FREQUENCY ANALYSIS OF LOW FLOWS: COMPARISON OF A PHYSICALLY BASED APPROACH AND HYPOTHETICAL DISTRIBUTION METHODS

by

Peter Paul Mattejat

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Civil Engineering

APPROVED:

G. V. Loganathan, Chairman

C. Y. Kuo

T. A. Dillaha

August 29, 1985

Blacksburg, Virginia

FREQUENCY ANALYSIS OF LOW FLOWS: COMPARISON OF A PHYSICALLY BASED APPROACH AND HYPOTHETICAL DISTRIBUTION METHODS

by

Peter Paul Mattejat

G. V. Loganathan, Chairman

Civil Engineering

(ABSTRACT)

Several different approaches are applied in low flow frequency analysis. Each method's theory and application is explained. The methods are (1) physically based recession model dealing with time series, (2) log-Pearson type III and mixed log-Pearson type III using annual minimum series, (3) Double Bounded pdf using annual minimum series, (4) Partial Duration Series applying truncated and censored flows.

Each method has a computer program for application. One day low flow analysis was applied to 15 stations, 10 perennial streams and 5 intermittent streams. The physically based method uses the exponential baseflow recession with duration, initial recession flow, and recharge due to incoming storm as random variables, and shows promise as an alternative to black box methods, and is appealing because it contains the effect of drought length. Log-Pearson is modified to handle zero flows by adding a point mass probability for zero flows. Another approach to zero flows is the Double Bounded probability density function which also includes a

point mass probability for zero flows. Maximum likelihood estimation is used to estimate distribution parameters. Partial Duration Series is applied due to drawbacks of using only one low flow per year in annual minimum series. Two approaches were used in Partial Duration Series (i) truncation, and (ii) censorship which represent different low flow populations. The parameters are estimated by maximum likelihood estimation.

ACKNOWLEDGMENTS

The author would like to extend his sincere appreciation to Dr. G. V. Loganathan for his help and knowledge in completing this study. With his patience and encouragement he helped the author complete the goals of this thesis. Also appreciated is the time and effort of Dr. T. A. Dillaha and C. Y. Kuo. Their insights and comments were very helpful in the completion of this study.

Special thanks are expressed to the author's parents, Guenther and Hedi Mattejat. Their love and attention were invaluable in motivating the author to accel in his studies.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

INTRODUCTION

Adequate supply of streamflow is necessary to satisfy the demands of agriculture, industry, human and wildlife needs. Streamflow becomes critical when dry periods occur leading to low flows, if severe, become droughts. Accurate prediction of low flows is required for effective management of streamflow to supply for all demands and water quality needs. Critical flows for specified return periods can be assessed by frequency analysis. A brief review of existing methods for the frequency analysis of low flows is presented.

Frequency analysis is the procedure for estimating the probability of occurrence of an event. A number of probability distributions are used in the frequency analysis of low flows. These distributions may be grouped into the following categories:

1. Transformation methods
2. Hypothetical distribution based models } Black box methods
3. Plotting position based methods
4. Physically based probability methods

The first three categories are collectively called black box methods. The transformation methods attempt to find a unique probability distribution in the transformed range regardless of the distribution of the original sample. Power transformation proposed by Box and Cox (1964) transforms the original sample to a near normal distribution. This transformation has been widely used in hydrologic time series analysis. Kumar and Devi (1982) applied it to the frequency analysis of low flows. Bethlahmy (1977) suggested the SMEMAX (Small, MEdian and MAXimum) transformation to transform the original data to a near normal distribution. Prakash (1981) applied it to the analysis of low flows. Rasheed et al. (1982) suggested a modified SMEMAX transformation in which the range can be chosen arbitrarily. In low flow analysis transformations to the normal distribution are widely used. There are also other transformations (Kruskal and Tunar,, 1978).

The hypothetical distribution method involves the fitting of a probability distribution directly to the sample data. The commonly used probability distributions for low flow analysis are the normal, log normal, gamma, log Pearson type III, Gumbel type III, and Weibull. McMahon and Arenas (1982), Matalas (1963), and McCormick (1984) provide extensive reviews of the application of the various probability distributions to low flow data. Kite (1977) and Kottegoda (1980) provide the fitting techniques for various probability distributions to hydrologic data. Johnson and Kotz (1970) provide detailed accounts of continuous distributions.

The plotting position based methods are probably the most straight forward of the various methods. For each observed discharge, a plotting position formula is used to estimate the exceedance probability. A curve is then fit through these probability points. Cunnane (1978) provided an excellent review of the various plotting positions available in the hydrologic literature. Boughton (1980, 1983) suggested a general curve fitting procedure for the plotting position data.

DRAWBACKS OF THE BLACK BOX METHODS

This section analyzes the drawbacks of the black box methods.

1. The black box techniques do not incorporate the physics of the hydrologic system.

2. The transformation methods are not exact. The exact convergence of probabilities to a specified distribution in the transformed range requires prior knowledge of the probability distribution of the original sample data which is unknown. For example Westphal (1984) has shown how SMEMAX could fail when the original data were log Pearson or Gumbel distributed. Also, it is known that the power transformation could lead to numerical errors if double precision is not used when the parameter used in the transformation is a negative number (McCormick 1984).

3. The hypothetical distribution based methods suffer from the nature of the distributions themselves. Almost all of these methods are used solely for the purpose of better curve fitting to the probabilities. The Gumbel distribution which is derived on the basis of maxima or minima suffers from the lower bound. Log Pearson has a

positive lower bound whenever the skew coefficient is positive. The discharge values must be above this lower bound. Many times the lower bound is physically unreasonable:

a. The lower bound can be greater than some of the observed values

b. The lower bound can be positive implying a probability of zero for the occurrence of low flows less than the lower bound.

c. The lower bound can be negative and thus the probability of negative flows can be positive.

4. The plotting position based methods require the correct plotting position formula which depends on the unknown underlying probability distribution of the original data (Cunnane,1978). Thus, the selection of a plotting position formula is arbitrary.

5. Intermittent streams are a major concern for two reasons: (1) many of the distributions require the calculation of the logarithms of the sample data and the logarithm of zero is negative infinity, and (2) continuous probability distributions have a zero probability that a specific flow will occur and thus, even if the stream is usually dry, the analysis produces a probability of a dry stream equal to zero. The first of these concerns can easily be rectified by adding a constant to all of the data points. After application of the selected distribution, the calculated low flows can then be found by subtracting the constant term from the low flows found from the distribution. The fact that the probability of the flow equal to the transformed zero flow is still zero is not addressed by this approach. A mixed distribution can be used to alleviate the problem. A point mass is placed at zero equal to the probability of occurrence of a zero discharge. The positive discharges are then represented by a truncated probability distribution. Aitchison (1955) provided an excellent account of this approach. Jennings and Benson (1969) used the mixed distribution for flood frequency analysis.

In light of the drawbacks of the black box methods, a physically based approach for the frequency analysis of low flows is presented. The physical model simulates stream recessions during dry weather, and a probability distribution is derived for the recession flow. In addition to the physical model, two distributions are discussed in handling of zero flows. One is a Double Bounded probability density function, and another being a mixed log-Pearson type III modified to handle zero flows. These are investigated because of inadequacies of other distributions in predicting zero flows. Lastly partial duration series (PDS) is investigated. Comparisons are made between classical truncated PDS and application of censoring to low flow which includes data the classical approach ignores. The motivation is to derive a more statistically sound approach to PDS.

Primarily this study explores the different methods for accuracy in prediction, consistency in performance for a variety of watersheds and theoretical justification. Application stresses both perennial and intermittent streams. Appendix C contains one day low flows for 15 stream gaging stations in Virginia.

CHAPTER 2: PHYSICALLY BASED METHOD

INTRODUCTION

The major disadvantage of the black box methods is they do not include the physics of the basin. It is well known that during low flow periods, streams derive their waters from the aquifer. The base flow is essentially the dry weather flow and is considered as the recession part of the hydrograph. Recent reviews relating to base flow are the ASCE Task Committee Report (1980), and McMahon and Arenas (1982).

The one dimensional, linearized, groundwater flow equation for an isotropic, homogeneous unconfined aquifer can be written as

$$\partial^2 h/\partial x^2 = (S_s/T)\partial h/\partial t \tag{1}$$

where: h = piezometric head

S_s = aquifer storativity

T = transmissivity

x = space coordinate

t = time

Eq. 1. is called the linearized Boussinesq equation. Recession modeling mainly involves the solution of the Boussinesq equation. A vast amount of literature is available on the solution of the groundwater flow equation. In this study the focus will be on those models which predict the base flow recession. Glover (1966), Aron and Borrelli (1973), Collins (1976), Brutsaert and Nieber (1977), McWhorter (1977), Naney et al. (1978), Birtles (1978) have solved the groundwater equation to find the recession flow. Unfortunately, these models are deterministic and assume perfect information which is questionable. These models being deterministic, also do not have the ability to predict the T year low flow. The recession equation can also be derived by using continuity and a linear reservoir relationship. The continuity equation is written as

$$dS/dt = I - Q \tag{2}$$

where: S = storage

I = inflow

Q = outflow

By assuming that groundwater storage can be represented as a linear reservoir:

$$S = KQ \tag{3}$$

Where K is the storage coefficient, or recession parameter. Base flow recessions occur during dry weather thus I=0 in Eq. 2 and combining Eqs. 2 and 3:

$$K(dQ/dt) = -Q \tag{4}$$

The solution of Eq. 4 is given as

$$Q(D) = Q(0)\exp(-D/K) \tag{5}$$

where: $Q(0)$ = start of recession flow

$Q(D)$ = Dth day flow

D = length of the recession

Eq. 5 is the recession equation. The solution of the linearized Boussinesq equation with suitable boundary conditions (Brutsaert and Nieber) for groundwater flow can also be derived in the form of Eq. 5. It is noted that zero flows in streams are due to the fall of the water table below the stream bed. Therefore the probability of zero flows is equal to the probability of the falling water table beneath the stream bed. In the present study instead of modeling the fall of the water table, bounds are used on low flows.

DISTRIBUTION OF LOW FLOWS

Eq. 5 is a basic procedure in the separation of base flow from the runoff hydrograph, and can be used to predict low flow during dry periods. Many times the recession parameter is used to characterize the hydrogeology of the basin. This basic equation will be exploited in this study. It is noted that Eq. 5 combines recession time periods with initial base flows and thus provides the drought length.

Fig. 1 illustrates a time series of discharge values. Because of long time intervals times between rainfall events, the discharge time series contains peaks followed by recessions. Consider $Q^n(0)$, the starting base flow discharge due to the nth rainfall event (see Fig. 1). From Fig. 1,

$$Q^n(0) = Q^n(D) + R^n \tag{6}$$

Where R^n is the recharge due to the n^{th} storm. Using Eq. 5, Eq. 6 can be written as:

$$Q^n(0) = A(n)Q^{n-1}(0) + R^n \tag{7}$$

Where $A(n)=\exp[(-D^n/k)]$, for $0 < A(n) < 1$; and D = number of time units (days) to reach the lowest flow on the recession side of the (n-1) rainfall event. Because the rainfall amounts and recession periods are random variables, Eq. 7 represents a stochastic relationship. Also, since the rainfall events are physically independent, it can be assumed that A(1), A(2),..., A(n) are independent identically

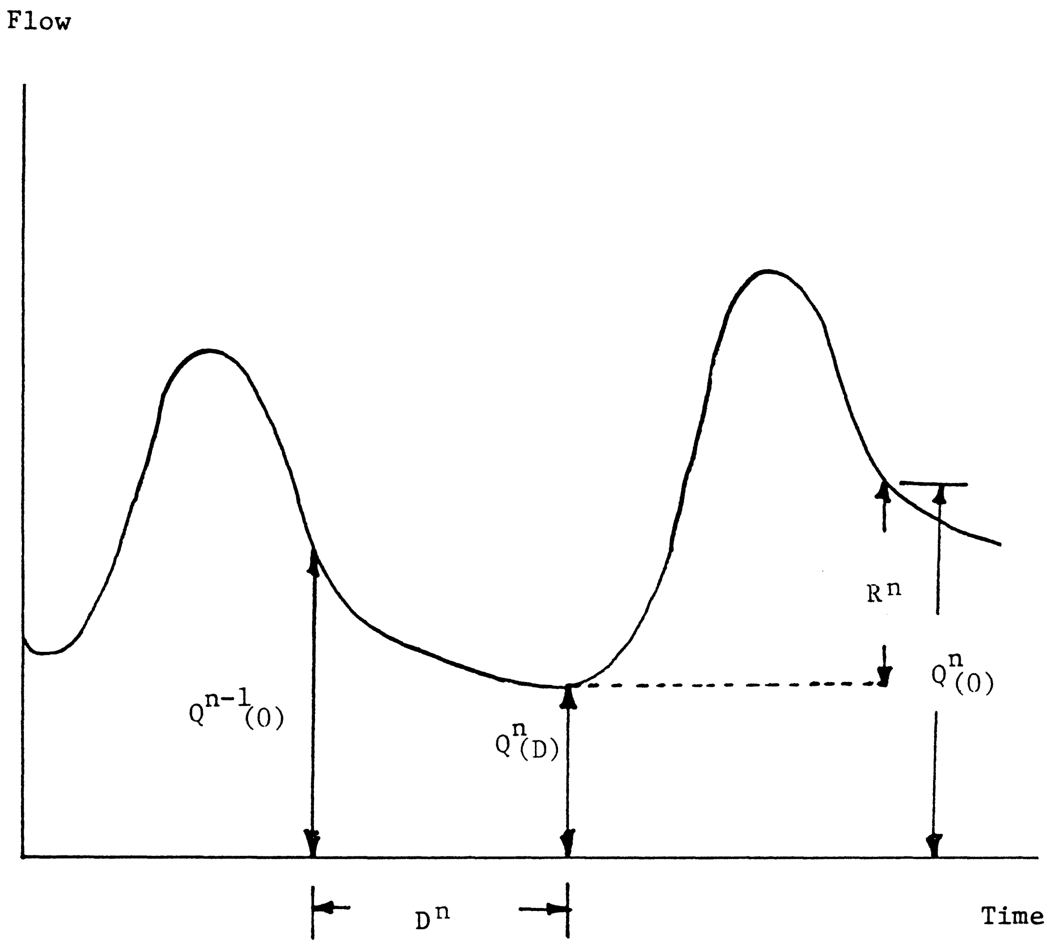

Fig. 1 - Time series of flows

distributed (i.i.d.) random variables and R(1), R(2),..., R(n) are i.i.d. random variables. Also, Q(0), A, and R are independent random variables. Thus, it can be written from Eq. 7.

$$P[Q^n(0) \le q | \; Q^{n-1}(0)=c] = P[Ac + R \le q] \tag{8}$$

Eq. 8 is the probability of $Q^n(0)$ being less than or equal to a specified discharge q given that, the previous initial base flow $Q^{n-1}(0)$ equals some constant c. It is noted that A is restricted between zero and one. Because of shape flexibility, and range (0,1), a beta distribution is fitted to the random variable A. Weibull distribution is fitted to R because of its shape flexibility and the closed form expression for its cumulative distribution function. For notational convenience hereafter,

$$Q^n(0) = Q(n) \quad \text{and } Q^{n-1}(0) = Q(n-1) \tag{9}$$

Thus Eq. 8 can be rewritten as

$$P[Q(n) \le q | \; Q(n-1)=c] = P[Ac + R \le q] \tag{10}$$

The probability density function of A is given as

$$f_A(a) = [\Gamma(\alpha+\beta)/\Gamma(\alpha)\Gamma(\beta)] \; a^{\alpha-1} \; (1-a)^{\beta-1} \tag{11}$$

for $0 < a < 1$

$= 0$ elsewhere

The probability function of R is given as

$$f_R(r) = \gamma\delta(r^{\delta-1}) \exp(-\gamma r^{\delta}) \quad \text{for } r > 0 \tag{12}$$

$$= 0 \quad \text{elsewhere}$$

In order to compute the probability from Eq. 10 the distribution of S must first be derived, where

$$S = Ac \tag{13}$$

$$f_S(s) = 1/c\ [\Gamma(\alpha+\beta)/\Gamma(\alpha)\Gamma(\beta)]\ (s/c)^{\alpha-1}\ (1-s/c)^{\alpha-1} \tag{14}$$

for $0 < s < c$

$$= 0 \quad \text{elsewhere}$$

Equation 10 can be written as

$$P[Q(n) \le q \mid Q(n-1)=c] = P[S + R \le q] \tag{15}$$

$$= P[R \le q - S] \tag{16}$$

Referring to Fig. 2, there are two ranges to be considered, i.e. $q > c$ and $q \le c$.

<u>$q < c$ case</u>:

From Fig. 2, the upper bound is q

$$P[S + R \le q] = P[R \le q - S] \tag{17}$$

$$= \int_0^q P[R \le q - s]\ f_S(s)ds \tag{18}$$

The cumulative probability distribution of R is given as

$$F_R(r) = \int_0^r \gamma\delta\ (x^{\delta-1})\ \exp(-\gamma x^{\delta})dx \tag{19}$$

$$= 1 - \exp[-\gamma r^{\delta}] \tag{20}$$

Thus Eq. 18 can be written as

$$P[S + R \le q] = \int_0^q F_R(q-s)\ f_S(s)ds \quad \text{for } 0 < q < c \tag{21}$$

<u>$q > c$ case</u>:

From Fig. 2, the upper bound is c

$$P[S + R \le q] = \int_0^c F_R(q-s)\ f_S(s)ds \quad \text{for } 0 < c < q \tag{22}$$

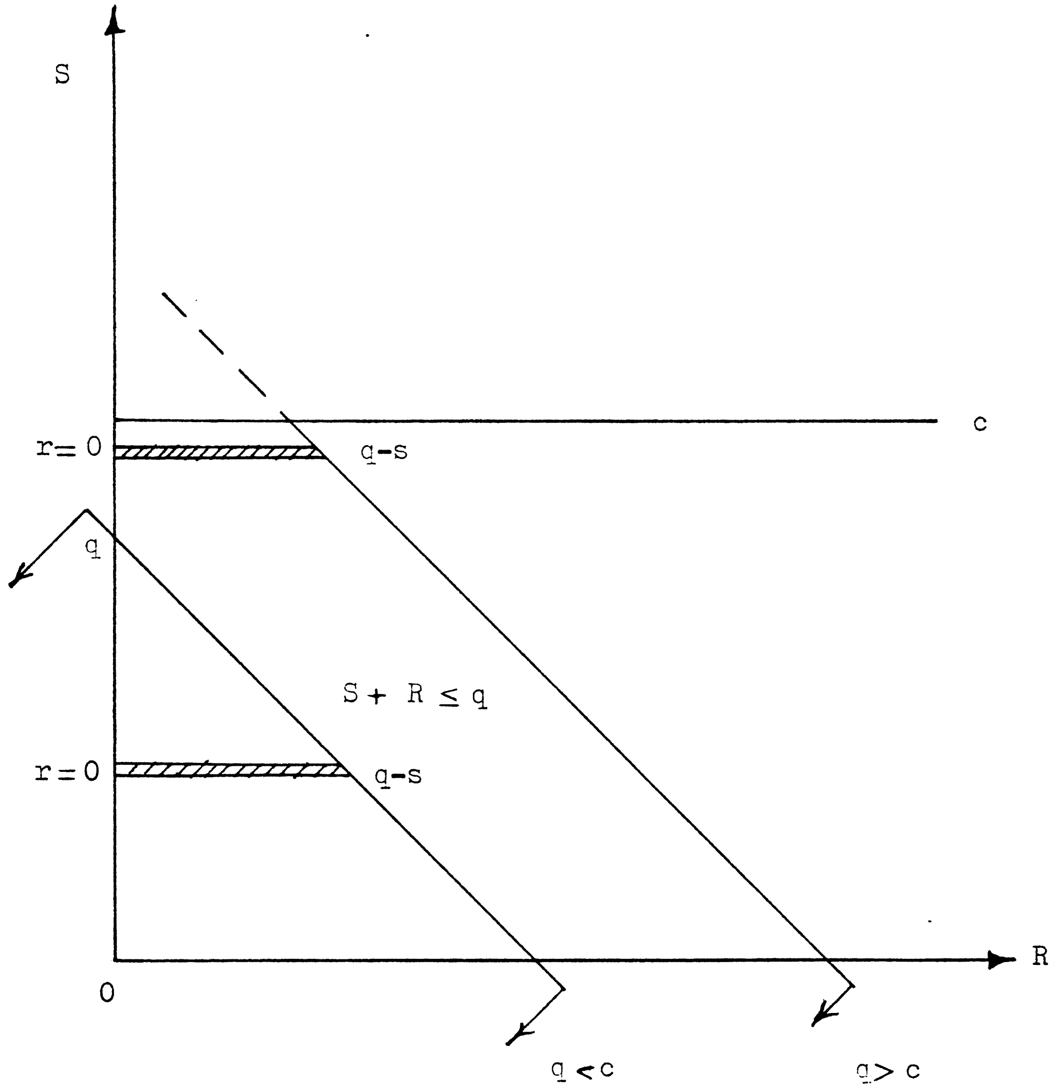

Fig. 2 - The event that $S + R \leq q$

Eqs. 21 and 22 can be used to generate the transition probabilities $P[Q(n) \leq q | Q(n-1) = c]$ in which the event $Q(n-1) = c$ is approximated as $c-\varepsilon < Q(n-1) < c+\varepsilon$. By using the transition probabilities, the long time steady state probabilities can be obtained from

$$\lambda = \lambda P \tag{23}$$

Where λ is the vector of steady state probabilities (i.e. $\lambda_i = P[Q(0)=q_i])$, P is the transition probability matrix. Eqs. 5 and 23 can be used to find the frequencies of one day low flow. Eq. 5 can be written as

$$Z = QA \tag{24}$$

Where Z is Q(D), Q is Q(O), and A is equal to exp(-D/k). The return period T in years is defined as the time interval between occurrences of flow q_T. In T years n recessions will occur where q_T occurs once. The expected number of recessions to occur are μT. The mean number of recessions per year μ is calculated by determining the number of recessions in the flow record and dividing by the number of years in the record. Then the modified return period gives

$$\mu T = 1/P[Z \leq z] \tag{25}$$

$$P[Z \le z] = P[QA \le z] \qquad (26)$$

$$P[Z \le z] = \int_q P[A \le z/q]\, dF(q) \qquad (27)$$

thus,

$$P[Z \le z] = \sum_{i=1}^{n} g(t_i)\, \lambda_i \qquad (28)$$

where: $\lambda_i = P[Q(O) = t_i]$ and (29)

$$g(t_i) = \int_0^{z/t_i} f_A(a)da \qquad (30)$$

It is noted that λ_i's are obtained from Eq. 23 and $g(t_i)$ from by direct integration of beta density. Therefore by using Eqs. 25 and 28 the T year low flow q_T can be found by trial and error.

BASEFLOW SEPARATION

INTRODUCTION

This section discusses the data aquisition and its input into the computer programs. A review of the recession model's application will explain the difficulties and possible improvements. Two approaches are given for data collection by baseflow separation: (i) graphing ln(Q) versus time and visual selection of recessions, and (ii) using a computer program to select recessions by criteria.

GRAPHICAL APPROACH

The Hydrologic Information Storage and Retrieval System (HISARS) is a good source for streamflow for Virginia and some of the neighboring states. For this research, all data came from HISARS.

For the graphing technique, two stations were chosen: Rappahannock River near Fredericksburg, Virginia (station number 01.6680.00) for April 1951 to March 1981, and James River near Richmond, Virginia (02.0375.00) for April 1963 to March 1983.

For each year the number of recession periods are determined by separating the base flow from the runoff for a given storm. Base flow separation is done by plotting the natural logarithm of flow [ln(Q)] versus time (see Fig. 3). Taking the natural logarithm of flows produces a straight line indicating recession occurred, as shown in Eq. 5. The recession line begins at Q=Q(0) and ends at Q=Q(D) when stormwater runoff occurs from the next storm starting the next hydrograph. The time difference between the occurrences

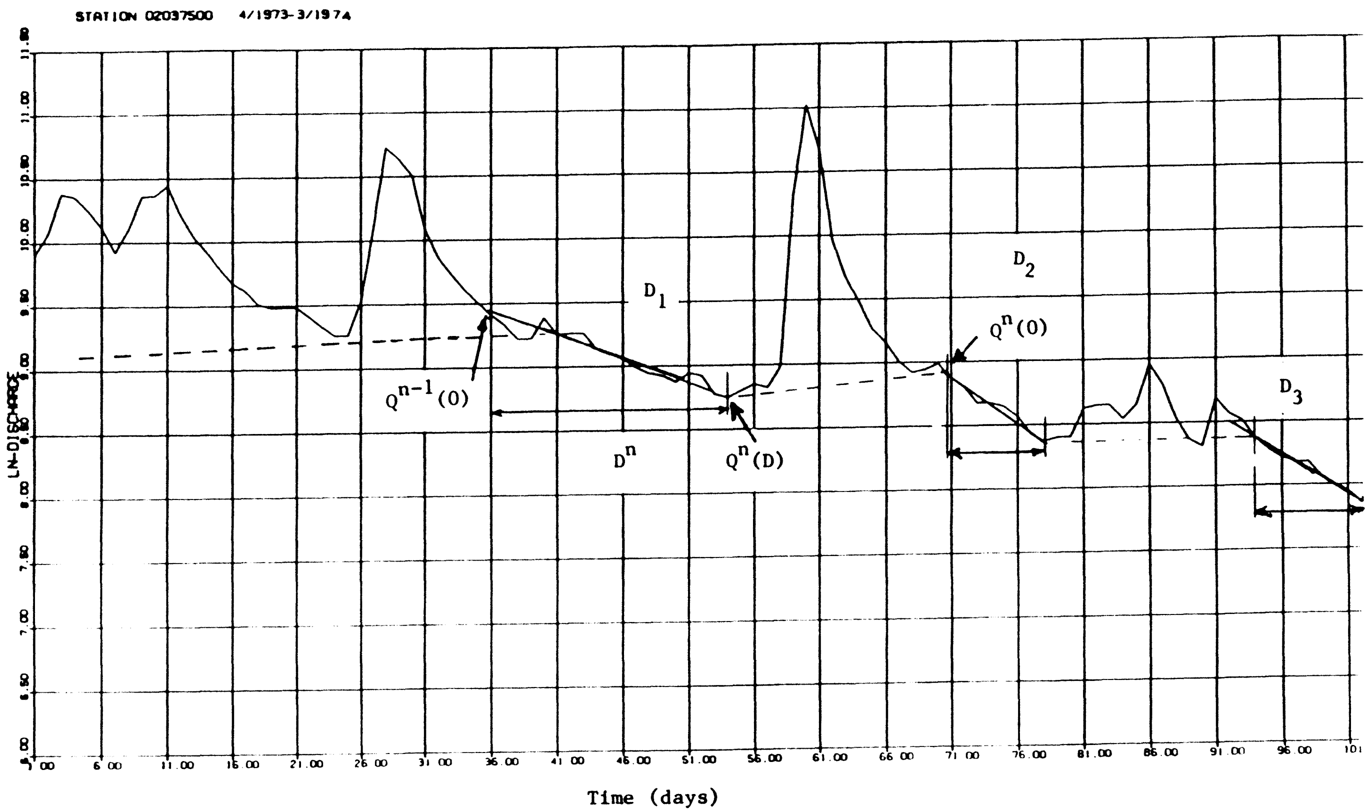

Fig. 3 - Base flow separation

of Q(0) and Q(D) is the recession period, also known as drought length, D.

When a storm occurs the hydrograph shows a sharp rise. After the stormwater runoff has flowed downstream the hydrograph sharply drops off, then gradually declines until base flow resumes, resulting in a straight line on the plot of ln(Q) vs. time. An actual straight line without any deviations does not always exist, but generally there is a distinct linear-like path for the recession limb on the semi-log plot. A recession does not always exist between storms, or peaks because another storm can occur before the recession begins from the previous storm. For example, in Fig. 3 before the first drought occurs there are several peaks but no recession is defined. Caution must be taken because the start of a recession where Q=Q(0) must be greater than the previous recession's lowest flow, Q(D). This is because recharge, R, must be greater than or equal to zero. It is noted that the start of recession is the nth Q(0), but the length of recession following the nth Q(0) is the (n+1)th D with its respective (n+1)th Q(D), which is related to the next Q(0) (following a storm) that is the (n+1)th Q(0). Also, at the end of a hydrograph plot for a particular year attention must be paid to the possible overlap of a drought period from one year to another, and that the next water year plot have a higher initial Q(0) than the previous Q(D) value from the previous ear.

Having calculated values of $\ln[Q(0)]$, $\ln[Q(D)]$, and D as shown in Fig. 3, recharge, R and A are calculated by

$$R = \exp\{\ln[Q(0)]\} - \exp\{\ln[Q(D)]\} \qquad (31)$$

$$A = \exp(-D/K) \qquad (32)$$

The Method of moments is used to calculate the distribution parameters for variables R and A, i.e. the calculation of the parameters α and β for Beta distribution, and γ and δ for the Weibull distribution. Weibull parameters γ and δ need D, $\ln[Q(0)]$, and $\ln[Q(D)]$. Beta parameters α and β need D and the recession coefficient K which needs to be estimated or found from previous testing. For the graphical application, this study used K values from McCormick (1984).

The parameters are used in the program MODD, found in Appendix D, which calculates low flows for return periods of 2, 5, 10, 25, 50, 100 years. Also inputted is an estimate of the upper limit for initial recession flow Q(0) and an initial value of the low flow sought. Since the Beta distribution does not have a closed form for the cdf, Eq. 22 is numerically integrated using subroutines from the International Mathematics Systems Library (IMSL). IMSL subroutine DCADRE was used to integrate Eq. 22 and gave good results for most stations. However, for three of the intermittent streams no convergence of the numerical analysis was possible. The steady-state probabilities of Q(0) in Eq. 23 are equal to the positive eigen vectors of the transposed transition probabilities, and were calculated using the IMSL subroutine EIGRF. To solve Eq. 30, the integrating of the Beta distribution the IMSL subroutine MDBETA is used.

COMPUTER APPLICATION

Computer application is similar to the graphical technique but a regression analysis is used for fitting recessions. Computer application does not require the time consuming process of plotting years of data, therefore more stations can be analyzed. A subroutine calculates the optimum recession coefficient by Lagrangian method from McCormick (1984). The daily flow records are reviewed to find N consecutive decreasing flows with a minimum of four days. Regressions are fit for ln(Q) versus time for different time lengths in the N days to represent potential recessions that may occur in the N days, starting with a length of three days to a maximum of N-1 increasing by increments of 1 day (the first flow is not considered in the recession because it is the peak flow and physically should not be considered in the recession). For each potential recession the reservoir coefficient K and the correlation coefficient R between ln(Q) and time (for this study, days) are calculated. Allowable ranges on K are determined by setting deviations from the optimum K value. Perennial streams generally have higher values of K than intermittent streams (they do not recede as fast) and should not have recessions with K smaller than two-thirds of the optimum K, and a K no larger than three times the optimum to avoid any successively decreasing flows that do not actually represent recession. For intermittent streams the lowest K is set as forty percent of the optimum because of the irregularity of the actual K and in general intermittent streams recede in flow at high rates following a storm. Any potential recession having K outside the acceptable K range is ignored, or if the correlation coefficient is less than a lower bound R_{min}, 0.98 being a good value. The longest potential recession with R still greater than R_{min} is the selected recession, however if no potential recession is found the consecutive decreasing flows may not represent a recession. If the length of days of decreasing flows is greater than ten days it is very likely a recession has occurred but none of the potential recessions have R greater than R_{min}. The restriction of R_{min} is relaxed slightly because slight deviations in flow may cause R to deviate below R_{min} but there may be a definite pattern of recession. The most relaxed R is for N greater than fourteen days, around 0.94, and for N greater than nine days, around 0.95.

The preceding represents the basic strategy for finding exponential recessions. However, the problem of obtaining positive recharge, that is the succeeding Q(0) equal or greater than the preceding Q(D), occurs in the computer application as in the graphical approach. Many times a selected recession is actually a section of a longer recession not accepted by the previous restrictions. Sometimes small peaks occur in the flow record disrupting the recession but do not significantly affect the recession. If the time difference between selected recessions is less than a minimum (normally less than or equal to 5 days) and the succeeding recession has a smaller Q(D), then the two recessions are added together to represent one recession. If selected recessions are not combined but Q(0) of the succeeding selected recession is less than the preceding Q(D) then a previous period of receding flows before Q(0) and after Q(D) not selected as a recession may be part of the succeeding selected recession. If the previous non-selected period of receding flows has a potential recession with the highest R being greater than 0.935 and Q(0) greater than Q(D), then the non-selected recession is combined with the succeeding selected recession to cause a positive recharge. In most cases this application was successful.

To illustrate this point, Fig. 4 shows an example of individual recessions combined to represent one recession. Table 1 in the following pages shows the step by step process of finding the individual recessions and how is each is combined by computer application. The first selected recession

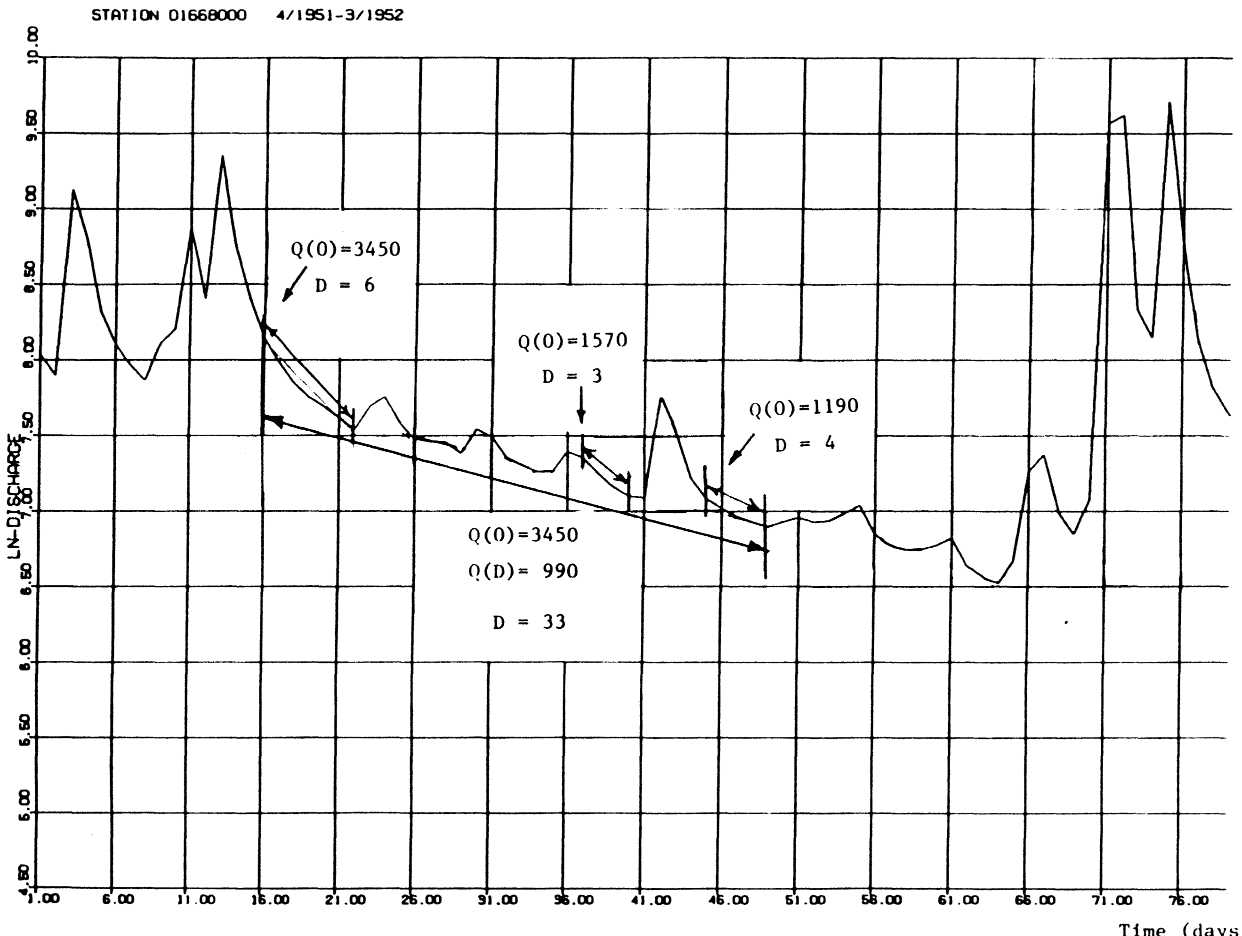

Fig. 4 - Example of recession selection

Table 1 - Example of computer selection of recessions

```
***************** SELECTION OF RECESSIONS *******************

          ACCEPTABLE RANGE FOR FLUCTUATIONS IS 41.133 CFS
             OPTIMUM RECECSSION CONSTANT, K IS 13.100 DAYS
                        MINIMUM ALLOWABLE K IS  8.734 DAYS
                        MAXIMUM ALLOWABLE K IS 39.301 DAYS
 DAY =   13     N =    10

FLOWS:  11600.00   6320.00   4480.00   3450.00   2940.00   2580.00
         2340.00   2190.00   2040.00   1890.00

                   THE "BEST" REGRESSIONS ARE AS FOLLOWS:

           LENGTH  START    END       K          R           S
              3      7       10    14.051377  -0.999494   0.003581
              4      6       10    13.164368  -0.997793   0.009230
              5      5       10    11.745235  -0.991359   0.023565
              6      4       10    10.369604  -0.984664   0.040434
              7      3       10     8.705017  -0.968419   0.078251
              8      2       10     7.206521  -0.953054   0.129074

                THE SELECTED REGRESSION IS FROM I =  4 TO J = 10
                 WHERE: K = 10.37   R = -0.9847      S = 0.0404
                 END OF RECESSION =    22     DURATION =  6.0

    SELECTED RECESSION:  1         Q0 = 3450.00     QD = 1890.00

 DAY =   24     N =     5

FLOWS:   2340.00   1960.00   1770.00   1740.00   1730.00

                   THE "BEST" REGRESSIONS ARE AS FOLLOWS:

           LENGTH  START    END       K          R           S
              3      2       5     25.538634  -0.864812   0.035945

NON-SELECTED RECESSION:  1       PQ0 = 1960.00     PQD = 1730.00

 DAY =   30     N =     5

FLOWS:   1890.00   1800.00   1570.00   1490.00   1430.00

                   THE "BEST" REGRESSIONS ARE AS FOLLOWS:

           LENGTH  START    END       K          R           S
              3      2       5     13.465544  -0.956689   0.035730

NON-SELECTED RECESSION:  2       PQ0 = 1800.00     PQD = 1430.00

 DAY =   36     N =     5

FLOWS:   1630.00   1570.00   1410.00   1290.00   1210.00

                   THE "BEST" REGRESSIONS ARE AS FOLLOWS:

           LENGTH  START    END       K          R           S
              3      2       5     11.490116  -0.993796   0.015400

                THE SELECTED REGRESSION IS FROM I =  2 TO J =  5
                 WHERE: K = 11.49   R = -0.9938      S = 0.0154

    SELECTED RECESSION:  2        Q0 = 1570.00     QD = 1210.00

                   <<<<<<<<<  Q0(II) < QD(II-1)  >>>>>>>>>>
              ----- ADDED PREVIOUS NON-SELECTED RECESSION -----

NEW SELECTED RECESSION:  2        Q0 = 1800.00     QD = 1210.00

             +++++++++  COMBINED RECESSIONS  ++++++++

                   END OF RECESSION =     40     DURATION = 24.0

NEW SELECTED RECESSION:  1        Q0 = 3450.00     QD = 1210.00

 DAY =   42     N =     9

FLOWS:   2350.00   1890.00   1370.00   1190.00   1120.00   1060.00
         1020.00    990.00   1020.00

                   THE "BEST" REGRESSIONS ARE AS FOLLOWS:

           LENGTH  START    END       K          R           S
              3      4       7     19.323230  -0.995224   0.008026
              4      4       8     21.666907  -0.989779   0.012141
              5      4       9     29.677195  -0.911167   0.031872
              6      3       9     20.793762  -0.908310   0.052410
              7      2       9     13.004633  -0.865181   0.117918

                THE SELECTED REGRESSION IS FROM I =  4 TO J =  8
                 WHERE: K = 21.67   R = -0.9898      S = 0.0121

    SELECTED RECESSION:  2        Q0 = 1190.00     QD =  990.00

             +++++++++  COMBINED RECESSIONS  ++++++++

                   END OF RECESSION =     50     DURATION = 33.0

NEW SELECTED RECESSION:  1        Q0 = 3450.00     QD =  990.00
```

occurs from day 16 to day 22, but the next selected recession occurs at day 36 to day 40. However, the succeeding selected recession has Q(0) less than the previous selected Q(D), therefore the non-selected recession before the second selected recession is combined with it. The second recession now starts at day 31 and is close to the first recession, and both are combined to make one recession starting at day 16 and end at day 40. Another selected recession occurs at day 45, but has Q(0) less than the previous selected Q(D). Since the time difference between selections is small the succeeding selected recession is combined with the preceding selected recession. No other recessions are selected before the storm runoff at day 71, thus the final selected recession starts at day 16 with Q(0)=3450 cfs and ends at day 49 with Q(D)=990 cfs. The end result is that the four recessions actually are part of a longer recession. Though fluctuations exist in the recession, the computer program is able to avoid the effects of the intervening peaks that do not contribute recharge to the basin. This is an example of the more difficult recession selections, other recessions normally require less modification.

Another possibility is the flow the day before Q(0) may be a better representative Q(0) if it is not too large. If the preceding flow is greater than the preceding Q(D) and does not surpass a maximum deviation depending on Q(0) then the recession is started with that Q (and is not the peak of the recession), and the recession duration is increased by one day. If all the above fails, then a arbitrary judgement is enforced. If the difference between Q(D) and Q(O) is not too large then Q(O) is set equal to Q(D) and the recession duration is increased by one day. Though arbitrary, it seems not to be very harmful and occurs infrequently, and in some cases storms provide little recharge to the ground water which is approximately zero.

The computer approach was applied to the Rappahanock and James Rivers and eight other perennial streams and five intermittent streams. Overall the computer application gave good results in determining recessions and their respective drought lengths and recharge. However, the results depend on the quality of the data: (i) the accuracy and precision of the stream flow, and (ii) that the flow be derived from the aquifer resulting in exponential recession. If inconsistencies in the data occur frequently then it is difficult to find representative recessions because the flows may not represent the true stream recession process. Especially critical is man-made obstructions to stream flow, abstraction or addition of stream flow, and any type of regulative actions altering stream flow. The recession model assumes the flow is totally derived from groundwater and any alteration of flow causes receding flow to deviate from an exponential pattern. Especially sensitive are intermittent streams because stream flow is estimated accurately to 1 or

0.1 cfs and flows below 0.1 cfs are not correctly estimated. Consistent calculations of duration were obtained for a few stations, but the recharge estimation may be susceptible to finding an approximate Q(0) and Q(D). Possibly the plotting technique may give more consistent results, but perennial streams showed good results in the computer application of finding recessions despite some data inconsistencies. Recessions done graphically for Rappahanock and James Rivers were compared with the computer selection and almost all recessions were identical except for slight deviations in recharge values and duration lengths. For good stream data the computer application may be more consistent because it does not require the drawing of a straight line to approximate the exponential recession, it only needs to compare the correlation coefficient and recession coefficient.

CHAPTER 3: LOG-PEARSON TYPE III DISTRIBUTION

INTRODUCTION

Log-Pearson type III (LPIII) is widely applied for flood frequency analysis, as well as for low flow analysis. LPIII uses annual minimum low flows which represents an annual series. For annual low flows a water year different from calendar year is used. Low flow generally occurs in summer, fall, or winter. A long duration of low flows may overlap into another year and influence the succeeding minimum which contradicts the assumption of independence in frequency analysis. In this study the year begins with April and goes to March.

LPIII is a black box method because it is a curve fitting scheme to the probabilities of streamflows. LPIII involves transforming the flows into natural logarithm reducing the skew of the data, and applying a Pearson distribution type III. The pdf is given by

$$f(x) = \{[\ln(x)-\gamma]/\alpha\}^{\beta-1} \exp\{-[\ln(x)-\gamma]/\alpha\}/[|\alpha|x\bullet\Gamma(\beta)] \quad (33)$$

where: x = low flow, Q

$f(x)$ = probability density of value x

α = scale parameter

β = shape parameter

γ = location parameter

$\Gamma(\beta)$ = a location parameter

The bounds on x are:

$\beta > 0$, $\alpha > 0$, then $\gamma \le x < \infty$

$\beta > 0$, $\alpha < 0$, then $-\infty < x \le \gamma$

LPIII has no closed form solution for the cumulative distribution function (cdf), however q_T, the flow occurring every T years, can be calculated by

$$q_T = \exp(M_y + KS_y) \tag{34}$$

where: K = frequency factor

M_y = mean of ln(Q)

S_y = standard deviation of ln(Q)

The frequency factor K represents the variation in low flow for different frequencies. Extensive tables for K are available for different values of return period, T, and skew (Bulletin 17A, Kite (1977)). Also Eq. 35 can be used if skew, Cs, is within the range -1<Cs<1 (Bulletin 17A).

$$K_T = 2/G \ \{[(x_T - G/6)G/6+1]^3 - 1\} \tag{35}$$

where: K_T = K for return period T

G = sample skew of $\ln(Q)$

x_T = standard normal deviate for $F(x_T)=1/T$

It is noted when skew=0, K becomes the standard normal deviate. A derivation for K is given in Appendix A. The major disadvantage is the logarithm of zero flows is negative infinity. To apply LPIII for data with zero flows, a small constant is added to each flow, then for each Q_T calculated the constant is subtracted. For perennial streams this is no concern because the probability of zero flow is zero. An alternative is to modify LPIII with a point mass probability for zero flow, called the conditional probability approach.

LPIII WITH A POINT MASS PROBABILITY FOR ZERO FLOW (MIXED LPIII)

Aitchison (1955) provided an excellent analysis of this approach. Jennings and Benson (1969) applied LPIII with point mass probability to flood flows. The incorporation of a positive probability for zero flow truncates LPIII for the positive flows. The probability of zero flow is estimated by

$$P(Q=0) = k/n \tag{36}$$

where: k = number of zero flows in record

n = number of flows in record

Then the probability of flow equal to or less than q_T is

$$P(Q \le q_T) = P(Q=0) + P(Q \le q_T|Q>0)[1-P(Q=0)] \qquad (37)$$

$P(Q \le q_T|Q>0)$ is the conditional probability given that flow is positive implied by LPIII.

T YEAR LOW FLOW PREDICTION

To determine q_T by mixed LPIII, Eq. 37 is solved for the conditional probability of positive flows.

$$P(Q_T \le q|Q>0) = [P(Q_T \le q) - P(Q=0)]/[1-P(Q=0)] \qquad (38)$$

where: $P(Q \le q_T) = 1/T$

Eq. 38 becomes

$$P(Q \le q_T|Q>0) = [(1/T)-(k/n)]/[1-(k/n)] \qquad (39)$$

$P(Q \le q_T|Q>0)$ is the probability estimated from LPIII fitted to positive flows only. The standard normal deviate for the $P(Q \le q_T|Q>0)$ and the skew of the positive low flow data are used to compute K in Eq. 35 or an expanded version of Eq. 35 in Kite (p.112, 1977), or by tables which is then used in Eq. 34 to calculate q_T. A drawback in Eq. 39 is P(Q

$\leq q_T|Q>0)$ is not at a known cut-off point when $k > 0$, and does not allow for easy application in finding the standard normal deviate for that cut-off point. Extensive tables of K for different probabilities are required, or an approximation series for standard normal deviate is needed. Abramowitz and Stegun (p. 932, 1972) give an approximation for the cdf, F(x), of normal distribution.

$$F(x) = 1-1/2(1+d_1x+d_2x^2+d_3x^3+d_4x^4+d_5x^5+d_6x^6)^{-16} \quad (40)$$

where: x = normal deviate for $F(x)$

$d_1 = .0498673470$ $\quad$ $d_4 = .0000380036$

$d_2 = .0211410061$ $\quad$ $d_5 = .0000488906$

$d_3 = .0032776263$ $\quad$ $d_6 = .0000053830$

Since $F(x) = P(Q \leq q_T|Q>0)$ given by Eq. 39, x is found by trial error. The approximation deviates significantly for return periods larger than 50 years, thus care must be taken in applying the approximation. Thus found x is used in Eq. 35 to obtain K_T which is then substituted in Eq. 34 along with the standard deviation of the logarithm of positive flows (zero flows are deleted) to obtain q_T.

CHAPTER 4: DOUBLE BOUNDED PDF WITH POINT A MASS AT ZERO

INTRODUCTION

In general low flows are limited by a lower bound, say A, and an upper bound, B. For intermittent streams the lower bound is zero and has a positive probability of occurrence. For perennial streams the lower bound will not reach zero, but some positive value, and zero flow has a zero probability of occurrence. Thus a comprehensive probability distribution for low flows requires a point mass at zero, bounded by an upper bound, and the ability to fit a variety of shapes for different stream data.

Based on the above characteristics an ideal probability distribution for low flows may have the form

$$f_X(x) = F_0 \qquad \text{for} \quad x = 0$$

$$= (1-F_0)\ f_1(x) \qquad 0 < x < 1 \tag{41}$$

where:

$$x = (z-A)/(B-A) \tag{42}$$

$$z = \text{low flow, } Q$$

Thus the distribution given in Eq. 41 involves the parameters F_0, A, B, and the parameters of the probability distribution for positive flows, $f_1(x)$. Because beta density assumes a variety of shapes and ranges between $0<x<1$, beta or a modified beta distribution seems to be an ideal choice for $f_1(x)$. In the following a maximum likelihood approach for the parameter estimation for f_X will be given. Appendix B contains an explanation of maximum likelihood estimator. The likelihood function, L, for the distribution in Eq. 41 may be written as

$$L = F_0{}^k \, (1-F_0)^{(n-k)} \prod_{i=1}^{n-k} f_1(x_i) \quad x_i > 0 \tag{43}$$

where: $k = \sum_{i=1}^{n} I_{\{0\}} x_i$

$= \text{number of zero flows}$

the indicator function, $I_{\{0\}}$ for zero flow is given as

$$\begin{aligned} I_{\{0\}} &= 1 \quad \text{if } x_i = 0 \\ &= 0 \quad \text{if } x_i \neq 0 \end{aligned}$$

Then the derivatives of L with respect to the parameters are set to zero to achieve maximum likelihood. Parameters are estimated by solution of

$$\partial L/\partial F_0 = 0 \quad \text{and} \quad \partial L/\partial \theta_i = 0 \qquad \text{for all } i$$

where: θ_i are the parameters of the f_1 distribution

DOUBLE BOUNDED PDF WITH POINT MASS

Kumaraswamy in 1980 suggested a particular form for $f_1(x)$.

A generalized probability density function (pdf), f, called Double Bounded pdf (DBPDF) which has lower and upper bounds with a point mass probability for the lower bound. In low flows, F_0 is positive only for A=0 since zero flow is the physical lower bound for intermittent streams and is recurring. Since F_0 is a point mass, the distribution of low flows is a mixed distribution that is, continuous and discrete distributions are applied simultaneously (see Fig. 5). Conversely for perennial streams F_0=0 and Eq. 41 becomes only a function of $f_1(x)$ which is continuous since all flows are non-zero, as shown in Fig 6. B is set equal to the largest annual minimum in the stream flow record. Denoting flows by z, Eq. 42 gives

$$z = x(B-A) + A \tag{44}$$

and the pdf $f_1(x)$ is given as

$$f(z) = ab/(B-A)[(z-A)/(B-A)]^{a-1}\{1-[(z-A)/(B-A)]^{a}\}^{b-1} \tag{45}$$

$$\text{for } A < z < B$$

where:

f(z) = probability density function of z

z = random variable, annual minimum low flow

A = lower bound of z

B = upper bound of z, largest annual minimum

Double Bounded pdf

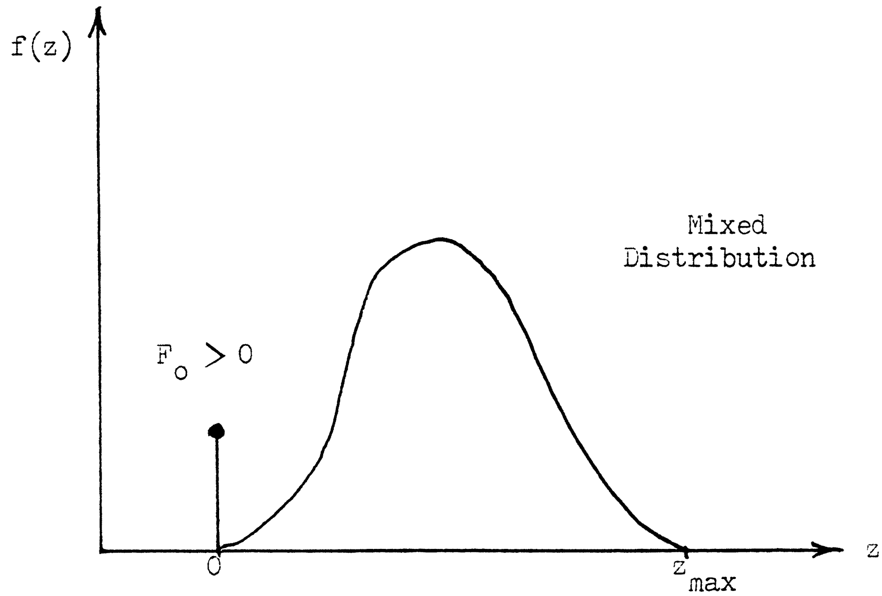

Fig. 5 - Intermittent Stream

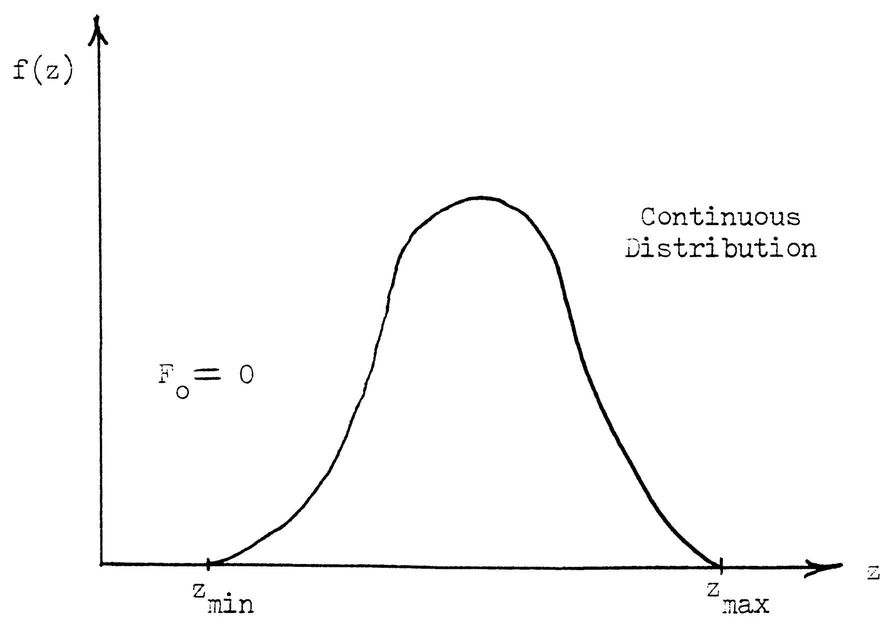

Fig. 6 - Perennial Stream

F_0 = point mass probability of A

a, b = shape parameters, > 0

The cumulative probability density function (cdf), F(z), is given by integrating the pdf in Eq. 45 which gives:

$$F(z) = F_0+(1-F_0)\{1-[1-[(z-A)/(B-A)]^a]^b\} \qquad (46)$$

The utility of DBPDF is the pdf's flexibility to change shape with various combinations of a and b. This is because the DBPDF is a modified beta distribution having bounds on X (0,1). The beta distribution is a special case of DBPDF when b = 1. Thus DBPDF tends to provide a better fit of data than beta since another parameter, b, is included. It is noted whenever a or b crosses the value of 1.0 a significant change occurs in the shape of the pdf. Thus DBPDF is a curve fitting scheme which depends on the values of a and b. Also the analytical form of Eq. 46 allows for straight forward application, however, estimating parameters a, b, F_0, and A (B is fixed) requires an iterative technique. Maximum likelihood method is used to estimate the parameters.

PARAMETER ESTIMATION

Applying maximum likelihood to DBPDF is slightly different because of the point mass probability, F_0 as set up in Eq. 43. For intermittent streams A remains equal to zero since it is recurring, but for perennial streams A may assume any value according to the fitting of parameters to data. For perennial streams this is a short coming since A may be negative thus giving the probability of negative flow a positive probability which is unsatisfactory, or a positive A which physically seems unacceptable since the probability of flow below the positive value is zero but the worst case of any stream is zero flow. The justification for the parameter A being different from zero is a better curve fitting of the probabilities to the physical data. Considering a sample set of n annual minima $Z_1, Z_2, \ldots, Z_n$ which contains k zero flows (note: k=0 for perennial streams) and n-k positive flows, the likelihood function L_1 from Eqs. 43 and 45 can be written for the DBPDF as

$$L_1 = (F_0)^k \prod_{i=1}^{n-k} (1-F_0)[ab/(B-A)]\ [(z-A)/(B-A)]^{(a-1)} \bullet \{1-[(z-A)/(B-A)]^a\}^{(b-1)} \tag{47}$$

F_0 represents the probability of zero flow, and $(1-F_0)$ is the probability of non-zero flow. Eq. 47 becomes

$$L_1 = F_0^{\,k}\ (1-F_0)^{n-k}\ [ab/(B-A)]^{n-k} \bullet \prod_{i=1}^{n-k} [(z_i-A)/(B-A)]^{a-1}\ \{1-[(z_i-A)/(B-A)]^a\}^{b-1} \tag{48}$$

Taking the natural logarithm of Eq. 47 simplifies the differentiation of the maximum likelihood function, given by $L=\ln(L_1)$, and Eq. 48 becomes

$$L = k\ \ln(F_0) + (n-k)\ln(1-F_0) + (n-k)[\ln(a)-\ln(b)]$$
$$- ab(n-k)\ln(B-A) + (a-1) \sum_{i=1}^{n-k} \ln(z_i-A)$$
$$+ (b-1) \sum_{i=1}^{n-k} \ln[(B-A)^a - (z_i-A)^a] \qquad (49)$$

Setting the derivatives of the log likelihood function to zero, gives the maximum likelihood estimates for a, b, F_0, and A.

$$\partial L/\partial F_0 = k/F_0+(n-k)/(1-F_0) = 0 \qquad (50)$$

simplifing Eq. 50 gives

$$F_0 = k/n \qquad (51)$$

$$\partial L/\partial a = (n-k)/a - b(n-k)\ln(B-A) + \sum_{i=1}^{n-k} \ln(z_i-A)$$
$$+ (b-1) \sum_{i=1}^{n-k} [(B-A)^a\ln(B-A)-(z_i-A)^a\ln(z_i-A)]$$
$$/[(B-A)^a - (z_i-A)^a] \qquad (52)$$

Note: For $y = x^a$, where: $x = (B-A)$ or $x = (z_i-A)$

Let $z = \ln(y) = a \ln(x)$ then

$dz/da = (1/y)\, dy/da = \ln(x)$

resulting in $dy/da = x^a \ln(x)$

$$\partial L/\partial b = (n-k)/b - a(n-k)\ln(B-A) + \sum_{i=1}^{n-k} \ln[(B-A)^a-(z_i-A)^a] \tag{53}$$

Simplifying gives

$$b = (n-k)/\{a[n-k]\ln(B-A)- \sum_{i=1}^{n-k} \ln[(B-A)^a-(z_i-A)^a]\} \tag{54}$$

$$\partial L/\partial A = ab(n-k)/(B-A) - (a-1) \sum_{i=1}^{n-k} 1/(z_i-A) + (b-1) \sum_{i=1}^{n-k} [(z_i-A)^{a-1} -(B-A)^{a-1}]/[(B-A)^a-(z_i-A)^a \tag{55}$$

where: $\partial L/\partial a = \partial L/\partial b = \partial L/\partial A = 0$

It is interesting to note that the maximum likelihood estimator of F_0 in Eq. 51 is the same as the frequency estimator. Eq. 54 gives an explicit solution for b and Eqs. 52 and 55 are solved to estimate the parameters a and A. The solution of a and A requires an iterative scheme and the following strategy is used for the solution.

STEP 1: F_0 is estimated by Eq. 51, but a, b, and A are dependent on each other, so Eqs. 52 and 55 must be solved implicitly for a and A, and b is computed from Eq. 54.

STEP 2: Initially assume A equal to zero (A remains zero for intermittent streams, thus requiring only a and b) then assume an initial estimate of a. Using these initial estimates of a and A, b is found directly by Eq. 54.

STEP 3: Using these values of a, b, and A, Eq. 52 is solved. If it does not approximately equal zero, then a is estimated by the method of bisection until Eq. 52 converges approximately to zero.

STEP 4: These values of a, b, and A are then used to solve Eq. 55. If it is approximately zero then the parameters are the maximum likelihood estimates. If Eq. 55 is not satisfied then A is decreased by an increment if the sign of the value given by Eq. 55 is negative, and increased by an increment if the sign is positive. This new A is then used in STEP 2 and follow the same procedure outlined for A=0 in STEP 2. This is repeated until Eq. 55 is close to zero.

T YEAR LOW FLOW PREDICTION

Annual minimum low flow is a random variable, and return period, T, is the expected number of years in which the annual minimum will be less than or equal to a specified q_T. The probability of Q being less than or equal to q_T is

$$P[Q \leq q_T] = 1/T \tag{56}$$

$P[Q \le q_T]$ is also the cumulative probability of $z=q_T$ given in Eq. 46, thus

$$F(q_T) = 1/T \tag{57}$$

Knowing the parameters and T, q_T is solved explicitly using Eq. 46 where $z=q_T$.

$$q_T = A + (B-A)\{1-\{1-[(1/T-F_0)/(1-F_0)]\}^{1/b}\}^{1/a} \tag{58}$$

CHAPTER 5: PARTIAL DURATION SERIES

INTRODUCTION

The methods predicting low flows using annual minimum low flows are called annual minimum series methods, for example log-Pearson type III and Double Bounded pdf. However for N years of record only N flows are used in prediction. If N is not sufficiently large, the estimated parameters may be biased. The main drawback in using only one low flow each year is that some years may have two or more droughts, and only the severest is used. The non-recorded drought may be less than several annual minima in the record.

Partial duration series (PDS) is a series of data selected by flow magnitudes being less than a pre-specified cut-off value, q_0 (see Fig. 7). PDS has a much larger record of flows than annual minimum series, however, each flow in PDS must be <u>independent</u>; an underlying assumption in low flow frequency analysis. The cut-off value must be such that the low flows are independent of each other which is a drawback in using PDS.

PDS can be viewed as two processes: (i) Truncation, and (ii) Censoring. For a chosen flow level q_0, let n flows occur below q_0, in which n is a random variable. This is called a truncation process. Traditional PDS is a form of truncation since any low flow occurring above q_0 is truncated.

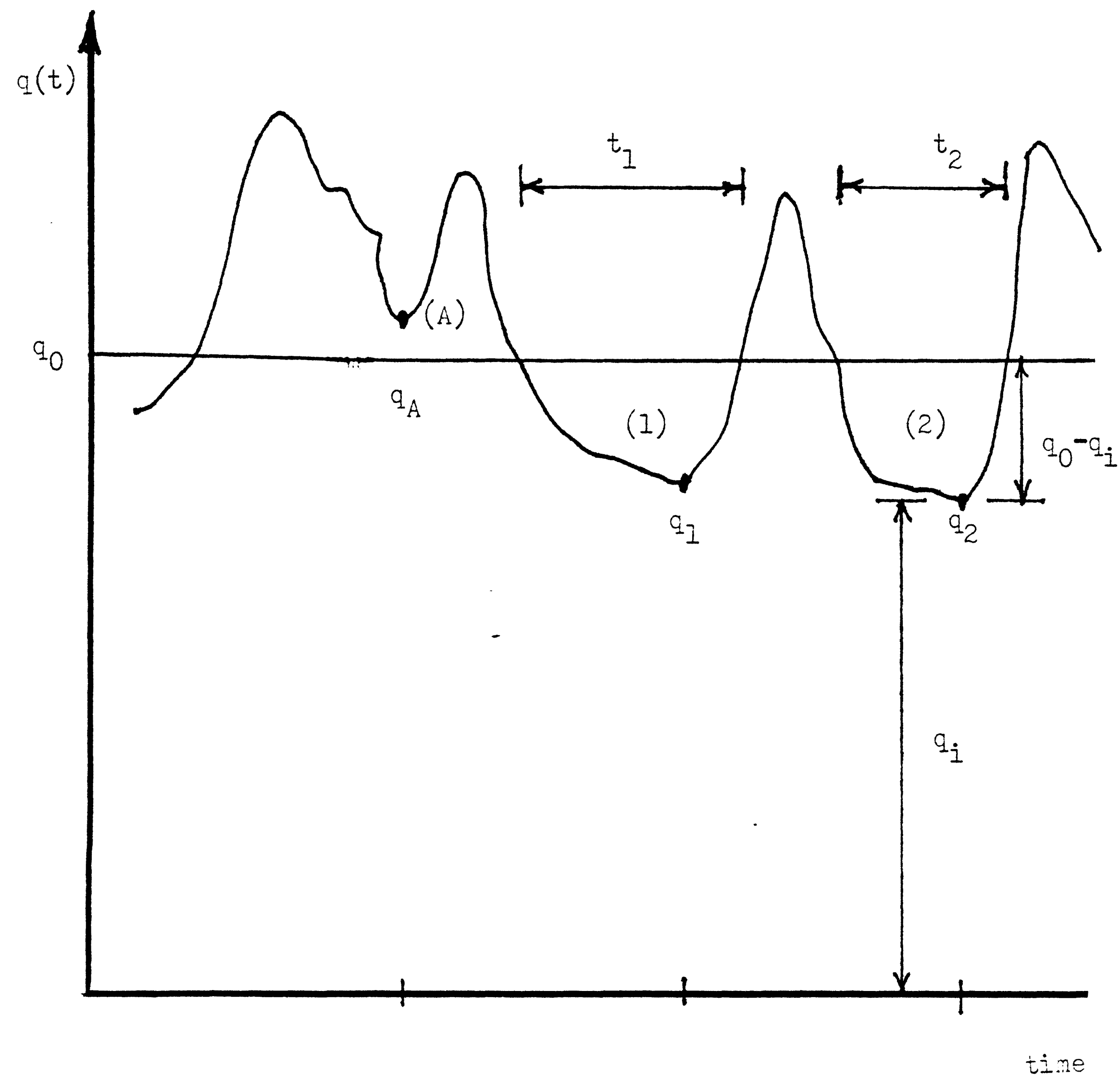

Fig. 7 - Partial Duration Series

q_0 = cut-off flow for low flow

q_1, q_2 = flows from truncation

q_A = flow from censorship

In truncation, flows above q_0 are ignored, while the censoring process requires the flows below q_0 and the number of flows above q_0 but not their values (their values being censored). Truncation is the process in which the population is assumed to be generated from a truncated range and the number of flows which exceed the cut-off value q_0 is not known. Let m flows occur above q_0 in addition to the n less than q_0. It is known that m flow values which exceed q_0 exist, and is a random variable. This is called Type I censoring. There is also Type II censoring in which a fixed number, k, flows are less than a cut-off value, and q_0 is a random variable. Therefore, in censoring the observations are curtailed above or below a certain level, but in truncation the entire population is generated from a curtailed range. In hydrology Type I censoring is typically used. Leese (1971) has applied censoring for floods. Ashkar and Rouselle (1981), Cunnane (1973, 1947) and Shane and Lynn (1964) have considered truncated PDS for floods.

PARTIAL DURATION SERIES BY TRUNCATION

The specified q_0 sets the criterion for the magnitude of the highest low flow allowed. However, q_0 traditionally is based on judgement. A common decision is to set q_0 to the largest annual minimum in the record, then each year will contribute at least one flow to the series.

A probability distribution is empirically applied to the series data. Chow (p.8-35,1964) suggests the exponential distribution. Although the exponential distribution allows for easy application, it is not flexible enough to fit data from different streams because it has only one parameter. A more flexible distribution is the Weibull which is used herein.

Two types of variables can describe the truncated PDS. One using the flow values Q, the other the differences between the cut-off flow, q_0 and the flow values, Q, called deviations (see Fig. 7). The deviation, X equals q_0-Q , in which X and Q are random variables.

Weibull is applied as a curve fitting scheme for the flow probabilities. Using Q as the random variable, the parameters may vary greatly, especially the shape parameter, between stations. The motivation of q_0-Q, as the random variable is as follows: Large low flows have a higher probability than small low flows, but the difference X=q_0-Q is small for the large low flows and reaches a maximum difference of q_0 (for zero flows). Thus large low flows have small differences and higher probability which suits the exponential distribution. The exponential distribution has a higher density at the lower values and tails off for the higher values; favoring X=q_0-Q (see Fig 8). It is noted that exponential is a special case of Weibull when the shape parameter, c, equals 1. To allow for more flexibility the two

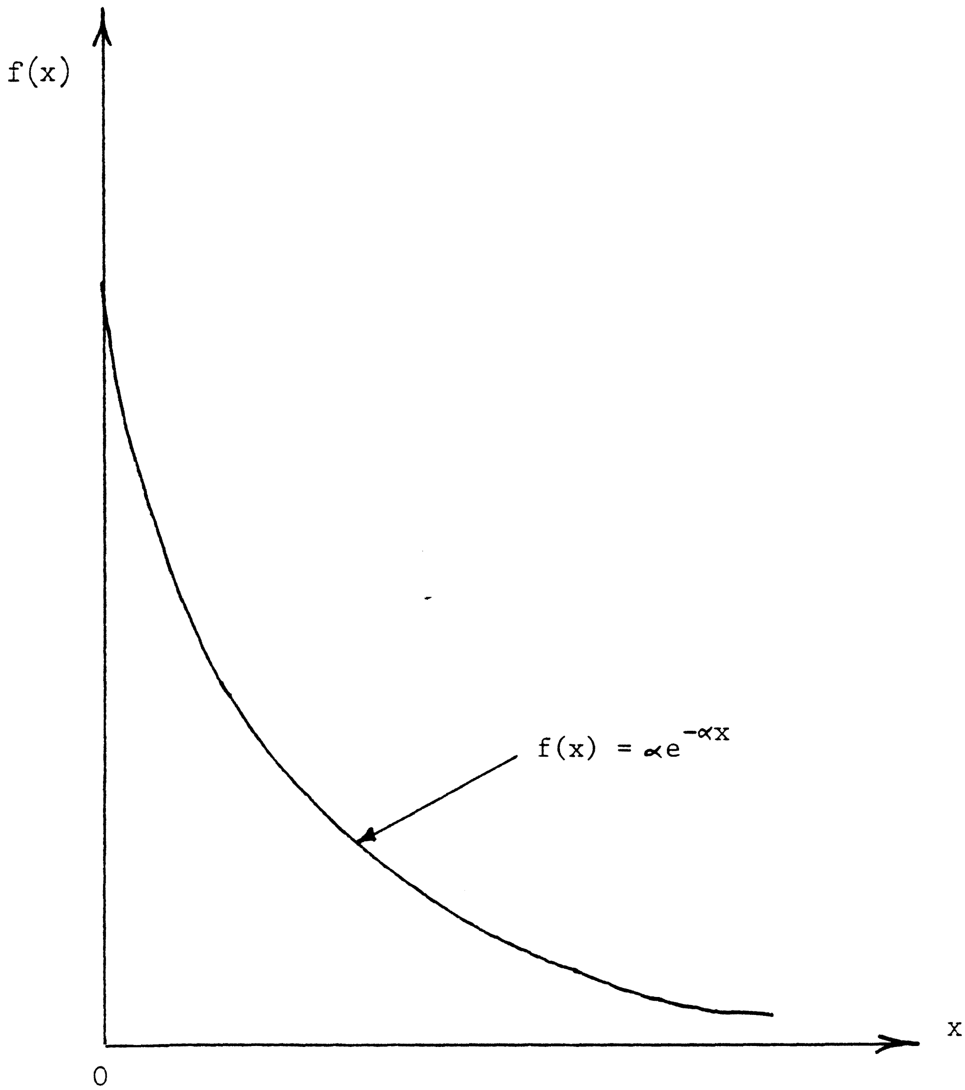

Fig. 8 - pdf for exponential distribution

parameter Weibull is used. It is found that c varies slightly from 1.0. The Weibull distribution is given by

$$f_Y(y) = c/b\ (y/b)^{c-1} \exp[-(y/b)^c] \quad 0 < y < \infty \tag{59}$$

where: y = random variable

b = scale parameter

c = shape parameter

For truncation, low flow Q is defined to be equal to or less than q_0 and thus the range of low flows is $0<q<q_0$, but Weibull ranges between $0<y<\infty$. Eq. 59 must be truncated so positive probability of Q only occurs between $(0,q_0)$. Also for partial duration series, as done in the recession model, the definition of return period is slightly different. Return period T, defines the time interval on the average that low flow Q will be less than or equal to a specified q_T. The mean number of flows per year below q_0, ξ, is calculated by dividing the number of occurrences below q_0 by the number of years of record. Recalling from PDS the time interval between the low flows is not a year, but there will be ξT events in T years in which q_T occurs once. Then $P[Q \le q_T | q<q_0]$ for PDS is defined as (see Cunane, 1979)

$$P[Q \le q_T | q<q_0] = 1/(\xi T) \tag{60}$$

PARTIAL DURATION SERIES BY TRUNCATION FOR DEVIATIONS

The probability for the deviations, $X=q_0-Q$ is given by $P[Q \leq q_T|q<q_0] = P[X \geq q_0-q_T|q<q_0]$. The pdf for X is

$$f(q) = \{(c/b)(x)^{c-1} \exp[-(x/b)^c]\}/F(q_0) \quad 0<x<q_0 \tag{61}$$

where:

$x = q_0-q$ = deviation

q_0 = cut-off flow

q = low flow

$F(q_0)$ = probability of non-exceedance of q_0

$$= \int_0^{q_0} (c/b)(x)^{c-1} \exp[-(x/b)^c]\, dx \tag{62}$$

$$= 1 - \exp[-(q_0/b)^c] \tag{63}$$

Dividing by $F(q_0)$ in Eq. 63 normalizes Eq. 59, then the pdf has positive probability only for $0<Q<q_0$. Parameters are calculated by maximum likelihood estimation (MLE). The likelihood function is given as

$$L_1 = \prod_{i=1}^{n} [(c/b)(x)^{c-1} \exp[-(x/b)^c]]/F(q_0) \tag{64}$$

Then L_1 is transformed by the natural logarithm, L and differentiated with respect to the parameters b and c.

$$\partial L/\partial b = (nc)[q_0{}^c/b^{c+1}]\exp[-(q_0/b)^c]/\{1-\exp[-(q_0/b)^c]\}$$

$$- (nc)/b + (c/b)\sum_{i=1}^{n}(x_i/b)^c \qquad (65)$$

$$\partial L/\partial c = -n(q_0/b)^c \ln(q_0/b)\exp[-(q_0/b)^c]/\{1-\exp[-(q_0/b)^c]\}$$

$$+ n/c - n\ln(b) + \sum_{i=1}^{n}\ln(x_i)$$

$$- (x_i/b)^c\ln(x_i/b) \qquad (66)$$

Maximum likelihood estimation leads to complex equations with no closed form solutions. An iterative scheme is applied to MLE for parameter solution. Parameter solution is discussed in the Results of Truncation Process section.

The cumulative probability of Q is given by integrating the pdf from 0 to q_T to compute flow q_T with return period T.

$$F(q_T) = \int_0^{q_T}\{c/b(q_0-q)^{c-1}\exp\{-[(q_0-q)/b]^c\}\}/F(q_0)\,dq \qquad (67)$$

$$= \{\exp\{-[(q_0-q_T)/b]^c - \exp[-(q_0/b)^c\}/F(q_0) \qquad (68)$$

$F(q_T)$ is the $P[Q \le q_T|q<q_0]$ given by Eq. 60. Since partial duration series is not for fixed time intervals the return period is modified by calculating the expected low flow events which are less than q_0 in period T, by ξT. For period

T, flow q_T will be equal or less once on the average in period T in years which is the same as once in ξT events and thus

$$P[Q \leq q_T | q < q_0] = 1/(\xi T) \tag{69}$$

where: ξ = mean number of low flow events below q_0 per year

A closed form solution for q_T is derived from Eqs. 60 and 68 as:

$$q_T = q_0 - b\{-\ln[(1/\xi T)+[1-(1/\xi T)]\exp[-(q_0/b)^c\}^{1/c} \tag{70}$$

PARTIAL DURATION SERIES BY TRUNCATION FOR FLOWS

The probability distribution of flows is derived the same as deviations except X = Q. MLE for the distribution of flows is set up the same as for the deviations. $P[Q \leq q_0]$ is the same as in Eq. 61 for the deviations, but the pdf is

$$f(q) = c/b\ (q/b)^{c-1}\exp[-(q/b)^c]/F(q_0) \tag{71}$$

A closed form solution for q_T is found by integrating Eq. 71 from 0 to q_T and solving for q_T

$$q_T = b\{-\ln[(1/\xi T)+[1-(1/\xi T)]\exp[-(q_0/b)^c\}^{1/c} \tag{72}$$

For solution of parameters b and c for deviations and flows, an iterative program applying MLE is used (Zutter, et al. 1982). Initial values for b and c are estimated by linear regression, and solution of maximum likelihood estimates by an iterative process given by Wingo (1972). The process known as quasilinearization searches for parameters over a bounded solution space.

PARTIAL DURATION SERIES BY CENSORSHIP

Truncation defines a limit for low flow by setting a cut-off flow, whereas censorship type I uses the data given by truncation plus the number of low flow occurrences above the cut-off flow. This is censorship above a cut-off flow, but censorship below a cut-off also occurs, for example when stream level goes below the lowest gage mark which is not addressed in this study. Point A in Fig. 6 illustrates this requirement. Censorship deals with PDS as a complete distribution but not knowing the values of the censored flows except for the number exceedances. The censored points give the PDS more information about the low flows than the truncated PDS because low flows may occur above the cut-off flow q_0. Knowing the number of censored points, the parameters of the probability distribution can represent Q with unrestricted range, thus avoiding the upper limit of q_0. Censorship increases the data from truncation to a more representative sample of low flows. As before, the Weibull distribution is applied to low flow data, but the pdf is not altered because of censorship. The influence of censoring occurs in parameter estimation. MLE is used as the parameter estimator. The probability of a flow being above q_0 is given as probability of exceeding q_0.

$$P[Q \geq q_0] = 1 - F(q_0) = \exp[-(q_0/b)^c] \qquad (73)$$

Where $F(q_0)$ is given by Eq. 63. Then MLE is setup as

$$L_1 = [1-F(q_0)]^k \prod_{i=1}^{n-k} f(q_i) \quad (74) \qquad (74)$$

$$0 < q_i < q_0$$

$$= \exp[-(q_0/b)^c]^k \prod_{i=1}^{n-k} c/b(q_i/b)^{c-1} \bullet \exp[-(q_i/b)^c] \qquad (75)$$

where: n = total number of low flows

k = number of low flows $> q_0$

L is differentiated with respect to b and c and the resulting equations are solved for b and c by equating them to zero. A closed form solution occurs for b and is a function of c. An iterative scheme is required for c.

By defining the log likelihood function, $L = \ln(L_1)$

$$\partial L/\partial b = -n(b)^c + k(q_0)^c + \sum_{i=1}^{n} (q_i)^c \tag{76}$$

$$b = \{[k(q_0)^c + \sum_{i=1}^{n} [(x_i)^c/n]\}^{1/c} \tag{77}$$

$$\partial L/\partial c = -n/c - k(q_0/b)^c \ln(q_0/b) - n\ln(b) + \sum_{i=1}^{n} \ln(x_i) - \sum_{i=1}^{n}(x_i)^c \ln(x_i) \tag{78}$$

where: $\partial L/\partial b = \partial L/\partial c = 0$

In applying the censored PDS it is noted that the number of low flows per period T is different in censored then in truncated. Censored has k more low flows occurring than truncated, thus ξ in censored PDS is defined by ξ_1 as the number of low flows below q_0 plus the number above q_0 divided by the number of years of record. Thus for period T, ξ_1 low flows occur. It is noted ξ_1 is always greater than or equal to ξ

The complete Weibull distribution in Eq. 59 is integrated from 0 to q_T for return period T. The closed form solution for q_T is given as

$$q_T = b\{-\ln[1-(1/\xi_1 T)]\}^{1/c} \tag{79}$$

Flows occurring above q_0 are found by finding successive declining flows for ten consecutive days (arbitrarily set ten to avoid dependence on other censored points) and ending with a rise in flow due to storm runoff. All the flows in the succession must be greater than q_0. Application of censored PDS is done by program CENSOR found in Appendix D.

CHAPTER 6: COMPARISON OF METHODS

RECESSION MODEL

The one day low flows are computed for return periods of 2, 5, 10, 25, 50, and 100 years. The recession model results are compared with the results from Log Pearson type III, (LPIII). For the Rappahannock River, using thirty years of data from April, 1951 to March, 1981, the recession model ten year low flow (q_{10}) is 48.86 cfs, and LPIII is 28.85 cfs. The mean number of recessions/yr was 9.70. Thus in ten years, 97 recessions are expected, with q_{10} having an expected probability of 1/97, however, since 48.86 cfs is significantly different from 28.85 cfs. The expected 97 recessions may be intuitively correct, but the actual probability is different because the number of recessions in a given period T is also a random variable. The flow q_T varies with recessions/yr, Fig. 9 shows the sensitivity between q_{10} and recessions/yr. For comparison purposes q_{10} for the recession model is set equal to the q_{10} of LPIII. For this q_{10} its probability is calculated by the recession model and is equal $1/\mu^* T$, where μ^* is the modified recessions/yr giving a recession model q_{10} equal to that of LPIII. The other return periods are calculated using the modified recessions/yr, μ^*.

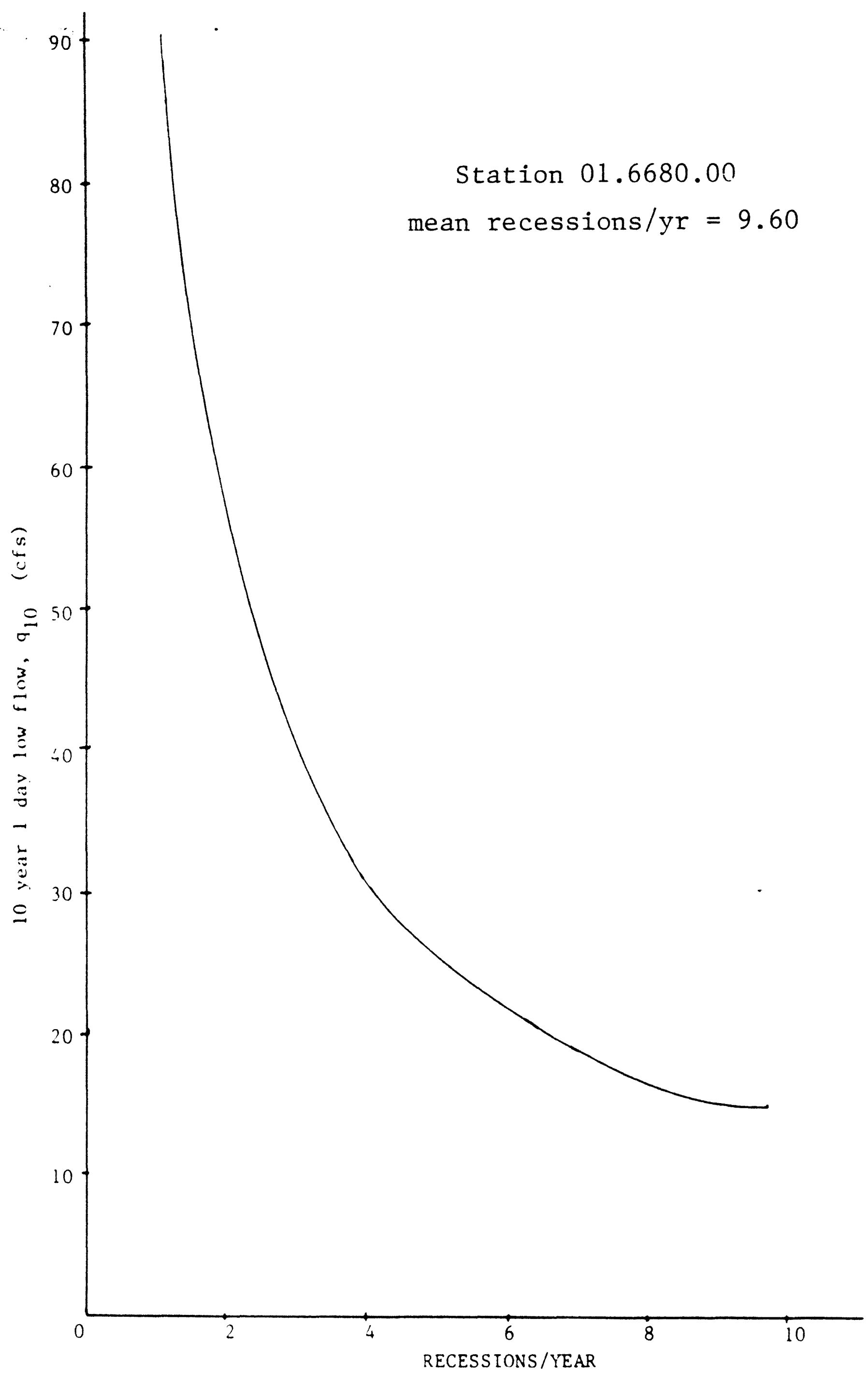

Fig. 9 - Sensitivity of q_{10} to recessions/yr

For Rappahanock River μ^* was 27.72. The results from calibrating μ to μ^* are

Calibration of Recessions/yr (μ)

Low Flows (cfs)

Return Period	Recession μ=9.60	Recession μ^*=27.72	LPIII
2	115.68	65.46	139.36
5	70.38	40.81	55.20
10	48.86	28.85*	28.85*
25	30.86	18.48	12.53
50	22.01	13.36	6.72

* flow used for calibration

Modifying μ decreases the flows for return periods less than ten years, but return periods higher than ten years have flows closer to the LPIII estimates than from the initial μ. However, this is not a general pattern for all streams.

The James River also has similar results. The recession model predicts a q_{10} of 190.00 cfs, while LPIII had 15.71 cfs. Unfortunately the problem of highly varying low flows for several years has occurred. Fourteen years have flow over 100 cfs with a high of 2050 cfs, while six years had low flows of 10 to 20 cfs. This data is more biased than the

Rappahannock stream data for use with LPIII, because some years of very low flow, some stream flow was diverted thus flows were not representative of recession from ground water, or true low flows for annual minimum series. The parameter μ was also modified and the results are in Appendix C.

The other thirteen stations were applied using the computer application of baseflow separation. Calibration of recessions/yr was also done on these stations. For the perennial streams, four of the stations had μ^* less than 4.0 while four had μ^* greater than 20.0, but two stations had μ^* approximately equal to μ. In the intermittent streams only two stations converged in the IMSL subroutine for numerical integration. One stream had a large μ^*, but the other stream had μ^* approximately the same as μ. More than half of the stations showed good comparison between recession model and other methods.

Application of the recession model shows promise even though results may be difficult to verify for precision since difficulties occur in data collection and estimation of μ. Separating base flow and runoff can be difficult, as straight line without deviations on the plot of $\ln(Q)$ versus time does not occur often. Generally a pattern occurs where between two distinct peaks that exponential recession occurs. The main problem is calculating positive recharge, and several remedies are applicable if $Q(0)$ is less than $Q(D)$, however, arbitrary judgement in some cases is required.

Another assumption is that Q(0) is dependent on the previous event of a drought which is dependent on the beginning of runoff from a storm. However, difficulties arise in differentiating what is a single rainfall event and several events. Many times several storms occur before recession takes place, so Q(0) may lose its relation with the previous recession due to the effect of several storms. Also, the stream may be affected by outside interference such as regulation of water (such as upstream reservoirs, water removal for irrigation) or local disturbances near the stream station. Another source of data inconsistencies is the actual determination of flow using approximate methods, for example gage heights with a rating curve. This is especially true for intermittent streams when trying to estimate flows between 0 and 1 cfs. Possibly to clarify events of storm runoff and recession, the hydrograph can be compared with rainfall records. Also using hourly flow records instead of daily flow will improve the accuracy because the start of recession will be clearer. Though distractions are possible, exponential recession is sound. All stations under study showed behavior of exponential decay following a peak from a storm. A possible explanation of slight deviations from the straight line of recession on the plot, is the occurrence of short rain showers or release of detained water which add no real contribution to groundwater.

Most difficulties occurred in the summer periods, July to September. Irregular fluctuations occurred in the hydrograph, as sharp increases and decreases in the flow possibly due to flash rain storms and removal of stream flow for man's purposes, the chief problem is that the value of K becomes very small and the receding side of the hydrograph is a steep curve.

The assumption of a constant recession coefficient, K, does not affect the results significantly. A deviation of +/- 1.00 day^{-1} will affect the calculated for a given return period slightly. Also K is found to be smaller in the summer than non-summer periods. The reservoir coefficient, K, is actually related to the time of year and groundwater supply, as Delleur and Kavvas (1984) found recession to be dependent on time from peak and storage state of the watershed. A correlation could be derived to have K vary with time and streamflow, or drought analysis could be conducted only on the summer periods and use a generalized K for the summer period. This would be appealing since droughts normally occur during the summer period, but this may not be practicable due to the erratic behavior of the summer flows. The use of an optimization scheme to calculate K given by McCormick (1984) seems reasonable, and it is more representative than using a mean value of K.

The only major drawback of the model is in determining the number of recessions per year. Using the mean number of recessions from the known selected recessions and estimating the number of possible recessions in return period T leads to underestimation or overestimation of flows for most streams. To calibrate the recession model, the ten year flow for LPIII is set as the ten year flow for the recession model. For the given flow the number of recessions/year can be found and then applied for the other return periods. In general this works well for prediction purposes, but at most stations the recessions/yr was less than 1.0 or was larger than 20. Over looking the calibrating by the ten year LPIII flow, the recession model works well for streams that show definite variations in flows between return periods. For streams with steady flow and small variations between flows for different return periods by LPIII, the recession model deviates significantly.

LOG PEARSON TYPE III

LPIII with a point mass, mixed LPIII, gives consistent results in predicting low flows for intermittent streams. These results were compared with Double Bounded pdf and LPIII by adding a small constant to the flows. No regular pattern of under or over estimation was found by using the point mass compared with the unconditional LPIII. However, to implement the point mass for $P(Q=0)$ the data should contain at least 20 or more years, to avoid a biased estimator of $P(Q=0)$. For long stream records of intermittent streams, both the unconditional LPIII and mixed LPIII with a point mass should be used. The advantage of the point mass probability is that the return period for zero flow can be roughly estimated, while other methods fall short in this aspect.

LPIII also gives good results for perennial streams. For return periods under 50 years, results showed good consistency, but variations occur for LPIII and other methods for long return periods. Also LPIII is sensitive to outliers especially if the stream record is short, because only one low flow is used for each year in the record.

Kolmogorov-Smirnov, (K-S), a goodness-of-fit test was applied to LPIII. K-S concentrates on the deviations between the hypothesized cumulative distribution function and the observed cumulative histogram. All stations passed the K-S test at level of significance of 0.01.

DOUBLE BOUNDED PDF

DBPDF was applied to perennial and intermittent streams. The results of intermittent streams are compared with those of mixed LPIII.

For large perennial streams the results are inconsistent. For streams with a derived lower bound, A equaling zero, the results compare favorably with LPIII. However, the other stations resulted in $A<0$. This occurred since DBPDF is a curve fitting technique. Unfortunately, this causes negative flows to have a positive probability of occurrence which is unacceptable. Generally for the return periods of 2 and 5 years the results are consistent, but for return periods of 10 years and above (especially 50 and 100 years) the results are extremely under estimated causing negative flows at two stations. For q_T's less than zero, they have been set to zero. Problems due to the tail of DBPDF converging to zero may cause the underestimation of flows.

Intermittent streams and small perennial streams gave much better and consistent results than larger perennial streams. For a good size data set (at least 20 years) DBPDF gives good estimates for all return periods (generally for 50 and 100 return periods, q_T is zero for intermittent streams). DBPDF compares favorably with LPIII and mixed LPIII with a point mass probability for zero flow. It is noted $P(Q=0)=F_0$ for mixed LPIII.

K-S test was applied for goodness of fit, and all stations passed for level of significance of 0.01. However, this may be due to small sample sizes used which give large allowable deviations.

PARTIAL DURATION SERIES

In the application of Partial Duration Series, q_0 is set to the largest annual minimum in the flow record, so that no year will be excluded from the series data. This makes the data compatible with other methods, that is all flows from the minimum annual series will be included in the partial duration series data. Partial Duration Series showed that two different sample populations are possible, truncated and censored.

Unfortunately, for X=Q in truncated PDS convergence was not obtained for parameter estimation, probably due to no general pattern fitting the maximum likelihood estimation equations. In contrast, deviation's parameter estimation converged for all stations, but zero data points are set to small increments to allow for convergence of the iterative scheme. The majority of the stations have the scale parameter c close to 1.0, and a few deviated from 1.0 by not more than ± 0.3. In most cases Weibull gave better prediction than exponential, though some stations results varied slightly from other methods possibly due to more data used. Generally for perennial streams truncated PDS predicts lower values for q_T likely due to the restrictive range of $0<Q<q_0$., but for intermittent streams PDS predicts slightly higher values for q_T than annual series methods. This may indicate the difficulty of Weibull to fit the extreme values of the stream's actual distribution. For more consistent results a three parameter Weibull should be applied, but may lead to more difficulty in parameter estimation.

Chi square goodness-of-fit test was applied to the truncated PDS. Program Weibull gives the chi-square deviate which is compared with the chi-square statistic for level of significance of 0.01. All stations passed, but for large sample sizes the allowable chi-square deviate is large thus not a rigorous test.

Censored PDS compared favorably with other methods such as LPIII and recession model with calibration. The few stations having inconsistent results over estimated q_T compared to the other methods. Differences may occur because of the data since PDS uses more flows to represent the low flow. Also ξ and ξ_1 are assumed constant, however the number of low flows occurring in period T is a random variable. Also, there is difficulty of finding the low flows above the cut-off flow and the inflexibility of two parameter Weibull distribution. For PDS to be applicable its results should be compared with other methods to check for any deviations.

K-S test was applied, and two perennial and two intermittent streams failed for level of significance of 0.01. For large sample sizes the K-S statistic is small thus a station might fail but still give good results.

CHAPTER 7: CONCLUSIONS

The recession model behaves as a low flow estimator given that the stream shows exponential recession. Most of all, the recession model physically represents low flows as exponential decay incorporating the drought period which the other methods do not include. Also, using time series of flows avoids the restriction of independence between flows. However, recession model did not work well for intermittent streams which are highly varying, and may draw water outside the supply of groundwater, contradicting the model assumptions. For consistent results for all streams, determining the probability for a return period needs to be defined more clearly.

Log Pearson type III in the theoretical application is restrictive, but the use of a frequency factor by an approximation scheme showed to be acceptable. Log-Pearson type III works well for streamflows because of its flexible curve fitting, but it is sensitive to outliers in the data. Mixed LPIII has better theoretical basis for intermittent streams than the unconditional LPIII which transforms the data by a constant.

Double Bounded pdf was found to give acceptable results for intermittent streams and small perennial streams, but generally underestimates flows for large perennial streams.

Problems due to the tail of DBPDF converging to zero quickly may cause the underestimation of flows. A drawback of DBPDF is when the lower bound, A is less than zero but that is physically unacceptable.

The advantage of Partial Duration Series is more than one flow per year is used which is beneficial if outliers in annual minimum series occur. Since truncated partial duration series had parameter c close to 1.0, this shows that $X=q_0-Q$ is approximately exponentially distributed, but c may vary to either side of 1.0. However, truncated PDS is restricted to the range of $0<Q<q_0$ which is a hindrance. Censored PDS is a more general case of PDS which makes censored PDS perform more consistently than truncated PDS.

BIBLIOGRAPHY

1. Abramowitz, M., and I. A. Stegun (1972) Handbook of Mathematical Functions. New York: Dove Publications Inc.

2. Aitchison, J. (1955) On the Distribution of a Positive Random Variable Having a Discrete Probability Mass at the Origin. Journal of the American Statistical Association, 50, pp. 901-908.

3. Aron, Gert, and J. Borrelli (1973). Stream Baseflow Prediction by Convolution of Antcedent Rainfall Effects. Water Resources Bulletin, 9, No. 2, pp. 360-365.

4. ASCE Task Committee on Low-Flow Evaluation, Methods, and Needs. (1980) Characteristics of Low Flows. Journal of the Hydraulics Division, ASCE, 105, No, HY5, pp. 717-731.

5. Ashkar, Fahim and Jean Rousselle (1981) Design Discharge as a Random Variable: A Risk Study, Water Resources Research, 17, No. 3, pp. 577-591.

6. Benjamin, J. R., and C. A. Cornell (1970) Probability, Statistics, and Decisions for Civil Engineers. New York: McGraw-Hill.

7. Bethlahmy, Nedavia. (1977) Flood Analysis by SMEMAX Transformation. Journal of the Hydraulics Division, ASCE, 103, No. HY1, pp. 69-78.

8. Birtles, A. B. (1978) Identification and Separation of Major Base Flow Components From a Stream Hydrograph. Water Resources Research , 14, No. 5, pp. 791-803.

9. Boughton, Walter (1980) A Frequency Distribution for Annual Floods, Water Resources Research, 16, No. 2, 1 pp. 347-354.

10. Boughton, Walter, and Edward Shirley. (1983) Fitting Procedures for the Log-Boughton Distribution. Journal of the Hydraulics Division , ASCE, 109,No. 4, pp. 579-589.

11. Box, G.E.P. and D.R. Cox (1964) An Analysis of Transformation. Journal of the Royal Statistical Society, B 26, pp. 211-252.

12. Brutsaert, Wilfred and John Nieber (1977) Regionalized Drought Flow hydrographs from a Mature Glaciated Plateau, Water Resources Research, 13, No. 3, pp. 637-643.

13. Carlslaw, H.S. and J. C. Jaeger (1959) Conduction of Heat in Solids, New Jersey: Oxford University Press.

14. Chow, Ven Te (1964) Hankbook of Applied Hydrology., New York : McGraw-Hill.

15. Cohen, A. (1965) Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and on Censored Samples. Technometrics, 7, No. 4, pp. 579-588.

16. Collins, Michael (1967) The Extended Boussinesq Problem. Water Resources Research, 12, No. 1, pp. 54-56.

17. Condie, Robert (1979) Drought Frequency Analysis for Intermittent Streams. International Symposium on Hydrologic Aspects of Droughts, New Delhi, India, pp. 242-248.

18. Crank, J. (1956) The Mathematics of Diffusion, New Jersey : Oxford University Press.

19. Cunnane, C. (1973) A Particular Comparison of Annual Maxima and Partial Duration Series Methods of Flood Frequency Prediction. Journal of Hydrology, No. 18, pp. 257-271.

20. Cunnane, C. (1978) Unbiased Plotting Position - A Review. Journal of Hydrology, No. 3, pp. 205-222.

21. Cunnane, C. (1979) A Note on the Poisson Assumption in Partial Duration Series. Water Resources Research, 15, No. 2, pp. 489-493.

22. Delleur, J.W. and M. L. Kavvas (1984) A Statistical Analysis of the Daily Stream flow Hydrograph. Journal of Hydrology, 71, pp. 253-275.

23. Eagleson, Peter (1970) Dynamic Hydrology, New York: McGraw-Hill.

24. Glover, Robert (1966) Use of the Mathematical Models in Drainage Design. Trans. of the ASAE, 2, pp. 210-212.

25. Hall, Francis (1968) Base Flow Recessions - A Review. Water Resources Research, 4, No. 5, pp. 973-983.

26. Hall, M. J. (1984) Urban Hydrology, New York: Elsevier, pp. 103-105.

27. Jackson, D.R. (1981) WRC Standard Flood Frequency Guidelines. Journal of Water Resources Planning and Management Division, 107, No. WR1, pp. 211-225.

28. Jennings, M. E., AND M. A. Benson (1969) Frequency Curves for Annual Flood Series with Some Zero Events or Incomplete Data. Water Resources Research, 5, No. 1, pp. 276-280.

29. Johnson, N. L., and S. Kotz (1970) Continuous Univariate Distributions-I Distributions in Statistics. John Wiley, New York.

30. Kelman, Jerson (1980) A Stochastic Model for Daily Streamflow. Journal of Hydrology, 47, pp. 235-249.

31. Kendall, M., and A. Stuart (1979) The Advanced Theory of Statistics. London 4 High Wycombe, Charles Griffin & Company Ltd.

32. Kite, G. W. (1977) Frequency and Risk Analysis in Hydrology Fort Collins: Water Resources Publications.

33. Kottegoda, N. T. (1980) Stochastic Water Resources Technology , London: Mc Millan Press Ltd.

34. Kruskal, W. H. and J. M. Tanur (1978) International Encyclopedia of Statistics, New York: The Free Press.

35. Kumar, Arun, and Rema, Devi (1982) Statistical Determination of Design Low Flows - A Comment. Journal of Hydrology, 58, pp. 175-177.

36. Kumaraswamy, P. (1980) A Generalized Density Function for Double- Bounded Fandomws Processes. Journal of Hydrology, 46, pp. 79-88.

37. Leese, Morven (1971) Use of Censored Data in the Estimation of Gumbel Distribution Parameters for Annual Flood Series. Water Resources Research, 9, No. 6, pp. 1534-1542.

38. Linsley, R. K., and M. A. Kohler and J. L. H. Paulhus (1982) Hydrology for Engineers, New York: McGraw-Hill.

39. Matalas, Nicholas (1963) Probability Distribution of Low Flows , U.S.G.S. Professional Paper 434-A, Washington, D. C.

40. McCormick, Timothy (1984) Frequency Analysis of Low Flows. M.S. Thesis, Virginia Tech.

41. McMahon, T. C. and A. D. Arenas (1982) Mehtods of Computaion of Low Stream flow., Paris: UNESCO.

42. McWhorter, David (1977) Drain Spacing Based on Dynamic Equilibrium. Journal of the Irrigation and Drainage Division, 103, No. IR2, pp. 259-271.

43. Naney, James et al. (1978) Predicting Base Flow Using Hydrogeologic Parameters. Water Resources Bulletin, 14, No. 3, pp. 640-649.

44. Prakash, Anand (1981) Statistical Determination of Design Low Flows. Journal of Hydrology, 51, pp. 109-118.

45. Raheed, Hameed, Marur Ramamoorthy, and A. S. Al-Dabbagh (1982) Modified SMEMAX Transformation for Frequency Analysis. Water Resources Bulletin, 18, No.3, pp. 509-511.

46. Shane, Richard M. and Walter Lynn (1964) Mathematical Model for Flood Risk Evaluation. Journal of the Hydraulics Division, ASCE, HY6.

47. Singh, Krishan P. and John Stall (1971) Derivation of Base Flow recession Curves and Parameters. Water Resources Research, 7, No. 2, pp. 292-303.

48. U. S. Water Resources Council, 1977 Guidelines for Determining Flood Flow Frequency, Bulletin 17A, Washington, D. C.: GPO.

49. Westphal, Jerome (1984) SMEMAX? Caution! Journal of Hydraulic Engineering, ASCE, 110, No. 3, pp. 209-218.

50. Wingo, D. R. (1972) Maximum Likelihood Estimation of the Parameters of the Weibull Distribution by Modified Quasilinearization. IEEE Trans. Rel., Vol. R-21(2), pp. 89-93.

51. Zutter, Bruce et al. (1982) WEIBUL A Program to Estimate Prameters of the Weibull Distribution Using Complete, Censored, and Truncated Data, Pub. No. FWS-3-82, School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University.

APPENDIX A: FREQUENCY FACTOR IN LOG-PEARSON TYPE III

This appendix presents a derivation for the frequency factor K for the log Pearson type III distribution. That is Pearson type III is fit to the logarithm of low flows.

Consider the Pearson type III distribution

$$f_X(x) = \{|a|/\Gamma(b)\}\ [a(x-c)]^{b-1}\ \exp[-a(x-c)]$$

$$\text{if } a > 0 \quad c \le x < \infty$$

$$\text{if } a < 0 \quad -\infty < x \le c \qquad \text{(A1)}$$

where: x = natural log transformed variate

Let $$Y = 2a(X-c) \qquad \text{(A2)}$$

and therefore $$X = Y/(2a) + c \qquad \text{(A3)}$$

also the Jacobian $$|J| = |1/(2a)| \qquad \text{(A4)}$$

Therefore the pdf of Y is given as

$$f_Y(y) = 1/[2^b\Gamma(b)]\ \ y^{b-1}\ \exp(-y/2) \qquad \text{(A5)}$$

$$\text{for } 0 < y < \infty$$

Which is a χ^2 distribution with 2b degrees of freedom.

Considering the following example

$$\alpha = P[X \le x] = P[Y/(2a)+c \le x] \qquad \text{(A6)}$$

$$= P[Y \le (x-c)2a] \tag{A7}$$

Because Y has a χ^2 distribution

$$(x-c)2a = \chi^2(\alpha,2b) \tag{A8}$$

$$x = \chi^2(\alpha,2b)/(2a) + c \tag{A9}$$

Also, based on the definition of frequency factor

$$x = \mu + K(\alpha)\sigma \tag{A10}$$

where: μ=mean of logarithm of low flows

σ=standard deviation of logarithm of low flows

Based on method of moments (Kite, 1977, pp. 105-107) the parameters of the Pearson type III distribution can be written as

$$c = \mu - \sigma b^{\frac{1}{2}} \tag{A11}$$

$$a = b^{\frac{1}{2}}/\sigma \tag{A12}$$

$$b = (2/g_1)^2 \tag{A13}$$

where: g_1=skew coefficient

By using Eqs. A11-A13 in Eq. A9

$$x = [\chi^2(\alpha,2b)\sigma]/(2b^{\frac{1}{2}}) + \mu - \sigma b^{\frac{1}{2}} \tag{A14}$$

$$= \mu + \sigma[\chi^2(\alpha,2b)/(2b^{\frac{1}{2}}) - b^{\frac{1}{2}}] \tag{A15}$$

Comparing Eqs. A10 and A14

$$K(\alpha) = \chi^2(\alpha,2b)/(2b^{\frac{1}{2}}) - b^{\frac{1}{2}} \qquad \text{(A16)}$$

By using Eq. A13 in Eq. A16

$$K(\alpha) = \chi^2(\alpha,2b)g_1/4 - 2/g_1 \qquad \text{(A17)}$$

This shows the frequency factor K for Pearson type III can be computed from the χ^2 tables and K is a function of return period T (=1/α) and the skew coefficient g_1. Consider the example:

α = 0.1 and g_1 = 1.0. Using the given data:

T = 10 years, b = 4, therefore degrees of freedom = 8
and $\chi^2(0.1,8)$ = 3.49 and thus K = -1.1275

The following approximation to the chi square distribution (Kendall and Stuart, 1969, pp. 371-373) is found to be very good.

$$P[Y \le \chi^2(\alpha,\nu)] \simeq \Phi\{[\chi^2(\alpha,\nu)/\nu]^{1/3} - 1 + 2/(9\nu)](9\nu/2)^{1/2}\} \qquad \text{(A18)}$$

where: Φ = standard normal pfd
$Y = \chi^2$ random variable with ν degrees of freedom

Therefore

$$Z = \{[Y/2b]^{1/3} - 1 + 1/(9b)\}3b^{1/2} \qquad \text{(A19)}$$

is approximately normally distributed with mean zero and unit variance. That is

$$[Y/(2b)]^{1/3} \sim N[1/(9b)-1, 1/(9b)] \qquad \text{(A20)}$$

Therefore it follows that chi square cutoff points can be estimated from the standard normal cutoff points. Thus,

$$\alpha = P[Z \leq z] = P\{Y \leq [z/(3b^{\frac{1}{2}})+1-1/(9b)]^3 2b\} \qquad \text{(A21)}$$

Therefore

$$\chi^2(\alpha, 2b) = [z_\alpha/(3b^{\frac{1}{2}})+1-1/(9b)]^3 2b] \qquad \text{(A22)}$$

Continuing with the example data $z_{.1} = -1.28$,

$$\chi^2(.1, 8) = 3.496$$

Therefore by using Eqs. 25 and 16 in Eq. 20 the frequency factor K for Pearson type III can be written as

$$K(\alpha) = [z_\alpha g_1/6+1-g_1{}^2/36]^3\ 2/g_1 - 2/g_1$$

$$= 2/g_1\{[(z_\alpha - g_1/6)g_1/6+1]^3 - 1\} \qquad \text{(A23)}$$

For $\alpha=0.2$ $z_\alpha=-0.84$ and $g_1=-1$: $K=-0.7517$

For $\alpha=0.2$ $z_\alpha=-0.84$ and $g_1=-9$: $K= 0.2222$

It is noted that Eq. A23 is precisely the same as Eq. 35.

APPENDIX B: METHOD OF MAXIMUM LIKELIHOOD

In the following the meaning of the maximum likelihood estimator is explained.

Suppose a random sample of variables $X_1, X_2, \ldots, X_n$ having the pdf $f(x|\theta)$ with the parameter θ. For any observed sample vector the joint pdf is denoted by $f_n(x|\theta)$. When $f_n(x|\theta)$ is regarded as a function of θ for a given sample vector it is called the likelihood function.

$$L = f_n(x|\theta) = \prod_{i=1}^{n} f(x_i)$$

If an estimate of θ must be selected, no value of θ should be considered which would make it impossible to obtain the sample vector that was actually observed. Furthermore, suppose that the probability $f_n(x|\theta)$ of obtaining the actual observed vector is a maximum when θ has a particular value, $\theta=\theta_0$, and is smaller for every other value of θ. Then it is natural to estimate the value of θ to be θ_0. For this reason for any observed sample to consider a value of θ for which the likelihood function $f_n(x|\theta)$ is a maximum and to use this value as an estimate of θ

If $f_n(x|\theta)$ is at a maximum, the derivatives of L with respect to the parameters should be zero.

$$\partial L/\partial \theta_i = 0 \quad \text{for all } i$$

These derivatives are easy to evaluate if the logarithm of the likelihood function could be used instead of the likelihood function itself. In such cases a log likelihood function is used.

However, for certain samples the maximum likelihood estimator (MLE) may not converge, and a MLE of θ may not exist. Also $f_n(x|\theta)$ may be a maximun at more than one point. For such a case MLE of θ is not unique, and any of these points may be chosen as the estimate of θ In many applications the MLE exists as a unique estimate, and gives a good estimate of θ. Also, MLE's are asymptotically efficient, i.e. minimum variance.

APPENDIX C: LOW FLOW ANALYSIS RESULTS

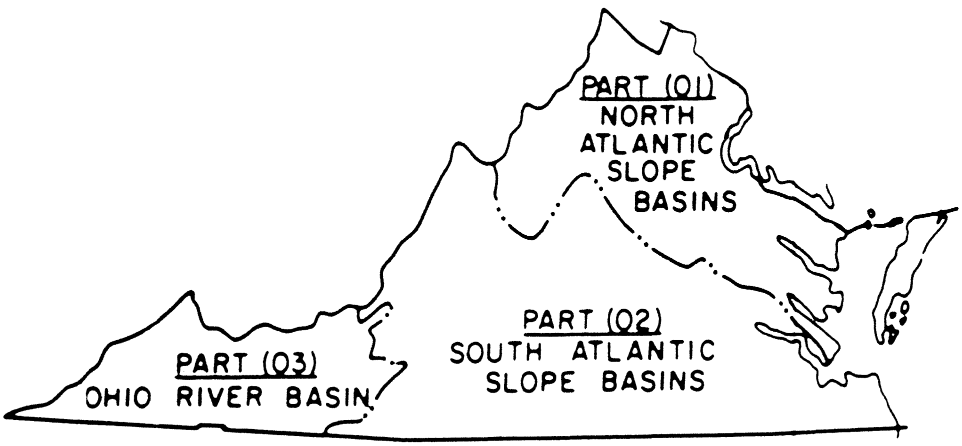

Table 2 - 1 Day Low Flows (cfs)

Perennial Streams

2 Year Return Period

STATION	RECESSION MODEL	LOG-PEARSON TYPE III	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	PARTIAL DURATION SERIES CENSORED
1.6260.00	116.92	29.32	32.38	26.29	26.73
1.6680.00	65.46	139.36	167.51	122.51	117.73
2.0157.00	103.37	28.32	31.21	27.64	26.79
2.0350.00	1229.69	1024.74	1137.27	881.75	966.77
2.0375.00	32.07	296.58	612.53	517.53	276.22
2.0545.00	101.10	40.04	39.16	32.03	34.17
3.1675.00	127.40	131.05	144.27	102.66	117.95
3.1680.00	1045.01	811.66	900.17	632.88	762.25 *
3.1700.00	101.29	97.92	111.35	65.46	88.41 *
3.2085.00	3.23	5.65	6.56	3.44	4.15

* Failure of K-S goodness of fit test at 0.01 significance level

Table 3 - 1 Day Low Flows (cfs)

Perennial Streams

5 Year Return Period

STATION	RECESSION MODEL	LOG-PEARSON TYPE III	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	PARTIAL DURATION SERIES CENSORED
1.6260.00	38.72	24.01	21.59	19.11	21.25
1.6680.00	70.38	55.20	52.13	63.01	79.57
2.0157.00	39.59	25.34	23.76	24.43	23.16
2.0350.00	740.70	665.80	634.62	529.94	740.41
2.0375.00	21.29	48.86	67.82	239.04	141.43
2.0545.00	39.72	30.26	27.59	24.84	28.88
3.1675.00	105.19	102.94	92.91	79.71	98.67
3.1680.00	742.03	654.13	615.12	438.68	654.51 *
3.1700.00	76.00	72.38	63.85	41.12	70.98 *
3.2085.00	1.64	1.95	1.69	1.45	2.26

* Failure of K-S goodness of fit test at 0.01 significance level

Table 4 - 1 Day Low Flows (cfs)

Perennial Streams

10 Year Return Period

STATION	RECESSION MODEL	LOG-PEARSON TYPE III	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	PARTIAL DURATION SERIES CENSORED
1.6260.00	21.75	21.75	15.72	14.26	17.89
1.6680.00	28.85	28.85	17.13	35.22	59.30
2.0157.00	24.29	24.29	19.59	22.19	20.76
2.0350.00	519.70	519.70	420.80	329.93	605.98
2.0375.00	15.69	15.69	0.00	126.29	85.55
2.0545.00	23.76	23.76	21.40	19.70	25.45
3.1675.00	91.23	91.23	69.51	63.38	86.26
3.1680.00	588.85	588.85	470.32	309.47	583.67 *
3.1700.00	62.35	62.35	42.20	26.59	60.17 *
3.2085.00	1.00	1.00	0.65	0.74	1.43

* Failure of K-S goodness of fit test at 0.01 significance level

Table 5 - 1 Day Low Flows (cfs)

Perennial Streams

25 Year Return Period

STATION	RECESSION MODEL	LOG-PEARSON TYPE III	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	CENSORED
1.6260.00	11.31	19.66	9.75	18.82	14.26
1.6680.00	30.86	12.53	0.00	15.23	40.25
2.0157.00	14.11	23.46	15.26	19.43	17.97
2.0350.00	333.49	392.07	247.64	158.07	465.36
2.0375.00	10.56	4.00	0.00	52.33	51.83
2.0545.00	13.35	16.80	15.35	13.45	21.54
3.1675.00	75.56	80.52	46.34	48.18	72.25
3.1680.00	444.06	529.36	332.59	173.80	501.85 *
3.1700.00	49.14	53.53	23.23	13.31	48.38 *
3.2085.00	0.53	0.44	0.19	0.30	0.78

* Failure of K-S goodness of fit test at the 0.01 significance level

Table 6 - 1 Day Low Flows (cfs)

Perennial Streams

50 Year Return Period

STATION	RECESSION MODEL	LOG-PEARSON TYPE III	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	PARTIAL DURATION SERIES CENSORED
1.6260.00	7.21	18.46	6.21	5.64	12.02
1.6680.00	22.01	6.72	0.00	7.83	30.04
2.0157.00	9.75	23.05	12.65	17.45	16.11
2.0350.00	242.05	323.57	166.50	85.10	381.21
2.0375.00	7.81	1.51	0.00	26.49	26.75
2.0545.00	9.05	12.70	11.96	9.36	18.99
3.1675.00	66.26	74.44	34.61	30.58	63.19
3.1680.00	364.02	495.66	256.59	102.77	447.73 *
3.1700.00	41.61	48.67	13.54	7.34	41.04 *
3.2085.00	0.32	0.25	0.08	0.15	0.50

* Failure of K-S goodness of fit test at 0.01 significance level

Table 7 - 1 Day Low Flows (cfs)

Perennial Streams

100 Year Return Period

STATION	RECESSION MODEL	LOG-PEARSON TYPE III	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	PARTIAL DURATION SERIES CENSORED
1.6260.00	4.70	17.47	3.32	3.34	10.13
1.6680.00	15.20	3.62	0.00	3.97	22.42
2.0157.00	6.89	22.76	10.50	15.55	14.45
2.0350.00	176.92	270.40	112.12	44.32	312.31
2.0375.00	5.82	0.59	0.00	13.33	16.22
2.0545.00	6.29	9.48	9.32	6.04	16.75
3.1675.00	58.17	69.47	25.89	19.92	55.27
3.1680.00	301.48	468.11	198.15	57.03	399.47 *
3.1700.00	35.54	44.79	6.54	3.88	34.80 *
3.2085.00	0.20	0.14	0.03	0.08	0.32

* Failure of K-S goodness of fit test at 0.01 significance level

Table 8 - 1 Day Low Flows (cfs)

Intermittent Streams

2 Year Return Period

STATION	RECESSION MODEL	LOG PEARSON TYPE III	LOG-PEARSON TYPE III W/ P(Q 0)	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	CENSORED
1.6560.00	0.04	0.47	0.54	0.91	0.80	0.26
1.6685.00	*	1.42	1.59	1.98	1.33	1.86**
1.6700.00	*	0.39	0.41	0.51	0.18	0.28**
2.0205.00	2.63	5.79	4.57	5.19	3.02	3.65
2.0440.00	*	0.82	0.78	1.03	0.50	0.59

5 Year Return Period

STATION	RECESSION MODEL	LOG PEARSON TYPE III	LOG-PEARSON TYPE III W/ P(Q 0)	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	CENSORED
1.6560.00	0.02	0.06	0.09	0.00	0.37	0.11
1.6685.00	*	0.18	0.16	0.13	0.54	1.05**
1.6700.00	*	0.08	0.11	0.06	0.07	0.16**
2.0205.00	1.10	1.86	1.86	1.61	1.28	2.19
2.0440.00	*	0.16	0.18	0.16	0.21	0.33

10 Year Return Period

STATION	RECESSION MODEL	LOG PEARSON TYPE III	LOG-PEARSON TYPE III W/ P(Q 0)	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	CENSORED
1.6560.00	0.01	0.01	0.00	0.00	0.20	0.05
1.6685.00	*	0.04	0.00	0.00	0.27	0.68**
1.6700.00	*	0.02	0.00	0.00	0.04	0.11**
2.0205.00	0.57	0.57	0.90	0.53	0.66	1.50
2.0440.00	*	0.05	0.06	0.02	0.11	0.21

* no convergance of IMSL subroutine for numerical integration

** Failure of K-S goodness of fit test at 0.01 significance level

Table 9 - 1 Day Low Flows (cfs)

Intermittent Streams

25 Year Return Period

STATION	RECESSION MODEL	LOG PEARSON TYPE III	LOG-PEARSON TYPE III W/ P(Q 0)	DOUBLE BOUNDED PDF	PARTIAL DURATION SERIES TRUNCATED	PARTIAL DURATION SERIES CENSORED
1.6560.00	0.00	0.00	0.00	0.00	0.08	0.02
1.6685.00	*	0.01	0.00	0.00	0.11	0.39**
1.6700.00	*	0.01	0.00	0.00	0.01	0.07**
2.0205.00	0.24	0.09	0.00	0.00	0.27	0.90
2.0440.00	*	0.01	0.00	0.00	0.04	0.12
50 Year Return Period						
1.6560.00	0.00	0.00	0.00	0.00	0.04	0.01
1.6685.00	*	0.00	0.00	0.00	0.05	0.25**
1.6700.00	*	0.00	0.00	0.00	0.01	0.05**
2.0205.00	0.13	0.02	0.00	0.00	0.13	0.62
2.0440.00	*	0.00	0.00	0.00	0.02	0.07
100 Year Return Period						
1.6560.00	0.00	0.00	0.00	0.00	0.02	0.01
1.6685.00	*	0.00	0.00	0.00	0.03	0.16**
1.6700.00	*	0.00	0.00	0.00	0.00	0.03**
2.0205.00	0.07	0.00	0.00	0.00	0.00	0.07
2.0440.00	*	0.00	0.00	0.00	0.01	0.05

* No convergence of IMSL subroutine for numerical integration

** Failure of K-S goodness of fit test at 0.01 significance level

Table 10 - Distribution Parameters

	Recession Model			Double Bounded pdf			
Station	K (days)	μ	μ^*	Z_{min} (cfs)	Z_{max} (cfs)	a	b
			Perennial Streams				
01.6260.00	13.6	8.71	0.70	-10.0	53.0	3.54	2.46
01.6680.00	15.4	9.70	27.72	-14.0	401.0	0.93	1.11
02.0157.00	11.8	8.30	0.66	0.0	43.0	3.73	1.92
02.0350.00	10.1	9.55	3.92	0.0	2350.0	1.76	2.12
02.0375.00	10.6	9.60	2411.00	-25.0	2050.0	0.45	0.78
02.0545.00	12.0	9.69	0.86	0.0	55.0	2.79	1.41
03.1675.00	28.8	5.05	20.19	0.0	276.0	2.40	2.92
03.1680.00	10.9	4.90	5.06	0.0	1440.0	2.69	2.09
03.1700.00	22.4	6.64	6.05	-12.0	244.0	2.18	3.03
03.2085.00	6.6	10.63	23.87	0.0	35.0	0.76	2.10
			Intermittent Streams				
01.6560.00	5.9	10.60	286.9	0.0	4.0	0.54	0.81
01.6685.00	10.2	10.44	*	0.0	16.0	0.54	1.49
01.6700.00	11.1	9.24	*	0.0	1.8	0.79	1.20
02.0205.00	7.4	12.09	13.32	0.0	13.0	0.95	1.19
02.0440.00	9.6	9.74	*	0.0	3.9	0.63	1.11

* No convergence for IMSL subroutine for numerical integration

Table 11 - Distribution Parameters

Station	q_0 (cfs)	Truncated PDS b	Truncated PDS c	Truncated PDS ξ	Censored PDS b	Censored PDS c	Censored PDS ξ_1
Perennial Streams							
01.6260.00	53.0	13.7	1.30	5.64	55.7	4.06	10.11
01.6680.00	401.0	127.7	0.98	6.77	449.5	2.37	12.23
02.0157.00	43.0	8.1	1.55	7.35	45.0	6.37	13.85
02.0350.00	2350.0	627.4	1.10	7.68	2454.4	3.48	13.04
02.0375.00	2050.0	831.1	0.83	6.35	2796.5	1.39	12.65
02.0545.00	55.0	10.1	1.12	6.20	62.7	5.51	14.36
03.1675.00	276.0	38.2	1.20	11.07	238.4	5.18	19.36
03.1680.00	1440.0	274.9	1.07	12.67	1330.9	6.08	15.07
03.1700.00	244.0	47.9	1.19	13.27	207.7	4.21	18.48
03.2085.00	35.0	24.2	1.42	8.50	35.7	1.53	13.59
Intermittent Streams							
01.6560.00	4.0	1.9	1.24	6.50	6.1	1.03	13.00
01.6685.00	16.0	30.6	0.75	10.70	14.4	1.62	14.20
01.6700.00	1.8	3.5	0.73	9.20	1.6	1.80	12.10
02.0205.00	13.0	22.0	0.65	4.30	20.6	1.82	11.90
02.0440.00	3.9	3.4	1.00	6.82	4.6	1.55	12.40

APPENDIX D: SELECTED COMPUTER PROGRAMS

```
C ******************************************************************
C *   PROGRAM MODD USES FORTRAN 77 TO FIND THE TR YEAR ONE DAY LOW  *
C *   FLOW BY APPLICATION OF A PHYSICALLY BASED APPROACH:           *
C *   EXPONENTIAL RECESSION MODEL BY DERIVED DISTRIBUTION           *
C ******************************************************************
        COMMON   Q,CT,CST,A1,B1,A2,B2
        REAL     X,P
        REAL     CQ(70),QO(70,70),CPQO(70,0:71)
        INTEGER  IER,N,IA,I,IJOB,IZ,IERR
        REAL     DCADRE
        REAL     F,A,B,CC,DD,AERR,RERR,ERROR,C
        EXTERNAL F
        REAL     PQO(70,70),TR(10),QE(10)
        REAL     D(70,70),WK(5040),SS(70)
        REAL     XX,AA,BB,PP
        COMPLEX  W(70),Z(70,70),ZN
C *******************************************************************
C  INPUT:     ISTA = STATION NUMBER (AS AN INTEGER)                   *
C              IY1 = FIRST YEAR OF DATA                               *
C              IY2 =  LAST YEAR OF DATA                               *
C              Q2  = ESTIMATE FOR   2 YEAR LOW FLOW (USE LPIII Q2)    *
C              Q5  = ESTIMATE FOR   5 YEAR LOW FLOW (USE LPIII Q5)    *
C              Q10 = ESTIMATE FOR  10 YEAR LOW FLOW (USE LPIII Q10)   *
C              Q25 = ESTIMATE FOR  25 YEAR LOW FLOW (USE LPIII Q25)   *
C              Q50 = ESTIMATE FOR  50 YEAR LOW FLOW (USE LPIII Q50)   *
C              Q100= ESTIMATE FOR 100 YEAR LOW FLOW (USE LPIII Q100)  *
C                                                                     *
C  INPUT PARAMETERS WHERE:                                            *
C                                                                     *
C            RY = AVERAGE NUMBER OF RECESSIONS PER YEAR               *
C            RK = RESERVOIR CONSTANT USED TO CALCULATE A AND B        *
C             A = ALPHA FOR BETA DISTRIBUTION                         *
C             B = BETA  FOR BETA DISTRIBUTION                         *
C           CST = INTEGRATION CONSTANT FOR THE BETA DISTRIBUTION      *
C            A2 = ALPHA FOR WEIBULL DISTRIBUTION                      *
C            B2 = BETA  FOR WEIBULL DISTRIBUTION                      *
C            UL = UPPERBOUND OF QO (INITIAL RECESSION FLOW)           *
C         TRLP3 = RETURN PERIOD TR, USED IN CALIBRATING THE RECESSION *
C                   MODEL'S RECESSIONS/YR                             *
C          QLP3 = FLOW FROM LPIII FOR RETURN PERIOD TRLP3 USED IN     *
C                   CALIBRATING THE RECESSION MODEL'S RECESSIONS/YR   *
C *******************************************************************
      READ(5,201) ISTA
      READ(5,201) IY1,IY2
      READ(5,202) Q2 ,Q5 ,Q10
      READ(5,202) Q25,Q50,Q100
      READ(5,202) RY,RK
      READ(5,202) A,B,CST
      READ(5,202) A2,B2
      READ(5,202) UL
      READ(5,202) TRLP3,QLP3
  201   FORMAT(2I10)
  202   FORMAT(3F10.0)
C ******************  RETURN PERIODS (YEARS)  **********************
      TR(1)=2
      TR(2)=5
      TR(3)=10
      TR(4)=25
      TR(5)=50
      TR(6)=100
C **************  INITIAL ESTIMATES FOR RECESSION MODEL  ************
      QE(1)=Q2
      QE(2)=Q5
      QE(3)=Q10
      QE(4)=Q25
      QE(5)=Q50
      QE(6)=Q100
```

```
      N=70
      A1=A-1
      B1=B-1
      AA=A
      BB=B
       WRITE(7,1)ISTA
C      WRITE(7,2)N
C ***********  DETERMINING POINTS FOR QO(N-1) AND QO(N) ************
        DO 100 I=0, N-1
      IF (I .LT. 15) THEN
       CT=3.5+7.0*I
       CQ(I+1)=CT
      ELSE
       CT=(UL-105)/(2*(N-15))+((UL-105)/(N-15))*(I-15.0)+105.0
       CQ(I+1)=CT
      END IF
        DO 110 K=0, N-1
      IF (K .LT. 15) THEN
       Q=3.5+7.0*K
       QO(I+1,K+1)=Q
      ELSE
       Q=(UL-105)/(2*(N-15))+((UL-105)/(N-15))*(K-15.0)+105.0
       QO(I+1,K+1)=Q
      END IF
C.....................................................................
C          DETERMINING THE LIMITS FOR THE INTEGRALS AND
C          EVALUATING THE LEFT HAND SIDE OF THE JOINT DISTRIBUTION
C              EQUALS 1.0 IF Q<C AND < 1.0 IF Q<C
C     MDBETA IS AN IMSL SUBROUTINE THAT EVALUATES THE CDF OF BETA
C.....................................................................
      IF (Q .GE. CT) THEN
       P=1.0
       DD=1.0
      ELSE
       X=Q/CT
       CALL MDBETA(X,A,B,P,IER)
       DD=Q/CT
      END IF
C ******** INTEGRATION OF RIGHT HAND SIDE OF JOINT DISTRIBUTION ****
C **** IMSL SUBROUTINE DCADRE INTEGRATES THE PROBABILITY FUNCTION ***
      CC=0.0
      DD=DD
      RERR=0.000
      AERR=0.00001
      C=DCADRE(F,CC,DD,AERR,RERR,ERROR,IER)
         IF (IER .GT. 66) WRITE(7,77)
         IF (IER .GT. 66) GO TO 106
C ********** CALCULATION OF P(QO(N) < Q| QO(N-1)=CT)* **************
C *****  CUMULATIVE TRANSITION PROBABILITIES FOR Q(O))=CT) **********
      PQ=P-C
      CPQO(I+1,K+1)=PQ
 110   CONTINUE
 100   CONTINUE
        DO 400 I=1,N
C     WRITE(7,35)CQ(I)
        DO 410 K=1,N
C     WRITE(7,31)QO(I,K),CPQO(I,K)
  410   CONTINUE
  400   CONTINUE
C ************** CALCULATION OF RELATIVE PROB   ****************
C **************  TRANSITION PROBABILITY OF Q(O) ***************
        DO 500 I=1, N
      CPQO(I,0)=0.0
        DO 510 K=1, N
      IF (K .LT. N) THEN
       PQO(I,K)=CPQO(I,K)-CPQO(I,K-1)
      ELSE
```

```
        PQO(I,N)=1.0-CPQO(I,N-1)
       END IF
 510     CONTINUE
 500     CONTINUE
          DO 600 K=1,N
C       WRITE (7,45)(PQO(I,K),I=1,N)
 600     CONTINUE
C ****** CALCULATION OF STEADY STATE PROBABILITIES OF Q(0) *********
C ****  EIGRF: IMSL SUBROUTINE CALCULATES STEADY-STATE P(Q(0)=Q) ***
       IA=N
       IZ=N
       IJOB=2
          DO 310 I=1,N
          DO 320 J=1,N
 320   D(I,J)=PQO(J,I)
 310     CONTINUE
       CALL EIGRF(D,N,IA,IJOB,W,Z,IZ,WK,IER)
          DO 700 J=1,N
       ZN=Z(1,J)
          DO 700 I=1,N
       Z(I,J)=Z(I,J)/ZN
 700   CONTINUE
C          WRITE(7,96)
C       WRITE(7,55)(Z(I,1),I=1,N)
C *************** SUMMING OF THE REAL POSITIVE EIGN VECTORS *********
         ZTOTAL=0.0
          DO 200 I=1,N
 200    ZTOTAL=ZTOTAL+Z(I,1)
C      WRITE(7,65)ZTOTAL
C ************* DETERMINATION OF STEADY STATE PROBABILITIES *********
C ***************** BY NORMALIZING EIGEN VECTORS BY ZTOTAL **********
       DO 220 I=1,N
 220    SS(I)=Z(I,1)/ZTOTAL
           IF (SS(N) .GT. 0.0025) THEN
             WRITE(7,221)
             GO TO 106
           END IF
         WRITE(7,75)
         WRITE(7,85)
         WRITE(7,250)
         WRITE(7,251)
          DO 240 K=1, N/2
 240    WRITE(7,230)CQ(K),SS(K),CQ(K+N/2),SS(K+N/2)
         WRITE(7,75)
C **********  FINDING RY TO CALIBRATE RECESSION MODEL TO LPIII ******
       PQLP3=0.0
         DO 801 K=1, N
       XX=QLP3/CQ(K)
       IF (XX .GT. 1.0)  XX=1.0
       CALL MDBETA(XX,AA,BB,PP,IERR)
C ******* SUMMING OF JOINT INTERSECTIONS OF P(A) AND P(QO = CT) *****
       PQLP3=PQLP3+SS(K)*PP
 801     CONTINUE
       TTRR=1/PQLP3
       ERY=TTRR/TRLP3
C ********** PROGRAM TO CALCULATE P(Q < 1 DAY LOW FLOW) *************
       WRITE(7,95) IY1,IY2,RK,RY,ERY
       WRITE(7,8)
         DO 103 I=1, 6
       TRQ=QE(I)
       QL=TRQ*0.25
       QH=TRQ*4.20
       IF (TRQ.LT.1) THEN
        QH=4
        QL=0
       END IF
       M=0
```

```
  101 PQ1=0.0
      M=M+1
      IF (M .EQ. 20) WRITE(7,102) TRQ
      IF (M .EQ. 20) GO TO 106
        DO 800 K=1, N
      XX=TRQ/(CQ(K))
      IF (XX .GT. 1.0)  XX=1.0
      CALL MDBETA(XX,AA,BB,PP,IERR)
C ******** SUMMING OF JOINT INTERSECTIONS OF P(A) AND P(Q0 = CT) ****
      PQ1=PQ1+SS(K)*PP
  800   CONTINUE
      TTR=1/PQ1
C     ...........................................................
C          ADJUSTMENT OF RETURN PERIOD FOR A YEAR TO RETURN
C          PERIOD FOR AN EVENT (EXPECTED NUMBER OF RECESSIONS IN TR(I))

C     ...........................................................
      TRR=ERY*TR(I)
      IF (TTR .LT. 0.995*TRR) THEN
       QH=TRQ
       TRQ=(QH+QL)/2.0
       GO TO 101
      END IF
      IF (TTR .GT. 1.005*TRR) THEN
       QL=TRQ
       TRQ=(QH+QL)/2
       GO TO 101
      END IF
      WRITE(7,105) TR(I),TRQ,QE(I)
 103    CONTINUE
   1  FORMAT(12X,'1 DAY LOW FLOW ANALYSIS FOR STATION: 0',I7)
   2  FORMAT(/19X,'THE TRANSITION TABLE SIZE: ',I3/)
   8  FORMAT(15X,'RETURN PERIOD',2X,'RECESSION',2X,'LOG-PEARSON III'/)
  31  FORMAT(1X,F6.1,2X,F8.6)
  35  FORMAT(1X,F6.1)
  45  FORMAT(14F9.6/14F9.6)
  55  FORMAT(14F9.1/14F9.1/14F9.1/14F9.1/14F9.1/14F9.1/14F9.1/14F9.1)
  65  FORMAT(1X,'***** ZTOTAL =',F10.1)
  75  FORMAT(10X,'**********************************************')
  77  FORMAT(2X,'ERROR HAS OCCURED IN THE INTEGRATION OF F !!!'/
     $2X,'NO CONVERGENCE FOR P[Q0(T) < Q| Q0(T-1)=CT]')
  85  FORMAT(19X,'STEADY-STATE PROBABILITIES FOR Q0')
  95  FORMAT(/25X,'PERIOD OF ',I4,' TO ',I4//20X,
     $'RECESSION CONSTANT, K = ',F5.2/23X,'MEAN RECESSIONS/YR = ',
     $F5.2/17X,'CALCULATED RECESSIONS/YR =',F7.3/)
  96  FORMAT(1X,'POSITIVE EIGEN VECTOR OF THE TRANSPOSED MARKOV TABLE')
  102 FORMAT(/1X,'QTR IS NOT CONVERGING, TRY ANOTHER INITIAL QTR,
     $ LAST TRQ = ',F6.1)
  105 FORMAT(20X,F4.0,6X,F7.2,7X,F7.2)
  221 FORMAT(1X,'UL IS TOO SMALL, TRY A LARGER VALUE')
  230 FORMAT(11X,F7.1,4X,F9.5,6X,F7.1,4X,F9.5)
  251 FORMAT(11X,'____________________________________________')
  252   FORMAT(19X,'______________________________')
  250   FORMAT(/15X,'Q0',7X,'P(Q0 = Q)',7X,'Q0',7X,'P(Q0 = Q)')
  106 STOP
      END
C ***************  SUBROUTINES/FUNCTIONS  *************************
C *** F(X) IS THE PROBABILITY FUNCTION FOR Q(0) TO BE INTEGRATED ***
      REAL FUNCTION F(X)
      REAL X
      COMMON Q,CT,CST,A1,B1,A2,B2
      F=CST*EXP(-A2*((Q-CT*X)**B2))*(X**A1)*((1-X)**B1)
      RETURN
      END
```

```
C *  PROGRAM CENSOR JCL FINDS THE PARTIAL DURATION SERIES  *
C *  FLOWS OCCURING BELOW CUT-OFF FLOW (LARGEST ANNUAL MIN-*
C *  MUM LOW FLOW) AND THE NUMBER OF LOW FLOWS ABOVE CUT-  *
C *  OFF FLOW.  CALCULATES WEIBULL PARAMETERS USING        *
C *  CENSORED DATA APPLICATION IN MAXIMUM LIKELIHOOD       *
C *  ESTIMATION                                            *
C *********************************************************
C *                                                        *
C *       SET THE PARTS OF THE PROGRAM TO BE RUN HERE      *
C *                                                        *
C *********************************************************
C FOR EACH PORTION OF THE PROGRAM, SET ITS PARAMETER AS
C FOLLOWS:
C     1 - IF THE PORTION IS TO BE EXECUTED WITH RESULTS
C         PRINTED ONLY IN THE SUMMARY
C     2 - IF THE PORTION IS TO BE EXECUTED PRINTING SOME
C         RESULTS AS THE PROGRAM IS RUN
C     3 - IF THE PORTION IS TO BE EXECUTED PRINTING ALL
C         OUTPUT (EXCEPT BISECTIONS)
C     4 - IF THE PORTION IS TO BE RUN PRINTING ALL OUTPUT
C
C THE PARAMETERS FOR EACH PORTION ARE DEFINED AS FOLLOWS:
C    IKAY - PROGRAM FOR LOW FLOWS ABOVE CUT-OFF FLOW
C *********************************************************
C *                                                        *
C *                 SET PARAMETERS HERE                    *
C *                                                        *
C *********************************************************
C  START IS THE START DATE FOR THE RANGE OF ACCEPTABLE
C  RECESSIONS.  FINISH IS THE END DATE FOR THE RANGE OF
C  ACCEPTABLE RECESSIONS.  EPSF IS THE FRACTION OF THE FIRST
C  YEARS'S AVERAGE FLOW ABOVE WHICH A FLUCTUATION IN THE
C  RECESSION LIMB IS UNACCEPTABLE.  BEGIN IS THE POINT
C  BEYOND WHICH ALL COMBINATIONS OF POINTS ON A RECESSION
C  ARE CHECKED TO FIND THE BEST STRAIGHT LINE (NOTE - THE
C  PEAK IS AT 1).  SHORT IS THE MINIMUM ACCEPTABLE LENGTH
C  LENGTH FOR RECESSIONS.
      IKAY=1
      IPR =1
      ITEMP=1
      START=1
      FINISH=365
      DUR=1
      EPSF=0.020
      BEGIN=4
      SHORT=7
C*********  RETURN PERIODS *********
      TR(1)=2
      TR(2)=5
      TR(3)=10
      TR(4)=20
      TR(5)=25
      TR(6)=50
      TR(7)=100
C *********************************************************
      DIMENSION DATA(38),TR(10),QTR(10),QCTR(10),QQMIN(5000)
      INTEGER DUR,START,FINISH,BEGIN,SHORT,RT(10)
      REAL MEAN,MDEV
      DOUBLE PRECISION X(100),Y(100),Q1(100),Q2(100)
      DOUBLE PRECISION Q3(100)
      COMMON/NUM1/Q(450)
      COMMON/NUM2/X,Y,RT,QOUT(20,10)
      COMMON/NUM9/QY(20000)
      COMMON/NUM12/QI(20000),QR(20000)
      IISAVE=0
      IZERO =0
      JJSAVE=0
```

```
      KK=0
      IIPNT=1
      IPOI=1
        DO 6 I=1,100
      IPOINT=IIPNT
      IIPNT=1
      IF (I.EQ.1) JMAX=14
      IF (I.NE.1) JMAX=12
         DO 2 J=1,JMAX
       READ (1,10,END=999) (DATA(K),K=1,38)
       IFLAG=0
       IF (I.EQ.1.AND.J.EQ.1) STAT=DATA(1)
       IF (DATA(1).NE.STAT) WRITE (7,20) I,J
       IF (DATA(1).NE.STAT) JFLAG=1
       IF (I.EQ.1.AND.J.LE.9.AND.DATA(3).NE.(FLOAT(J)+3.0)) IFLAG=1
       IF (I.EQ.1.AND.J.GT.9.AND.DATA(3).NE.(FLOAT(J)-9.0)) IFLAG=1
       IF (I.NE.1.AND.J.LE.7.AND.DATA(3).NE.(FLOAT(J)+5.0)) IFLAG=1
       IF (I.NE.1.AND.J.GT.7.AND.DATA(3).NE.(FLOAT(J)-7.0)) IFLAG=1
       IF (IFLAG.EQ.1) WRITE (7,30) I,J
       IF (IFLAG.EQ.1) JFLAG=1
       NUM=DATA(6)
       JMAXM1=JMAX-1
       IF (J.GE.JMAXM1) IIPNT=IIPNT+NUM
       IF (I.EQ.1.AND.J.EQ.1) NYEAR=DATA(2)-1
       LIMIT=NUM+6
         DO 1 K=7,LIMIT
       L=IPOINT+K-7
       LL=IPOI+K-7
         Q(L)=DATA(K)
         QY(LL)=DATA(K)
         IF (QY(LL).EQ.0) IZERO=IZERO+1
    1    CONTINUE
         IPOI  =IPOI  +NUM
    2    IPOINT=IPOINT+NUM
       III=IPOINT-IIPNT
       IISAVE=IISAVE+II
       JJSAVE=JJSAVE+JJ
       IF (ITEMP.GT.0) CALL SMALL (I,DUR,III,X(I),KFLAG)
       IIPM1=IIPNT-1
       DO 5 L=1,IIPM1
       IIIPL=III+L
    5  Q(L)=Q(IIIPL)
    6  CONTINUE
  999  IF (ITEMP.GT.0) CALL SMALL (I,DUR,III,X(I),KFLAG)
         LEND=I
         QMAX=1
            DO 967 K=1, LEND
         IF (X(K) .GT. QMAX) QMAX = X(K)
C        WRITE(7,968) K, X(K)
 967        CONTINUE
C 968     FORMAT(5X,I4,5X,'QMIN = ',F9.2)
         QC=QMAX

         IEND=IPOI-1
C          WRITE(7,969) IEND
C 969      FORMAT(/1X,'IEND =',I6,/)
       IF=FINISH-50
        IF (IKAY.GT.0) CALL KAY (IKAY,START,
      1III,II,EPSF,BEGIN,SHORT,JJ,EPS,QC,IEND)
       IISAVE=IISAVE+II
       JJSAVE=JJSAVE+JJ
       II=IISAVE
       JJ=JJSAVE
       NNN=I
       IITEMP=0
       STAT1=STAT/100.0
       IF (IKAY.GT.1) WRITE (7,40) STAT1
       IF (IITEMP.EQ.1) WRITE (7,41) DUR
```

```
      IF (IITEMP.EQ.1) WRITE (7,42)
      IF (IITEMP.EQ.1) WRITE (7,43) (X(I),I=1,NNN)
      IF (ITEMP.GT.0) CALL RANK (NNN)
      M=0
      DO 9 I=1,10
    9 IF (RT(I).GT.0) M=M+1
      NYEAR=NYEAR+1
      NYPN=NYEAR+NNN
      IF (KFLAG.EQ.1) WRITE (7,45)
      WRITE (7,50) STAT1
      WRITE (7,51) NYEAR,NYPN
      IF (IKAY.GT.0) WRITE (7,54) START,FINISH
      IF (IKAY.GT.0) WRITE (7,55) BEGIN
      IF (IKAY.GT.0) WRITE (7,56) JJ,II
          KN=II
      IF (IPR .EQ. 0)  WRITE(7,122)
122   FORMAT(/16X,' PARTIAL DURATION SERIES '/)
C PARTIAL DURATION SERIES ANALYSIS
C           QC = CUT-OFF FLOW TO DEFINE A RECESSION
C         QMIN = LOWEST FLOW OCCURING IN A RECESSION BELOW QC
C         IEND = NUMBER OF DAYS IN PERIOD OBSERVED
C                (ALSO NUMBER OF CONTINUOUS FLOWS)
C.........................................................
C       IEND=L
          QMIN=QC+1.0
        MFLAG=0
        SUMQ=0
        ILEN=0
        IF (IPR .EQ. 0)   WRITE(7,21)
  21    FORMAT(13X,'N',10X,'M',7X,'LEN',9X,'Q(M)'/11X,'____________________
     $____________________')
C ..........................................................
         DO 11 M=1,IEND
        IF (QY(M) .GT. QC) GO TO 11
        IF (QY(M) .LE. QC) QM=QY(M)
        IF (QM .LT. QMIN) MTIME=M
        IF (QM .LE. QMIN) QMIN=QM
        ILEN=ILEN+1
          IF (QY(M+1) .GT. QC ) GO TO 12
          IF (M .EQ. IEND .AND. ILEN .GT. 0) GO TO 12
         GO TO 11
C
  12    MFLAG=MFLAG+1
        NN=MFLAG
        SUMQ=SUMQ+QMIN
        QQMIN(NN)=QMIN
        IF (IPR .EQ.0 ) WRITE(7,13) MTIME,NN,ILEN,QQMIN(NN)
        ILEN=0
        QMIN=QC+1
  11    CONTINUE
C
C
  13    FORMAT(10X,I5,5X,I5,5X,I5,5X,F8.1)
  276   FORMAT(4X,F8.2,5X,F8.2)
C..... CALCULATION OF PARAMETERS ..............................
      QN=FLOAT(NN)
      NY=NYPN-NYEAR
      RMEAN=QN/FLOAT(NY)
      RCMEAN=(FLOAT(II)+FLOAT(NN))/FLOAT(NY)
      QMEAN=SUMQ/FLOAT(NN)
       WRITE(7,14) IEND,NN,KN
  14   FORMAT(/32X,'TOTAL DAYS IN PERIOD =',I6/28X,
     $'NUMBER OF FLOWS BELOW Q0 =',I5/
     $28X,'NUMBER OF FLOWS ABOVE Q0 =',I5)
       WRITE(7,15) RMEAN,RCMEAN,QMEAN,QC
  15  FORMAT(/12X,'AVERAGE NUMBER OF LOW FLOWS (TRUNCATION) =',F7.2/17X,
```

```
     $'AVERAGE NUMBER OF LOW FLOWS (TOTAL) =',F7.2/17X,
     $'  MEAN VALUE OF THE LOW FLOWS, QAVG =',F7.2/40X,
     $'CUT-OFF FLOW =',F7.2)
      IF (IZERO .GT. 0) WRITE(7,277) IZERO
  277 FORMAT(32X,'NUMBER OF ZERO FLOWS =',I5//
     $17X,'NOTE: THIS IS AN INTERMITTENT STREAM'/)
C............ PARAMETER ESTIMATION FOR CENSORED DISTRIBUTION ......
      FFLAG=0
      C=4.0
      CH=10.0
      CL=0.0
 81   FFLAG=FFLAG+1
      IF (FFLAG .GT. 25) GO TO 85
      SUMB=0.0
      SUMC=0.0
      CKN=FLOAT(KN)
        DO 99 I=1, NN
 99   SUMB=SUMB+(QQMIN(I)**C)
      B=((CKN*(QC**C)+SUMB)/QN)**(1/C)
        DO 98 I=1, NN
      QMI=QQMIN(I)
      IF (QMI.LE.0) QMI=0.01
      QIB=QMI/B
      SUMC=SUMC+ALOG(QMI)-(QIB**C)*ALOG(QIB)
 98     CONTINUE
      QCB=QC/B
      QCBC=QCB**C
      DLDC=QN/C-CKN*QCBC*ALOG(QCB)-QN*ALOG(B)+SUMC
      IF (DLDC .LT. -0.01) GO TO 82
      IF (DLDC .GT.  0.01) GO TO 83
      GO TO 84
 82   CH=C
      C =(CH+CL)/2
      GO TO 81
 83   CL=C
      C =(CH+CL)/2
      GO TO 81
 85   WRITE(7,86)
 86   FORMAT(5X,'++++ C IS NOT CONVERGING, TRY ANOTHER INITIAL C')
 84   WRITE(7,87) C,B
 87   FORMAT(/34X,'WEIBULL DISTRIBUTION'//30X,'SHAPE PARAMETER, C =',
     $F7.3/30X,'SCALE PARAMETER, B =',F7.2/)
C............. CALCULATION OF FLOWS FOR RETURN PERIODS ......
      CC=1/C
        DO 89 I=1, 7
      RTR=RCMEAN*TR(I)
      PTR=1/RTR
      QCTR(I)=B*((-1*ALOG(1-PTR))**CC)
 89      CONTINUE
      IF (ITEMP.GT.0) WRITE (7,60) DUR
      IF (ITEMP.GT.0) WRITE (7,70) (TR(I),I=1,7)
      IF (JFLAG.EQ.1) WRITE (7,140)
      IF (ITEMP.GT.0) WRITE (7,92) (QCTR(I),I=1,7)
C..............  RANK AND ORDER OF FLOWS................
        DO 100 I=1, NN
 100  QI(I)=QQMIN(I)
      N=1
 105  QR(N)=100000
        DO 110 I=1, NN
      IF (QI(I).LT.QR(N))    NI=I
      IF (QI(I).LT.QR(N)) QR(N)=QI(I)
 110    CONTINUE
C     WRITE(7,120) N, QR(N)
C120  FORMAT(5X,I5,5X,'Q(N) =',F7.1)
      QI(NI)= 100001
      N=N+1
      IF (N.LE.NN) GO TO 105
```

```
C............  GOODNESS OF FIT (KOLMOGOROV-SMIRNOV TEST) .......
      DEVM=0
      TN=QN+CKN
        DO 91 I=1, NN
      FLOW=QR(I)
      IF (FLOW.EQ.0) GO TO 91
      AX=(FLOW/B)**C
      FQ=1-EXP(-1*AX)
      FR=FLOAT(I)/TN
      DEV=ABS(FR-FQ)
      IF (DEV.GT.DEVM) DEVM=DEV
 91     CONTINUE
      CRIT05=1.36/SQRT(TN)
      CRIT01=1.63/SQRT(TN)
      WRITE(7,97) DEVM
      IF (DEVM.LE.CRIT05) WRITE(7,93)
     IF (DEVM.LE.CRIT05) GO TO 96
     IF (DEVM.LE.CRIT01) WRITE(7,94)
     IF (DEVM.LE.CRIT01) GO TO 96
     IF (DEVM.GT.CRIT01) WRITE(7,95) CRIT01
 97  FORMAT(26X,'MAXIMUM DEVIATION =',F8.5)
 93  FORMAT(/15X,'PASSED K-S TEST FOR LEVEL OF SIGNIFICANCE = 0.05'/)
 94  FORMAT(12X,'** FAILED K-S TEST FOR LEVEL OF SIGNIFICANCE = 0.05 **
    $'/15X,'PASSED K-S TEST FOR LEVEL OF SIGNIFICANCE = 0.01'/)
 95  FORMAT(12X,'** FAILED K-S TEST FOR LEVEL OF SIGNIFICANCE = 0.01 **
    $'/,25X,'CRITICAL DEVIATION =',F8.5/)
96   WRITE (7,150)
C..................................
   10 FORMAT (38E10.0)
   20 FORMAT (3X,39HERROR - THE STATION IDENTIFICATION NUMB,
     121HER IS WRONG WHERE I =,I3,8H AND J =,I3,/)
   30 FORMAT (5X,39HERROR - THE PERIOD OF THE ACCESSED DATA,
     130H MUST BE FOR FULL WATER YEARS.,/,4X,
     236HWATER YEARS ARE APRIL 1 TO MARCH 31.,2X,
     325HTHE ERROR OCCURRED AT I =,I3,3HJ =,I3)
   40 FORMAT (17X,37HTHE STATION IDENTIFICATION NUMBER IS ,
     1F8.2)
   41 FORMAT (/,21X,28HTHE DURATION OF LOW FLOWS IS,I3,
     15H DAYS,//)
   42 FORMAT (5X,31HTHE YEARLY LOW FLOW VALUES ARE:)
   43 FORMAT (10X,7F9.2)
   45 FORMAT (/,20X,36H***** WARNING - THIS IS AN INTERMITT,
     116HENT STREAM *****,/)
   50 FORMAT (/,4X,37H********************************* SUM,
     138HMARY *******************************,//,30X,
     215HSTATION NUMBER ,F8.2)
   51 FORMAT (16X,38HTHE PERIOD OF DATA ACCESSED IS FROM 4/,
     1I4,6H TO 3/,I4,/)
   54 FORMAT (18X,'FOR THE RECESSIONS USED, PEAK FLOWS OCCURRED'/
     $23X,7HBETWEEN,I4,4H AND,I4,6H DAYS ,
     213HAFTER APRIL 1)
   55 FORMAT (13X,38HRECESSIONS WERE TAKEN TO BEGIN AT LEAS,
     11HT,I2,20H DAYS AFTER THE PEAK)
   56 FORMAT (16X,I3,39H RECESSIONS WERE IDENTIFIED AS HAVING L,
     $14HONG RECESSIONS,/,18X,I3,27H OF THESE RECESSIONS WERE U,
     $22HSED AS CENSORED FLOWS /)
   60 FORMAT (/30X,'PARTIAL DURATION SERIES'//
     $23X,28HTHE DURATION OF LOW FLOWS IS,I3,
     $5H DAYS,//,30X,22HRETURN PERIODS (YEARS))
   70 FORMAT (4X,15HMETHOD        | ,F7.2,6F8.2,
     1/,4X,44H-------------+------------------------------,
     227H---------------------------)
   92 FORMAT (4X,15HCENSORED      | ,F7.2,6F8.2,
     1/,4X,44H-------------+------------------------------,
     227H-------------------------//)
  140 FORMAT (22X,38H***** CHECK ABOVE OUTPUT FOR ERROR ***,
     12H**,/)
```

```
  150 FORMAT (4X,39H***************************************,
     136H************************************)
 275  STOP
      END
      SUBROUTINE KAY (IKAY,START,III,II,EPSF,BEGIN,SHORT,JJ,EPS,QC,IEND)
      DOUBLE PRECISION Q(200)
      DIMENSION MEND(1000)
      INTEGER START,FINISH,BEGIN,SHORT,FP80
      COMMON/NUM1/QQ(450),ISSS(10)
      COMMON/NUM9/QY(40000)
C.............................
          MINR=10
      II=0
      JJ=0
      SUM=0
      DO 1 J=1, 365
    1 SUM=SUM+QY(J)
      EPS=EPSF*SUM/FLOAT(III)
      IF (IKAY.GT.2) WRITE (7,20) EPS
    2 N=0
        DO 13 J=1,IEND
      TE1=QY(J+1)-EPS
      TE2=QY(J+2)-EPS
      TE3=QY(J+3)-EPS
      IF (QY(J).GE.TE1.AND.QY(J).GE.TE2.AND.QY(J).GE.TE3) THEN
       N=N+1
       IF (N.EQ.1) JSAVE=J
       GO TO 13
      END IF
      N=N+2
      JFLAG=0
      IF (JSAVE.GE.START.AND.JSAVE.LE. IEND ) JFLAG=1
      IF (N.GE. MINR .AND.JFLAG.EQ.1) GO TO 3
      N=0
      GO TO 13
    3 II=II+1
      JJ=JJ+1
      LOWER=JSAVE
      LIMIT=JSAVE+N-1
      ISP1=ISAVE+1-1900
      IF (IKAY.GT.2) WRITE (7,30) ISAVE,ISP1,JSAVE,N
      DO 4 I=LOWER,LIMIT
      JI=I-LOWER+1
    4 Q(JI)=QY(I)
      IF (IKAY.GT.2) WRITE (7,80) (Q(K),K=1,N)
      MEND(II)=JSAVE+N
      IDIF=MEND(II)-MEND(II-1)
      RATIO=Q(1)/Q(N)
           IF (Q(N).LT.QC) II=II-1
           IF (RATIO.LT.1.2) II=II-1
           IF (IDIF.LT.10) II=II-1
   12 CONTINUE
      N=0
   13 CONTINUE
C..........................................................
   20 FORMAT (16X,38HTHE ACCEPTABLE RANGE FOR FLUCTUATIONS ,
     12HIS,F6.3)
   30 FORMAT (//,5X,6HYEAR =,I5,1H-,I2,5X,5HDAY =,I5,5X,

     13HN =,I5,/)
   60 FORMAT (/,22X,36HTHE "BEST" REGRESSIONS ARE AS FOLLOW,
     12HS:,//,13X,6HLENGTH,2X,5HSTART,4X,3HEND,8X,1HK,11X,
     21HR,10X,1HS)
   70 FORMAT (9X,3I8,2X,3F11.6)
   80 FORMAT (9X,6F10.2)
```

```
   90 FORMAT (/,15X,35HTHE SELECTED REGRESSION IS FROM I =,
     1I3,7H TO J =,I3,/,21X,3HK =,F6.2,3X,3HR =,F8.4,3X,
     23HS =,F7.4)
      RETURN
      END
C
      SUBROUTINE SMALL (I,DUR,III,AAA,LFLAG)
      INTEGER DUR,DURP1
      COMMON/NUM1/Q(450)
      AAA=0
      DO 1 J=1,DUR
    1 AAA=AAA+Q(J)
      TEMP=AAA
      DURP1=DUR+1
      DO 2 J=DURP1,III
      JMDUR=J-DUR
      TEMP=TEMP-Q(JMDUR)+Q(J)
    2 IF (TEMP.LT.AAA) AAA=TEMP
      AAA=AAA/FLOAT(DUR)
      IF (AAA.LE.0) AAA=0
      RETURN
      END
C
      SUBROUTINE RANK (N)
      DOUBLE PRECISION X(100),Y(100),W(100),Z(100)
      INTEGER RT(10)
      COMMON/NUM2/X,Y,RT,Q(10,10)
      DO 1 I=1,N
    1 Z(I)=X(I)
      M=N
      DO 5 I=1,N
      Y(I)=Z(1)
      DO 2 J=1,M
      JFLAG=0
      IF (Z(J).LE.Y(I)) JFLAG=1
      IF (JFLAG.EQ.1) Y(I)=Z(J)
    2 IF (JFLAG.EQ.1) JSAVE=J
      MM=0
      DO 3 J=1,M
      IF (J.NE.JSAVE) MM=MM+1
      IF (J.NE.JSAVE) W(MM)=Z(J)
    3 CONTINUE
      M=MM
      DO 4 J=1,M
    4 Z(J)=W(J)
    5 CONTINUE
      RETURN
      END
```

The vita has been removed from the scanned document