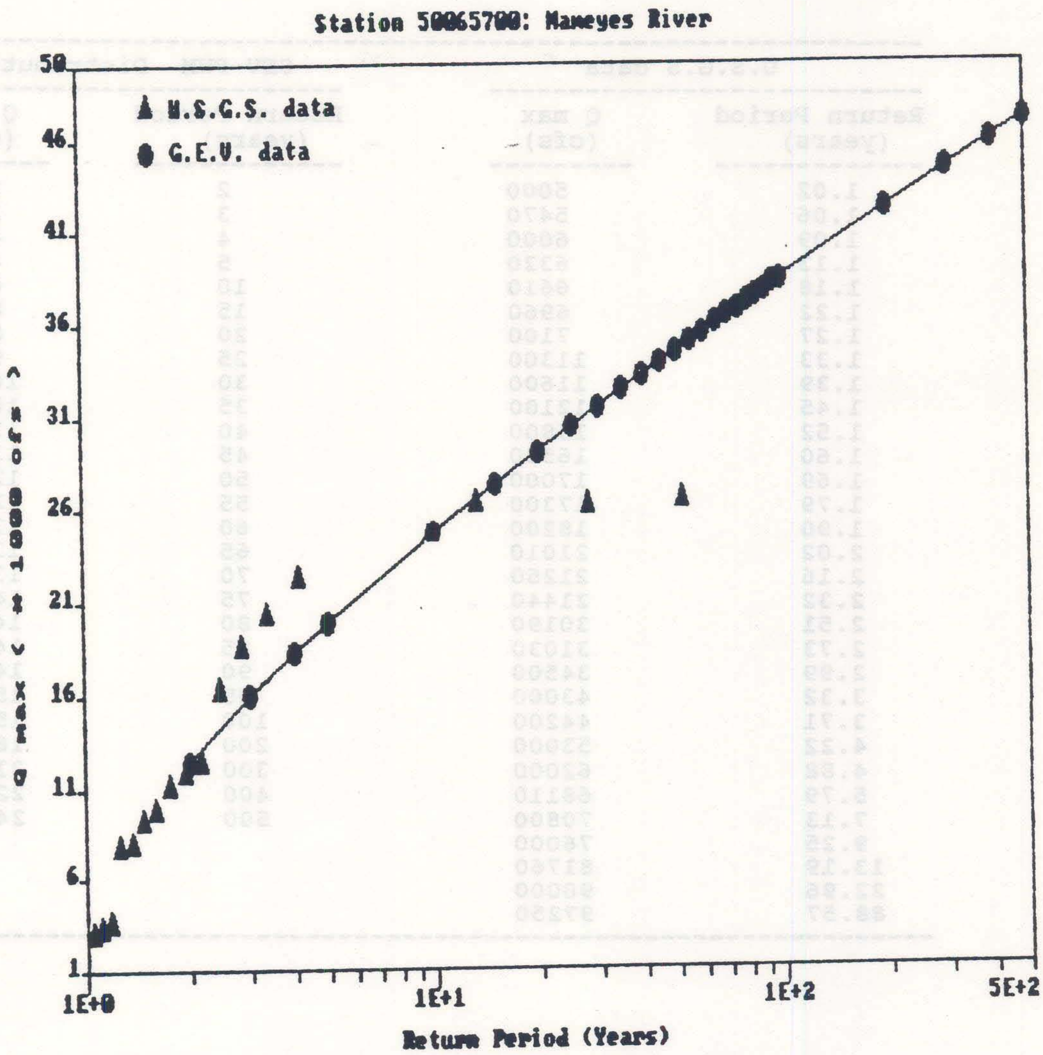

Figure B-20
Flood Frequency Curve for Río Mameyes basin

Table B-21

Return Period for Historical data and GEV-PWM Distribution
for Station 50038100 at Río Grande de Manatí

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.02	5000	2	24593
1.06	5470	3	35181
1.09	6000	4	42623
1.13	6320	5	48495
1.18	6610	10	67828
1.22	6960	15	80186
1.27	7100	20	89530
1.33	11300	25	97143
1.39	11600	30	103616
1.45	12100	35	109277
1.52	15800	40	114325
1.60	16500	45	118893
1.69	17000	50	123074
1.79	17300	55	126935
1.90	18200	60	130527
2.02	21010	65	133888
2.16	21250	70	137051
2.32	21440	75	140040
2.51	30190	80	142875
2.73	31030	85	145573
2.99	34500	90	148148
3.32	43000	95	150613
3.71	44200	100	152977
4.22	53000	200	187564
4.88	62000	300	210278
5.79	68110	400	227623
7.13	70800	500	241830
9.25	76000		
13.19	81760		
22.96	90000		
88.57	97250		

Station 50038100: Grande de Manati River

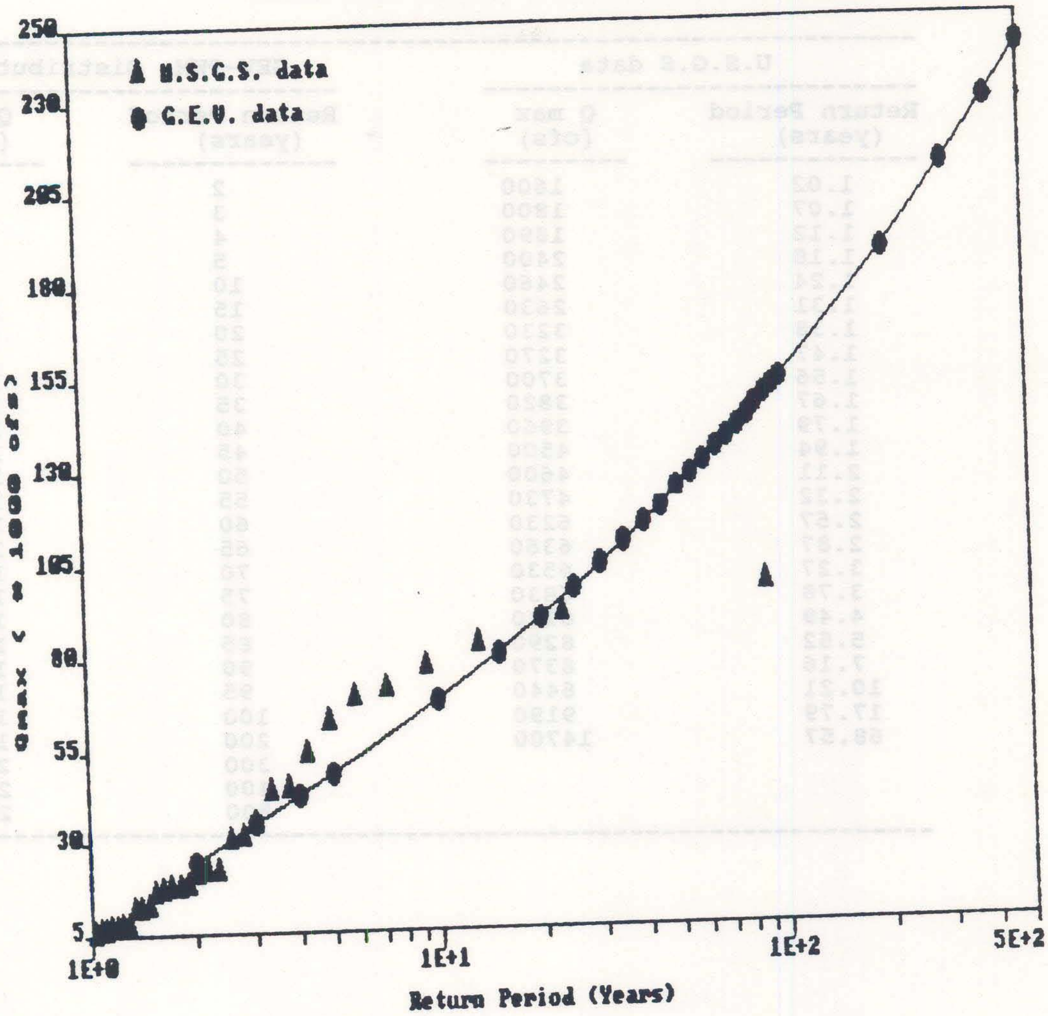

Figure B-21
Flood Frequency Curve for Río Grande de Manatí basin

Table B-22

Return Period for Historical data and GEV-PWM Distribution
for Station 5092000 at Río Grande de Patillas

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.02	1600	2	4664
1.07	1800	3	5947
1.12	1890	4	6796
1.18	2400	5	7440
1.24	2460	10	9415
1.31	2630	15	10581
1.38	3230	20	11420
1.47	3270	25	12080
1.56	3700	30	12625
1.67	3820	35	13091
1.79	3960	40	13498
1.94	4500	45	13860
2.11	4600	50	14187
2.32	4730	55	14484
2.57	6230	60	14757
2.87	6360	65	15009
3.27	6530	70	15244
3.78	6830	75	15464
4.49	8130	80	15671
5.52	8290	85	15866
7.16	8370	90	16051
10.21	8440	95	16226
17.79	9190	100	16394
68.57	14700	200	18714
		300	20127
		400	21157
		500	21971

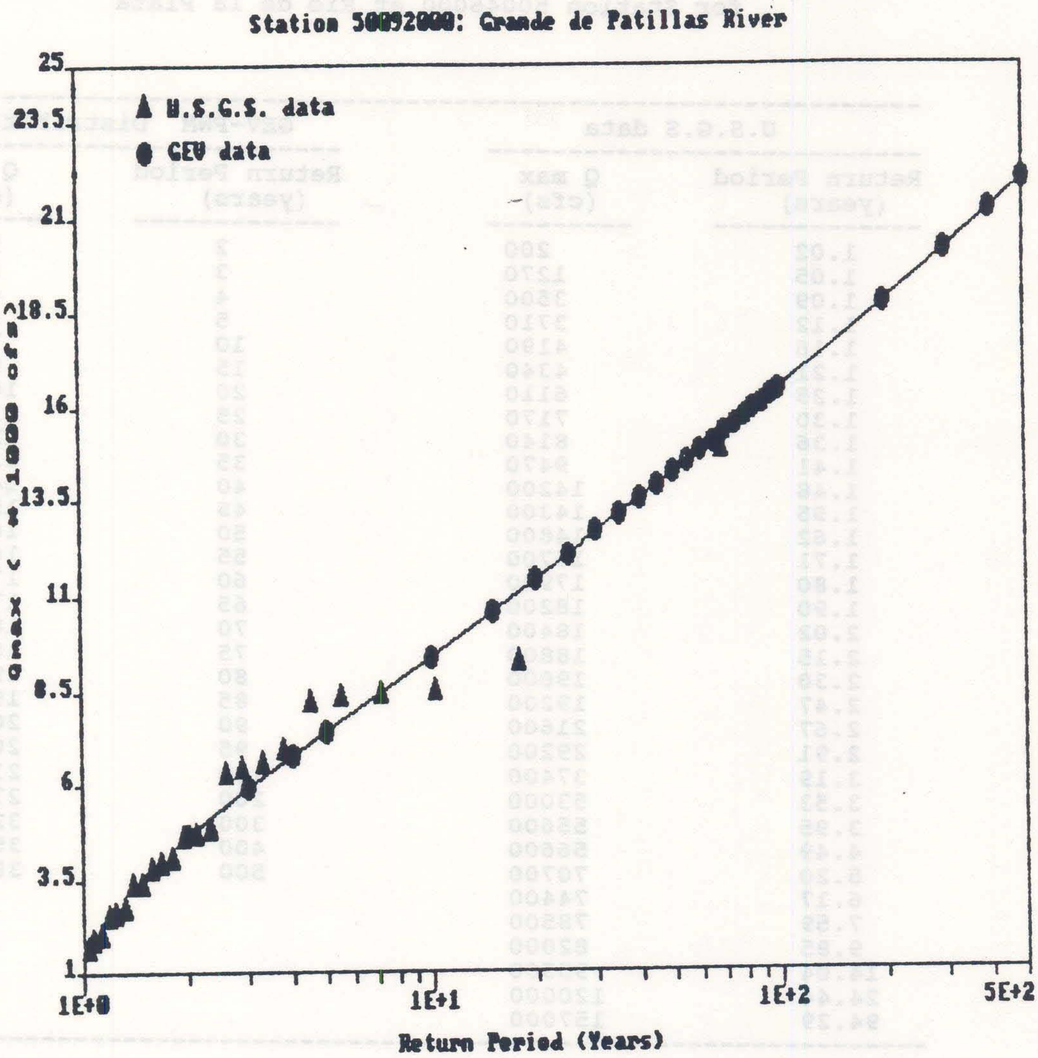

Figure B-22
Flood Frequency Curve for Río Grande de Patillas basin

Table B-23

Return Period for Historical data and GEV-PWM Distribution for Station 50046000 at Río de la Plata

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.02	200	2	22805
1.05	1270	3	34980
1.09	3500	4	43961
1.12	3710	5	51283
1.16	4190	10	76800
1.21	4340	15	94188
1.25	6110	20	107866
1.30	7170	25	119336
1.36	8140	30	129316
1.41	9470	35	138208
1.48	14200	40	146268
1.55	14300	45	153664
1.62	14800	50	160519
1.71	16700	55	166920
1.80	17900	60	172936
1.90	18200	65	178619
2.02	18400	70	184012
2.15	18800	75	189149
2.30	19000	80	194058
2.47	19200	85	198763
2.67	21600	90	203283
2.91	29200	95	207636
3.19	37400	100	211836
3.53	53000	200	275943
3.95	55600	300	320654
4.49	56600	400	356127
5.20	70700	500	386011
6.17	74400		
7.59	78500		
9.85	82000		
14.04	95500		
24.44	120000		
94.29	157000		

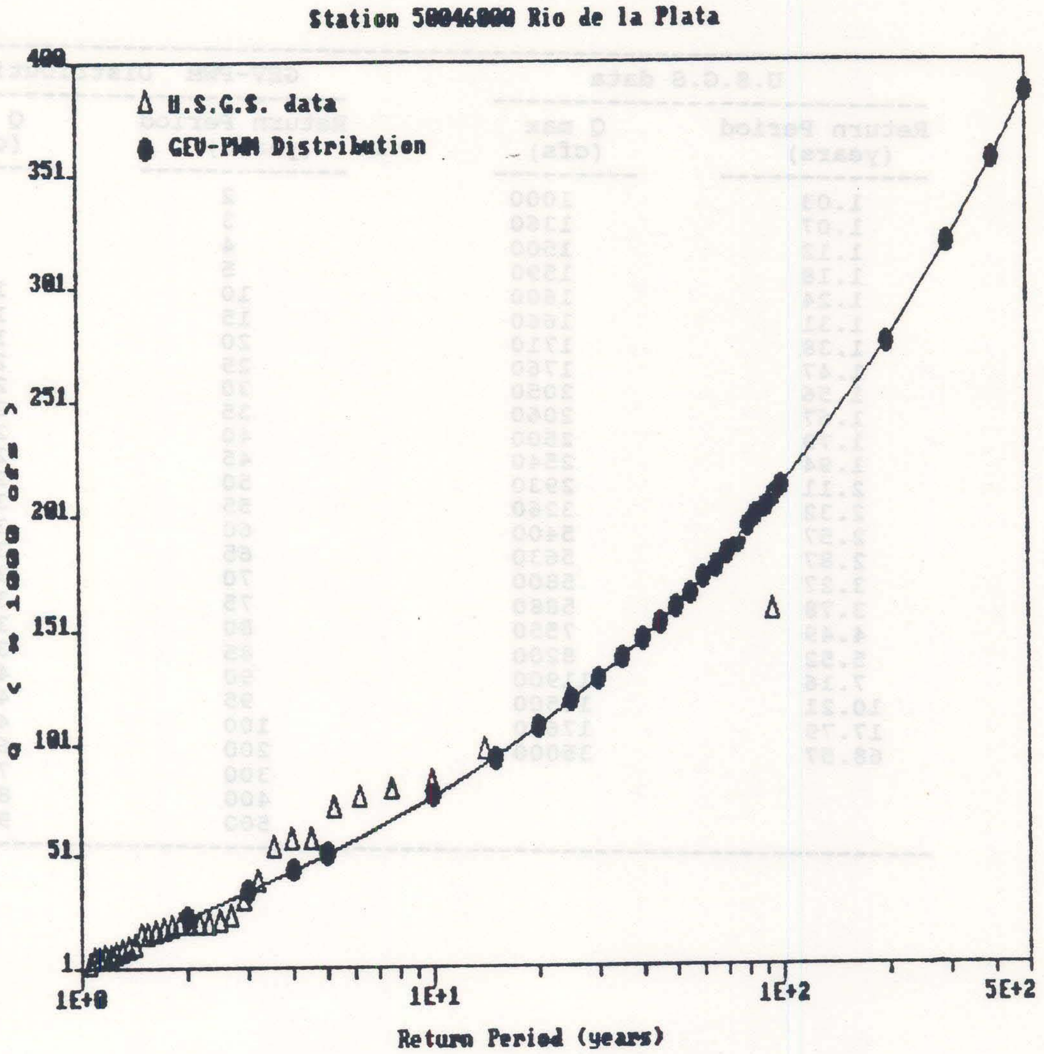

Figure B-23
Flood Frequency Curve for Río de La Plata basin

Table B-24

Return Period for Historical data and GEV-PWM Distribution for Station 50115900 at Río Portugués

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	1000	2	3478
1.07	1360	3	5121
1.12	1500	4	6431
1.18	1590	5	7556
1.24	1600	10	11861
1.31	1660	15	15110
1.38	1710	20	17834
1.47	1760	25	20228
1.56	2050	30	22388
1.67	2060	35	24373
1.79	2500	40	26220
1.94	2540	45	27953
2.11	2930	50	29592
2.32	3260	55	31151
2.57	5400	60	32640
2.87	5630	65	34068
3.27	5800	70	35442
3.78	5880	75	36768
4.49	7550	80	38050
5.52	8200	85	39292
7.16	11900	90	40499
10.21	14500	95	41673
17.79	17600	100	42816
68.57	35000	200	61524
		300	75893
		400	88015
		500	98703

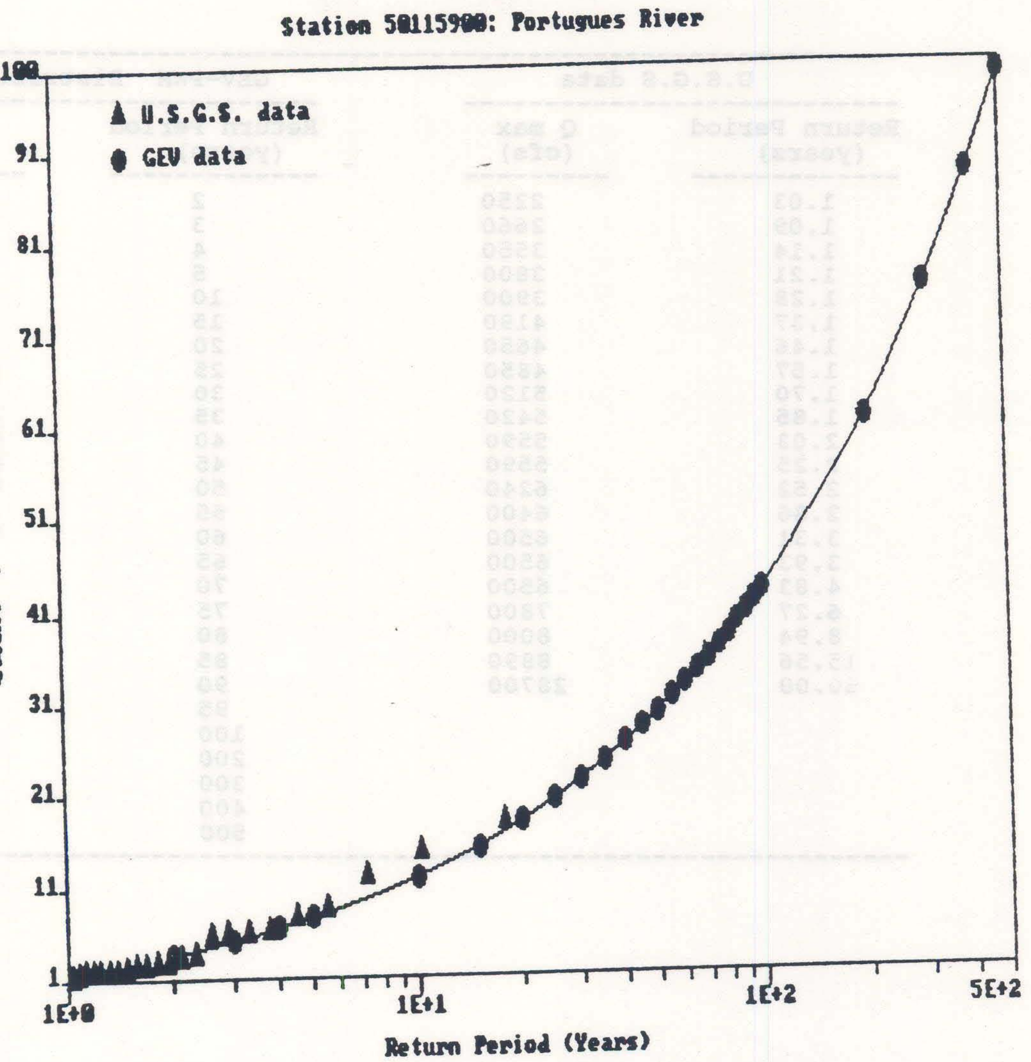

Figure B-24
Flood Frequency Curve for Río Portugués basin

Table B-25

Return Period for Historical data and GEV-PWM Distribution for Station 50136000 at Río Rosario

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	2250	2	5127
1.09	2660	3	6311
1.14	3550	4	7213
1.21	3800	5	7963
1.28	3900	10	10667
1.37	4190	15	12604
1.46	4650	20	14158
1.57	4850	25	15486
1.70	5120	30	16659
1.85	5420	35	17717
2.03	5590	40	18685
2.25	5590	45	19582
2.52	6240	50	20420
2.86	6400	55	21208
3.31	6500	60	21954
3.93	6500	65	22662
4.83	6500	70	23339
6.27	7800	75	23986
8.94	8000	80	24607
15.56	8890	85	25206
30.00	28700	90	25783
		95	26341
		100	26882
		200	35365
		300	41513
		400	46593
		500	50801

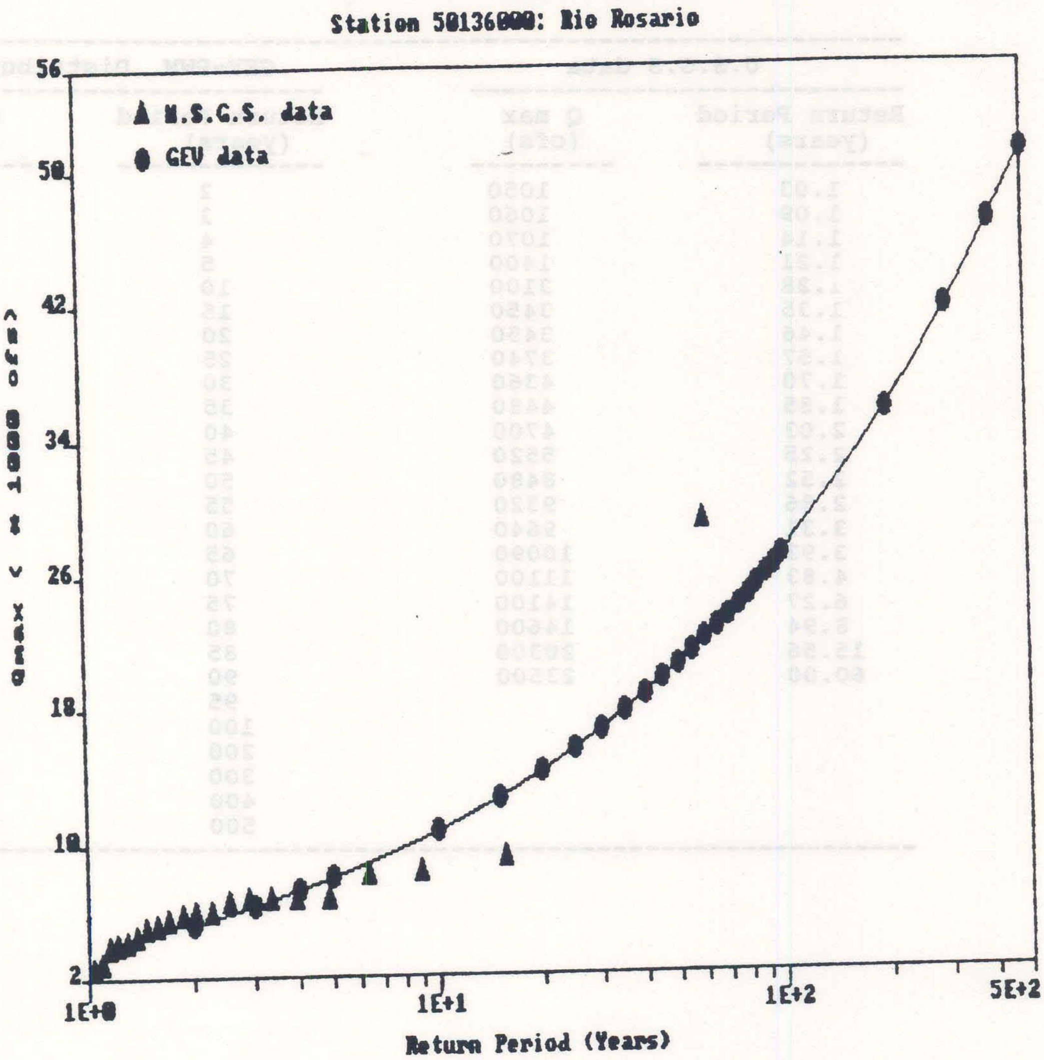

Figure B-25
Flood Frequency Curve for Rio Rosario basin

Table B-26

Return Period for Historical data and GEV-PWM Distribution for Station 50121000 at Río Tallaboa

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.03	1050	2	5831
1.09	1060	3	8234
1.14	1070	4	9909
1.21	1400	5	11223
1.28	3100	10	15509
1.35	3450	15	18219
1.46	3450	20	20255
1.57	3740	25	21905
1.70	4360	30	23302
1.85	4480	35	24520
2.03	4700	40	25604
2.25	5520	45	26582
2.52	8480	50	27475
2.86	9320	55	28298
3.31	9640	60	29063
3.93	10090	65	29778
4.83	11100	70	30449
6.27	14100	75	31082
8.94	14600	80	31682
15.56	20300	85	32253
60.00	23500	90	32797
		95	33317
		100	33815
		200	41052
		300	45756
		400	49325
		500	52234

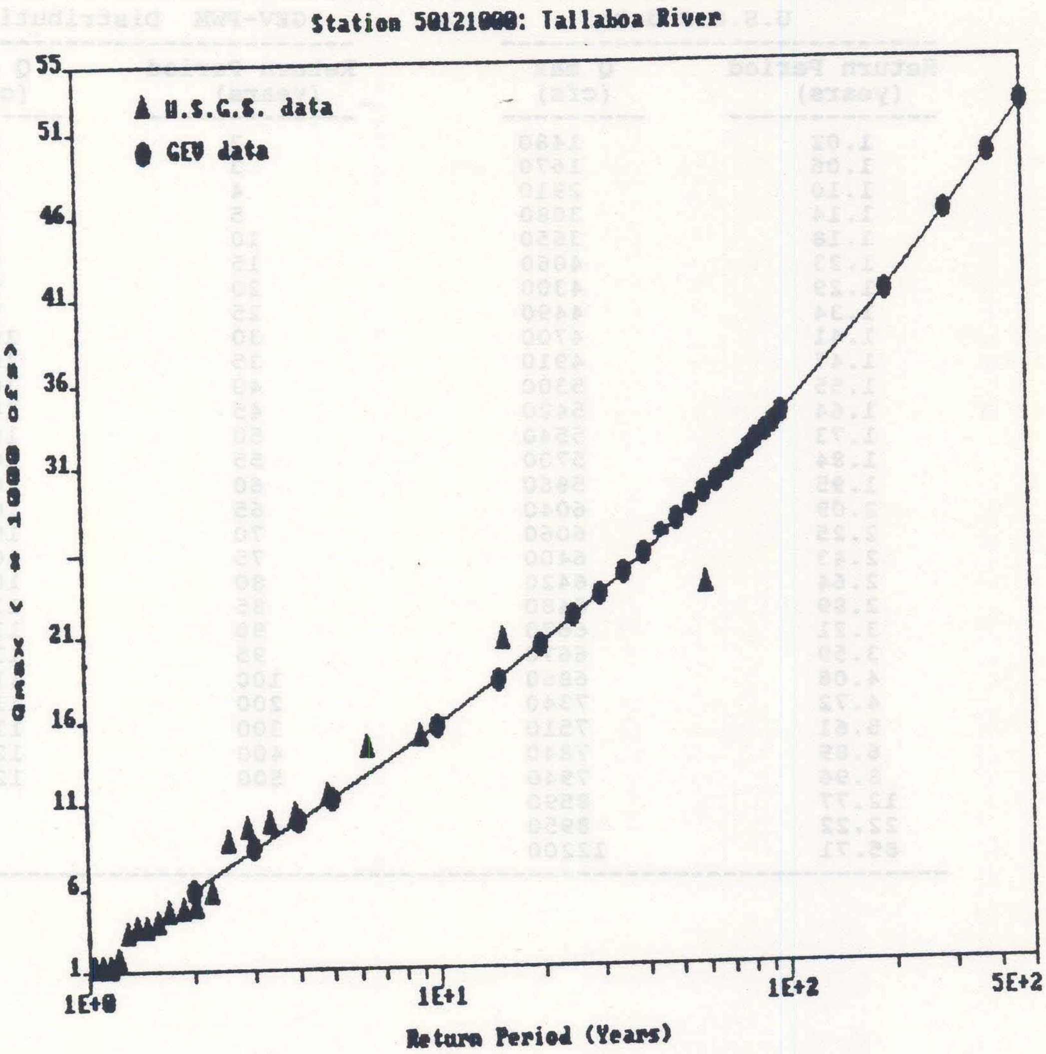

Figure B-26
Flood Frequency Curve for Rio Tallaboa basin

Table B-27

Return Period for Historical data and GEV-PWM Distribution
for Station 50028000 at Río Tanamá

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.02	1480	2	5741
1.06	1670	3	6733
1.10	2910	4	7309
1.14	3080	5	7709
1.18	3650	10	8768
1.23	4060	15	9293
1.29	4300	20	9633
1.34	4490	25	9881
1.41	4700	30	10073
1.47	4910	35	10229
1.55	5300	40	10360
1.64	5420	45	10472
1.73	5540	50	10570
1.84	5700	55	10656
1.95	5860	60	10733
2.09	6040	65	10803
2.25	6060	70	10866
2.43	6400	75	10924
2.64	6420	80	10977
2.89	6480	85	11027
3.21	6670	90	11073
3.59	6670	95	11116
4.08	6860	100	11156
4.72	7340	200	11658
5.61	7510	300	11918
6.89	7840	400	12089
8.96	7940	500	12214
12.77	8590		
22.22	8950		
85.71	12200		

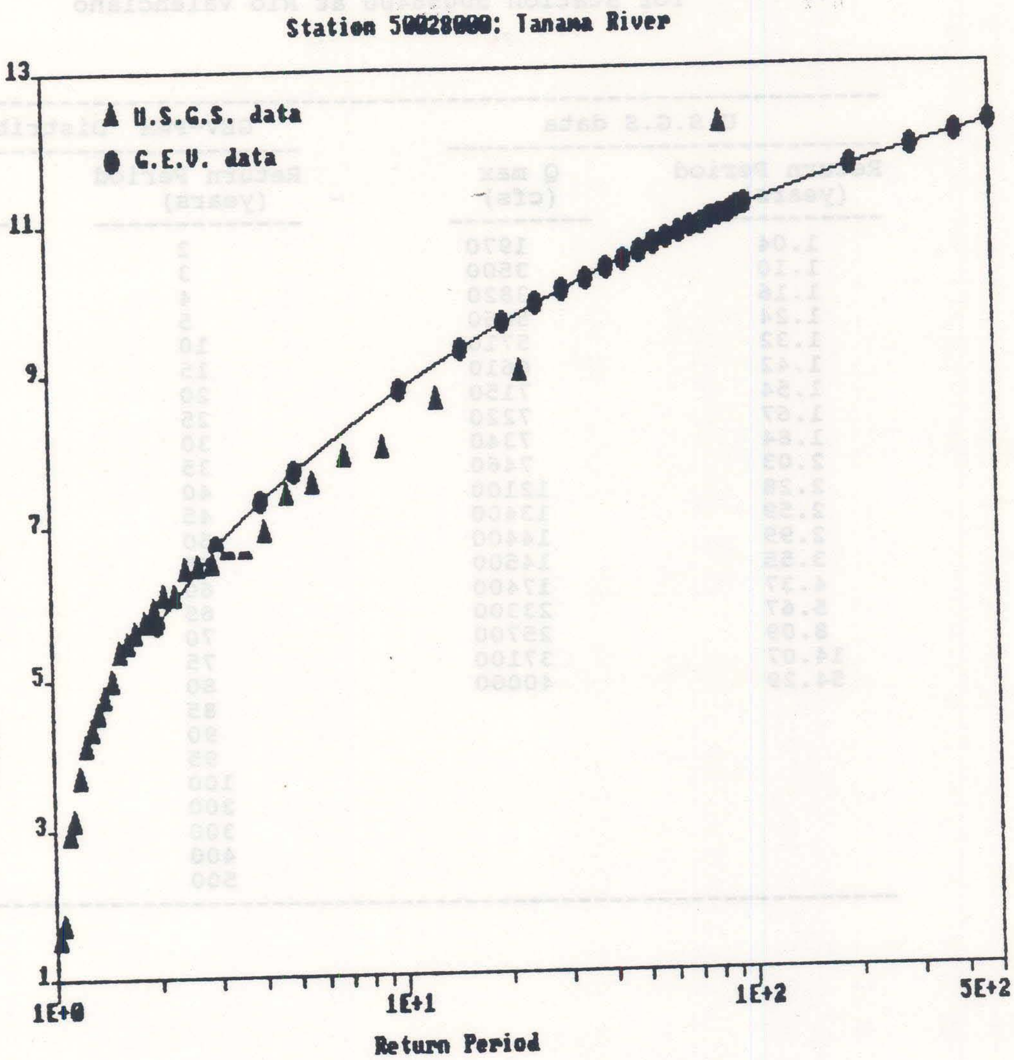

Figure B-27
Flood Frequency Curve for Río Tanamá basin

Table B-28

Return Period for Historical data and GEV-PWM Distribution
for Station 50056400 at Rio Valenciano

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.04	1970	2	10014
1.10	3500	3	13852
1.16	3820	4	16619
1.24	5060	5	18841
1.32	5710	10	26375
1.42	6610	15	31356
1.54	7150	20	35201
1.67	7220	25	38382
1.84	7340	30	41119
2.03	7460	35	43537
2.28	12100	40	45711
2.59	13400	45	47693
2.99	14400	50	49520
3.55	14500	55	51216
4.37	17400	60	52802
5.67	23300	65	54295
8.09	25700	70	55705
14.07	37100	75	57043
54.29	40000	80	58317
		85	59534
		90	60700
		95	61820
		100	62897
		200	79008
		300	89928
		400	98437
		500	105510

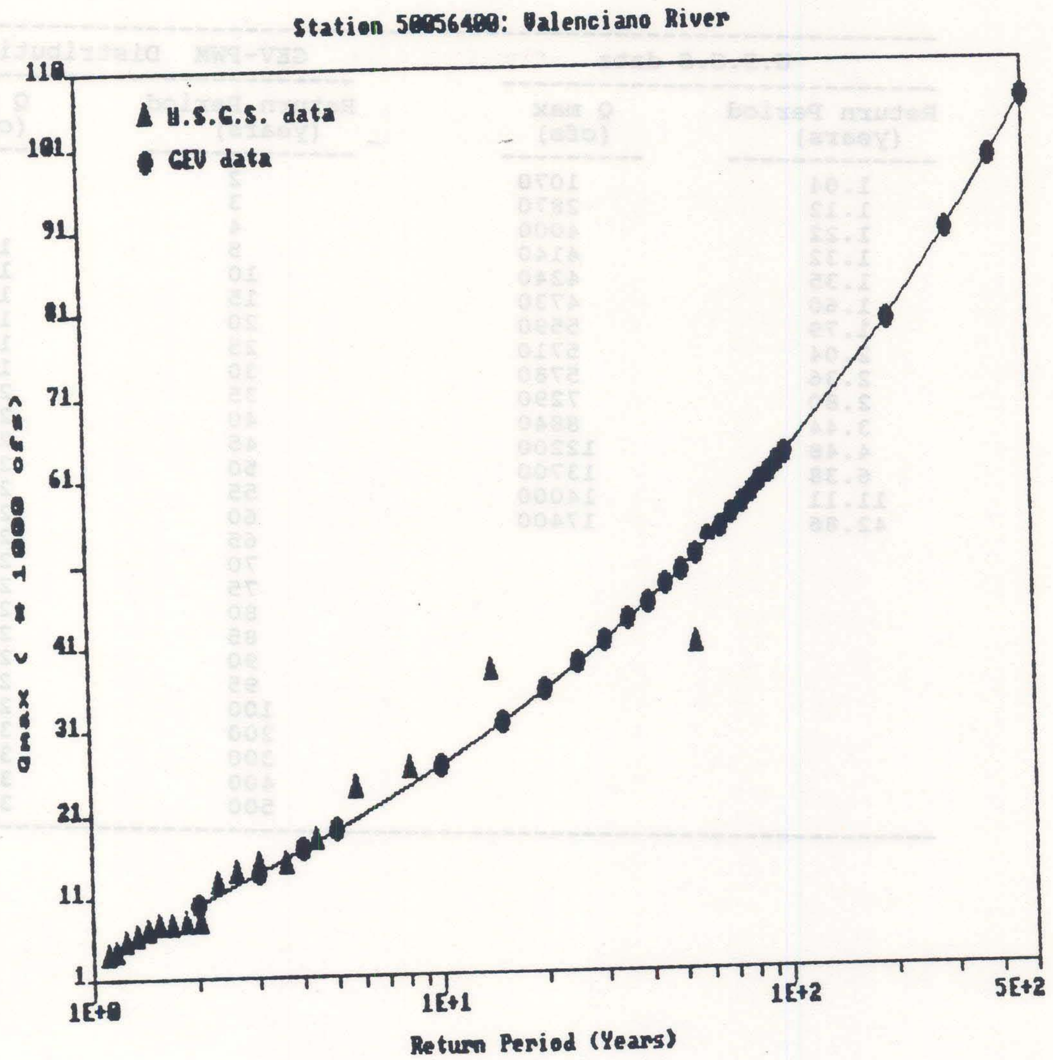

Figure B-28
Flood Frequency Curve for Río Valenciano basin

Table B-29

Return Period for Historical data and GEV-PWM Distribution
for Station 50064200 at Rio Grande

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.04	1070	2	6256
1.12	2870	3	8190
1.22	4000	4	9505
1.32	4140	5	10519
1.35	4240	10	13730
1.60	4730	15	15694
1.79	5590	20	17139
2.04	5710	25	18292
2.36	5780	30	19258
2.80	7290	35	20091
3.44	8840	40	20826
4.48	12200	45	21485
6.38	13700	50	22082
11.11	14000	55	22630
42.86	17400	60	23136
		65	23606
		70	24046
		75	24460
		80	24850
		85	25220
		90	25571
		95	25906
		100	26226
		200	30769
		300	33631
		400	35759
		500	37469

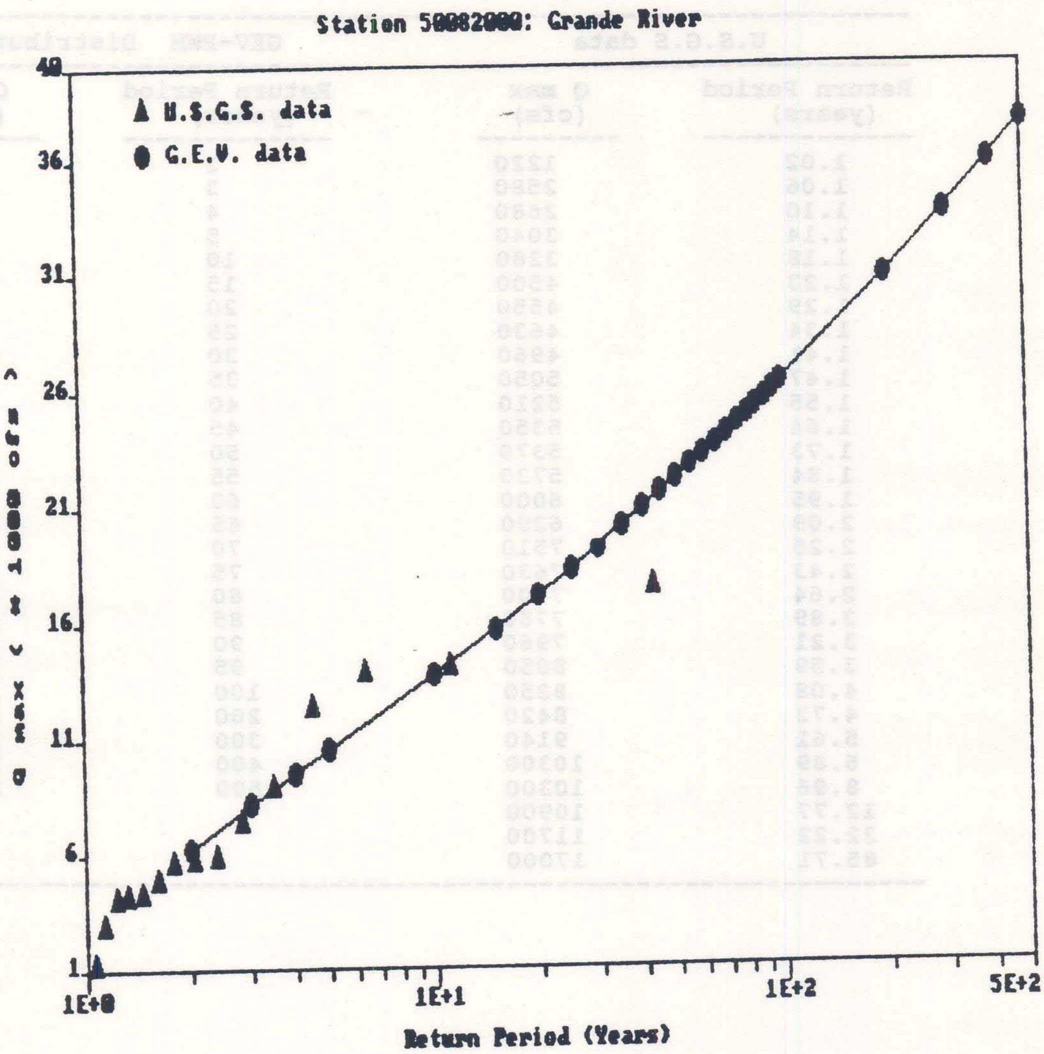

Figure B-29
Flood Frequency Curve for Río Grande basin

Table B-30

Return Period for Historical data and GEV-PWM Distribution
for Station 50141000 at Río Yahuecas

U.S.G.S data		GEV-PWM Distribution	
Return Period (years)	Q max (cfs)	Return Period (years)	Q max (cfs)
1.02	1220	2	6288
1.06	2580	3	7695
1.10	2680	4	8581
1.14	3040	5	9231
1.18	3280	10	11115
1.23	4500	15	12157
1.29	4550	20	12877
1.34	4630	25	13427
1.41	4960	30	13870
1.47	5050	35	14242
1.55	5210	40	14562
1.64	5350	45	14843
1.73	5370	50	15092
1.84	5730	55	15317
1.95	6000	60	15521
2.09	6290	65	15708
2.25	7510	70	15881
2.43	7630	75	16041
2.64	7700	80	16191
2.89	7780	85	16331
3.21	7960	90	16463
3.59	8050	95	16587
4.08	8250	100	16705
4.72	8420	200	18273
5.61	9140	300	19172
6.89	10300	400	19801
8.96	10300	500	20284
12.77	10900		
22.22	11700		
85.71	17000		

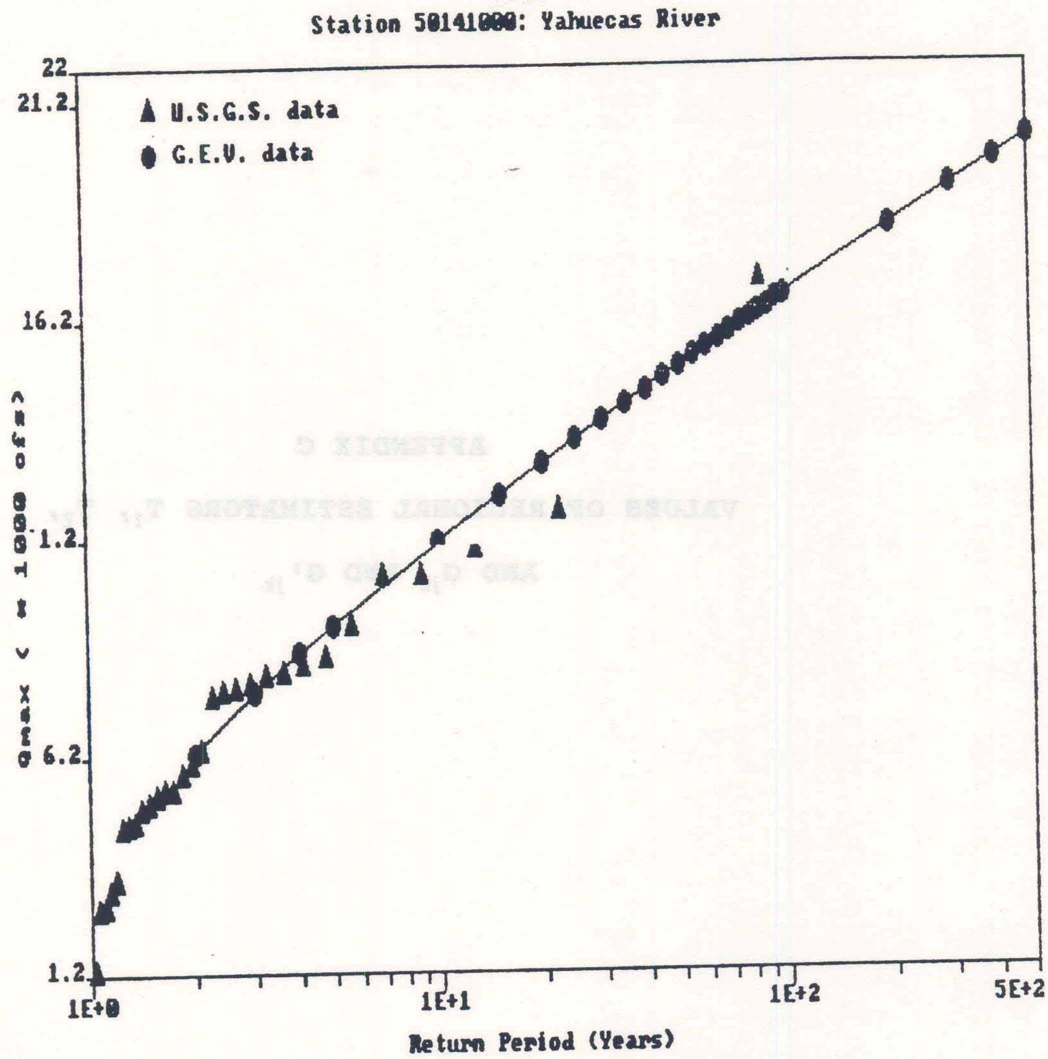

Figure B-30
Flood Frequency Curve for Río Yahuecas basin

APPENDIX C

VALUES OF REGIONAL ESTIMATORS T_1, T_2, AND G_{jk} AND G'_{jk}

Table C-1
Regional Estimators T_1, and T_2, for Clusters 1 and 2

Cluster # 1

Station	Bo	B1	B2	t(1)	t(2)	n	t(1)*n(i)	t(2)*n(i)
50029000	29200.550	24624.030	21887.300	0.843273	0.749551	18	15.1789	13.4919
50112500	2544.846	1916.991	1620.724	0.753284	0.636865	26	19.5854	16.5585
50115900	6040.833	4661.939	3949.509	0.771738	0.653802	24	18.5217	15.6912
50144000	23118.520	17332.540	14604.490	0.749725	0.631723	27	20.2426	17.0565
50124500	8844.348	6928.040	5861.710	0.783330	0.662763	23	18.0166	15.2436
50082000	8605.333	6438.586	5360.275	0.748209	0.622902	15	11.2231	9.3435
50111300	14781.720	10945.040	9025.069	0.740444	0.610556	29	21.4729	17.7061
50046000	34881.820	27043.230	22496.350	0.775282	0.644931	33	25.5843	21.2827
50106500	10820.000	7861.825	6469.774	0.726601	0.597946	20	14.5320	11.9589
50048000	19199.380	14582.320	12156.910	0.759520	0.633193	16	12.1523	10.1311
50011440	9278.182	6719.666	5520.606	0.724244	0.595009	22	15.9334	13.0902
50114000	4809.810	3385.258	2733.130	0.7038236	0.5682407	21	14.7803	11.9331
						274	207.2235	173.4874

T(1)= 0.756290056

T(2)= 0.633165514

Cluster # 2

Station	Bo	B1	B2	t(1)	t(2)	n	t(1)*n(i)	t(2)*n(i)
50057000	25358.670	18343.800	14482.430	0.72337389	0.57110368	30	21.7012	17.1331
50121000	7548.095	5475.990	4388.791	0.7254797	0.5814435	21	15.2351	12.2103
50059000	66053.840	48386.710	38682.910	0.7325344	0.5856270	13	9.5229	7.6132
50039500	10198.060	7628.284	6234.731	0.7480132	0.6113644	31	23.1884	18.9523
50038100	32541.940	23868.200	19240.900	0.7334597	0.5912647	31	22.7372	18.3292
						126	92.3849	74.2381

T(1)= 0.733213481

T(2)= 0.589191092

Table C-2
Regional Estimators T_1, and T_2, for Clusters 3 and 4

Cluster # 3

Station	Bo	B1	B2	t(1)	t(2)	n	t(1)*n(i)	t(2)*n(i)
50092000	5319.167	3511.724	2687.751	0.6602019	0.5052955	24	15.8448	12.1271
50147800	26467.270	16658.480	12467.170	0.6293993	0.4710410	22	13.8468	10.3629
50028000	5834.670	3541.030	2576.410	0.6068946	0.4415691	30	18.2068	13.2471
53136000	6528.571	4276.406	3340.378	0.6550294	0.5116553	21	13.7556	10.7448
						97	61.6541	46.4818

T(1)= 0.635609129

T(2)= 0.479194107

Cluster # 4

Station	Bo	B1	B2	t(1)	t(2)	n	t(1)*n(i)	t(2)*n(i)
50056400	13354.74	9559.947	7676.762	0.71584673	0.57483425	19	13.6011	10.9219
50064200	7437.333	5041.663	3917.219	0.67788587	0.52669673	15	10.1683	7.9005
50061800	6351.273	4298.623	3292.467	0.67681282	0.51839481	22	14.8899	11.4047
50063800	7916.957	5173.93	3938.531	0.65352508	0.49748040	23	15.0311	11.4420
50065700	13583.33	9081.714	6924.706	0.66859260	0.50979443	18	12.0347	9.1763
50062500	2300.556	1458.909	1076.243	0.63415496	0.46781864	18	11.4148	8.4207
50075000	1426.444	883.9027	659.7668	0.61965467	0.46252555	27	16.7307	12.4882
50071000	9061.667	6096.591	4698.581	0.67278912	0.51851177	30	20.1837	15.5554
50141000	6769.333	4285.569	3201.683	0.63308585	0.47296875	30	18.9926	14.1891
						202	133.0467	101.4987

T(1)= 0.658647114

T(2)= 0.502468704

Table C-3
Values of G_{jk} , and G'_{jk} , for Cluster 1 for Stations 50029000, 50112500, and 50115900

Station 50029000			Station 50112500			Station 50115900		
Q(x)	Gjk(x)	G'jk	Q(x)	Gjk(x)	G'jk	Q(x)	Gjk(x)	G'jk
2000	0.013594	0.972811	643	0.145743	0.708513	1000	0.066354	0.867291
3410	0.034171	0.931656	691	0.165245	0.669508	1360	0.118406	0.763186
3430	0.034546	0.930907	705	0.171022	0.657954	1500	0.141317	0.717365
3680	0.039422	0.921154	786	0.204980	0.590038	1590	0.156579	0.686841
3780	0.041472	0.917055	881	0.245312	0.509374	1600	0.158296	0.683407
5200	0.076292	0.847415	895	0.251250	0.497498	1660	0.168672	0.662654
6790	0.125591	0.748816	945	0.272369	0.455260	1710	0.177407	0.645185
7770	0.159570	0.680858	1080	0.328125	0.343748	1760	0.186206	0.627586
8070	0.170321	0.659356	1100	0.336169	0.327661	2050	0.237929	0.524141
8680	0.192516	0.614967	1120	0.344146	0.311707	2060	0.239718	0.520562
10200	0.248679	0.502640	1120	0.344146	0.311707	2500	0.317243	0.365513
15800	0.439823	0.120354	1200	0.375338	0.249322	2540	0.324094	0.351810
21000	0.574960	0.149921	1230	0.386722	0.226555	2930	0.388351	0.223297
23700	0.629662	0.259325	1310	0.416186	0.167627	3260	0.438567	0.122864
26100	0.670950	0.341901	1460	0.467787	0.064425	5400	0.670995	0.341991
76000	0.939712	0.879424	1510	0.483915	0.032168	5630	0.688081	0.376162
105000	0.967138	0.934276	1830	0.574919	0.149838	5800	0.699959	0.399919
195000	0.990324	0.980649	2300	0.675541	0.351083	5880	0.705341	0.410682
			2800	0.750868	0.501736	7550	0.793013	0.586026
			2800	0.750868	0.501736	8200	0.817427	0.634854
			3410	0.813748	0.627497	11900	0.900941	0.801882
			3480	0.819496	0.638992	14500	0.930154	0.860308
			3960	0.852913	0.705826	17600	0.951060	0.902120
			4190	0.865872	0.731744	35000	0.987124	0.974249
			5720	0.921420	0.842841			
			19000	0.992282	0.984565			
	Sum =	12.3934859		Sum =	12.15893		Sum =	14.04990

Table C-4
Values of G_{jk}, and G'_{jk}, for Cluster 1 for Stations 50144000, 50124500, and 50082000

Station 50144000

Q(x)	Gjk(x)	G'jk
5950	0.150555	0.698889
6980	0.197521	0.604957
7450	0.219432	0.561134
7830	0.237207	0.525585
8000	0.245154	0.509691
8020	0.246088	0.507823
8090	0.249357	0.501285
8800	0.282307	0.435385
8960	0.289660	0.420679
9300	0.305166	0.389666
9320	0.306073	0.387852
10700	0.366816	0.266366
11600	0.404158	0.191683
12500	0.439476	0.121046
14400	0.507150	0.014301
15000	0.526605	0.053211
15600	0.545175	0.090351
16000	0.557077	0.114154
19300	0.641965	0.283931
20300	0.663528	0.327056
28600	0.789829	0.579659
29700	0.801396	0.602793
32300	0.825421	0.650843
38700	0.869494	0.738989
53600	0.925657	0.851314
77200	0.962148	0.924297
140000	0.988217	0.976434
		Sum = 12.32938

Station 50124500

Q(x)	Gjk(x)	G'jk
1120	0.039783	0.920433
1380	0.059241	0.881516
1530	0.072186	0.855626
1800	0.098213	0.803572
2210	0.142905	0.714189
2280	0.150989	0.698021
2400	0.165068	0.669863
2400	0.165068	0.669863
2590	0.187798	0.624402
3240	0.267008	0.465983
3250	0.268220	0.463558
3390	0.285125	0.429749
4530	0.413823	0.172353
5000	0.460749	0.078501
5400	0.497600	0.004798
5600	0.514976	0.029952
7800	0.665357	0.330714
16000	0.885396	0.770792
18500	0.910769	0.821539
23000	0.939620	0.879240
23000	0.939620	0.879240
28000	0.958124	0.916248
39000	0.977878	0.955756
		Sum = 14.03592

Station 50082000

Q(x)	Gjk(x)	G'jk
1500	0.073241	0.853517
1740	0.097024	0.805951
2500	0.185318	0.629363
3500	0.309823	0.380353
3600	0.321898	0.356202
4500	0.423621	0.152757
4700	0.444281	0.111437
5500	0.519445	0.038891
6000	0.560551	0.121103
6400	0.590423	0.180846
9200	0.740625	0.481251
9940	0.767588	0.535177
13000	0.845725	0.691451
17000	0.901425	0.802851
40000	0.980054	0.960109
		Sum = 7.101263

Table C-5

Values of G_{jk} , and G'_{jk} , for Cluster 1 for Stations 50111300, 50046000, and 50106500

Station 50111300

Q(x)	Gjk(x)	G'jk
1720	0.033943	0.932112
3000	0.097702	0.804594
3500	0.129775	0.740448
4000	0.164278	0.671443
4130	0.173520	0.652958
4940	0.232348	0.535302
5000	0.236738	0.526523
5000	0.236738	0.526523
5200	0.251352	0.497294
5740	0.290454	0.419090
6360	0.334143	0.331712
7500	0.409385	0.181228
9780	0.535871	0.071743
10000	0.546380	0.092761
10400	0.564771	0.129543
10400	0.564771	0.129543
11200	0.598906	0.197812
12300	0.640551	0.281102
14300	0.702933	0.405867
15600	0.735924	0.471848
16000	0.745071	0.490142
16400	0.753791	0.507582
19600	0.810724	0.621448
22000	0.842000	0.684000
22300	0.845385	0.690771
22500	0.847586	0.695172
24400	0.866437	0.732874
57400	0.971662	0.943325
78000	0.984483	0.968966
		Sum = 14.93374

Station 50046000

Q(x)	Gjk(x)	G'jk
200	0.002097	0.995805
1270	0.005916	0.988167
3500	0.025875	0.948248
3710	0.028758	0.942483
4190	0.036021	0.927957
4340	0.038481	0.923036
6110	0.073926	0.852146
7170	0.100055	0.799889
8140	0.126405	0.747188
9470	0.165202	0.669594
14200	0.310203	0.379592
14300	0.313193	0.373612
14800	0.328025	0.343948
16700	0.382329	0.235340
17900	0.414714	0.170571
18200	0.422561	0.154876
18400	0.427737	0.144525
18800	0.437951	0.124096
19000	0.442991	0.114017
19200	0.447984	0.104031
21600	0.504347	0.008694
29200	0.643148	0.286296
37400	0.741666	0.483332
53000	0.847145	0.694290
55600	0.858550	0.717101
56600	0.862616	0.725232
70700	0.905718	0.811436
74400	0.913776	0.827552
78500	0.921593	0.843186
82000	0.927480	0.854961
95500	0.945071	0.890142
120000	0.964234	0.928469
157000	0.978750	0.957500
		Sum = 19.96732

Station 50106500

Q(x)	Gjk(x)	G'jk
1790	0.066272	0.867454
2770	0.149157	0.701684
3500	0.220754	0.558491
4600	0.328897	0.342204
5120	0.376949	0.246101
5200	0.384093	0.231812
5400	0.401645	0.196709
5700	0.427113	0.145772
6790	0.510713	0.021427
7100	0.531962	0.063924
7340	0.547675	0.095351
8080	0.592288	0.184576
8510	0.615722	0.231445
9400	0.659027	0.318054
10100	0.688715	0.377430
11600	0.741631	0.483263
16300	0.845032	0.690064
21100	0.899194	0.798388
22000	0.906236	0.812473
54000	0.982667	0.965335
		Sum = 8.331965

Table C-6
Values of G_{jk}, and G'_{jk}, for Cluster 1 for Stations 50048000, 50011440, and 50114000

Station 50048000

Q(x)	Gjk(x)	G'jk
3680	0.087732	0.824535
3680	0.087732	0.824535
3860	0.095990	0.808019
6010	0.209476	0.581047
7300	0.281853	0.436293
7700	0.303886	0.392226
8450	0.344159	0.311681
9110	0.378207	0.243585
11100	0.471474	0.057050
12000	0.508791	0.017582
13300	0.557513	0.115027
27000	0.827192	0.654385
29000	0.845688	0.691377
35000	0.886873	0.773747
58000	0.954264	0.908529
72000	0.969688	0.939377
	Sum =	8.579002

Station 50011440

Q(x)	Gjk(x)	G'jk
2190	0.129043	0.741913
2490	0.161946	0.676107
3000	0.220607	0.558785
3370	0.263660	0.472678
3560	0.285550	0.428899
4200	0.356759	0.286480
4890	0.427330	0.145339
4900	0.428298	0.143402
4960	0.434076	0.131847
5100	0.447330	0.105338
5680	0.498875	0.002248
6180	0.539044	0.078088
6900	0.590395	0.180791
7200	0.609703	0.219407
8800	0.695107	0.390215
10000	0.743540	0.487080
10200	0.750576	0.501153
11000	0.776219	0.552438
12500	0.815295	0.630590
20000	0.915369	0.830738
20000	0.915369	0.830738
47000	0.983171	0.966343
	Sum =	9.360628

Station 50114000

Q(x)	Gjk(x)	G'jk
821	0.070363	0.859272
865	0.077737	0.844524
1420	0.190349	0.619300
1750	0.264328	0.471342
3150	0.531035	0.062070
3190	0.537014	0.074029
3200	0.538495	0.076991
3280	0.550143	0.100287
3280	0.550143	0.100287
3360	0.561442	0.122885
3580	0.590784	0.181568
3620	0.595856	0.191712
3770	0.614189	0.228379
3950	0.634823	0.269647
4060	0.646739	0.293479
4400	0.680514	0.361028
4600	0.698410	0.396821
8270	0.875175	0.750350
8900	0.889706	0.779412
9140	0.894607	0.789215
22400	0.980129	0.960259
	Sum =	8.532866

Table C-7
Values of G_{jk} , and G'_{jk} , for Cluster 2 for Stations 50121000, 50059000, and 50039500

Station 50121000

Q(x)	Gjk(x)	G'jk
1050	0.102219	0.795560
1060	0.102870	0.794259
1070	0.103522	0.792954
1400	0.126123	0.747752
3100	0.265966	0.468067
3450	0.297238	0.405524
3450	0.297238	0.405524
3740	0.323229	0.353541
4360	0.378306	0.243387
4480	0.388807	0.222384
4700	0.407872	0.184256
5520	0.476206	0.047586
8480	0.674886	0.349773
9320	0.717106	0.434214
9640	0.731709	0.463419
10090	0.750952	0.501906
11100	0.789090	0.578181
14100	0.869656	0.739314
14600	0.879422	0.758846
20300	0.947618	0.895237
23500	0.965758	0.931517
		Sum = 11.1132015

Station 50059000

Q(x)	Gjk(x)	G'jk
2000	0.056067	0.887866
7900	0.092874	0.814251
15800	0.156457	0.687085
23600	0.230454	0.539090
30600	0.301425	0.397149
51800	0.507614	0.015229
59500	0.571876	0.143752
62000	0.591211	0.182422
79000	0.703058	0.406117
85600	0.737927	0.475854
98000	0.792468	0.584937
161000	0.931582	0.863164
182000	0.951063	0.902127
		Sum = 6.899048

Station 50039500

Q(x)	Gjk(x)	G'jk
1570	0.109620	0.780759
2050	0.134628	0.730743
2160	0.140665	0.718669
2210	0.143443	0.713112
2220	0.144002	0.711995
2400	0.154193	0.691612
2500	0.159966	0.680066
2500	0.159966	0.680066
2500	0.159966	0.680066
2500	0.159966	0.680066
2600	0.165815	0.668369
3900	0.247071	0.505856
4600	0.293178	0.413643
4640	0.295830	0.408338
6000	0.385393	0.229212
6050	0.388625	0.222749
6740	0.432454	0.135091
6800	0.436188	0.127622
9100	0.567453	0.134907
9700	0.597446	0.194893
10100	0.616425	0.232851
12000	0.695805	0.391611
15000	0.789168	0.578336
16600	0.826099	0.652198
20000	0.883151	0.766302
20200	0.885788	0.771576
22200	0.908762	0.817525
25000	0.932612	0.865225
28000	0.950573	0.901146
30300	0.960638	0.921277
34000	0.972242	0.944485
		Sum = 17.95037

Table C-8
Values of G_{jk} , and G'_{jk} , for Cluster 2 for Stations 50038100, and 50057000

Station 50038100			Station 50057000		
Q(x)	Gjk(x)	G'jk	Q(x)	Gjk(x)	G'jk
5000	0.109466	0.781066	900	0.057932	0.884134
5470	0.116893	0.766212	1620	0.068719	0.862560
6000	0.125530	0.748938	4320	0.118078	0.763842
6320	0.130874	0.738251	4450	0.120780	0.758439
6610	0.135797	0.728405	5170	0.136216	0.727566
6960	0.141837	0.716325	5350	0.140195	0.719608
7100	0.144282	0.711435	6140	0.158181	0.683636
11300	0.223880	0.552239	6440	0.165218	0.669562
11600	0.229915	0.540168	7340	0.186934	0.626131
12100	0.240042	0.519914	7840	0.199343	0.601313
15800	0.316493	0.367013	8700	0.221163	0.557672
16500	0.331030	0.337939	8790	0.223477	0.553045
17000	0.341388	0.317222	16200	0.418325	0.163349
17300	0.347589	0.304820	19300	0.494047	0.011905
18200	0.366106	0.267786	25300	0.619869	0.239739
21010	0.422692	0.154614	27200	0.653609	0.307219
21250	0.427414	0.145171	28600	0.676649	0.353299
21440	0.431138	0.137723	29400	0.689144	0.378289
30190	0.585719	0.171438	29400	0.689144	0.378289
31030	0.598620	0.197240	32300	0.730561	0.461122
34500	0.648206	0.296413	34800	0.761743	0.523486
43000	0.746125	0.492251	36500	0.780779	0.561559
44200	0.757533	0.515067	38200	0.798213	0.596426
53000	0.826291	0.652582	41600	0.828766	0.657533
62000	0.875247	0.750495	42100	0.832818	0.665636
68110	0.899645	0.799290	48600	0.876887	0.753775
70800	0.908635	0.817271	50600	0.887708	0.775416
76000	0.923534	0.847068	54900	0.907523	0.815047
81760	0.936885	0.873771	64100	0.937907	0.875814
90000	0.951579	0.903159	74600	0.959473	0.918947
97250	0.961307	0.922614			
	Sum =	17.07391		Sum =	17.84437

Table C-9
Values of G_{jk} , and G'_{jk} , for Cluster 3 for Stations 50147800, and 50028000

Station 50147800

Q(x)	Gjk(x)	G'jk
4850	0.003544	0.992911
12000	0.088335	0.823328
15100	0.180074	0.639850
15700	0.200928	0.598143
17130	0.253311	0.493377
18600	0.309640	0.380719
18700	0.313519	0.372961
20100	0.367917	0.264165
22500	0.459020	0.081959
24600	0.533458	0.066916
24700	0.536841	0.073683
25800	0.572987	0.145974
26700	0.601042	0.202085
26800	0.604073	0.208146
30000	0.691776	0.383553
31400	0.724593	0.449186
32800	0.754224	0.508449
33000	0.758209	0.516418
33800	0.773549	0.547098
34700	0.789703	0.579407
44300	0.904829	0.809659
69000	0.985923	0.971846
	Sum =	10.10984

Station 50028000

Q(x)	Gjk(x)	G'jk
1480	0.011095	0.977809
1670	0.017239	0.965521
2910	0.120210	0.759578
3080	0.143249	0.713500
3650	0.231944	0.536111
4060	0.302551	0.394897
4300	0.344838	0.310323
4490	0.378271	0.243456
4700	0.414815	0.170369
4910	0.450619	0.098761
5300	0.514290	0.028581
5420	0.532991	0.065983
5540	0.551227	0.102454
5700	0.574785	0.149570
5860	0.597449	0.194898
6040	0.621850	0.243701
6060	0.624489	0.248979
6400	0.667131	0.334263
6420	0.669509	0.339019
6480	0.676558	0.353117
6670	0.698035	0.396071
6670	0.698035	0.396071
6860	0.718253	0.436507
7340	0.763990	0.527981
7510	0.778459	0.556919
7840	0.804165	0.608331
7940	0.811369	0.622738
8590	0.852224	0.704448
8950	0.870906	0.741813
12200	0.960720	0.921441
	Sum =	13.14322

Table C-10
Values of G_{jk} , and G'_{jk} , for Cluster 3 for Stations 50136000, and 50092000

Station 50136000			Station 50092000		
Q(x)	Gjk(x)	G'jk	Q(x)	Gjk(x)	G'jk
2250	0.033943	0.93211267	1600	0.020678	0.958643
2660	0.061475	0.87704853	1800	0.031771	0.936457
3550	0.156567	0.68686541	1890	0.037891	0.924217
3800	0.190592	0.61881418	2400	0.086927	0.826145
3900	0.204864	0.59027066	2460	0.094309	0.811380
4190	0.247920	0.50415981	2630	0.116969	0.766060
4650	0.319398	0.36120398	3230	0.214226	0.571546
4850	0.350924	0.29815166	3270	0.221447	0.557105
5120	0.393274	0.21345066	3700	0.302302	0.395394
5420	0.439450	0.12109842	3820	0.325470	0.349059
5590	0.464981	0.07003706	3960	0.352558	0.294882
5590	0.464981	0.07003706	4500	0.455006	0.089987
6240	0.556718	0.11343781	4600	0.473253	0.053493
6400	0.577633	0.15526793	4730	0.496502	0.006995
6500	0.590342	0.18068575	6230	0.715536	0.431073
6500	0.590342	0.18068575	6360	0.730031	0.460063
6500	0.590342	0.18068575	6530	0.747973	0.495946
7800	0.729493	0.45898793	6830	0.776961	0.553923
8000	0.746745	0.49349045	8130	0.869343	0.738687
8890	0.811730	0.62346065	8290	0.877664	0.755329
28700	0.999303	0.99860635	8370	0.881620	0.763240
			8440	0.884974	0.769948
	Sum =	8.72855857	9190	0.915319	0.830638
			14700	0.989527	0.979054
				Sum =	14.31927

Table C-11
Values of G_{jk} , and G'_{jk} , for Cluster 4 for Stations 50075000, 50071000, and 50141000

Station 50075000

Q(x)	Gjk(x)	G'jk
641	0.149737	0.700524
752	0.205271	0.589456
780	0.220283	0.559432
840	0.253490	0.493018
859	0.264250	0.471498
859	0.264250	0.471498
893	0.283734	0.432530
1000	0.346241	0.307517
1040	0.369764	0.260470
1180	0.450945	0.098108
1220	0.473478	0.053042
1230	0.479049	0.041901
1260	0.495594	0.008810
1280	0.506479	0.012959
1300	0.517240	0.034481
1330	0.533138	0.066277
1350	0.543567	0.087135
1400	0.569017	0.138035
1500	0.617076	0.234153
1650	0.681662	0.363325
1790	0.733762	0.467524
1900	0.769401	0.538802
2200	0.845713	0.691426
2460	0.891806	0.783613
2460	0.891806	0.783613
2480	0.894733	0.789466
2680	0.920008	0.840017

	Sum =	10.31864

Station 50071000

Q(x)	Gjk(x)	G'jk
1430	0.022358	0.955283
2300	0.048940	0.902118
2600	0.061556	0.876887
2830	0.072488	0.855023
2940	0.078103	0.843792
4670	0.196400	0.607198
5040	0.227570	0.544859
5220	0.243233	0.513533
5320	0.252053	0.495892
5530	0.270813	0.458372
5590	0.276225	0.447549
6000	0.313661	0.372676
6080	0.321034	0.357931
7340	0.436923	0.126153
7850	0.482217	0.035565
7900	0.486570	0.026859
7950	0.490905	0.018188
9200	0.592677	0.185354
10500	0.682804	0.365609
10600	0.689034	0.378069
10700	0.695164	0.390329
10800	0.701195	0.402390
11200	0.724333	0.448666
14500	0.862090	0.724181
14560	0.863854	0.727709
15100	0.878777	0.757554
15600	0.891163	0.782327
19400	0.952077	0.904155
19600	0.954093	0.908186
23500	0.979950	0.959900

	Sum =	16.37232

Station 50141000

Q(x)	Gjk(x)	G'jk
1220	0.027325	0.945348
2580	0.107646	0.784706
2680	0.116187	0.767625
3040	0.149636	0.700727
3240	0.169890	0.660218
4500	0.316022	0.367954
4550	0.322195	0.355608
4630	0.332090	0.335818
4960	0.372988	0.254023
5050	0.384118	0.231763
5210	0.403826	0.192346
5350	0.420954	0.158090
5370	0.423390	0.153218
5730	0.466642	0.066714
6000	0.498164	0.003670
6290	0.530946	0.061893
7510	0.653938	0.307876
7630	0.664617	0.329234
7700	0.670726	0.341452
7780	0.677599	0.355199
7960	0.692644	0.385288
8050	0.699949	0.399898
8250	0.715667	0.431335
8420	0.728478	0.456956
9140	0.777346	0.554692
10300	0.839543	0.679087
10300	0.839543	0.679087
10900	0.864886	0.729772
11700	0.892718	0.785436
17000	0.976591	0.953183

	Sum =	13.42823

Table C-12
Values of G_{jk} , and G'_{jk} , for Cluster 4 for Stations 50063800, 50065700, and 50062500

Station 50063800

Q(x)	Gjk(x)	G'jk
1400	0.026518	0.946963
2580	0.078785	0.842428
3400	0.136737	0.726524
3710	0.162806	0.674387
4100	0.198277	0.603444
4710	0.258364	0.483271
5140	0.302929	0.394141
5400	0.330340	0.339318
5960	0.389616	0.220766
6800	0.476375	0.047248
7400	0.534869	0.069739
7460	0.540511	0.081023
7880	0.578832	0.157665
7960	0.585889	0.171779
8440	0.626524	0.253049
8870	0.660375	0.320750
9400	0.698743	0.397487
9880	0.730343	0.460687
10200	0.749799	0.499599
11800	0.829479	0.658959
12400	0.852711	0.705422
17200	0.954942	0.909884
20000	0.977216	0.954432
		Sum = 10.91897

Station 50065700

Q(x)	Gjk(x)	G'jk
2810	0.034160	0.931678
3100	0.040440	0.919118
3500	0.050313	0.899372
7610	0.230744	0.538510
7780	0.240617	0.518764
9000	0.314034	0.371930
9600	0.351027	0.297944
10900	0.430737	0.138524
11100	0.442777	0.114445
11600	0.472454	0.055091
12200	0.507114	0.014229
16200	0.701631	0.403262
18400	0.779018	0.558037
20200	0.828374	0.656749
22200	0.870972	0.741945
26000	0.925355	0.850710
26100	0.926424	0.852848
26200	0.927477	0.854954
		Sum = 9.718119

Station 50062500

Q(x)	Gjk(x)	G'jk
240	0.013213	0.973573
580	0.048341	0.903317
1110	0.172518	0.654963
1200	0.201075	0.597849
1580	0.334287	0.331425
2170	0.541230	0.082460
2190	0.547646	0.095292
2300	0.581933	0.163867
2350	0.596937	0.193875
2380	0.605759	0.211519
2390	0.608669	0.217339
2710	0.693606	0.387213
2720	0.696004	0.392008
2740	0.700752	0.401504
2820	0.719133	0.438266
3800	0.875279	0.750558
3920	0.887353	0.774707
4210	0.911982	0.823965
		Sum = 8.393708

Table C-13
Values of G_{jk} , and G'_{jk} , for Cluster 4 for Stations 50056400, 50064200, and 50061800

Station 50056400

Q(x)	Gjk(x)	G'jk
1970	0.020315	0.959368
3500	0.051916	0.896167
3820	0.061196	0.877607
5060	0.106319	0.787361
5710	0.135529	0.728940
6610	0.181434	0.637131
7150	0.211505	0.576990
7220	0.215518	0.568962
7340	0.222455	0.555090
7460	0.229458	0.541084
12100	0.513178	0.026357
13400	0.584466	0.168932
14400	0.634341	0.268684
14500	0.639078	0.278157
17400	0.756490	0.512981
23300	0.895984	0.791968
25700	0.926846	0.853692
37100	0.985855	0.971711
40000	0.990561	0.981122
		Sum = 11.9823059

Station 50064200

Q(x)	Gjk(x)	G'jk
1070	0.019626	0.960746
2870	0.110296	0.779407
4000	0.213368	0.573263
4140	0.227931	0.544137
4240	0.238506	0.522986
4730	0.291939	0.416120
5590	0.388614	0.222771
5710	0.402059	0.195881
5780	0.409872	0.180254
7290	0.568110	0.136220
8840	0.699437	0.398875
12200	0.872476	0.744952
13700	0.914038	0.828077
14000	0.920570	0.841141
17400	0.967398	0.934797
		Sum = 8.279634

Station 50061800

Q(x)	Gjk(x)	G'jk
258	0.006385	0.987228
2010	0.074387	0.851224
2070	0.078803	0.842393
2300	0.097115	0.805769
2510	0.115702	0.768595
3320	0.201914	0.596171
3440	0.216280	0.567438
3590	0.234671	0.530656
3630	0.239648	0.520702
5260	0.451714	0.096571
6080	0.551563	0.103127
6220	0.567496	0.134993
6910	0.640469	0.280939
6960	0.645384	0.290769
8200	0.751057	0.502115
8870	0.795908	0.591816
9400	0.826015	0.652030
10200	0.863645	0.727290
10400	0.871752	0.743504
10600	0.879390	0.758780
12500	0.932844	0.865689
15000	0.968763	0.937527
		Sum = 13.15533