NEW PLOTTING POSITION FORMULAS FOR SOME WELL-KNOWN DISTRIBUTIONS IN ENGINEERING HYDROLOGY

by

NOPHADOL IN-NA

Department of Civil Engineering and Applied Mechanics

McGill University

Montreal, Canada

March 1989

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

The main objective of this research is to develop new unbiased plotting position formulas for two general probability distributions which are widely used in hydrologic frequency analyses: the Pearson type III (P3) and the General Extreme Value (GEV) distributions. The research study is divided into three parts. First, using the Probability Weighted Moment (PWM) theory, an analytical method is proposed to derive the exact plotting positions for *systematic* flood records (i.e., complete flood samples which occurred during the period of systematic gauging). Second, for the convenience of practical application, simple unbiased plotting position formulas representing a very reliable approximation to the exact plotting positions are developed. Third, new unbiased plotting position formulas for P3 and GEV distributions are proposed for *historical* flood records (i.e., data on very large floods which occurred outside or within the systematic gauging period). The incorporation of historical flood information in plotting position formulas would significantly improve the estimation of flood quantiles.

The analytical method and plotting position formulas proposed in the present study are verified and compared with various existing techniques and formulas. The suggested analytical method was found to be preferable to the conventional direct numerical integration and the Monte Carlo simulation procedure in the estimation of expected values of P3 and GEV order statistics. Results of the numerical and graphical comparisons have also demonstrated that the plotting position formulas developed in this study provided a better agreement to the exact plotting positions than several existing formulas. In particular, the suggested formulas are more flexible because they can take explicitly into account the skewness coefficient of the underlying distribution. Moreover, for illustration purposes, the proposed formulas were applied to observed flow data of various rivers. It was found that the proposed formulas provided better estimates of flood quantiles than many existing formulas including the well-known

Weibull formula. Finally, special probability papers for various skewness values are developed for the P3 and GEV distributions. It can be concluded that the development of new plotting position formulas and probability papers for the P3 and GEV distributions in the present study has provided a convenient and practical tool for the application of these distributions in engineering practice.

RESUME

La présente étude a pour objet de développer de nouvelles formules non-biaisées de probabilité empirique pour les deux lois de probabilité qui sont couramment utilisées dans l'analyse fréquentielle en hydrologie: la loi de Pearson type III (P3) et la loi générale des valeurs extrêmes (GVE). Cette étude se divise en trois parties. Premièrement, en se basant sur la théorie des moments pondérés de probabilité (MPP) une méthode analytique est suggérée pour calculer les valeurs exactes de probabilité empirique pour des séries de données *systématiques* de crue (i.e., enregistrements continus des données de crue durant la période de jaugeage systématique). Deuxièmement, pour les applications pratiques, de nouvelles formules plus simples sont développées qui permettent toutefois d'obtenir une bonne approximation des valeurs exactes de probabilité empirique. Troisièmement, de nouvelles formules non-biaisées pour les deux lois P3 et GVE sont proposées pour des séries de données *historiques* de crue (i.e., données sur les crues extrêmement grandes qui apparaissent avant ou durant la période de jaugeage systématique). L'introduction de l'information historique des crues dans le développement des formules de probabilité empirique pourrait améliorer considérablement l'estimation des quantiles de crue.

La méthode analytique et les formules de probabilité empirique développées dans la présente étude sont vérifiées et comparées aux diverses techniques et formules présentement disponibles. On a trouvé que la méthode analytique proposée est préférable à la technique traditionnelle d'intégration numérique et à la procédure de simulation de Monte Carlo dans le calcul de l'espérance mathématique des statistiques d'ordre pour les deux lois P3 et GVE. Les résultats des comparaisons numérique et graphique ont également démontré que les formules développées dans la présente étude permettent une meilleure estimation de la probabilité empirique que celle donnée par plusieurs formules existantes. En particulier, les formules suggérées sont plus flexibles

parce qu'elles peuvent tenir compte, d'une façon explicite, du coefficient d'asymétrie de la distribution considérée. De plus, afin d'illustrer l'utilisation des nouvelles formules dans la pratique, ces formules sont appliquées aux donnees réelles de crue de diverses rivières. On a trouvé que les formules proposées donnent une meilleure estimation des quantiles de crue que celle obtenue par plusieurs formules existantes, incluant la fameuse formule de Weibull. Finalement, des nouveaux papiers de probabilité pour différentes valeurs d'asymétrie sont developpés pour les deux lois P3 et GVE. On peut conclure que l'élaboration des nouvelles formules de probabilité empirique et le développement des nouveaux papiers de probabilité dans la présente étude pour les deux lois P3 et GVE fournissent un outil simple et pratique qui facilite l'utilisation de ces deux lois dans les applications pratiques en ingénierie.

ACKNOWLEDGEMENTS

The author wishes to thank Prof. V.T.V. Nguyen, Associate Professor, Department of Civil Engineering and Applied Mechanics, for directing this research program and for his useful suggestions, constructive criticism and constant encouragement throughout the study.

The author would also like to thank Prof. J.E. Nash, Editor of Journal of Hydrology, Dr. C. Cunnane, Prof. J.C.I. Dooge and Dr. K.M. O'Connor, faculty members of Department of Engineering Hydrology, University College Galway, Ireland who first contributed the knowledge of hydrology to the author when he studied for his M.Sc degree.

Much appreciation is expressed for the support and assistance given by the committee members of this thesis during the completion of this research work and McGill University Computing Center for running the computer programs. After three years of Academic pursuit at the McGill University, the author wishes to avail himself of the opportunity to express his most sincere thanks and appreciation to the Natural Science and Engineering Research Council of Canada for providing the financial support covering the entire study program.

Finally, a special debt is owed to his beloved parents, brother and sisters for their continuous inspiration, supports and encouragement and to Yuwaree Phadungchewit for her understanding and help in different ways.

TABLE OF CONTENTS

TABLE OF CONTENTS *(Continued)*

LIST OF FIGURES

LIST OF FIGURES *(Continued)*

LIST OF TABLES

LIST OF SYMBOLS

a	constant
b	constant
C_1, C_2, C_3	regression coefficients
C_4, C_5, C_6	regression coefficients
C_7, C_8	regression coefficients
C_s	biased estimate of the population skewness coefficient
C_{su}	unbiased estimate of the population skewness coefficient
C_v	coefficient of variation
CDF	cumulative distribution function
e	number of larger floods that occurred during systematic flood records
e_1	relative difference between estimations by Harter and PWM
e_2	relative difference between estimations by Xuewu and PWM
$E[\cdot]$	expected value operator
$EV1$	Extreme value type I
$EV2$	Extreme value type II
$EV3$	Extreme value type III
$f(\cdot)$	probability density function (PDF)
$F(\cdot)$	cumulative distribution function (CDF)
$g(\cdot)$	PDF of order statistics
g	number of observed floods in complete flood records
GEV	General extreme value distribution
GG	Generalized gamma distribution
k	numbers of floods that larger than the threshold
$LP3$	Log Pearson type III distribution
m	rank of data in an ordered sample
$M_{l,i,j}$	probability weighted moment of order l, i, j
N	sample size
$P3$	Pearson type III distribution
P_m	plotting position expressed in term of non-exceedance probability
$\hat{P}_m$	plotting position expressed in term of exceedance probability

LIST OF SYMBOLS *(Continued)*

$P(\lambda, y)$	incomplete gamma function
PDF	probability density function
PWM	Probability weighted moments
Q	flood discharge
$RMSE$	root mean square error
s	s-year period of systematic flood records
T	return period
W_0	location parameter for GG distribution
x_0	location parameter for P3 distribution
x_m	ordered observation
y	standardized variate
y_m	m^{th} order statistic
Z_0	location parameter for LP3 distribution
α	correction coefficient
α'	correction coefficient
β	scale parameter
$\Gamma(\cdot)$	complete gamma function
γ	skewness coefficient
δ	Euler's constant
λ	shape parameter
μ	mean
$\mu'_{Z\,r}$	r^{th} moment about the origin
σ	standard deviation

CHAPTER 1

INTRODUCTION

1.1 General

Information on flood magnitudes and their associated probabilities of occurrence is important for planning and design of many hydraulics structures. In practice, hydrologists use flood data collected at a river gauging station to establish a flood-frequency relationship which applies at that location. Given a series of flood events, it is necessary to assign to each event an empirical probability or recurrence interval. These empirical probabilities are often estimated based on "plotting position" formulas. The subject of plotting positions or probability plots has been discussed for several decades by hydrologists and statisticians (e.g., Hazen, 1914; Blom, 1958; Kimball, 1960; Benson, 1962; Barnett, 1975; Cunnane, 1978; Harter, 1984; Xuewu et al., 1984; Arnell et al., 1986; Hirsch and Stedinger, 1987; Nguyen et al., 1989; In-na and Nguyen, 1989). In particular, plotting positions have been used widely in hydrologic frequency analysis in a variety of ways, e.g., to estimate the magnitude of hydrologic events and their corresponding probability of occurrence, to detect outliers, to fit distributions to data, and to evaluate the adequacy of the fit. Recently, some analytical procedures for estimating distribution parameters (such as probability weighted moments and maximum likelihood) have been considered, in theory, more efficient than the graphical fitting methods. The use of probability plots in engineering practice, however, is not diminished. Many hydrologists would not make engineering decisions without the use of graphical displays.

Probability plots were recently recommended as a means for extrapolation of flood frequency curves in dam safety evaluations (U.S. National Research Council, 1985). The use of probability plots was also suggested in the determination of the probability distribution of annual maximum flood elevations which occur due to the combined effects of ice jam and storm-induced flooding (U.S. Federal Emergency Management Agency, 1982). Although the U.S. Water Resources Research Council (Interagency Advisory Committee on Water Data, 1982) recommended the use of method of moments to fit the Log-Pearson type III distribution to floodflow data, their recommendations included also the use of probability plots. Probability plots, therefore, are still playing an important role in engineering practice.

A probability plot is defined as a graphical representation of the ordered observations of a hydrologic event (e.g., flood magnitudes) versus their associated empirical probabilities which are estimated using a plotting position formula. Hence, it is necessary to select an appropriate plotting formula in order to provide a reliable estimation of hydrologic event magnitudes for a given exceedance probability. However, in practice, given a large number of various plotting positions available, the choice of a suitable formula is not an easy task. Detailed reviews and discussion on this subject have been provided by Cunnane (1978) and Harter (1984). It is clear from the review of the existing literature that the optimum choice of a plotting position formula should be based on the purpose of the investigation or should depend on the use which is to be made of the results. For example, Kimball (1960) investigated the choice of plotting positions for the normal and extreme-value (Gumbel) distributions using various statistical criteria (test of fit, estimation of parameters, extrapolation at one of the extremes). Benson (1962) studied the effects of a selection of plotting position method on the determination of the economic feasibility of some engineering works (bridge design, reservoir design and flood insurance).

Much of the confusion and disagreement concerning the choice of plotting positions is probably due to the fact that the cumulative distribution function $F[E(y_m)]$ of the expected value of the reduced variate mth order statistic y_m is not equal to the expected value of the cumulative distribution function at y_m. That is, except in the case of the uniform distribution,

$$F[E(y_m)] \neq E[F(y_m)] = \frac{m}{N+1}$$

in which the expression $m/(N+1)$ is the familiar Weibull (1939) formula. The plotting positions defined by $F[E(y_m)]$ provide unbiased quantile estimates (Cunnane, 1978), while those based on $E[F(y_m)]$ give unbiased estimates of the cumulative probabilities associated with particular values of y_m. Hence, if the objective of the probability plot is to obtain an unbiased estimate of the probability corresponding to a particular value of the variable under consideration hydrologists have usually favored the familiar Weibull (1939) formula. However, if the purpose of the plot is to test whether a set of data conforms to a hypothetical distribution, or to estimate either quantiles or distribution parameters the *unbiased* plotting positions were considered to be preferable as shown by Kimball (1960) and Cunnane (1978).

Given the attractive features of unbiased plotting positions in hydrologic frequency analyses (see, e.g., NERC, 1975a; Harter, 1984; Arnell et al., 1986), the present study was undertaken to develop exact and approximate unbiased plotting positions for two widely known distributions in hydrology: the Pearson type III (P3) and the General Extreme Value (GEV) distributions. The overall objective of this study is to propose new plotting formulas which would be more suitable for assessing the adequacy of the hypothetical distributions and, especially, can provide better estimates of flood quantiles than those given by existing formulas.

1.2 Statement of Problem

In hydrology four main distributions for flood frequency analysis are most often used : Log Pearson type III (LP3) distribution, Pearson type III (P3) distribution, General Extreme Value (GEV) distribution, and Generalized Gamma (GG) distribution. The LP3 distribution has been selected for use in flood frequency analysis in North America by the U. S. Water Resources Council (1967). The P3 distribution which contains the exponential distribution (skewness coefficient $\gamma = 2$) and normal distribution ($\gamma = 0$) as special cases, has also been frequently used in the U.S.A. and many other countries (e.g., Canada, Japan, and Thailand). The GEV distribution, which has Extreme value type I distribution (EV1, with shape parameter $\lambda = 0$), Extreme value type II distribution (EV2, $\lambda < 0$) and Extreme value type III distribution (EV3, $\lambda > 0$) as special cases, was recommended for use in Britain by NERC (1975a) and has recently been selected as a model in regional flood frequency analysis tests (Hosking et al., 1985). In the U.S.S.R., the GG distribution is still the most popular distribution used in flood frequency analysis (UNESCO, 1987).

For practical applications, it would be preferable to have a specific plotting position formula for each distribution mentioned above. Since the GG, P3 and LP3 distributions are all members of the gamma distribution family, a plotting position formula derived for the P3 can also be used for the other two distributions. The GEV distribution which represents a family of probability distributions for extreme values, does not however belong to the gamma group. Therefore, in the present study it is necessary to develop plotting formulas for two distinct distribution families represented respectively by the P3 and the GEV.

As mentioned above, several plotting position formulas have been proposed in the hydrological and statistical literature but very few were derived for three-parameter distributions, especially for the P3 and GEV distributions. Recent studies have provided some solutions to the problems, but these solutions are incomplete and have limited practical applications. For example, in the case of GEV and P3 distributions (Xuewu et al., 1984; Arnell et al., 1986), the parameters in the suggested plotting position formulas were shown to vary with the sample size, the skewness coefficient and the order number of the arranged sample. These parameters, however, were estimated only for some particular sample sizes and for some particular values of the skewness coefficient. The use of these formulas in practice is thus somewhat limited depending on the availability of the parameter values computed by Xuewu et al. (1984) and Arnell et al. (1986). No explicit relations between the parameters of the proposed formulas and the characteristics of the data sample were given.

Furthermore, most of the existing formulas are concerned with *systematic* flood records (i.e., complete flood samples which occurred during the period of systematic gauging), but little has been reported on the plotting position formulas for *historical* or *non-systematic* flood records (i.e., data on very large floods which occurred either outside or within the systematic gauging period). However, as shown by many recent investigations (see, e.g., Condie and Lee, 1982; Stedinger and Cohn, 1986; Hirsch and Stedinger, 1987; Sutcliffe, 1987) the incorporation of historical information about some extraordinary floods into formal frequency analysis would significantly improve the estimates of flood quantiles, especially for the quantiles with return periods of 50 years, 100 years or even 1000 years, which are often of greatest interest for purposes of design of various hydraulic structures (spillways, bridges, ...) or for purposes of flood plain zoning or insurance.

In view of the important role of the P3 and GEV distributions in engineering practice, and due to various limitations and problems concerning their application as described above, the present study is thus undertaken to develop new plotting position formulas for systematic and non-systematic flood records for these two distributions. The study consists of three parts. First, based on the Probability Weighted Moment (PWM) theory introduced by Greenwood et al. (1979) an analytical method will be proposed to derive the exact plotting positions for systematic flood records. Second, for the convenience of practical application, simple unbiased plotting formulas representing a very good approximation will be developed. Third, new unbiased plotting positions considering historical flood information will be proposed. The plotting formulas developed in this study can provide a better agreement to the exact plotting positions than several existing formulas, and they are, furthermore, conceptually more flexible and computationally more convenient, as will be shown in the following chapters.

Chapter 2 presents a literature review of existing plotting position formulas. Chapter 3 gives the theoretical consideration of P3, LP3, GG and GEV distributions. Chapter 4 describes the methodology used to develop new plotting position formulas. Chapter 5 presents the verification and comparison of various plotting positions. Chapter 6 contains some applications of the new formulas to observed flood data. Chapter 7 provides the conclusions of the present study.

1.3 Objectives of Study

The main objectives of this research are :

1. To develop new plotting position formulas for systematic flood records for the P3 and GEV distributions.

2. To develop new plotting position formulas for historical flood records for the above distributions.

3. To develop new probability papers for various skewness values for the two distributions considered.

The new plotting formulas and the special probability papers developed in the present study would provide a convenient and practical tool for the application of the P3 and GEV distributions in practice.

CHAPTER 2
LITERATURE REVIEW

The plotting position problem has been discussed in numerous research studies during the past 50 years. In the following, a brief review of the previous works concerning the plotting formulas for both systematic and historical flood records are presented.

2.1 Plotting Position Formulas for Systematic Flood Records

The first reference to "probability paper" is found in an article by Galton (1899). Galton compressed together or stretched apart, laterally, the ordinates on a sheet of ordinary graph paper so as to transform a normal ogive into a straight line. He termed the process by which a proper transformation of one variable enables us to represent a given curve by a straight line, anamorphic geometry. Probability paper was also used by hydrologists as early as 1896, but apparently was not mentioned in the literature on hydrology until 1914, in the papers by Fuller and Hazen and in the discussion on those papers. Hazen (1914) wrote (pp. 626-628) about the contribution to the discussion of the paper by Fuller (1914) " This is a most important paper, because, as far as the writer knows, it is the first attempt to apply the principles of probabilities to the flood problem." Probability paper was mentioned in the literature more than 30 times before 1950, mainly by hydrologists.

The cumulative probability or cumulative distribution function (CDF) of a sample of size N was usually defined as a step function which jumped from $(m-1)/N$

to m/N at the mth order statistic of the sample. If the plotting position m/N was used, the largest value could not be plotted, while if $(m-1)/N$ was used, the smallest value could not be plotted, since the probabilities 1 and 0 were off the scale of probability papers constructed for the normal distribution or for any other distributions unlimited in both extreme ends ($-\infty$ to $+\infty$). Hazen (1914) therefore suggested the compromise plotting position $(2m-1)/2N = (m-1/2)/N$, as an alternative means of including all values on the graph. Most hydrologists and other users during the next quarter century followed Hazen's suggestion. Some (e.g., Gerson, 1975 and Mage, 1982), however, persisted in using m/N, the so-called California method, because of its use by the California Department of Public Works (1923). Gumbel (1943) disagreed with Hazen's formula because the largest value of plotting position was plotted at $(1-1/2)/N$ which corresponded to a return period of $2N$, i.e., an artificial lengthening of the period of record. This statement was again pointed out by Benson (1962).

Kimball (1946) chose the only case for which $m/(N+1)$ was unbiased, the uniform distribution, to support a general inference that it was unbiased for all distributions. He recommended the plotting position $m/(N+1)$ for general use but noted that if $F(E(x_m))$, the population CDF at the expected value of the mth order statistic of the sample, could be estimated independently of the unknown parameters, such point might prove more desirable for graphical fitting near the extremes. After discussion by Gumbel (Kimball, 1947), he showed more favor to $F(E(y_m))$ as plotting position in which y_m was the reduced variate $y_m = (x_m - \mu)/\sigma$; where x_m was the mth order statistic of the sample and μ and σ were mean and standard deviation, respectively. In 1960, he tried to clarify the problem of choosing plotting positions using the normal and extreme value type I (EV1) papers. He considered several plotting positions, including (i) $m/(N+1)$; (ii) $(m-3/8)/(N+$

1/4); (iii) $(m - 1/2)/N$; (iv) $F(E[y_m])$; (v) F(median of y_m); and (vi) F(mode of y_m). After examination of the plotting positions in question, he showed that method (iv) was preferable for testing the fit of the hypothetical distribution to data and for estimating either quantiles or distribution parameters.

Gumbel (1947) stated four postulates to be satisfied by any plotting formula but did not offer any proofs as to their necessity. They were: (1) The plotting positions must be such that all members may be plotted. (2) The return period of a value equal to or larger than the largest observation and that of a value equal to or smaller than the smallest observation should converge towards N, the number of observations. He noted that this condition was not fulfilled in Hazen's method. (3) The observations should be equally placed on the frequency scale; that is the difference between the plotting positions of the $(m + 1)th$ and mth observation should be independent of m. (4) The plotting position ought to be simple and have an intuitive meaning. Gumbel (1958) repeated these postulates and added a fifth: (5) The plotting position should lie between the observed frequencies $(m - 1)/N$ and m/N and should be universally applicable, i.e., it should be distribution-free and this excluded the probabilities at the mean, median and modal values of y_m.

The above postulates have been cited in support of the Weibull formula $(m/(N + 1))$ by many papers (e.g., Singh and Sinclair, 1972; Yevjevich, 1972). Cunnane (1978) wrote (pp. 217-218) about these postulates : "No exception can be taken to postulate (1); in fact it is necessary. Postulate (2) is not in keeping with statistical fact. As already noted this postulate is most misleading and also appears to have played a major part in the adoption of the Weibull formula. Kimball (1947) questioned both the necessity and desirability of postulate (3), and one would tend to agree with him. Postulate (4), although desirable is not reconcilable with any mathematical derivation, as simplicity cannot be used in the same way as can, for instance, a boundary or initial condition, and consequently can play no part in the rational development of a formula. The simplicity and distribution-free postulates have obviously been rated highly in the opinion of many users."

Johnson (1951) tabulated the median plotting positions, which he called "median ranks", for $N < 20$, and gave an approximate formula for $N > 20$. Bernard and Bos-Levenbach (1953) showed that the median rank was closely approximated by the plotting position $(m - 0.30)/(N + 0.4)$. At about the same time, this plotting position was also advocated in Russian publications by Lebedev (1952), Chegodayev (1953), and others.

Chernoff and Lieberman (1956) used optimization methods, namely Lagrange multipliers, to estimate the standard deviation and the percentile. They concluded that estimates of standard deviation based on the plotting position $m/(N + 1)$ were much less efficient than those based on the position $(m - 1/2)/N$. This conclusion was advocated by Barnett (1975) but it was argued by Cunnane (1978). He wrote (pp. 218) : "This work was a fresh start to an old problem, and was novel in that it approached the problem from a statistical point of view setting out to find a formula not from probability arguments but rather from desirable statistical properties of the plot. They did, however, lose sight of the original problem and in fact derived coefficients for use with regression analysis rather than true plotting positions."

Blom (1958) introduced the α, α'-correction to the generalized mean value formula and applied these corrections to normal, EV1, and Weibull distributions. He proposed $P_m = (m - \alpha)/(N - \alpha' - \alpha + 1)$ as the general formula. If, for symmetry, one took $\alpha = \alpha'$, this became $(m - \alpha)/(N - 2\alpha + 1)$. Many previously suggested plotting positions were special cases of this general formula, e.g., $\alpha = 0$ gave the mean position $m/(N + 1), \alpha = 1$ gave the modal position $(m - 1)/(N - 1), \alpha = 0.5$ gave Hazen's compromise position $(m - 1/2)/N$ and $\alpha = 0.3$ gave a good approximation to the median position. Finally, he proposed $P_m = (m - 3/8)/(N + 1/4)$ as the plotting formula for the normal distribution. Gringorten (1963) modified also the general Blom's formula to obtain a formula for the double exponential distribution, $P_m = (m - 0.44)/(N + 0.12)$.

Chow (1964) summarized some plotting position formulas and demonstrated theoretically that the California method was suitable for plotting annual exceedance series or partial-duration series. Since this method could not be plotted on a probability paper for a probability of 100 percent thus, it was gradually replaced by the Hazen formula, which plotted data at the centers of group intervals. He found that all methods of determining plotting positions gave practically the same results in the middle of a distribution but produced different positions near the "tail" of the distribution. He came to the conclusion that the choice of a plotting position formula had become important in engineering practice.

Stipp and Young (1971) selected 20-year samples from 37 U.S. Geological Survey gauging stations providing regional coverage of the United States. Using these data, they computed the mean, standard deviation, and skewness coefficient, and then fitted the LP3 curve to data at each station. The frequencies of the highest and lowest discharges were determined from the curve, and the corresponding values of the constant α in the formula $P_m = (m - \alpha)/(N - 2\alpha + 1)$ were computed. He found that the best expression for the data used was $P_m = (m - 0.4)/(N + 0.2)$. This formula is very subjective in the sense that it is derived from the assumption of the symmetrical parent distribution (see Blom, 1958) but it was used to fit LP3 distribution which is not symmetric.

Cunnane (1978) pointed out that any quantile estimate made from the plot should be unbiased and should have smallest mean square error among all such estimates. He suggested $(m-\alpha)/(N-2\alpha+1)$ as the general form of plotting position formulas with $\alpha = 0$ for the Weibull case, $\alpha = 3/8$ for the normal, and $\alpha = 0.44$ for the EV1 and exponential distributions. He concluded that: (1) The expected value of the reduced variate order statistic, $E[y_m]$, depended on the form of the distribution being considered. (2) The Weibull formula was the exact probability corresponding to $E[y_m]$ when the distribution was uniform. (3) If the reduced variate depended on a shape parameter then the unbiased plotting position $E[y_m]$ depended on that parameter. Finally, he proposed $\alpha = 2/5$ as the best compromise for a single simple distribution free formula. This formula was adopted by Reich and Renard (1981) and Srikanthan and McMahon (1981) among many others.

Adamowski (1981) proposed the form of plotting formula as $P_m = (m-a)/(N+b)$ in which P_m were ordered plotting probability values, m was the rank of the mth value in an ordered sample of size N, and a and b were constants. The values of the constants a and b were derived by mean square error criterion. He found that $a = 0.25$ and $b = 0.50$ were the constants for his plotting position formula. This formula gave quite good result for EV1 distribution, especially for exceedance probabilities at high values. He recommended this formula for use in case of P3 distribution but the results might not be reliable for high floods.

King (1981) found that as data samples increased above a size of 20, the differences among the plotting positions determined by any method of estimation decrease to the point where they were practically unimportant. To evaluate the criteria for optimum choice of plotting positions proposed by Cunnane (1978) he has conducted a series of Monte Carlo experiments using also, for comparison purposes, Weibull formula, $m/(N+1)$, and Hazen formula $(m-0.5)/N$. Test results indicated that there is no practical difference in the estimates of the mean obtained by any of the three methods, but there is a highly significant difference among the several estimates of the sample standard deviation. The Weibull formula consistently overestimates the standard deviation; this represents, for some purposes, a conservative error. On the other hand, the Hazen formula consistently underestimates the standard deviation which results in unconservative errors. The Cunnane formula give estimates which average 1 % to 2 % high for the standard deviation. These estimates are, therefore, slightly conservative and practically irrelevant. Cunnane estimates are also consistently closer to the true population parameters than either of the other formulas.

Xuewu et al. (1984) showed that, for P3 distribution, the parameter α inside Blom's formula was mainly dependent on the order number m and the values of skewness coefficient γ of the parent distribution. They employed the Monte Carlo method to develop a plotting position formula for the P3 distribution for γ in the range of between 0 and 2. It was found that the formula could be applied to both symmetrical and unsymmetrical distributions and would also provide unbiased estimations of quantiles. Nevertheless, as mentioned in section 1.2, the convenience in the use of this formula in practice is somewhat limited because it cannot take explicitly into account the skewness coefficient of the underlying distribution, and cannot be applied for skewness values outside the range from 0 to 2. Further, the use of Monte Carlo simulation procedure could provide results less accurate than those given by the analytical method proposed in the present study.

Arnell et al. (1986) presented exact plotting positions for the GEV distribution, and a simple plotting formula such as would be suitable for a scientific calculator but yet which provided good approximation to the exact values. They used the PWM theory to derive the exact plotting positions. However, probably due to some errors in the mathematical derivation, the proposed approximate formulas did not give very good results, as will be shown in the present study. Moreover, the plotting relation suggested by Arnell et al. (1986) had limited practical applications as did Xuewu's formula, because it was also developed for some limited values of parameters and skewness coefficient.

Sinclair and Ahmad (1988) attempted to improve Arnell formula and proposed a new plotting relation for the GEV distribution based on the location-invariant concept. The new formula appears to be more covenient than the Arnell formula. The results given by this formula, however, are biased at both extreme ends of the probability plot, as will be shown in the present study.

From the review of the existing literature, it is clear that the optimum choice of a plotting position formula should be based on the purpose of the investigation and should also depend upon the distribution of the variable under consideration. It would be therefore preferable to have a specific plotting position formula for each particular distribution considered. In addition, the unbiased plotting positions advocated by Cunnane (1978) have been found to be suitable for various purposes of hydrologic frequency analyses.

2.2 Plotting Position Formulas for Historical Flood Records

As described in the previous section, a variety of plotting position formulas have been proposed for systematic flood records. However, very few papers have been reported on the subject of plotting formulas for historical flood data. Major citations on this subject include Benson (1950), World Meteorological Organization (1969), U.S. Water Resources Council (1977), Gerard and Karpuk (1979), Zhang (1982), Hirsch and Stedinger (1987), and Hirsch (1987).

Benson (1950) proposed a Weibull type plotting formula for floods above a threshold which was assumed to be known precisely. He recognized from his approach the possibility of a substantial discontinuity in the probabilities assigned to floods above and below the threshold. He wrote " In order to arrive at consistent results, it necessary to obtain an array of peaks properly representative of the single long period. The known peaks, historical and recent, must be combined in the proper proportions in order to obtain such an array."

NERC (1975a) proposed a formula based on Gringorten (1963) plotting positions. There were at least two problems with this formula. One was non-monotonicity, i.e., a large flood may be assigned a higher exceedance probability than a smaller flood. Another problem was that large gaps could occur between the probabilities assigned to the largest floods and the floods below a threshold level. This formula was used by Beable and McKerchar (1982) for some particular cases. Therefore, more guidances were needed in order to evaluate the proposed formula.

Gerard and Karpuk (1979) described a situation in which all floods greater than a threshold over a period of N years were known. Their method allowed a systematic analysis of all historical data available, but did not present a universally applicable formula.

Zhang (1982) derived a generalized plotting position formula based on order statistic theory. His formula could account for both large and small historical floods. The assumption involved in the derivation was that if the k largest historical floods were observed, they were observed because they were the k largest. Bernier et al. (1986) proposed a similar plotting position formula for partial duration series. The assumption given by Zhang was later on disagreed by Hirsch and Stedinger (1987), and Hirsch (1987). It was argued that historical and paleoflood discharges were often observed because they were large enough to exceed some *perception threshold* and hence to be recorded, but they were not necessarily the largest.

Hirsch and Stedinger (1987), and Hirsch (1987) proposed a general model for plotting position formulas which can combine both systematic and historical information in a consistent and statistically efficient manner. More specifically, the plotting formulas proposed can account for the rank of observation, the number of historical observed floods, and the lengths of the historical period and the systematic record. These formulas have been tested for bias in terms of discharges. It was found that none of the above formulas were highly accurate. Hirsch (1987) wrote "The exact magnitude of the bias depends, of course, on the family of distribution, the skewness coefficient and the expected number of largest floods, and consequently one could develop special, optimal plotting positions for each situation." It is clear from this study that an optimal plotting position formula should be able to take into account the skewness coefficient. Moreover, it would be preferable to have a specific plotting position formula for each particular distribution considered.

In summary, some plotting positions for historical or extraordinary floods have been proposed, but no consensus has been reached. It is noted from the literature review that several such formulas did not provide very accurate estimates of the largest floods because they were based on biased plotting positions (e.g., Hazen and Weibull formulas) developed for systematic flood records. It will be shown in the present study that new plotting positions proposed could give flood estimates much better than those estimated by several existing formulas.

CHAPTER 3
THEORETICAL CONSIDERATIONS

3.1 Fundamental Equations

The fundamental equations involved in the present study will be summarized in this chapter. The main distributions that will be described are P3, LP3, GG and GEV distributions. The most important parameters for these distributions (mean, variance, skewness coefficient) will be presented. The expected values of P3 and GEV order statistics will also be shown. Finally, the exact plotting positions for both P3 and GEV distributions are analytically derived using the PWM method.

3.1.1 P3 distribution

The P3 distribution involves three parameters. It can adopt every shape from the extremely skewed reverse-J shape to the symmetrical normal shape depending on the value of the shape parameter. The probability density function (PDF) of a random variable x which follows a P3 distribution may be expressed as follows :

$$f(x) = \begin{cases} \frac{1}{\beta^\lambda \Gamma(\lambda)} e^{-\frac{(x-x_0)}{\beta}} (x - x_0)^{\lambda-1}, & \text{for } x_0 \leq x < \infty \\ 0, & \text{for } x < x_0 \end{cases} \tag{3-1}$$

in which x_0 $(-\infty < x_0 < \infty)$ is the location parameter, β $(\beta > 0)$ is the scale parameter, and λ $(\lambda > 0)$ is the shape parameter. Its cumulative distribution function (CDF) is then given by :

$$F(x) = \int_{x_0}^{x} f(t)dt \tag{3-2}$$

where f(·) is defined by eqn. (3-1) above. The distribution parameters are related to the mean, μ, variance, σ^2, and skewness coefficient, γ, of the random variable x by the following relations :

$$\mu = x_0 + \beta\lambda \tag{3-3}$$

$$\sigma^2 = \beta^2\lambda \tag{3-4}$$

$$\gamma = \frac{2}{\sqrt{\lambda}} \tag{3-5}$$

The linear transformation $y = (x - x_0)/\beta$ reduces the P3 distribution to a standard form with $x_0 = 0$ and $\beta = 1$. That is,

$$F(y) = \frac{1}{\Gamma(\lambda)} \int_0^y e^{-y} y^{\lambda-1} dy = P(\lambda, y) \qquad \text{for } 0 \leq y < \infty \tag{3-6}$$

where $P(\lambda, y)$ is the incomplete gamma function. Hence, the standardized variate y has the PDF :

$$f(y) = \frac{1}{\Gamma(\lambda)} e^{-y} y^{\lambda-1} \tag{3-7}$$

If $\lambda = 1$, the P3 distribution reduces to the exponential. As $\lambda \to \infty$ or $\gamma \to 0$ the distribution tends to the normal distribution.

3.1.2 LP3 distribution

A random variable Z follows a LP3 distribution if its PDF is :

$$f(Z) = \frac{1}{\beta^{\lambda-1}|\beta|\Gamma(\lambda)Z} e^{-\frac{(lnZ - Z_0)}{\beta}} (lnZ - Z_0)^{\lambda-1} \tag{3-8}$$

where Z_0, β, and λ are respectively location, scale, and shape parameters for the LP3 distribution.

The population moments of the LP3 distribution are given by :

$$\mu'_{Zr} = exp(rZ_0)(1 - r\beta)^{-\lambda}, \qquad 1 - r\beta > 0, r = 1, 2, 3, \ldots \tag{3-9}$$

where μ'_{Zr} is the rth moment about the origin.

The mean, coefficient of variation, and skewness coefficient are expressed as :

$$\mu = exp(Z_0)(1-\beta)^{-\gamma} \tag{3-10}$$

$$C_v = A^{1/2}/B \tag{3-11}$$

$$\gamma = C/A^{3/2} \tag{3-12}$$

where

$$A = (1-2\beta)^{-\lambda} - (1-\beta)^{-2\lambda} \tag{3-13}$$

$$B = (1-\beta)^{-\gamma} \tag{3-14}$$

$$C = (1-3\beta)^{-\lambda} - 3[(1-\beta)(1-2\beta)]^{-\lambda} + 2(1-\beta)^{-3\lambda} \tag{3-15}$$

Note that except for the mean, any normalized higher statistics of the LP3 distribution such as coefficients of variation and skewness coefficient are independent of the location parameter, Z_0.

3.1.3 GG distribution

If it is supposed that $[(W-W_0)/\beta]^c = y$ (with $c > 0$) has the standard gamma distribution, then the PDF of a random variable W which follows a GG distribution is

$$f(W) = \frac{c(W-W_0)^{c\lambda-1}}{\beta^{c\lambda}\Gamma(\lambda)} exp\left[-\left(\frac{W-W_0}{\beta}\right)^c\right] \qquad (W \geq W_0) \tag{3-16}$$

where W_0, β, and λ are location, scale, and shape parameters for the GG distribution respectively.

This was defined (with $W_0 = 0$) by Stacy (1962) as one family of generalized gamma distribution. It includes Weibull distribution ($\lambda = 1$), half-normal distribution ($\lambda = \frac{1}{2}, c = 2, W_0 = 0$), and of course, P3 distribution ($c = 1$).

Stacy and Mihram (1965) proposed a method of estimation based on the moments of ln W. The results of the mean, coefficient of variation and skewness coefficient are expressed as :

$$\mu(lnW) = c^{-1}\psi(\lambda) + ln\beta \tag{3-17}$$

$$C_v = c^{-2}\psi'(\lambda) \tag{3-18}$$

$$\gamma = \psi''(\lambda)/\psi'(\lambda)^{3/2} \tag{3-19}$$

where

$$\psi^{(S)}(\lambda) = (-1)^{S-1}(S-1)!(\lambda - \frac{1}{2})^{-S} \qquad (S \geq 1) \tag{3-20}$$

3.1.4 GEV distribution

The GEV distribution involves 3 parameters and includes the Extreme Value type I (EV1) or Gumbel distribution as a special case (Jenkinson, 1955). A random variable x has a GEV distribution if its PDF has the following form :

$$f(x) = \frac{1}{\beta}\left[1 - \frac{\lambda(x - x_0)}{\beta}\right]^{\frac{1}{\lambda}-1} e^{-\left[1-\frac{\lambda(x-x_0)}{\beta}\right]^{\frac{1}{\lambda}}} \tag{3-21}$$

The CDF is

$$F(x) = \begin{cases} exp\left\{-\left[1 - \frac{\lambda(x-x_0)}{\beta}\right]^{\frac{1}{\lambda}}\right\}, & \lambda \neq 0 \\ exp\left\{-exp\left[-\frac{(x-x_0)}{\beta}\right]\right\}, & \lambda = 0 \end{cases} \tag{3-22}$$

where $x_0, \beta,$ and λ are respectively location, scale and shape parameters.

As λ tends to zero, the extreme value type 1 (EV1) distribution is obtained, and its reduced variate y_1 is related to the type 1 variate x_1 by

$$y_1 = (x_1 - x_0)/\beta \tag{3-23}$$

for which the CDF is

$$F_1 = F(y_1) = exp\{-exp(-y_1)\} \tag{3-24}$$

The mean, variance, and skewness coefficient of y_1 are expressed as :

$$\mu_{y_1} = 0.5772 \tag{3-25}$$

$$\sigma^2_{y_1} = \frac{\pi^2}{6\beta} \tag{3-26}$$

$$\gamma_{y_1} = 1.139 \tag{3-27}$$

Whereas, the mean, and variance of x_1 are :

$$\mu_{x_1} = x_0 + 0.5772\beta \tag{3-28}$$

$$\sigma_{x_1}^2 = \frac{\pi^2\beta}{6} \tag{3-29}$$

The skewness coefficient of x_1 is the same as that of y_1.

If $\lambda < 0$, the distribution becomes the Extreme Value type II (EV2) distribution, the reduced variable y_2 is $y_2 = 1 - \frac{\lambda(x_2 - x_0)}{\beta}$, and the CDF is

$$F_2 = F(y_2) = exp\left\{-(y_2^{1/\lambda})\right\} \tag{3-30}$$

The mean, variance, and skewness coefficient of y_2 are expressed as :

$$\mu_{y_2} = \Gamma(1+\lambda) \tag{3-31}$$

$$\sigma_{y_2}^2 = \Gamma(1+2\lambda) - \Gamma^2(1+\lambda) \tag{3-32}$$

$$\gamma_{y_2} = \mu_3/\mu_2^{3/2} \tag{3-33}$$

where $\mu_2 = \sigma_{y_2}^2$ and $\mu_3 = \Gamma(1+3\lambda) - 3\Gamma(1+2\lambda)\Gamma(1+\lambda) + 2\Gamma^3(1+\lambda)$

Consequently, the mean, and variance of x_2 are :

$$\mu_{x_2} = x_0 + \frac{\beta}{\lambda} - \frac{\beta}{\lambda}\mu_{y_2} \tag{3-34}$$

$$\sigma_{x_2}^2 = (\frac{\beta}{\lambda})^2\sigma_{y_2}^2 \tag{3-35}$$

The skewness coefficient of x_2 is the same as that of y_2.

If $\lambda > 0$, the distribution becomes the Extreme Value type III (EV3) distribution, the reduced variable y_3 is $y_3 = -\{1 - \frac{\lambda(x_3 - x_0)}{\beta}\}$, and the CDF is :

$$F_3 = F(y_3) = exp\left\{-(-y_3)^{1/\lambda}\right\} \tag{3-36}$$

The mean, variance, and skewness coefficient of y_3 are expressed as :

$$\mu_{y_3} = -\Gamma(1+\lambda) \tag{3-37}$$

$$\sigma_{y_3}^2 = \Gamma(1+2\lambda) - \Gamma^2(1+\lambda) \tag{3-38}$$

$$\gamma_{y_3} = \mu_3/\mu_2^{3/2} \tag{3-39}$$

where $\mu_2 = \sigma_{y_3}^2$ and $\mu_3 = -\Gamma(1+3\lambda) + 3\Gamma(1+2\lambda)\Gamma(1+\lambda) - 2\Gamma^3(1+\lambda)$

According to eqns. (3-37) and (3-38) the mean, and variance of x_3 can be written as :

$$\mu_{x_3} = x_0 + \frac{\beta}{\lambda} + \frac{\beta}{\lambda}\mu_{y_3} \tag{3-40}$$

$$\sigma_{x_3}^2 = (\frac{\beta}{\lambda})^2 \sigma_{y_3}^2 \tag{3-41}$$

The skewness coefficient of x_3 is independent of the location and scale parameters and equals the skewness of the y_3 variate.

It can be seen from eqns. (3-1), (3-8) and (3-16) that the LP3 and GG distributions are related to the P3 by means of a transformation of variable. The LP3 distribution can be transformed to the P3 by taking the logarithm of the variable. For the GG distribution, its random variable W is related to the P3 random variable y by the relation $y = (W - W_0)/\beta$. The GEV distribution, however, does not belong to the gamma distribution family. Hence, in the present study it is necessary to develop new plotting position formulas for the P3 and GEV distributions only. The development of these formulas will be illustrated in the next chapter.

3.2 Expected Values of P3 Order Statistics

Consider an ordered random sample of N observations, $y_1 \geq y_2 \geq \ldots \geq y_N$. The probability density function $g(y_m)$ of the mth element of the ordered sample is then given by :

$$g(y_m) = \frac{N!}{(m-1)!(N-m)!} F(y)^{N-m} [1 - F(y)]^{m-1} f(y) \tag{3-42}$$

where $F(y)$ and $f(y)$ are respectively the cumulative distribution function and the density function of the random variate y. The expectation of the mth order statistic, y_m, is therefore :

$$E[y_m] = \int_0^\infty y g(y_m) dy \tag{3-43}$$

or from eqn. (3-42) :

$$E[y_m] = \frac{N!}{(m-1)!(N-m)!} \int_0^1 yF(y)^{N-m}[1-F(y)]^{m-1} dF(y) \qquad (3\text{-}44)$$

Substitute eqns. (3-6) and (3-7) into (3-42) :

$$g(y_m) = \frac{N!}{(m-1)!(N-m)!} \left[\frac{1}{\Gamma(\lambda)} \int_0^y y^{\lambda-1} e^{-y} dy\right]^{N-m}$$
$$\left[1 - \frac{1}{\Gamma(\lambda)} \int_0^y y^{\lambda-1} e^{-y} dy\right]^{m-1} \frac{1}{\Gamma(\lambda)} y^{\lambda-1} e^{-y} \qquad (3\text{-}45)$$

and substitute eqn. (3-45) into (3-43) :

$$E(y_m) = \frac{N!}{(m-1)!(N-m)!} \int_0^\infty \frac{1}{\Gamma(\lambda)} y^\lambda e^{-y} \left[\frac{1}{\Gamma(\lambda)} \int_0^y y^{\lambda-1} e^{-y} dy\right]^{N-m}$$
$$\left[1 - \frac{1}{\Gamma(\lambda)} \int_0^y y^{\lambda-1} e^{-y} dy\right]^{m-1} dy \qquad (3\text{-}46)$$

Note that eqn. (3-46) produces the exact plotting positions for the P3 distribution. However, it can be observed that it is impossible to integrate explicitly the integral involved in the equation, and it is a formidable task to evaluate numerically this expression. By means of numerical integration, Harter (1964) provided tables for the expected value of the mth-order statistic y_m for $m = 1, 2, \ldots, N; N = 1, 2, \ldots, 40$; and for $\lambda = 0.5, 1, \ldots, 4.0$. It was noted that, probably due to the formidability of the numerical integration task, the results for $E(y_m)$ were given for sample sizes not greater than 40. More recently, using the Monte Carlo method, values of $E(y_m)$ for larger samples, $N = 50$ and $N = 100$, have been computed by Xuewu et al.(1984) but only for some particular values of the skewness coefficient γ (γ was limited to values below 2). Nevertheless, the Monte Carlo method could consume considerable computer resources and could provide less accurate results. Therefore, in section 3.4, it will be shown that a simpler procedure for evaluating $E(y_m)$ for the P3 distribution can be achieved using the PWM theory.

3.3 Expected Values of GEV Order Statistics

From eqn. (3-24), the reduced variate y_1 of the EV1 distribution can be expressed as :

$$y_1 = -ln[-ln(F_1)] \tag{3-47}$$

Hence, according to eqn. (3-44) the expected value of the mth order statistic for the EV1 distribution can be written as :

$$E[y_{1m}] = \frac{N!}{(m-1)!(N-m)!} \int_0^1 -ln(-lnF_1)F_1^{N-m}(1-F_1)^{m-1}dF_1 \tag{3-48}$$

Similarly, on the basis of eqns. (3-30), (3-36), and (3-44), the corresponding expressions for the EV2 and EV3 distributions are :

$$E[y_{2m}] = \frac{N!}{(m-1)!(N-m)!} \int_0^1 (-lnF_2)^{\lambda} F_2^{N-m}(1-F_2)^{m-1}dF_2 \tag{3-49}$$

$$E[y_{3m}] = \frac{-N!}{(m-1)!(N-m)!} \int_0^1 (-lnF_3)^{\lambda} F_3^{N-m}(1-F_3)^{m-1}dF_3 \tag{3-50}$$

Equations (3-48), (3-49) and (3-50) provide the exact plotting positions for the GEV distribution. However, it is not possible to integrate analytically the above equations, and consequently numerical integration techniques must be employed. The numerical integration requires normally a very substantial amount of computer time. In the following section we will present a simpler procedure for evaluating $E(y_m)$ for the GEV distribution using the PWM theory.

3.4 Exact Plotting Positions from Probability Weighted Moments

The probability weighted moments are defined as (Greenwood et al., 1979) :

$$M_{l,i,j} = \int_0^1 y^l F(y)^i [1-F(y)]^j dF(y) \tag{3-51}$$

where l, i, and j are real numbers, and $F(y) = P(Y \leq y)$. If $i = j = 0$ and l is a nonnegative integer, then $M_{l,0,0}$ is the conventional lth moment about the origin. In particular, if l, i and j are nonnegative integers, $M_{l,i,j}$ is proportional to the lth moment about the origin of the $(j+1)th$ order statistic for a sample of size $(i+j+1)$ (Greenwood et al., 1979). That is,

$$M_{l,i,j} = \frac{i!j!}{(i+j+1)!} E(y^l_{j+1,i+j+1}) \tag{3-52}$$

Suppose $N = i + j + 1$ and $m = j + 1$, then from eqn. (3-52):

$$E(y^l_{m,N}) = \frac{N!}{(m-1)!(N-m)!} M_{l,N-m,m-1} \tag{3-53}$$

For the first moment $l = 1$, eqn. (3-53) becomes

$$E[y_m] = \frac{N!}{(m-1)!(N-m)!} M_{1,N-m,m-1} \tag{3-54}$$

Furthermore, $M_{1,N-m,m-1}$ can be expressed as a sum of simpler moments as follows (Greenwood et al., 1979) :

$$M_{1,N-m,m-1} = \sum_{s=0}^{m-1} \binom{m-1}{s} (-1)^s M_{1,N-m+s,0} \tag{3-55}$$

which is substituted into eqn. (3-54) to give :

$$E[y_m] = \frac{N!}{(m-1)!(N-m)!} \sum_{s=0}^{m-1} \binom{m-1}{s} (-1)^s M_{1,N-m+s,0} \tag{3-56}$$

In the case of P3 distribution, given the expressions in (3-6), (3-7) and (3-52), the probability weighted moments $M_{1,N-m+s,0}$ can be expressed as :

$$M_{1,N-m+s,0} = \frac{1}{\Gamma(\lambda)} \int_0^{\infty} [P(\lambda, y)]^{N-m+s} y^{\lambda} e^{-y} dy \tag{3-57}$$

in which $P(\lambda, y)$ is the incomplete gamma function.

By comparing eqn. (3-56) with eqn. (3-46), it is noted that the relation (3-56) has a simpler analytical structure and thus requires a simpler computation scheme because it involves a finite summation. More specifically, eqns. (3-56) and (3-57) are finite series and thus can be evaluated without difficulty to provide exact plotting positions for the P3 distribution. The computations of the expected values of y_m, eqn. (3-56), were performed in double precision on the mainframe IBM 4381 computer for $m = 1, 2, \ldots, N$; $N = 5, 10, \ldots, 100$; and for skewness coefficients $\gamma = 0, 0.1, \ldots, 3.0$.

For purposes of illustration, Table 3.1 shows a comparison of expected values estimated by the direct numerical integration procedure (Harter, 1964), by the Monte Carlo technique (Xuewu et al., 1984), and by the PWM theory, eqn. (3-56), for $N = 30$ and $\gamma = 1.265$ (or $\lambda = 2.5$). It can be observed that the results obtained by the random sampling technique are less accurate than those given by the direct numerical integration and PWM methods, even though a very large number of random samples (30,000 samples of size $N = 30$) have been generated (Xuewu et al., 1984). Furthermore, as

Table 3.1: Comparison of expected values of P3 order statistics, $E(y_m)$, for skewness coefficient $\gamma = 1.265, N = 30$.

Rank (m)	Harter	Xuewu	PWM	e_1 (%)	e_2 (%)
1	6.76301	6.75855	6.76309	0.0012	0.067
2	5.51268	5.51051	5.51269	0.0002	0.040
3	4.85972	4.85204	4.85968	0.0008	0.157
4	4.41026	4.40302	4.41022	0.0009	0.163
5	4.06379	4.05732	4.06375	0.0010	0.158
10	2.96017	2.96097	2.96017	0.0000	0.027
15	2.25378	2.25263	2.25378	0.0000	0.051
20	1.68273	1.68547	1.68274	0.0006	0.162
25	1.14056	1.13869	1.14056	0.0000	0.164
30	0.42069	0.42097	0.42069	0.0000	0.067

e_1 = relative difference between estimations by Harter and PWM

e_2 = relative difference between estimations by Xuewu and PWM compared with the numerical integration technique, the PWM procedure produced comparable results, and its simpler computational scheme permitted to evaluate the expected values of P3 order statistics for large sample sizes without any difficulty. This suggests that the PWM procedure is preferable to the direct numerical integration and Monte Carlo methods in the computation of exact plotting positions for the P3 distribution. Therefore, the PWM method will be used in this study to develop an approximate and unbiased plotting position formula for the P3 distribution.

In the case of GEV distribution, the expected values of the order statistics y_{1m}, y_{2m} and y_{3m} can be computed more simply using the PWM theory (Arnell et al., 1986) :

$$E[y_{1m}] = \frac{N!}{(m-1)!(N-m)!} \sum_{s=0}^{m-1} \binom{m-1}{s} (-1)^s \left[\delta + ln(N-m+s+1)\right] \cdot (N-m+s+1)^{-1} \tag{3-58}$$

$$E[y_{2m}] = \frac{N!}{(m-1)!(N-m)!} \Gamma(1+\lambda) \sum_{s=0}^{m-1} \binom{m-1}{s} (-1)^s \cdot (N-m+s+1)^{-(1+\lambda)} \tag{3-59}$$

$$E[y_{3m}] = -E[y_{2m}] \tag{3-60}$$

where $\Gamma(\cdot)$ and δ are complete gamma function and Euler's constant ($\delta = 0.5772$) respectively.

By comparing with eqns. (3-48), (3-49) and (3-50), eqns. (3-58), (3-59) and (3-60) are finite series and thus can be evaluated without difficulty to provide exact plotting positions for the GEV distribution. The computations of the expected values of y_m, eqns. (3-58), (3-59) and (3-60), were performed in double precision on the mainframe IBM 4381 computer for $m = 1, 2, \ldots, N; N = 5, 10, \ldots, 100$; and for shape parameter $\lambda = -0.2, -0.1, \ldots, 1.5$ (skewness coefficient $\gamma = 3.535, 1.903, \ldots, -3.802$).

CHAPTER 4

METHODOLOGY

4.1 Development of New Plotting Position Formulas for P3 and GEV Distributions for Systematic Flood Records

This chapter is intended to illustrate the development of new plotting position formulas and probability papers for P3 and GEV distributions. In this study the plotting position, P_m is defined as :

$$P_m = F[E(y_m)] \tag{4-1}$$

where $F[E(y_m)]$ is the cumulative distribution function of the expected value of y_m. If $E(y_m)$ is computed using eqns. (3-56) and (3-57), the resulting values of P_m represent the exact plotting positions for the P3 distribution. Similarly, if $E(y_m)$ is estimated using eqns. (3-58), (3-59), and (3-60), the exact plotting positions for EV1, EV2, and EV3 are respectively obtained. However, this method is cumbersome in engineering practice because it requires evaluation of the summation and the complete and incomplete gamma functions in those equations. Consequently, for the convenience of practical applications, it is desirable to develop simpler plotting position formulas which will be derived using the exact plotting positions calculated by the PWM method.

Blom (1958) proposed a general form for plotting position formulas as :

$$P'_m = (m - \alpha)/(N - \alpha' - \alpha + 1), \qquad \alpha, \alpha' < 1 \tag{4-2}$$

in which P'_m is the probability of mth order statistic; and α, α' are coefficients depending on parent distribution and sample size N. If the parent distribution is symmetric, Blom (1958) proved also that $\alpha' = \alpha$. Hence, eqn. (4-2) becomes :

$$P_m = (m - \alpha)/(N - 2\alpha + 1) \tag{4-3}$$

Note that the condition of symmetry imposed by eqn. (4-3) is not a theoretical requirement. The expression given by eqn. (4-3) can also be used, by a proper choice of α, for non-symmetric distributions. More specifically, if the form of the parent distribution is characterized by a shape parameter, the value of α also depends on that parameter (see, e.g., Cunnane, 1978; Xuewu et al., 1984; Harter, 1984; Nguyen et al., 1989; In-na and Nguyen, 1989).

In this study, an approximation to the exact P3 and GEV plotting positions can be achieved by restating the general formula, eqn. (4-3), as :

$$P_m = \frac{m + b}{N - a} \tag{4-4}$$

where a and b are parameters which vary with the sample size N and the skewness coefficient γ of the parent distribution. From Table 4.1 it is noted that most plotting position formulas used by hydrologists can be expressed as special cases of this general expression. The values of the coefficients a and b will be estimated in this study using the least-squares technique and based on the exact plotting positions previously computed by the PWM method, as will be shown in the following.

Equation (4-4) can be written as :

$$NP_m - m = aP_m + b \tag{4-5}$$

Table 4.1: Plotting position formulas (after Cunnane, 1978; Harter, 1984; Xuewu et al., 1984).

Method	Formula	a and b in $P_m = \frac{m+b}{N-a}$
Hazen(1914)	$\frac{m-0.5}{N}$	$a = 0, b = -0.5$
California(1923)	$\frac{m}{N}$	$a = 0, b = 0$
Foster(1936)	$\frac{2m-1}{2N}$	$a = 0, b = -1/2$
Weibull(1939)	$\frac{m}{N+1}$	$a = -1.0, b = 0$
Beard(1943)	$\frac{m-0.31}{N+0.38}$	$a = -0.38, b = -0.31$
Benard and Bos-Levenbach(1953)	$\frac{m-0.3}{N+0.2}$	$a = -0.2, b = -0.3$
Chegodayev(1955)	$\frac{m-0.3}{N+0.4}$	$a = -0.4, b = -0.3$
Blom(1958)	$\frac{m-3/8}{N+1/4}$	$a = -1/4, b = -3/8$
Tukey(1962)	$\frac{m-1/3}{N+1/3}$	$a = -1/3, b = -1/3$
Gringorten(1963)	$\frac{m-0.44}{N+0.12}$	$a = -0.12, b = -0.44$
Cunnane(1978)	$\frac{m-0.4}{N+0.20}$	$a = -0.2, b = -0.4$
Adamowski(1981)	$\frac{m-0.25}{N+0.5}$	$a = -0.5, b = -0.25$

It can be observed that the right hand side of eqn. (4-5) is a linear function of P_m with a and b considered as unknown coefficients. Hence, for known values of $m(m = 1, 2, \ldots, N)$; $N(N = 5, 10, \ldots, 100)$; and P_m (computed by the PWM method) for different values of skewness coefficient γ (or shape parameter λ), a regression of $(NP_m - m)$ on P_m yields the least-squares solution for a and b. Some values of these coefficients are shown in Tables 4.2 and 4.3 for the P3 and the GEV distributions, respectively. It can be observed that a and b vary systematically and, in particular, are more sensitive to the sample size N than to the skewness

Table 4.2: Example of some values of coefficients a and b for P3 distribution.

N	γ	a	b
	3.00	10.987	-6.428
	2.80	10.898	-6.332
	2.50	10.756	-6.202
	2.20	10.641	-6.087
	2.00	10.574	-6.018
	1.80	10.515	-5.957
5	1.50	10.441	-5.878
	1.20	10.384	-5.812
	1.00	10.354	-5.774
	0.80	10.331	-5.741
	0.50	10.380	-5.723
	0.00	10.288	-5.644
	3.00	101.024	-51.499
	2.80	100.952	-51.414
	2.50	100.772	-51.243
	2.20	100.606	-51.089
	2.00	100.519	-51.006
	1.80	100.447	-50.934
50	1.50	100.361	-50.845
	1.20	100.298	-50.773
	1.00	100.268	-50.734
	0.80	100.243	-50.699
	0.50	100.407	-50.716
	0.00	100.213	-50.606
	3.00	200.738	-101.318
	2.80	200.898	-101.384
	2.50	200.795	-101.262
	2.20	200.608	-101.094
	2.00	200.513	-101.005
	1.80	200.438	-100.931
100	1.50	200.351	-100.840
	1.20	200.287	-100.768
	1.00	200.256	-100.728
	0.80	200.231	-100.693
	0.50	200.575	-100.762
	0.00	200.149	-100.575

Table 4.3: Example of some values of coefficients a and b for GEV distribution.

N	λ	γ	a	b
	-0.2	3.535	10.254	-5.769
	-0.1	1.903	10.288	-5.758
	0.0	1.139	10.321	-5.746
5	0.1	0.638	10.384	-5.744
	0.5	-0.631	10.478	-5.682
	1.0	-2.000	10.600	-5.600
	1.1	-2.309	10.625	-5.583
	1.5	-3.802	10.728	-5.515
	-0.2	3.535	100.119	-50.720
	-0.1	1.903	100.152	-50.703
	0.0	1.139	100.185	-50 686
50	0.1	0.638	100.213	-50.667
	0.5	-0.631	100.347	-50.600
	1.0	-2.000	100.515	-50.514
	1.1	-2.309	100.548	-50.497
	1.5	-3.802	100.683	-50.428
	-0.2	3.535	200.099	-100.713
	-0.1	1.903	200.133	-100.696
	0.0	1.139	200.167	-100.678
100	0.1	0.638	200.199	-100.661
	0.5	-0.631	200.335	-100.592
	1.0	-2.000	200.507	-100.600
	1.1	-2.309	200.542	-100.490
	1.5	-3.802	200.680	-100.422

coefficient γ. Therefore, for a given value of the skewness γ, it can be assumed that

$$a = C_1 N + C_2 \tag{4-6}$$

and

$$b = C_3 N + C_4 \tag{4-7}$$

where C_1, C_2, C_3 and C_4 are parameters.

Since a, b and N are known (Tables 4.2 and 4.3), the least-squares method can be applied to eqns. (4-6) and (4-7) to evaluate C_1, C_2, C_3, and C_4. Tables 4.4 and

4.5 show for selected values of the skewness coefficient γ the estimated values of these coefficients for P3 and GEV distributions. It is found that only C_2 and C_4 vary with γ, while C_1 and C_3 are approximately constant ($C_1 \approx 2$ and $C_3 \approx -1$). Therefore, the least-squares method again can be used to derive linear relations between γ and the parameters C_2 and C_4. Results are shown in the following equations for the P3 distribution :

$$C_2 = 0.30\gamma + 0.05 \tag{4-8}$$

$$C_4 = -0.30\gamma - 0.47 \tag{4-9}$$

and for the GEV distribution :

$$C_2 = -0.08\gamma + 0.38 \tag{4-10}$$

$$C_4 = -0.05\gamma - 0.65 \tag{4-11}$$

Table 4.4: Example of some values of constants C_1, C_2, C_3, and C_4 for P3 distribution.

γ	C_1	C_2	C_3	C_4
3.00	1.997	1.112	-0.999	-1.527
2.80	2.000	0.929	-1.000	-1.377
2.50	2.000	0.748	-1.000	-1.212
2.20	2.000	0.620	-1.000	-1.086
2.00	2.000	0.547	-1.000	-1.012
1.80	2.000	0.482	-1.000	-0.946
1.50	1.999	0.403	-1.000	-0.862
1.20	1.999	0.343	-1.000	-0.793
1.00	1.999	0.313	-1.000	-0.755
0.80	1.999	0.289	-1.000	-0.722
0.50	2.002	0.314	-1.001	-0.696
0.00	1.999	0.251	-1.000	-0.625

Table 4.5: Example of some values of constants C_1, C_2, C_3, and C_4 for GEV distribution.

λ	γ	C_1	C_2	C_3	C_4
-0.2	3.535	1.9988	0.1998	-0.9995	-0.7490
-0.1	1.903	1.9988	0.2309	-0.9995	-0.7340
0.0	1.139	1.9988	0.2618	-0.9995	-0.7188
0.1	0.638	1.9988	0.2936	-0.9995	-0.7036
0.2	0.254	1.9989	0.3221	-0.9994	-0.6874
0.3	-0.069	1.9989	0.3505	-0.9994	-0.6709
0.4	-0.359	1.9990	0.3791	-0.9994	-0.6544
0.5	-0.631	1.9991	0.4078	-0.9994	-0.6378
0.6	-0.896	1.9991	0.4367	-0.9994	-0.6211
0.7	-1.160	1.9992	0.4657	-0.9994	-0.6039
0.8	-1.430	1.9993	0.4949	-0.9994	-0.5875
0.9	-1.708	1.9993	0.5240	-0.9994	-0.5705
1.0	-2.000	1.9994	0.5536	-0.9994	-0.5536
1.1	-2.309	1.9994	0.5831	-0.9994	-0.5365
1.2	-2.640	1.9995	0.6128	-0.9994	-0.5193
1.3	-2.996	1.9995	0.6427	-0.9994	-0.5021
1.4	-3.382	1.9996	0.6727	-1.0002	-0.4595
1.5	-3.802	1.9997	0.7029	-0.9994	-0.4673

Given the results obtained by the least-squares method, the parameters a and b for the P3 distribution can be expressed as functions of the skewness coefficient γ and the sample size N as follows :

$$a = 2N + 0.30\gamma + 0.05 \qquad (4\text{-}12)$$

$$b = -N - 0.30\gamma - 0.47 \qquad (4\text{-}13)$$

Similarly, for the GEV distribution expressions for the parameters a and b are :

$$a = 2N - 0.08\gamma + 0.38 \qquad (4\text{-}14)$$

$$b = -N - 0.05\gamma - 0.65 \qquad (4\text{-}15)$$

Hence, by substituting eqns. (4-12) and (4-13) into eqn. (4-4) the unbiased plotting position formula for the P3 distribution (called *P3 formula* in this study) has the following form :

$$P_m = \frac{N - m + 0.3\gamma + 0.47}{N + 0.3\gamma + 0.05} \tag{4-16}$$

Similarly, on the basis of eqns. (4-14), (4-15), and (4-4), a new unbiased plotting formula for the GEV distribution (hereafter referred to as *GEV formula*) can be written as :

$$P_m = \frac{N - m + 0.05\gamma + 0.65}{N - 0.08\gamma + 0.38} \tag{4-17}$$

It can be seen that the new plotting position formulas proposed here can take explicitly into account the skewness coefficient of the parent distribution, and in addition, they have a simple structure as do most existing formulas. In particular, the simple expressions given by eqns. (4-16) and (4-17) appear to be preferable to the ones suggested by Xuewu et al. (1984) for the P3 distribution, and by Arnell et al. (1986) for the GEV distribution because, to use Xuewu and Arnell formulas, the parameter a and b in eqn. (4-4) for a given set of γ, N and m must be estimated, and the values of these parameters were tabulated only for some selected values of γ, N and m. No explicit relation between α and the set γ, N and m was given. The convenience in the application of Xuewu and Arnell formulas in practice is thus somewhat limited, depending on the availability of the estimated values of the parameters a and b. The formulas suggested in this study, however, can be readily used for various sample sizes $N(5 \leq N \leq 100)$ and for a wide range of skewness values $\gamma(-3.0 \leq \gamma \leq +3.0)$. Noted is the validity of the proposed P3 formula for negative skewness values because of the symmetrical property of P3 order statistics for positive and negative skewness coefficients (Bobee and Morin, 1973). Note also that the limitations on the values of N and γ were selected to represent most conditions frequently encountered in hydrologic frequency analyses. The exact plotting positions given by the PWM theory however should be, in theory, valid for any sample size or skewness value.

The new plotting position formulas developed in this section can be used only for the case of systematic flood records. More general formulas which can take into account the historical information of very large floods will be developed in the next section.

4.2 Development of New Plotting Position Formulas for P3 and GEV Distributions for Historical Flood Records

As mentioned previously, most hydrometric records are available only for a relatively short period of time. The probability estimates of rare events are therefore unreliable. Obviously, any historic information which effectively enlarges the sample size would significantly improve the frequency analysis of such events. Therefore, the objective of this section is to develop new plotting position formulas which can take into account the historic information of extraordinary floods

In this study, to describe plotting positions for historical flood records, some standard notations are introduced. Let N (in years) be the length of the historical period (need not be continuous). This N-year period contains some systematic flood record period of s years ($s \leq N$). Let g be the number of observed floods in complete flood records where $s \leq g < N$. Among these floods, k of them are known to be the k largest in a period of N years. Some of these k largest floods, e, may have occurred during the systematic flood records ($e \leq k$ and $e \leq s$). Note that $g = s + k - e$.

The assumption involved in this study is that there is a perception threshold or *base level* Q_0 such that the k largest floods are larger than or equal to it and the remainder are smaller than it (Hirsch, 1987). In addition, we have records of k floods because they were large, but not because they were the k largest. On the basis of this assumption Hirsch and Stedinger (1987) proposed a general form for plotting formulas with historical information as :

$$\hat{P}_m = \begin{cases} \frac{m-\alpha}{k+1-2\alpha} \cdot P_e & m = 1 \ldots, k \\ P_e + (1 - P_e) \cdot \frac{m-k-\alpha}{s-e+1-2\alpha} & m = k+1, \ldots, g \end{cases} \qquad (4\text{-}18)$$

in which $\hat{P}_m$ is the probability of mth order statistic; α is a coefficient depending on the parent distribution and the sample size N; and P_e is the probability that flood will equal or exceed the base level Q_0. The maximum likelihood (and the method of moments) estimator of P_e is k/N. This formula is applicable only for symmetric parent distribution. If the parent distribution is not symmetric, then eqn. (4-18) may be rewritten (Blom, 1958) as :

$$\hat{P}_m = \begin{cases} \frac{m-\alpha}{k+1-\alpha'-\alpha} \cdot \frac{k}{N} & m = 1, \ldots, k \\ \frac{k}{N} + \frac{N-k}{N} \cdot \frac{m-k-\alpha}{s-e+1-\alpha'-\alpha} & m = k+1, \ldots, g \end{cases} \qquad (4\text{-}19)$$

Note that α' is a coefficient depending on the parent distribution, and note also that if the parent distribution is symmetric ($\alpha' = \alpha$), eqn. (4-19) will become eqn. (4-18). The condition of symmetry imposed by eqn. (4-18) is not a theoretical requirement as indicated in section 4.1. Therefore, the expression given by eqn. (4-18) can also be used, by a proper choice of α, for non-symmetric distributions. Hence, if the form of the parent distribution is characterized by a shape parameter, the value of α also depends on that parameter.

In this research, approximation to the exact P3 and GEV plotting positions with historical information can be achieved by restating the general formula, eqn. (4-19), as :

$$\hat{P}_m = \begin{cases} \frac{m+b}{k-a} \cdot \frac{k}{N} & m = 1, \ldots, k \\ \frac{k}{N} + \frac{N-k}{N} \cdot \frac{m-k+b}{s-e-a} & m = k+1, \ldots, g \end{cases} \tag{4-20}$$

where a and b are parameters. We call this the *E formula* (for exceedances). From this general relation, one can form an Exceedance-Weibull (E-W) formula by setting $a = -1, b = 0$, an Exceedance-Blom (E-B) formula with $a = -0.25, b = -0.375$, an Exceedance-Cunnane (E-C) formula with $a = -0.20, b = -0.40$, an Exceedance Adamowski formula (E-A) with $a = -0.50, b = -0.25$, and an Exceedance-Gringorten (E-G) formula with $a = -0.22, b = -0.44$. Since an objective of this research is to develop unbiased (in terms of discharges) plotting position formulas when flood data follow P3 and GEV distributions, the parameters a and b, therefore, must be selected in such a way that the above formula yields a very good estimates of these floods.

The P3 and GEV formulas (eqns. (4-16) and (4-17)) introduced in the previous section for systematic flood records will be used to develop new plotting position formulas for historical floods. These formulas can be re-written in term of exceedance probability as follows :

$$\hat{P}_m = 1 - P_m = \frac{m - 0.42}{N + 0.3\gamma + 0.05} \tag{4-21}$$

for the P3 distribution, and

$$\hat{P}_m = 1 - P_m = \frac{m - 0.13\gamma - 0.27}{N - 0.08\gamma + 0.38} \tag{4-22}$$

for the GEV distribution. It can be seen that the formulas given by eqns (4-21) and (4-22) can take explicitly into account the skewness coefficient of the parent distribution. Note from eqn. (4-21) that $a = -0.3\gamma - 0.05$, and $b = -0.42$ and from eqn. (4-22) that $a = 0.08\gamma - 0.38$, and $b = -0.13\gamma - 0.27$. Substituting these expressions for a and b into eqn. (4-20) to obtain for the P3 ditribution :

$$\hat{P}_m = \begin{cases} \frac{m-0.42}{k+0.3\gamma+0.05} \cdot \frac{k}{N} & m = 1, \ldots, k \\ \frac{k}{N} + \frac{N-k}{N} \cdot \frac{m-k-0.42}{s-e+0.3\gamma+0.05} & m = k+1, \ldots, g \end{cases} \tag{4-23}$$

and for the GEV:

$$\hat{P}_m = \begin{cases} \frac{m-0.13\gamma-0.27}{k-0.08\gamma+0.38} \cdot \frac{k}{N} & m = 1, \ldots, k \\ \frac{k}{N} + \frac{N-k}{N} \cdot \frac{m-k-0.13\gamma-0.27}{s-e-0.08\gamma+0.38} & m = k+1, \ldots, g \end{cases} \tag{4-24}$$

Equations (4-23) and (4-24) represent plotting position formulas for historical flood records for the P3 and GEV distributions. These equations are respectively called Exceedance P3 (E-P3) formula and Exceedance GEV (E-GEV) formula.

4.3 Probability Papers

As mentioned previously, an important advantage in the use of plotting position formulas is the possibility of plotting and visual comparison of the cumulative frequency curve of the data and the assumed probability law. This plot permits an immediate assessment of the closeness of the observed frequency distribution and the assumed theoretical model. Practically, both the plotting and the comparison of cumulative curves can be conveniently simplified by using special plotting paper called *probability paper.* This special paper provides properly scaled axes such that the CDF of the probability law plots as a straight line. With such paper, comparison between the assumed model and the data is reduced to a comparison between the cumulative frequency plot of the data and a straight line. Further, the straight line plot would make the extrapolation task easier and safer.

However, no single probability paper for the P3 distribution is available because the graduation of the probability scale depends on the skewness coefficient (or the shape parameter) of the distribution. For example, Cunnane (1978) recommended the use of normal- or exponential-probability paper for the P3 distribution when the skewness coefficient γ is either close to 0 or 2. In addition, if γ is very different from these two particular values, the plotting procedure proposed by Wilk et al. (1962) was suggested because there were no appropriate probability papers available. Similarly, in the case of GEV distribution, only probability paper for the EV1 ($\gamma = 1.139$) distribution is available in the commercial market but not for EV2 and EV3 distributions (GEV distribution with $\gamma \neq 1.139$). Therefore, for practical applications, it would be preferable to develop probability papers for the P3 and GEV distributions for a wide range of skewness values.

Recent advent of microcomputer drafting capability provides new possibilities in the development of probability papers. In particular, for the P3 and GEV distributions, with the ease of computing exact plotting positions by PWM method as shown above, and the availability of simple drafting software packages such as Prodesign II (Webster, 1985), it could be possible, in theory, to plot probability papers for any skewness value. However, for purposes of comparison, probability papers for the P3 and GEV distributions have been developed in the present study for a number of selected values of the skewness coefficient γ ($\gamma = 0, 0.1, 0.2, \ldots, 3.0$). These skewness values are chosen because most of the flood data normally have the skewness coefficient between 0 and 3.0 (NERC, 1975a; Matalas et al., 1975; Beable and McKerchar, 1982). It has been observed that a difference of less than 10% between two skewness values produced no significant difference in the resulting probability papers. Hence, the probability papers developed in this study might be used for any skewness value in the range from 0 to 3 (see Appendix A for P3 papers, and Appendix B for GEV papers). An illustrative application of these papers will be presented in the following chapter.

CHAPTER 5
VERIFICATION AND COMPARISON OF PLOTTING POSITION FORMULAS

5.1 Systematic Flood Records

It is possible to verify and compare various plotting position formulas for systematic flood records by graphical and numerical methods. According to the graphical procedure (Cunnane, 1978), a plotting position formula can be judged by plotting the expected value, $E[y_m]$, $m = 1, 2, \ldots, N$, as ordinates against the plotting positions P_m under consideration. Because $E[y_m]$ depends on the form of the parent distribution, this type of judgment on a particular plotting position must be performed separately for each underlying distribution. If the formula is correct a linear plot (or straight line) will be obtained on an appropriate probability paper. The graphical technique provides therefore a simple tool to verify and compare immediately the adequacy of the plotting formulas considered. Moreover, to obtain a more objective judgment on the performance of various formulas a numerical comparison should be carried out using as comparing criteria the root mean square error (RMSE) and the absolute maximum difference between exact and approximate plotting positions.

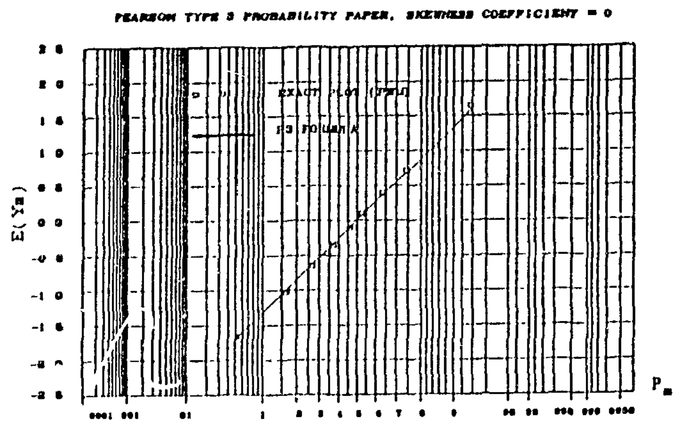
Verification and comparison of the new plotting position formulas for P3 and GEV distributions are shown in the following sections.

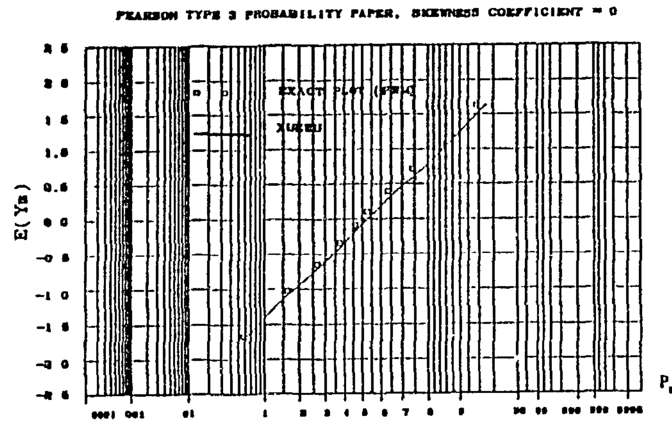
5.1.1 P3 Distribution

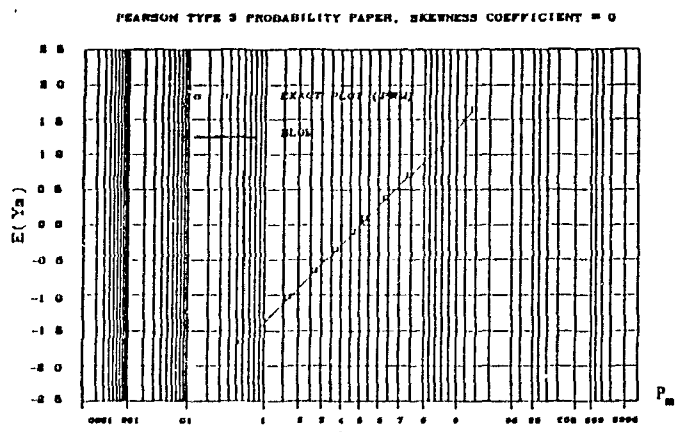
For purposes of illustration, comparisons will be performed for two sample sizes $N = 10$ and $N = 30$ which represent the data samples commonly available in practice. Further, only those formulas that were recommended for use with the P3 distribution are chosen. Since the symmetrical normal ($\gamma = 0$) and skewed exponential ($\gamma = 2$) distributions are special cases of the P3 distribution and, in particular, there exist special probability papers and plotting formulas specifically derived for these distributions, the normal and exponential distributions are selected for the graphical and numerical comparisons.

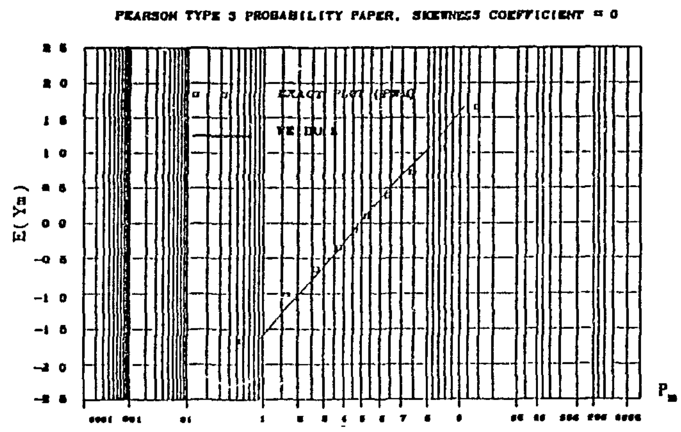
More specifically, for the skewness coefficient $\gamma = 0$ the formulas proposed by Blom (1958), Adamowski (1981), and Xuewu et al.(1984) are considered. Note that Blom formula was selected because it was derived specifically for the normal distribution. For $\gamma = 2$, it is preferable to replace Blom formula by the formula suggested especially for the P3 distribution by Cunnane (1978). Moreover, due to its popularity in engineering practice the Weibull formula will be considered in these comparisons, although this formula was not specifically derived for the P3 distribution as indicated above.

Results of the verification of the P3 formula are shown in Figs. 5.1 and 5.2 for the normal distribution ($\gamma = 0, N = 10$ and $N = 30$); and in Figs. 5.3 and 5.4 for the exponential distribution ($\gamma = 2, N = 10$ and $N = 30$). It can be seen that the plots of $E[y_m]$ against the plotting positions P_m from the proposed formula are straight lines in both exponential and normal distributions and for both small and large samples. This indicates that the new plotting position formula performs very well for symmetric and skewed distributions.









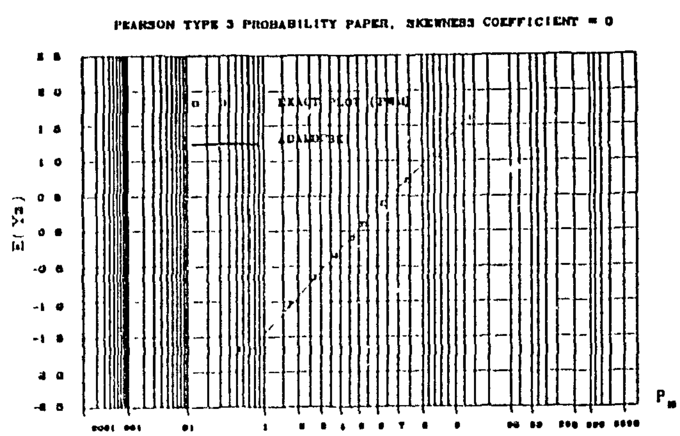
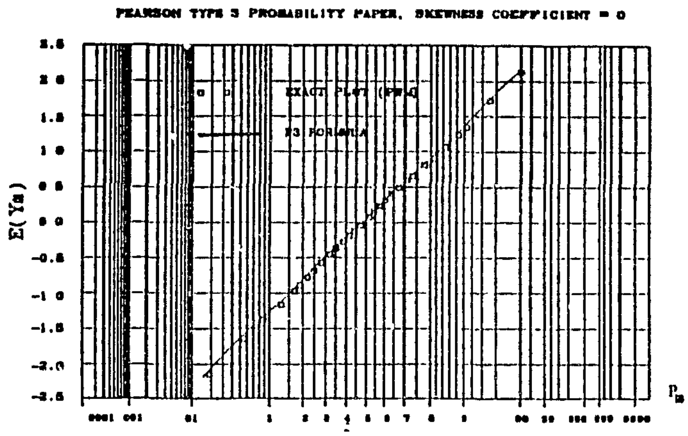
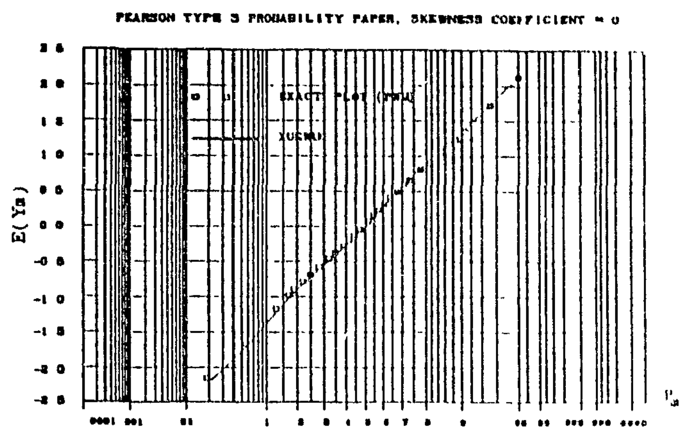
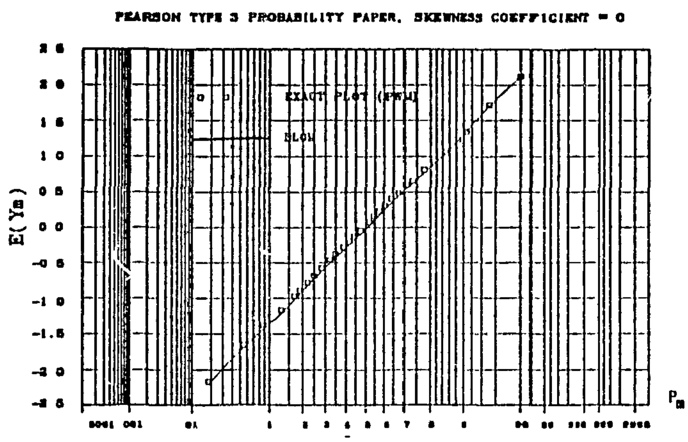
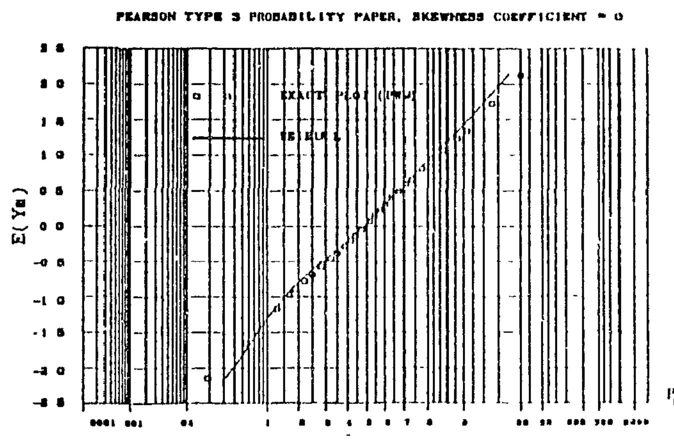


Figure 5.1 Comparison of probability plots from different plotting formulas for the Normal distribution (N=10)









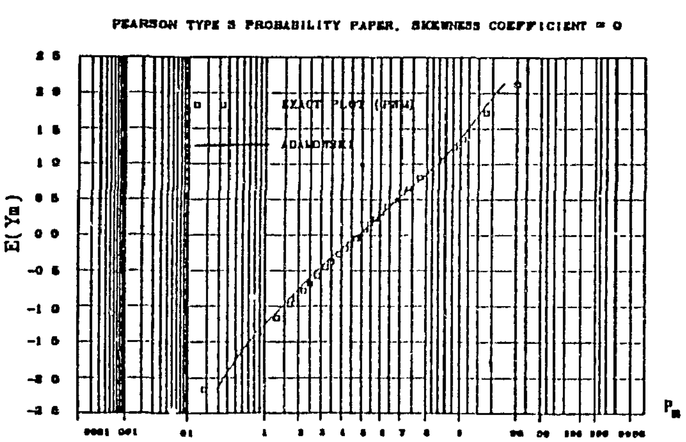
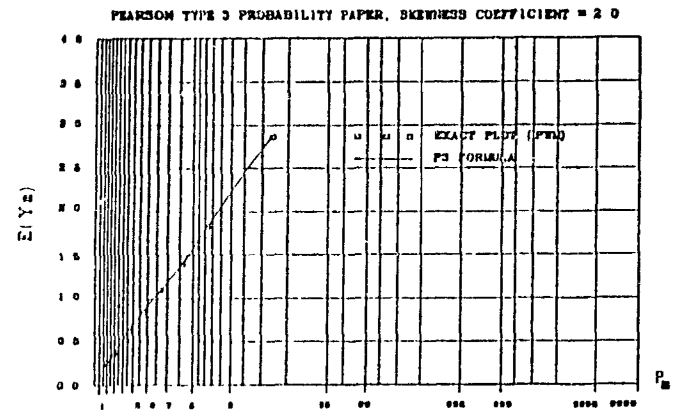
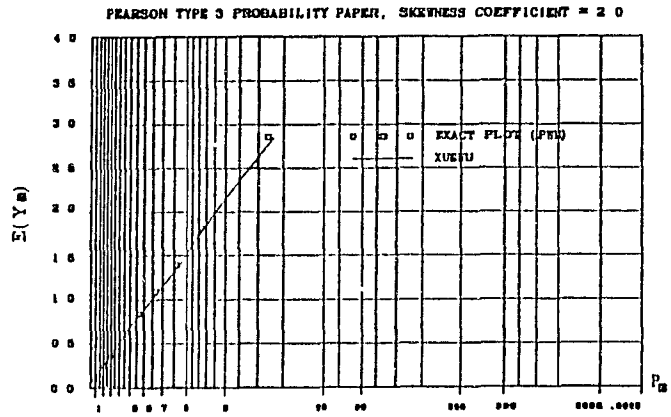
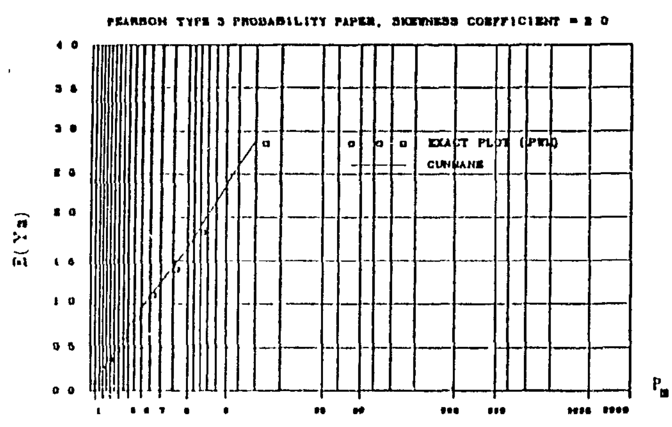
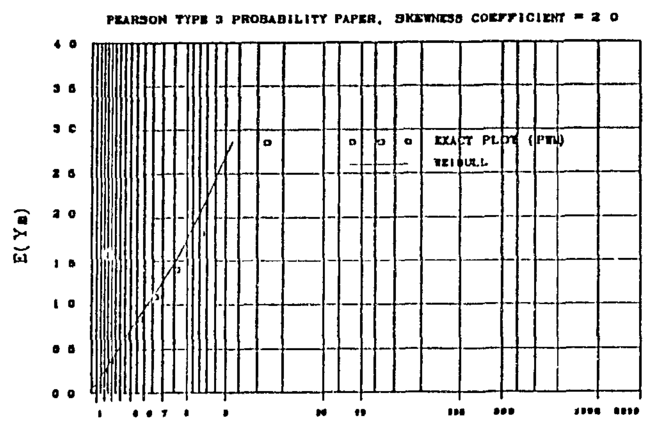


Figure 5.2 Comparison of probability plots from different plotting formulas for the Normal distribution (N=30)









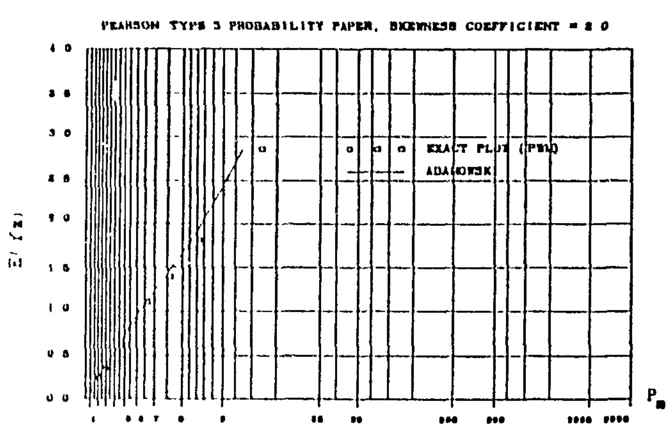
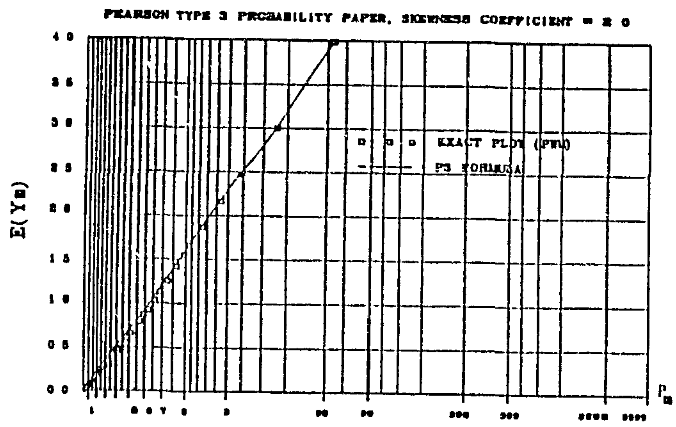
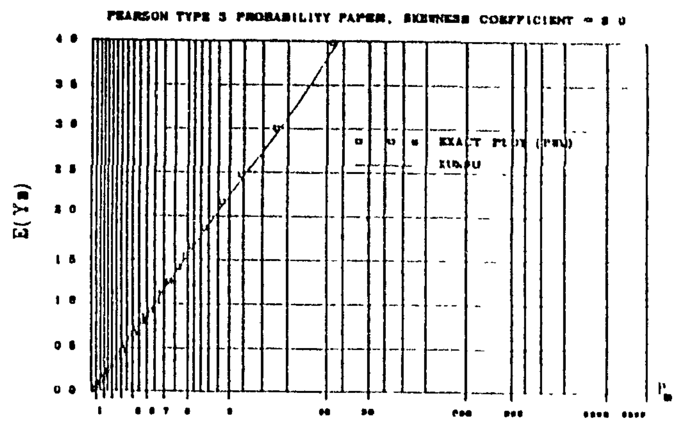
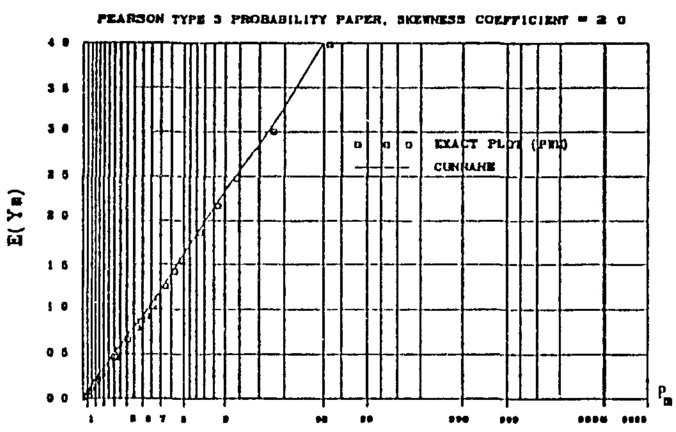
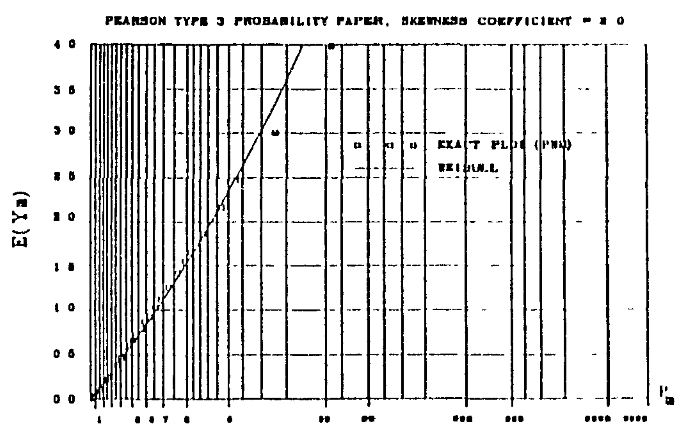


Figure 5.3 Comparison of probability plots from different plotting formulas for the Exponential distribution (N=10).









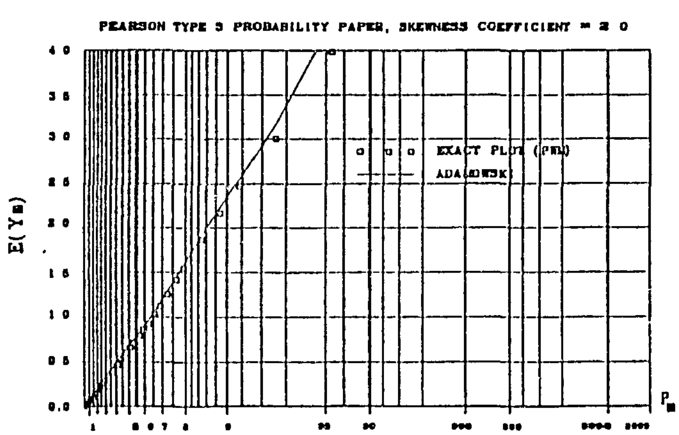


Figure 5.4 Comparison of probability plots from different plotting formulas for the Exponential distribution (N=30).

Figures 5.1–5.2 and Tables 5.1–5.2 show respectively graphical and numerical comparisons between different plotting formulas in the normal case ($\gamma = 0$). As expected, it can be seen that Blom formula performs very well, especially for the small sample considered, $N = 10$ (Fig. 5 1, Table 5.1), because it was specifically derived for the normal distribution. Moreover, as compared with Blom formula, P3 and Xuewu formulas perform equally well. Results obtained by these three formulas are better than those given by Adamowski formula which is biased at both extreme ends. The Weibull formula was found to be the most biased as compared with the other formulas.

In the exponential case ($\gamma = 2$), the proposed P3 formula gives the best performance as clearly indicated in Figs. 5.3–5.4 and Tables 5.3–5.4. Further, it is noted that, although Xuewu and Cunnane formulas were specifically recommended for the P3 distribution, results given by these two formulas were found to be as biased as those provided by the Weibull formula in terms of the RMSE and the maximum absolute difference (Tables 5.3 and 5.4). The bias is most pronounced at the upper end of the plot for Weibull formula, and at the lower end for Xuewu and Cunnane formulas. The formula proposed by Adamowski performs slightly better than Xuewu, Weibull, and Cunnane formulas in this case.

In summary, for the symmetrical normal distribution the proposed P3 formula gave a comparable performance as the well-known Blom formula. However, for a skewed distribution the P3 formula performs much better than other existing formulas. Therefore, it can be concluded that the proposed P3 formula developed in this study is the most appropriate for the P3 distribution, and for both small and large data samples.

Table 5.1: Comparison of plotting position formulas for the Normal distribution ($\gamma = 0.0, N = 10$).

		Plotting Position Formulas					
m	$E(y_m)$	PWM	P3	Xuewu	Weibull	Blom	Adamowski
1	-1.53875	0.056	0.047	0.055	0.090	0.061	0.071
2	-1.00136	0.155	0.146	0.152	0.182	0.158	0.166
3	-0.65606	0.254	0.246	0.250	0.272	0.256	0.262
4	-0.37576	0.351	0.345	0.350	0.363	0.353	0.357
5	-0.12267	0.450	0.445	0.450	0.454	0.451	0.452
6	0.12267	0.546	0.544	0.550	0.545	0.548	0.547
7	0.37576	0.646	0.644	0.650	0.636	0.646	0.643
8	0.65606	0.745	0.743	0.750	0.727	0.743	0.738
9	1.00136	0.842	0.843	0.850	0.818	0.841	0.833
10	1.53875	0.938	0.942	0.950	0.909	0.939	0.928
RMSE			0.006	0.005	0.021	0.002	0.008
Max$\lvert P_{m(exact)} - P_{m(formula)} \rvert$			0.009	0.012	0.034	0.005	0.015

Table 5.2: Comparison of plotting position formulas for the Normal distribution ($\gamma = 0.0, N = 30$).

		Plotting Position Formulas					
m	$E(y_m)$	PWM	P3	Xuewu	Weibull	Blom	Adamowski
1	-2.04276	0.019	0.016	0.021	0.032	0.021	0.025
2	-1.61560	0.052	0.049	0.053	0.065	0.054	0.057
3	-1.36481	0.084	0.082	0.086	0.097	0.087	0.090
4	-1.17885	0.117	0.115	0.119	0.129	0.120	0.123
5	-1.02609	0.151	0.149	0.152	0.161	0.153	0.156
6	-0.89439	0.183	0.182	0.185	0.194	0.185	0.189
7	-0.77666	0.217	0.215	0.219	0.226	0.219	0.221
8	-0.66885	0.250	0.249	0.252	0.258	0.252	0.254
9	-0.56834	0.283	0.282	0.285	0.290	0.285	0.287
10	-0.47329	0.317	0.315	0.318	0.323	0.318	0.320
11	-0.38235	0.350	0.348	0.351	0.355	0.351	0.352
12	-0.29449	0.382	0.382	0.384	0.387	0.384	0.385
13	-0.20885	0.416	0.415	0.417	0.419	0.417	0.418
14	-0.12473	0.449	0.448	0.450	0.452	0.450	0.451
15	-0.04148	0.482	0.482	0.483	0.484	0.483	0.484
16	0.04148	0.515	0.515	0.517	0.516	0.517	0.516
17	0.12473	0.548	0.548	0.550	0.548	0.550	0.549
18	0.20885	0.582	0.581	0.583	0.581	0.583	0.582
19	0.29449	0.615	0.615	0.616	0.613	0.616	0.615
20	0.38235	0.648	0.648	0.649	0.645	0.649	0.648
21	0.47329	0.681	0.681	0.682	0.677	0.682	0.680
22	0.56834	0.715	0.715	0.715	0.710	0.715	0.713
23	0.66885	0.748	0.748	0.748	0.742	0.748	0.746
24	0.77666	0.781	0.781	0.781	0.774	0.781	0.779
25	0.89439	0.814	0.814	0.815	0.806	0.814	0.811
26	1.02609	0.848	0.848	0.848	0.839	0.847	0.844
27	1.17855	0.881	0.881	0 881	0.871	0.880	0.877
28	1.36481	0.914	0.914	0.914	0.903	0.913	0.910
29	1.61560	0.947	0.947	0.947	0.935	0.946	0.943
30	2 04276	0.980	0.981	0.980	0.968	0.979	0.975
	RMSE		0.001	0.001	0.008	0.002	0.004
	$\text{Max}\lvert P_{m(exact)} - P_{m(formula)}\rvert$		0 003	0.002	0.013	0.003	0.006

Table 5.3: Comparison of plotting position formulas for the Exponential distribution ($\gamma = 2.0, N = 10$).

		Plotting Position Formulas					
m	$E(y_m)$	PWM	P3	Xuewu	Weibull	Cunnane	Adamowski
1	0.10004	0.096	0.100	0.055	0.090	0.058	0.071
2	0.21111	0.190	0.194	0.152	0.182	0.156	0.166
3	0.33611	0.285	0.288	0.250	0.272	0.255	0.262
4	0.47897	0.381	0.382	0.350	0.363	0.353	0.357
5	0.64564	0.476	0.476	0.450	0.454	0.451	0.452
6	0.84564	0.571	0.570	0.550	0.545	0.549	0.547
7	1.09564	0.666	0.664	0.650	0.636	0.647	0.643
8	1.42897	0.760	0.758	0.750	0.727	0.745	0.738
9	1.92897	0.855	0.852	0.850	0.818	0.843	0.833
10	2.92896	0.947	0.946	0.950	0.909	0.941	0.928
RMSE			0.003	0.026	0.026	0.025	0.023
Max$\lvert P_{m(exact)} - P_{m(formula)} \rvert$			0.004	0.041	0.038	0.038	0.025

Table 5.4: Comparison of plotting position formulas for the Exponential distribution ($\gamma = 2.0, N = 30$).

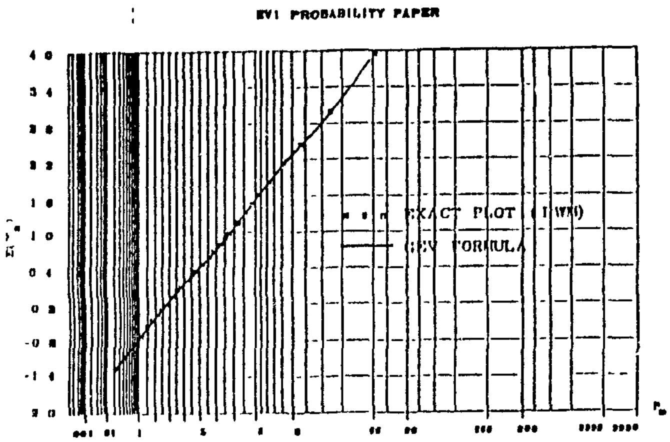
		Plotting Position Formulas					
m	$E(y_m)$	PWM	P3	Xuewu	Weibull	Cunnane	Adamowski
1	0.03346	0.033	0.033	0.019	0.032	0.020	0.025
2	0.06782	0.066	0.068	0.051	0.065	0.053	0.057
3	0.10353	0.098	0.100	0.083	0.097	0.086	0.090
4	0.14057	0.131	0.133	0.117	0.129	0.119	0.123
5	0.17903	0.164	0.165	0 150	0.161	0.152	0.156
6	0.21903	0.197	0.198	0.183	0.194	0.185	0.189
7	0.26070	0.229	0.231	0.217	0.226	0.219	0.221
8	0.30417	0.262	0.263	0.250	0.258	0.252	0.254
9	0.34963	0.295	0.296	0.283	0.290	0.285	0.287
10	0.39725	0.328	0.329	0.317	0.323	0.318	0.320
11	0.44725	0.361	0.361	0.350	0.355	0.351	0.352
12	0.49988	0.393	0.394	0.383	0.387	0.384	0.385
13	0.55543	0.426	0.426	0.417	0.419	0.417	0.418
14	0.61426	0.459	0.459	0.450	0.452	0.450	0.451
15	0.67676	0.492	0.492	0.483	0.484	0.483	0.484
16	0.74343	0.525	0.524	0.517	0.516	0.517	0.516
17	0.81485	0.557	0.556	0.541	0.548	0.550	0.549
18	0.89178	0.590	0.590	0.583	0.581	0.583	0.582
19	0.97511	0.623	0.622	0.617	0.613	0.616	0.615
20	1.06602	0.656	0.655	0.650	0.645	0.649	0.648
21	1.16602	0.688	0.687	0.683	0.677	0.682	0.680
22	1.27713	0.721	0.720	0.717	0.710	0.715	0.713
23	1.40213	0.754	0.753	0.750	0.742	0.748	0.746
24	1.54499	0.787	0.785	0.783	0.774	0.781	0.779
25	1.71165	0.819	0.818	0.817	0.806	0.815	0.811
26	1.91165	0.852	0.851	0.850	0.839	0.848	0.844
27	2.16165	0.885	0.883	0.883	0.871	0.881	0.877
28	2.49498	0.918	0.916	0.917	0.903	0.914	0.910
29	2.99497	0.950	0.949	0.950	0.935	0.947	0.943
30	3.99497	0.982	0.981	0.983	0.968	0.980	0.975
RMSE			0.003	0.026	0.026	0.025	0.023
$\text{Max}\lvert P_{m(exact)} - P_{m(formula)}\rvert$			0.002	0.015	0.015	0.013	0.009

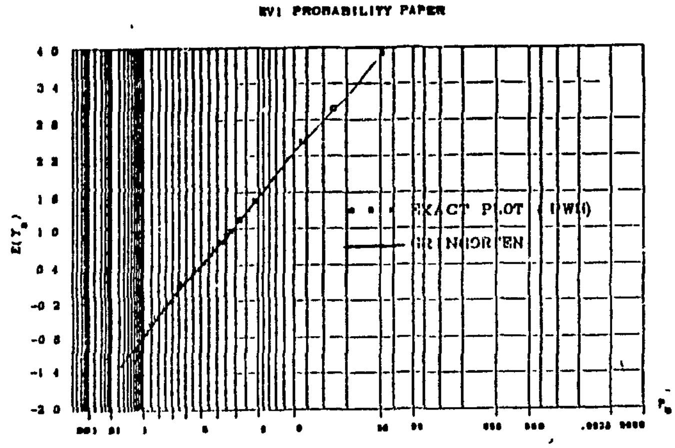
5.1.2 GEV Distribution

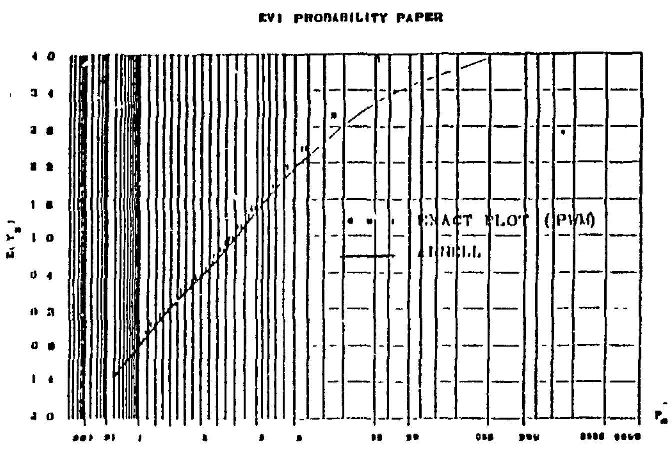
For the GEV distribution, comparisons between plotting formulas will be performed for the sample size $N = 30$. Further, only those formulas that were recommended for use with the GEV distribution are chosen. Since the EV1 ($\gamma = 1.139$), EV2 ($\gamma = 2$), and EV3 ($\gamma = 1$) distributions are special cases of the GEV distribution and, in particular, there exist appropriate formulas derived for them, these distributions are therefore selected for graphical and numerical comparisons.

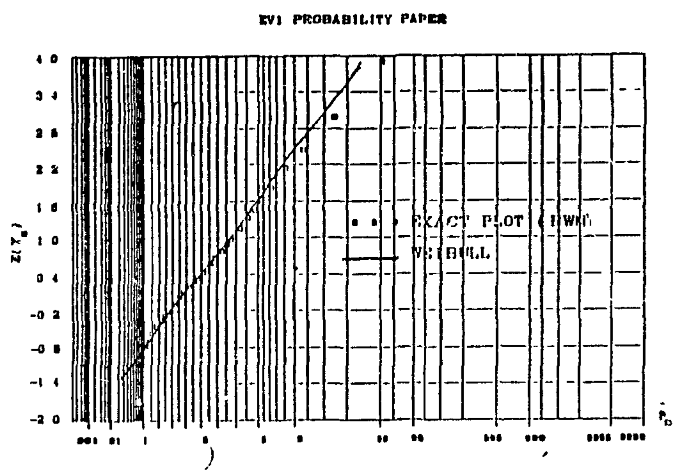
More specifically, for skewness coefficient $\gamma = 1.139$ the formulas proposed by Gringorten (1963), Arnell et al. (1986), and Sinclair and Ahmad (1988) are considered. Note that Gringorten formula was selected because it was derived specifically for the EV1 distribution. For $\gamma = 2$ and $\gamma = 1$ it is preferable to replace Gringorten formula by the compromise Cunnane formula (Cunnane, 1978) because there are no specific formulas for the EV2 and EV3 distributions. Moreover, due to its popularity in engineering practice the Weibull formula will be considered in these comparisons, although this formula was not specially derived for the GEV distribution as indicated above.

Results of the verification of the GEV formula are shown in Fig. 5.5 for EV1 distribution ($\gamma = 1.139, N = 30$); in Fig. 5.6 for EV2 distribution ($\gamma = 2, N = 30$); and in Fig. 5.7 for EV3 distribution ($\gamma = 1, N = 30$). It can be seen that the plots of $E[y_m]$ against the plotting positions P_m from the proposed formula are straight lines in all cases. This indicates that the new plotting position formula performs very well for the GEV distribution









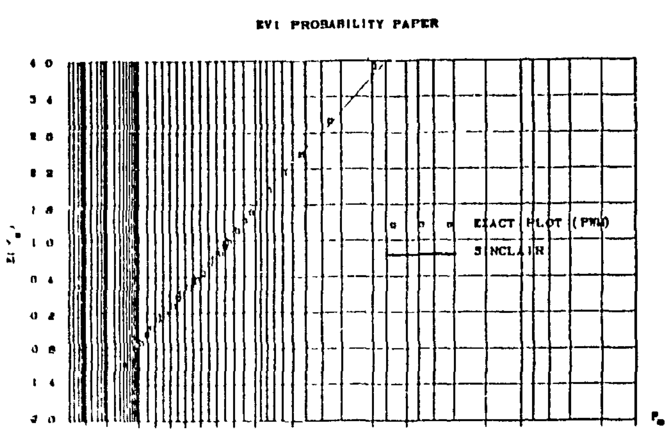
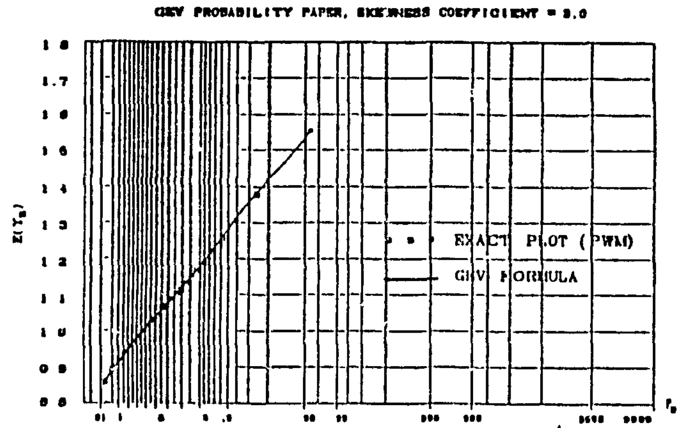
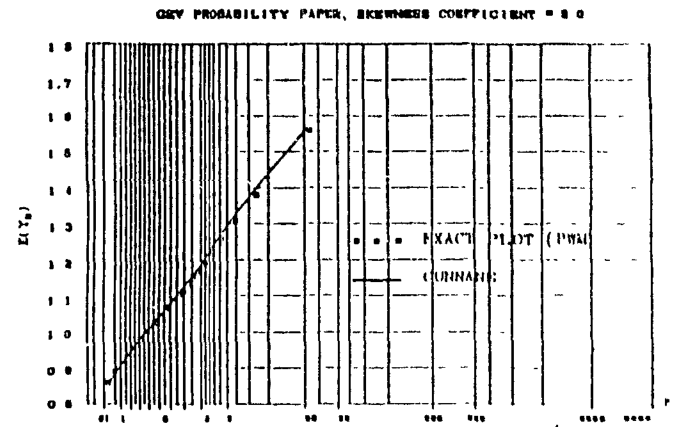
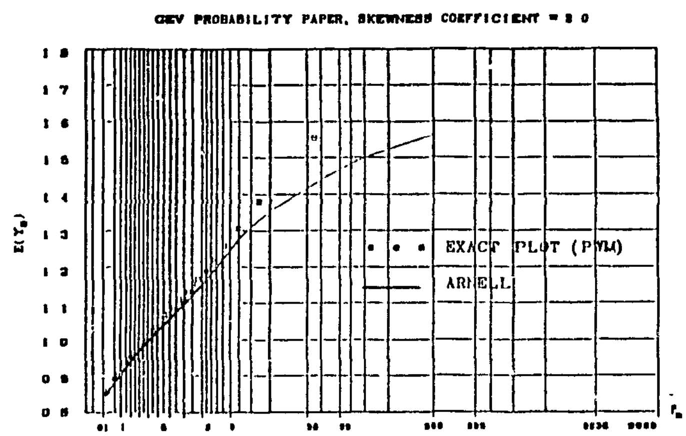
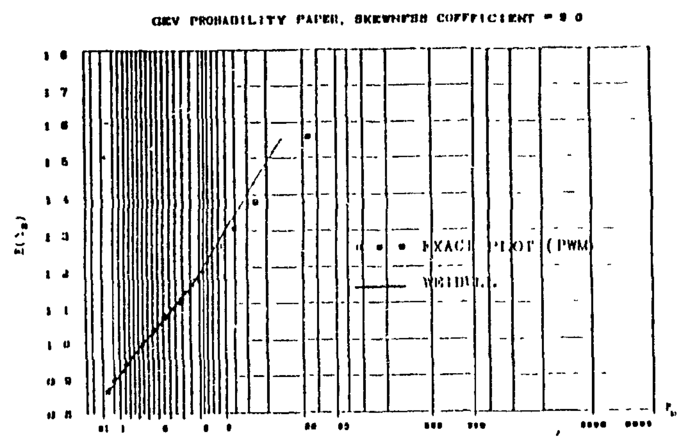


Figure 5.5 Comparison of probability plots from different plotting formulas for EV1 distribution (N=30).









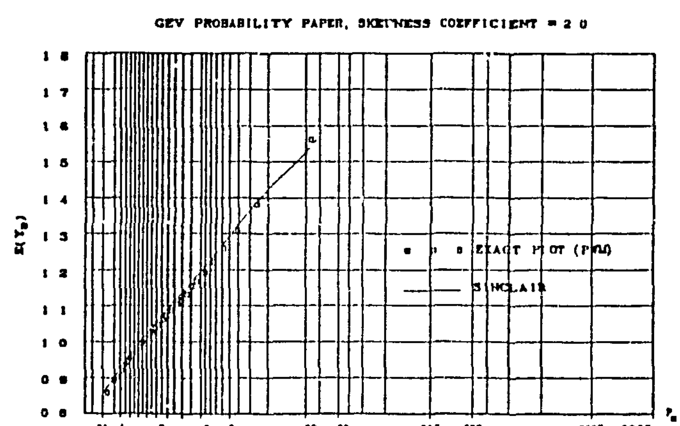
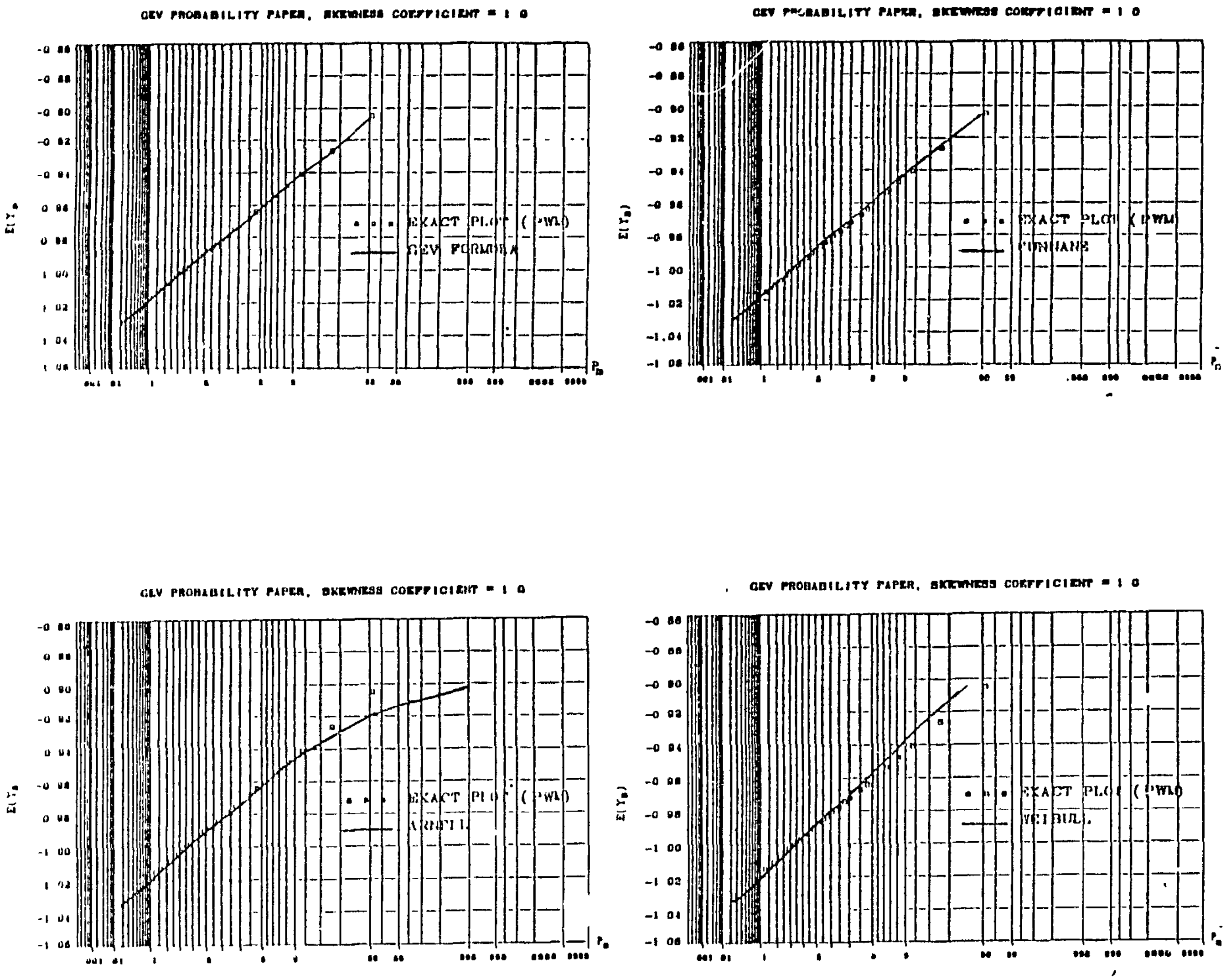


Figure 5.6 Comparison of probability plots from different plotting formulas for EV2 distribution (N=30).



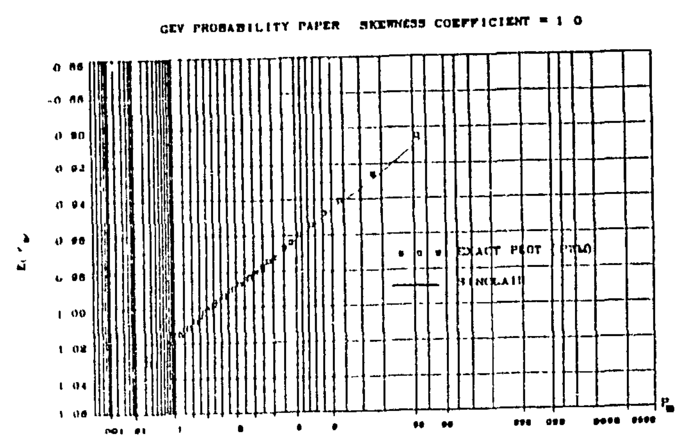


Figure 5.7 Comparison of probability plots from different plotting formulas for EV3 distribution (N=30).

Figure 5.5 and Table 5.5 illustrate respectively graphical and numerical comparisons between different plotting formulas in the EV1 case. It can be seen that the GEV formula yields the least bias in terms of the RMSE and the maximum absolute difference (Table 5.5). As expected, the Gringorten formula performs equally well, because it was specially derived for the EV1 distribution. Results obtained by these two formulas are better than those given by Sinclair, Arnell, and Weibull formulas which are biased at the upper end of the plot (Fig. 5.5).

For the EV2 and EV3 distributions, the proposed GEV formula gives the best performance as clearly indicated in Figs. 5.6-5.7 and Tables 5.6-5.7. Cunnane and Sinclair formulas seem to perform equally well in these cases. However, it is noted that, although Arnell formula was specifically recommended for the GEV distribution, results given by this formula was found to be as biased as those provided by the Weibull formula in terms of the RMSE and the maximum absolute difference (Tables 5.6-5.7).

In summary, for the EV1 distribution the Gringorten formula gave a comparable performance as the proposed GEV formula. However, for EV2 and EV3 distributions the GEV formula performs better than other existing formulas. Therefore, it can be concluded that the proposed GEV formula developed in this study is the most appropriate for the GEV distribution.

The above verification and comparison of P3 and GEV formulas were carried out for the case of systematic flood data. In the following section we will focus on the formulas for historical flood records.

Table 5.5: Comparison of plotting position formulas for the EV1 distribution ($\gamma = 1.139, N = 30$).

		Plotting Position Formulas					
m	$E(y_m)$	PWM	GEV	Arnell	Weibull	Gringorten	Sinclair
1	-1.33845	0.022	0.023	0.023	0.032	0.019	0.016
2	-1.06410	0.055	0.056	0.057	0.065	0.052	0.050
3	-0.88585	0.088	0.089	0.090	0.097	0.085	0.083
4	-0.74442	0.122	0.122	0.124	0.129	0.118	0.116
5	-0.62239	0.155	0.155	0.158	0.161	0.151	0.150
6	-0.51229	0.188	0.188	0.191	0.194	0.185	0.183
7	-0.40959	0.222	0.221	0.225	0.226	0.218	0.217
8	-0.31262	0.255	0.254	0.258	0.258	0.251	0.250
9	-0.21859	0.288	0.287	0.292	0.290	0.284	0.283
10	-0.12730	0.321	0.320	0.326	0.323	0.317	0.317
11	-0.03653	0.354	0.353	0.359	0.355	0.351	0.350
12	0.05359	0.388	0.387	0.393	0.387	0.384	0.383
13	0.14456	0.421	0.420	0.426	0.419	0.417	0.417
14	0.23599	0.454	0.453	0.460	0.452	0.450	0.450
15	0.32939	0.487	0.486	0.494	0.484	0.483	0.483
16	0.42523	0.520	0.519	0.527	0.516	0.517	0.517
17	0.52457	0.553	0.552	0.561	0.548	0.550	0.550
18	0.62789	0.586	0.585	0.594	0.581	0.583	0.583
19	0.73643	0.620	0.618	0.628	0.613	0.616	0.617
20	0.85148	0.653	0.651	0.662	0.645	0.649	0.650
21	0.97462	0.686	0.684	0.695	0.677	0.683	0.683
22	1.10797	0.719	0.717	0.729	0.710	0.716	0.717
23	1.25438	0.752	0.750	0.762	0.742	0.749	0.750
24	1.41787	0.785	0.783	0.796	0.774	0.782	0.783
25	1.60447	0.818	0.816	0.830	0.806	0.815	0.817
26	1.82374	0.851	0.849	0.863	0.839	0.849	0.850
27	2.09240	0.884	0.882	0.897	0.871	0.882	0.884
28	2.44382	0.917	0.915	0.930	0.903	0.915	0.917
29	2.96137	0.950	0.948	0.964	0.935	0.948	0.950
30	3.97841	0.981	0.981	0.998	0.968	0.981	0.984
RMSE			0.001	0.008	0.008	0.003	0.004
$\text{Max}\lvert P_{m(exact)} - P_{m(formula)}\rvert$			0.002	0.017	0.015	0.004	0.006

Table 5.6: Comparison of plotting position formulas for the EV2 distribution ($\gamma = 2.0, N = 30$).

		Plotting Position Formulas					
m	$E(y_m)$	PWM	GEV	Arnell	Weibull	Cunnane	Sinclair
1	0.86522	0.022	0.025	0.024	0.032	0.020	0.016
2	0.89124	0.056	0.058	0.058	0.065	0.053	0.050
3	0.90861	0.089	0.091	0.091	0.097	0.086	0.083
4	0.92264	0.123	0.124	0.125	0.129	0.119	0.116
5	0.93494	0.156	0.157	0.159	0.161	0.152	0.150
6	0.94619	0.189	0.190	0.192	0.194	0.185	0.183
7	0.95680	0.223	0.223	0.226	0.226	0.219	0.217
8	0.96676	0.255	0.256	0.259	0.258	0.252	0.250
9	0.97674	0.289	0.290	0.293	0.290	0.285	0.283
10	0.98643	0.322	0.323	0.327	0.323	0.318	0.317
11	0.99631	0.355	0.356	0.360	0.355	0.351	0.350
12	1.00631	0.389	0.389	0.394	0.387	0.384	0.383
13	1.01586	0.421	0.422	0.427	0.419	0.417	0.417
14	1.02652	0.456	0.455	0.461	0.452	0.450	0.450
15	1.03680	0.488	0.488	0.494	0.484	0.483	0.483
16	1.04773	0.521	0.521	0.528	0.516	0.517	0.517
17	1.05906	0.555	0.554	0.562	0.548	0.550	0.550
18	1.07106	0.588	0.587	0.595	0.581	0.583	0.583
19	1.08380	0.621	0.620	0.629	0.613	0.616	0.617
20	1.09748	0.654	0.654	0.662	0.645	0.649	0.650
21	1.11231	0.687	0.687	0.696	0.677	0.682	0.683
22	1.12862	0.720	0.720	0.730	0.710	0.715	0.717
23	1.14681	0.753	0.753	0.763	0.742	0.748	0.750
24	1.16749	0.786	0.786	0.797	0.774	0.781	0.783
25	1.19157	0.820	0.819	0.830	0.806	0.815	0.817
26	1.22057	0.853	0.852	0.864	0.839	0.848	0.850
27	1.25717	0.886	0.885	0.898	0.871	0.881	0.884
28	1.30694	0.919	0.918	0.931	0.903	0.914	0.917
29	1.38459	0.951	0.951	0.965	0.935	0.947	0.950
30	1.55610	0.983	0.984	0.998	0.968	0.980	0.984
RMSE			0.001	0.008	0.009	0.004	0.005
$\text{Max}\lvert P_{m(exact)} - P_{m(formula)}\rvert$			0.003	0.015	0.016	0.006	0.007

Table 5.7: Comparison of plotting position formulas for the EV3 distribution ($\gamma = 1.0, N = 30$).

		Plotting Position Formulas					
m	$E(y_m)$	PWM	GEV	Arnell	Weibull	Cunnane	Sinclair
1	-1.03433	0.022	0.023	0.023	0.032	0.020	0.016
2	-1.02720	0.055	0.056	0.057	0.065	0.053	0.050
3	-1.02259	0.088	0.089	0.090	0.097	0.086	0.831
4	-1.01896	0.122	0.122	0.124	0.129	0.119	0.116
5	-1.01584	0.155	0.155	0.158	0.161	0.152	0.150
6	-1.01301	0.188	0.188	0.191	0.194	0.185	0.183
7	-1.01043	0.221	0.221	0.225	0.226	0.219	0.217
8	-1.00793	0.255	0.254	0.258	0.258	0.252	0.250
9	-1.00536	0.290	0.287	0.292	0.290	0.285	0.283
10	-1.00293	0.325	0.320	0.326	0.323	0.318	0.317
11	-1.00112	0.352	0.353	0.359	0.355	0.351	0.350
12	-0.99863	0.388	0.386	0.393	0.387	0.384	0.383
13	-0.99617	0.424	0.419	0.426	0.419	0.417	0.417
14	-0.99405	0.454	0.452	0.460	0.452	0.450	0.450
15	-0.99171	0.487	0.485	0.494	0.484	0.483	0.483
16	-0.98938	0.520	0.518	0.527	0.516	0.517	0.517
17	-0.98689	0.553	0.551	0.561	0.548	0.550	0.550
18	-0.98433	0.586	0.584	0.594	0.581	0.583	0.583
19	-0.98164	0.619	0.617	0.628	0.613	0.616	0.617
20	-0.97880	0.652	0.650	0.662	0.645	0.649	0.650
21	-0.97577	0.685	0.683	0.695	0.677	0.682	0.683
22	-0.97250	0.718	0.716	0.729	0.710	0.715	0.717
23	-0.96893	0.751	0.749	0.762	0.742	0.748	0.750
24	-0.96495	0.785	0.782	0.796	0.774	0.781	0.783
25	-0.96043	0.818	0.815	0.830	0.806	0.815	0.817
26	-0.95515	0.851	0.848	0.863	0.839	0.848	0.850
27	-0.94872	0.884	0.881	0.897	0.871	0.881	0.884
28	-0.94039	0.916	0.914	0.930	0.903	0.914	0.917
29	-0.92828	0.949	0.947	0.964	0.935	0.947	0.950
30	-0.90507	0.981	0.980	0.998	0.968	0.980	0.984
RMSE			0.002	0.009	0.008	0.003	0.004
$\mathrm{Max}\lvert P_{m(exact)} - P_{m(formula)}\rvert$			0.005	0.017	0.014	0.007	0.006

5.2 Historical Flood Records

Cunnane (1978) found that a plotting position should be relatively unbiased in terms of discharges rather than in terms of probabilities. Bias in discharge is defined as the difference between the expected values of the mth largest flood, $E[Q_m]$ and the expected value of the flood evaluated at the exceedance probability plotting position $\hat{P}_m$ for a given distribution $F(\cdot)$, $E[F^{-1}(1 - \hat{P}_m)]$. Hence, the bias in discharge is dependent on the distribution of floods (Hirsch, 1987).

In the present study, the bias in discharge is estimated only for $k \geq m$. The expectation of the magnitude of the mth largest flood, Q_m, for given values of P_e, N and m, is thus given by :

$$E[Q_m] = \sum_{k=m}^{N} E[Q_m/P_e, k] Prob[k/k \geq m] \qquad (5\text{-}1)$$

where

$$Prob[k/k \geq m] = \binom{N}{k} P_e^k (1 - P_e)^{N-k} \Bigg/ \left[\sum_{j=m}^{N} \binom{N}{j} P_e^j (1 - P_e)^{N-j} \right] \qquad (5\text{-}2)$$

Note that $E[Q_m/P_e, k]$ depends on the distribution of Q_m, while on the other hand $Prob[k/k \geq m]$ is distribution free. $E[Q_m/P_e, k]$ was estimated by Monte Carlo simulation using a total of 20,000 repetitions over the set of k values which have non-negligible probabilities. The expected value of the discharge for the given $\hat{P}_m$ values is found by :

$$E[F^{-1}(1 - \hat{P}_m)] = \sum_{k=m}^{N} F^{-1}[1 - \hat{P}_m(k)] Prob[k/k \geq m] \qquad (5\text{-}3)$$

where $F^{-1}[1 - \hat{P}_m(k)]$ is the discharge with exceedance probability of $\hat{P}_m(k)$ which is estimated using a plotting position formula.

In the following, comparison of various plotting position formulas for P3 and GEV distributions is performed by comparing the bias in the estimation of the largest flood discharge from a sample of size $N = 150$, a historical flood sample commonly available in practice.

5.2.1 P3 Distribution

For P3 distribution, only those formulas that were recommended for use with that distribution are chosen. As indicated in section 5.1.1, the symmetrical normal ($\gamma = 0$) and skewed exponential ($\gamma = 2$) distributions are special cases of the P3 distribution and, in particular, there exist probability papers and plotting formulas specifically derived for these distributions, the normal and exponential distributions are also selected for the comparison.

More specifically, the E-type plotting positions considered for the case of skewness coefficient $\gamma = 0$ include (see section 4-2) the E-B based on Blom (1958) formula, the E-A given by Adamowski (1981) formula, the E-W suggested by Hirsch and Stedinger (1987), and the E-P3 proposed in this study. For $\gamma = 2$, it is more appropriate to replace the E-B formula by the E-C derived from Cunnane (1978) formula for the P3 distribution. The Weibull formula for systematic flood records will also be considered in this comparison because of its popularity in engineering practice.

Results of the comparison of plotting position formulas are shown in Figs. 5.8, 5.9, and 5.10, respectively, for normal distribution ($\gamma = 0$), for exponential distribution ($\gamma = 2$), and for the case of $\gamma = 1$. In all cases the coefficient of variation is $C_v = 0.5$. Note that Figs. 5.8-5.10 present $E[Q_1]$ as a function of $E[k]$ where $E[k] = NP_e$. Note also that the expected value of the largest flood discharge, $E(Q_1)$, estimated

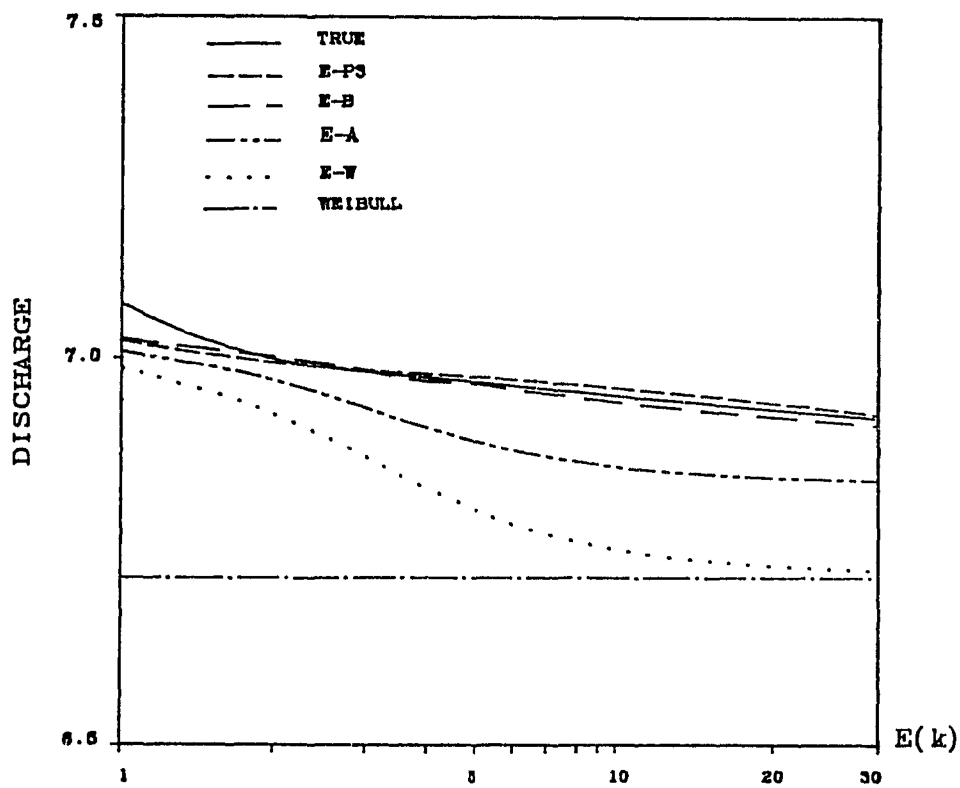


Figure 5.8 Bias in discharge for Normal distribution, $\gamma = 0$

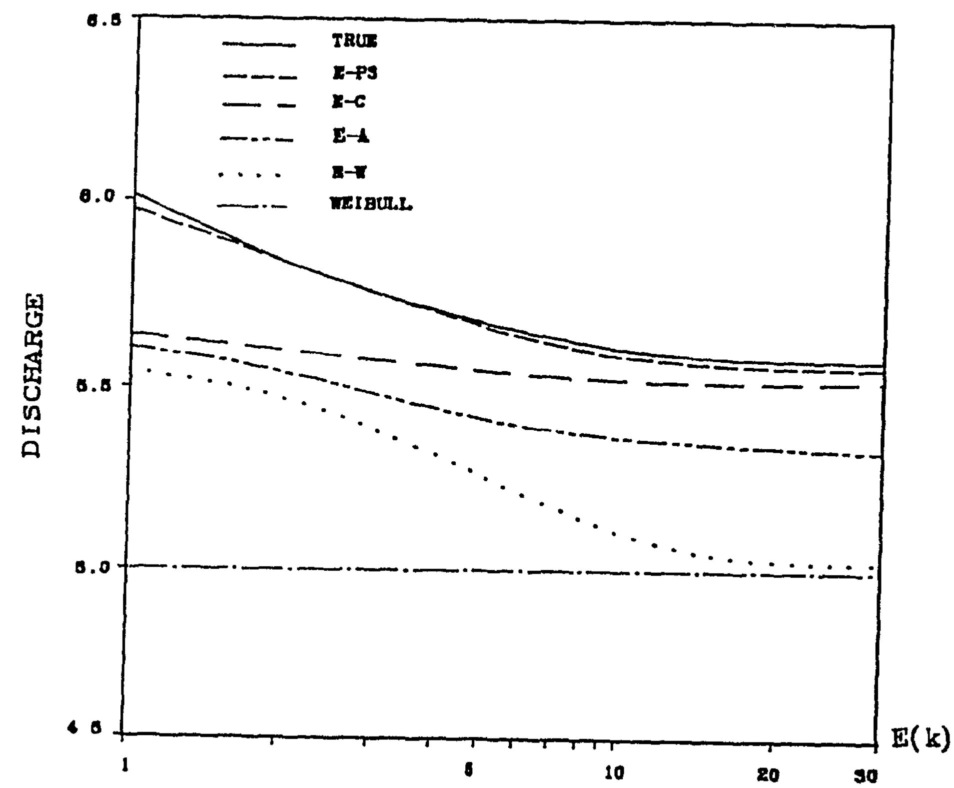


Figure 5.9 Bias in discharge for Exponential distribution, $\gamma = 2$.

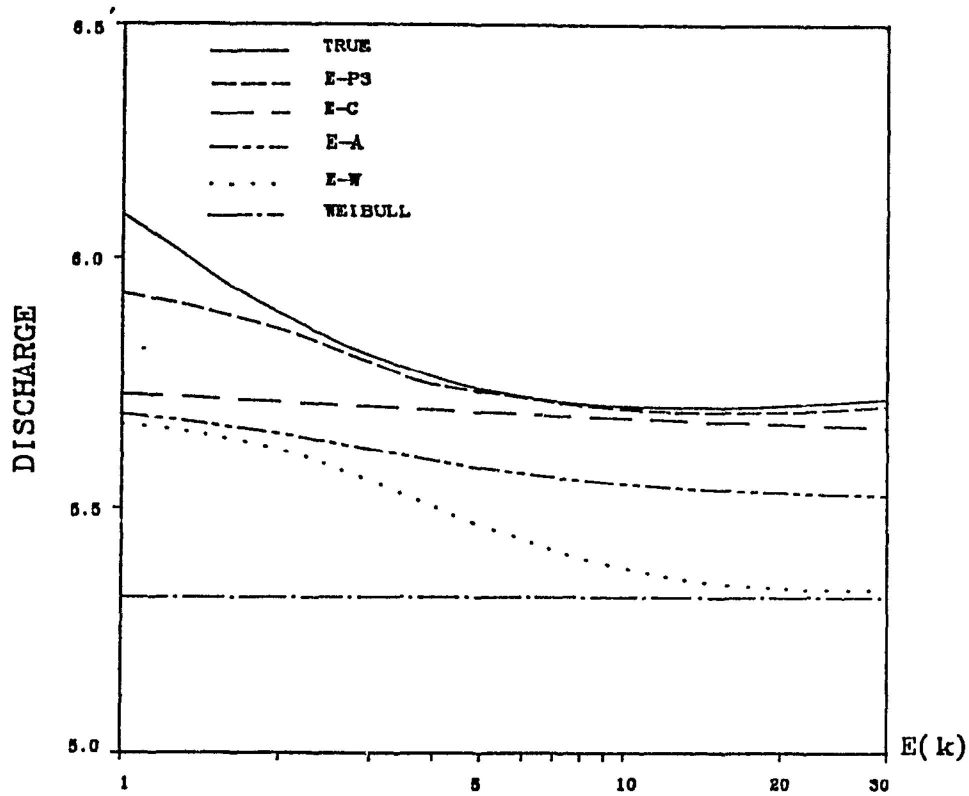


Figure 5.10 Bias in discharge for P3 distribution, $\gamma = 1$.

by the Monte Carlo simulation procedure is represented by the curve marked "true" in the above figures.

From Fig. 5.8 for the normal distribution ($\gamma = 0$), as expected, it can be seen that the E-B formula performs very well, especially for low values of $E[k]$, because Blom formula was specifically derived for the normal distribution. Further, as compared with the E-B, the proposed E-P3 formula performs equally well. Results obtained by these two formulas are better than those given by the E-A, E-W, and Weibull formulas. The well-known Weibull formula was found to be the most biased as compared with the other formulas.

In the exponential case ($\gamma = 2$), Fig. 5.9 indicates clearly the best performance of the E-P3 formula. Moreover, it is noted that, although E-C formula was especially recommended for the P3 distribution, this formula performs slightly better than the E-A, E-W, and Weibull formulas, especially for low value of $E[k]$. Finally, results for the P3 distribution with $\gamma = 1$ are similar to those for the exponential case as indicated in Fig. 5.10.

In summary, for a symmetrical normal distribution the E-P3 formula gave comparable performance as the E-B formula. However, for a skewed distribution the E-P3 formula performs much better than other existing formulas. Therefore, it can be concluded that the E-P3 formula developed in this study is the most appropriate for the P3 distribution for the analysis of historical flood data.

The same technique is used to verify and compare the developed E-GEV formula for the case of GEV distribution. Results are shown in the following section.

5.2.2 GEV Distribution

Similar to the P3 case, the proposed E-GEV formula is compared only with those formulas that were recommended for use with the GEV distribution. In particular, the EV1 ($\gamma = 1.139$), EV2 ($\gamma = 2$), and EV3 ($\gamma = 1$) distributions which are special cases of the GEV, are selected because there exist appropriate formulas especially derived for these distributions. More specifically, the E-AR and E-S formulas are considered because they are derived respectively from the formulas proposed by Arnell et al. (1986), and Sinclair and Ahmad (1988) for the GEV distribution. For the case of EV1 distribution, the E-G formula which is obtained from the plotting positions developed by Gringorten (1963) for this distribution is selected. For the EV2 and EV3 distributions, it is preferable to replace the E-G formula by the compromise E-C formula (Cunnane, 1978) as indicated in the previous section. For the same reason mentioned earlier, the Weibull formula will also be considered in the present comparison.

Results of comparison among selected plotting position formulas are shown in Fig. 5.11 for EV1 distribution ($\gamma = 1.139$), in Fig. 5.12 for EV2 distribution ($\gamma = 2$), and in Fig. 5.13 for EV3 distribution ($\gamma = 1$). In all cases the population mean is $\mu = 1.0$. As in previous section, Figs. 5.11-5.13 present $E[Q_1]$ as a function of $E[k]$ where $E[k] = NP_e$. The results indicate clearly that the E-GEV formula developed in the present study gives the best estimate of the true value of $E[Q_1]$ than all other formulas. Moreover, it is noted that, although E-AR and E-S formulas were also recommended for the GEV distribution, results given by these formulas are not as good as those provided by the E-GEV formula. As expected, it can be seen from Fig. 5.11 that the E-G formula performs well, especially for high values of $E[k]$, because it was specially derived for the EV1 distribution. The E-C formula

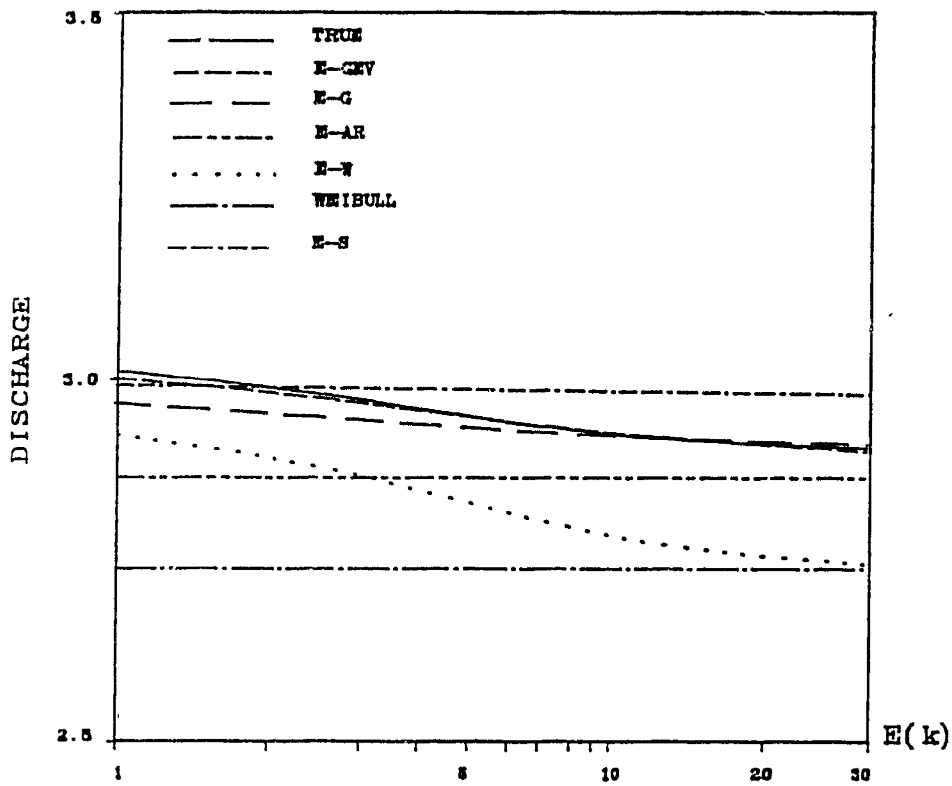


Figure 5.11 Bias in discharge for EV1 distribution, $\gamma = 1.139$.

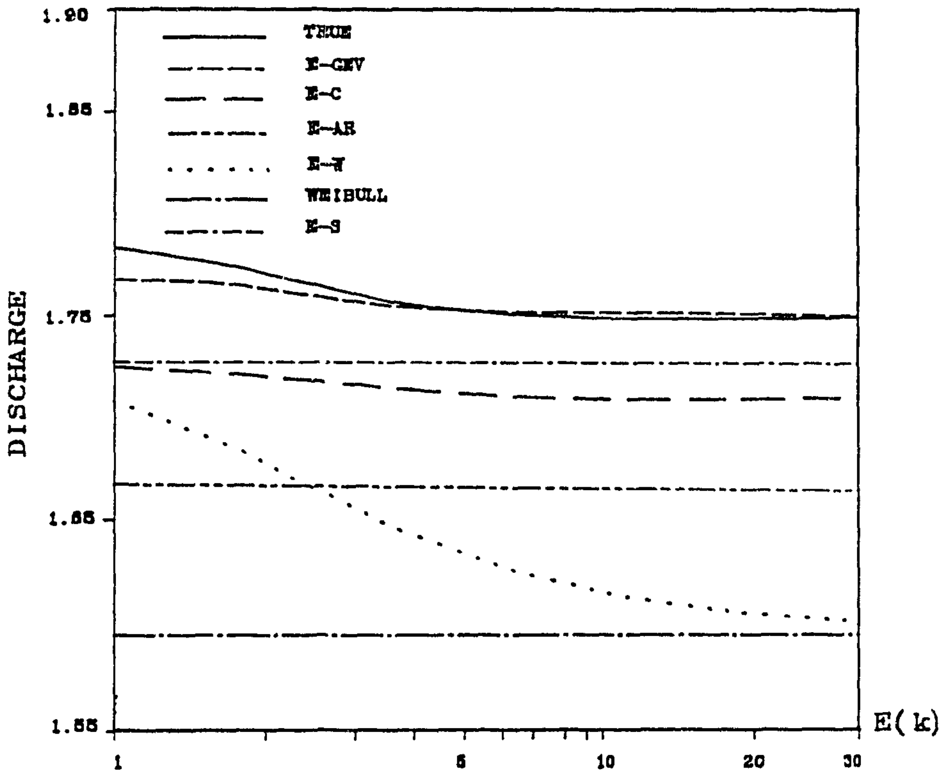


Figure 5.12 Bias in discharge for EV2 distribution, $\gamma = 2$.

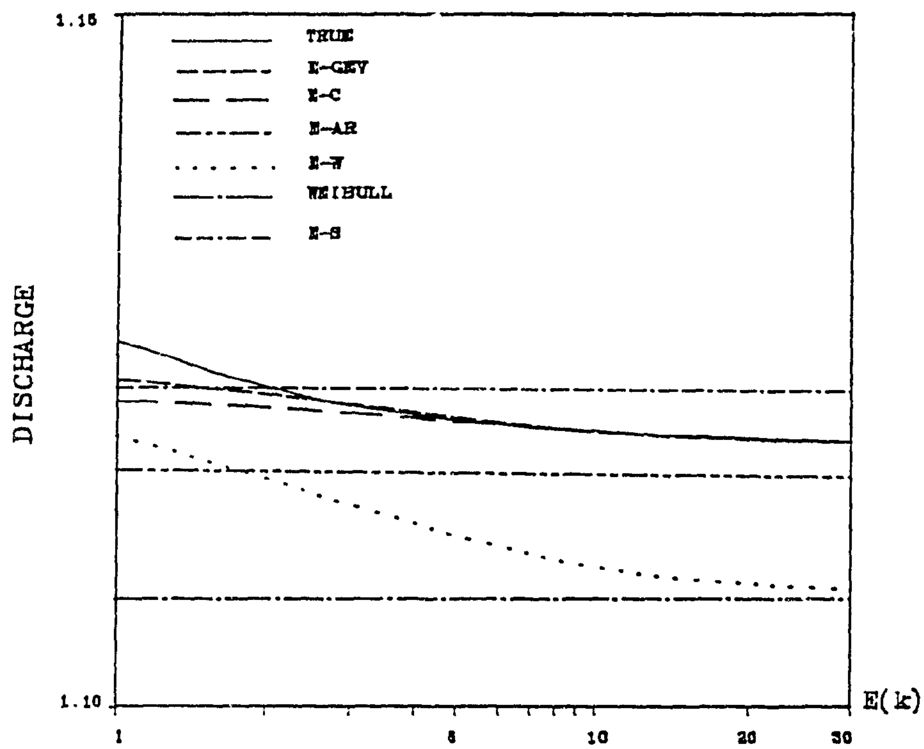


Figure 5.13 Bias in discharge for EV3 distribution, $\gamma = 1$.

has also a good performance for the EV3 distribution (Fig. 5.13). The E-W and Weibull formulas were found to be the most biased as compared with other formulas, especially for large values of $E[k]$.

In summary, results of the comparison have indicated that the E-GEV formula developed in this study performs much better than other existing formulas. Therefore, it can be concluded that the proposed E-GEV formula is the most appropriate for the GEV distribution for the analysis of flood records considering historical information.

CHAPTER 6

APPLICATION OF PLOTTING POSITION FORMULAS

6.1 Systematic Flood Records

As described in section 1.1, an advantage in the use of unbiased plotting positions involves the possibility of providing the best estimates of flood quantiles, which are unbiased. In the previous chapter, it has been demonstrated that the proposed P3 and GEV formulas are the most suitable for the P3 and GEV distributions as compared with several existing formulas. The present section therefore will present results involving the estimation of flood quantiles by the P3, GEV, and Weibull formulas as compared to those given by some theoretical distribution fitting methods. Note that the Weibull formula is also considered in this comparison because of its popularity in engineering practice as mentioned previously.

The comparison is performed using graphical and numerical methods, and using observed flood records from various geographical regions. The graphical method is carried out by plotting as ordinates the flood quantiles estimated by plotting position formulas against the flood quantiles which are computed by fitting a theoretical distribution to the observed flood data. The graphical technique, "quantile-quantile plot" (Q-Q plot), provides therefore a graphical assessment of the adequacy of the plotting position formulas considered as compared with the fitted theoretical distribution. If the plotting formulas provide the same flood estimates as the fitted distribution, the flood quantiles obtained should fall on the 45^o-line on the Q-Q plot. Moreover, to obtain a more objective judgment on the performance of plotting formulas a numerical comparison is also carried out using as criteria the root mean square error (RMSE) and the absolute maximum difference between the flood estimates from plotting formulas and fitted distribution.

The application of the proposed formulas to the observed flood data lies in the determination of the skewness coefficient γ of the parent distribution. Problems involved in the estimation of this parameter have been examined in several previous studies, especially by Bobee and Robitaille (1975, 1977). Concerning the P3 distribution an unbiased estimate of γ, denoted C_{su}, for $0.5 \leq \gamma \leq 2.0$ and $20 \leq N \leq 90$ can be computed from the following equation (Bobee and Robitaille, 1977):

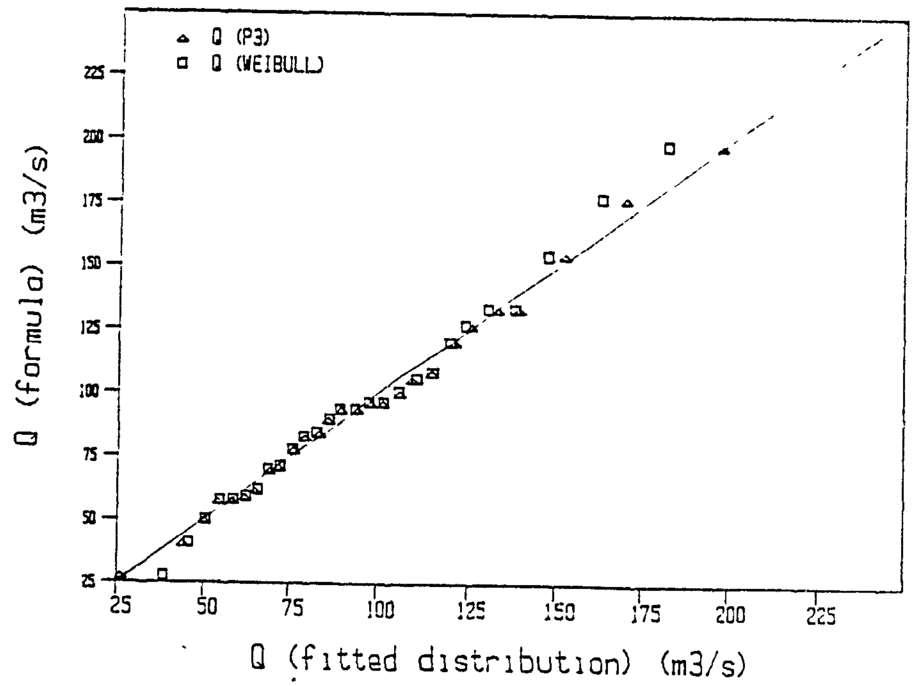
$$C_{su} = \frac{[N(N-1)]^{1/2}}{N-2}\left(1 + \frac{8.5}{N}\right)C_s \tag{6.1}$$

where C_s is a biased estimate of the population skewness coefficient and is usually computed by the ratio of the unbiased estimates of the third central moment and standard deviation. The above formula has been also recommended for use with the GEV distribution (World Meteorological Organization, 1969).

The application of the plotting positions considered will be illustrated by three examples using flood data of Madawaska River at Madawaska and Missinaibi River at Mattice in Ontario (Canada) (Environment Canada, 1983), and Dee River at Cairnton (U.K.) (NERC, 1975b). For Madawaska River, the continuous record of annual maximum floods for the 1916-1942 period ($N = 27$) is considered. In the case of Missinaibi River, the flood record from 1930 to 1979 ($N = 50$) is used. Finally, for Dee River, the flood record considered is from 1930 to 1953 ($N = 24$). These three rivers were selected to represent the sample sizes commonly available in practice.

In order to fit the P3 distribution to the observed flood data, the method of moments has been frequently recommended for use in the estimation of the parameters of this distribution (Bobee and Robitaille, 1977; UNESCO, 1987). For GEV distribution, the probability weighted moments is used as suggested by Hosking et al. (1985). Results of the graphical comparison (Q-Q plots) between P3, GEV, Weibull formulas and the fitted P3 and GEV distributions are shown in Figs. 6.1, 6.2, and 6.3, respectively, for Madawaska River ($C_{su} = 1.0, N = 27$); Missinaibi River ($C_{su} = 1.4, N = 50$); and Dee River ($C_{su} = 0.7, N = 24$). Further, Tables 6.1, 6.2, and 6.3 present results of the numerical comparison in terms of the RMSE and the maximum absolute difference of estimated flood quantiles, respectively for the three rivers considered.

In general, it can be seen from Figs. 6.1-6.3 and Tables 6.1-6.3 that the P3 and GEV formulas perform much better than the Weibull formula in terms of flood quantile estimates. The flood estimates given by P3 and GEV formulas are closer to those computed by the fitted P3 and GEV distributions as compared with the values provided by the Weibull formula. In particular, for Madawaska and Missinaibi rivers, results shown in Figs. 6.1-6.2 and Tables 6.1-6.2 demonstrate clearly the best performance of the P3 formula. However, for Dee River the GEV formula is the most suitable as indicated by Fig. 6.3 and Table 6.3. In summary, it was found that the P3 and GEV formulas proposed in this study gave a comparable performance as the conventional distribution fitting techniques. The well-known Weibull formula, however, is the most biased in the estimation of flood quantiles.



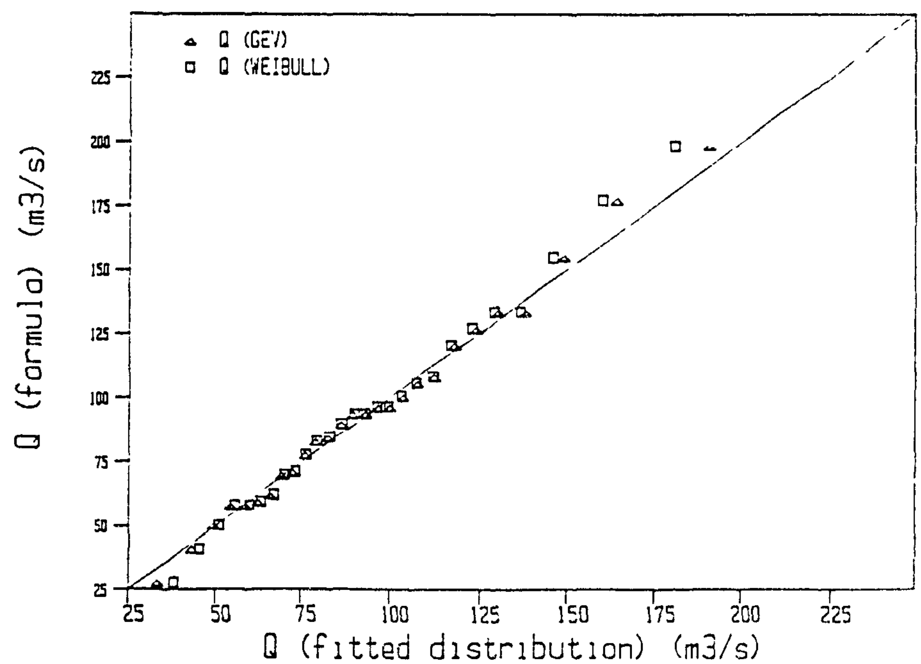
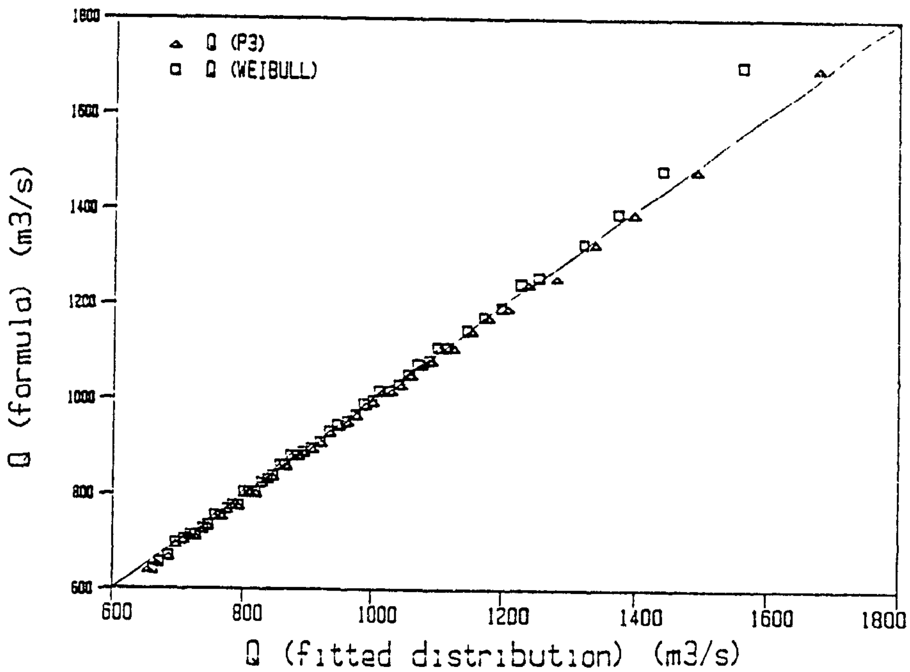


Figure 6.1 Quantile-quantile plot of P3, GEV and Weibull formulas (Madawaska River, $C_{su} = 1.0, N = 27$).

Table 6.1 Numerical comparison of P3, GEV, and Weibull formulas ($C_{su} = 1.0, N = 27$).

Madawaska River					P3 distribution		GEV distribution	
m	Q_{fitted} (m^3/s)	P3	GEV	Weibull	$\hat{Q}$ (P3)	$\hat{Q}$ (Weibull)	$\hat{Q}$ (GEV)	$\hat{Q}$ (Weibull)
1	27.2	0.028	0.026	0.036	26.0	38.0	33.0	38.0
2	41.3	0.065	0.062	0.071	44.0	45.0	43.0	45.0
3	50.1	0.101	0.099	0.107	50.0	51.0	49.0	51.0
4	58.0	0.138	0.136	0.143	54.0	55.0	54.0	56.0
5	58.3	0.174	0.172	0.179	58.0	59.0	58.0	60.0
6	59.7	0.211	0.209	0.214	62.0	62.0	62.0	63.0
7	62.0	0.248	0.245	0.250	65.0	66.0	66.0	67.0
8	70.2	0.284	0.282	0.286	69.0	69.0	69.0	70.0
9	71.1	0.321	0.319	0.321	72.0	72.0	73.0	73.0
10	77.6	0.357	0.355	0.357	76.0	76.0	76.0	76.0
11	83.3	0.394	0.392	0.393	79.0	79.0	79.0	80.0
12	84.4	0.430	0.429	0.429	83.0	82.0	83.0	83.0
13	90.3	0.467	0.465	0.464	85.0	85.0	86.0	86.0
14	93.2	0.503	0.502	0.500	89.0	89.0	89.0	89.0
15	94.0	0.540	0.538	0.536	94.0	93.0	93.0	93.0
16	96.0	0.577	0.575	0.571	98.0	98.0	96.0	96.0
17	96.3	0.613	0.612	0.607	102.0	102.0	100.0	100.0
18	101.0	0.650	0.648	0.643	107.0	106.0	104.0	104.0
19	106.0	0.686	0.685	0.679	111.0	111.0	108.0	108.0
20	108.0	0.723	0.722	0.714	115.0	115.0	113.0	112.0
21	120.0	0.759	0.758	0.750	121.0	120.0	118.0	117.0
22	127.0	0.796	0.795	0.786	126.0	124.0	124.0	123.0
23	133.0	0.833	0.832	0.821	134.0	131.0	131.0	129.0
24	134.0	0.869	0.868	0.857	140.0	139.0	139.0	137.0
25	155.0	0.906	0.905	0.893	152.0	148.0	149.0	147.0
26	177.0	0.942	0.941	0.929	170.0	163.0	164.0	160.0
27	198.0	0.979	0.978	0.964	198.0	182.0	191.0	181.0
				RMSE	3.6	5.8	4.3	6.0
				Max $\left\|Q(fitted) - \hat{Q}(formula)\right\|$	7.0	16.0	13.0	17.0

$\hat{Q}(\ )$ = Flood estimated from plotting position formula for the same non-exceedance probability



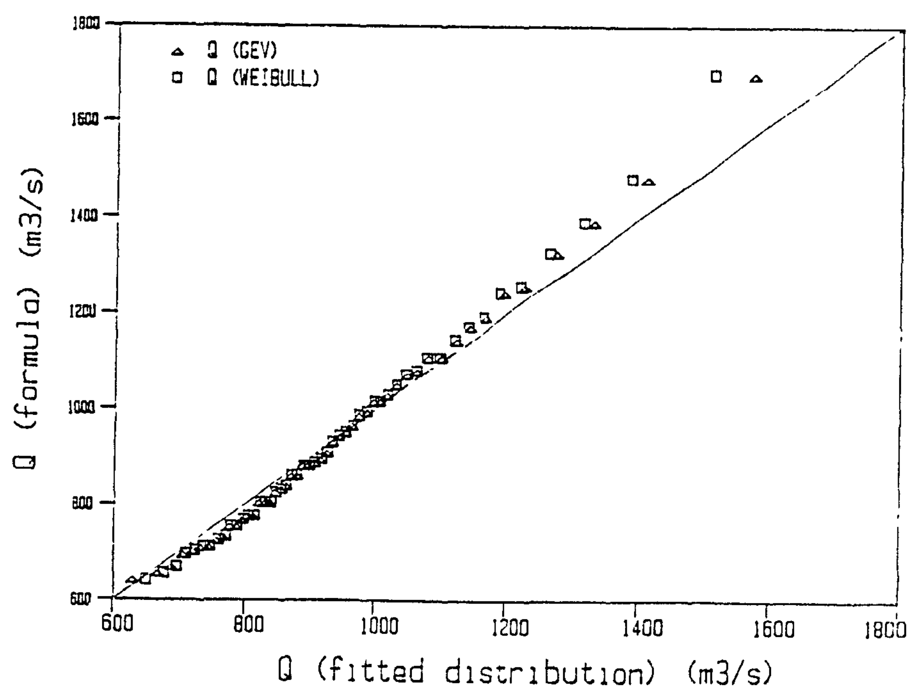
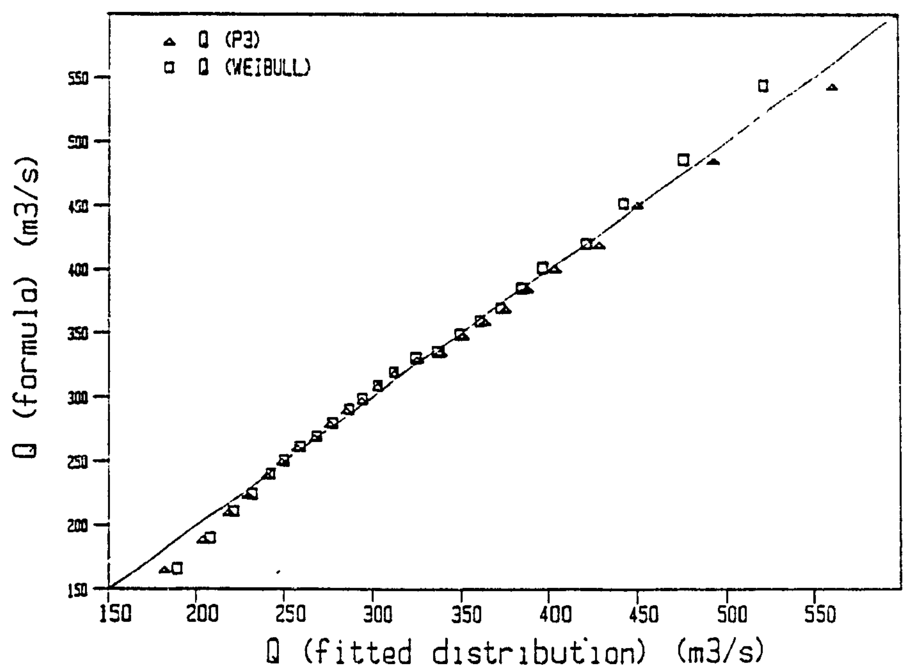


Figure 6.2 Quantile-quantile plot of P3, GEV and Weibull formulas (Missinaibi River, $C_{su} = 1.4, N = 50$).

Table 6.2. Numerical comparison of P3, GEV, and Weibull formulas ($C_{su} = 1.4, N = 50$)

	Missinaibi River				P3 distribution		GEV distribution	
m	Q_{fitted} (m^3/s)	P3	GEV	Weibull	Q (P3)	Q (Weibull)	Q (GEV)	Q (Weibull)
1	640.0	0.018	0.014	0.020	655.0	663.0	629.0	649.0
2	660.0	0.037	0.034	0.039	670.0	673.0	667.0	677.0
3	674.0	0.057	0.054	0.059	687.0	687.0	689.0	697.0
4	699.0	0.077	0.074	0.078	697.0	698.0	707.0	712.0
5	705.0	0.097	0.094	0.098	709.0	710.0	721.0	726.0
6	711.0	0.117	0.114	0.118	719.0	720.0	734.0	738.0
7	711.0	0.137	0.134	0.137	728.0	728.0	746.0	749.0
8	728.0	0.156	0.154	0.157	737.0	738.0	757.0	760.0
9	732.0	0.176	0.173	0.176	747.0	747.0	767.0	770.0
10	753.0	0.196	0.193	0.196	756.0	756.0	777.0	779.0
11	753.0	0.216	0.213	0.216	765.0	765.0	786.0	789.0
12	770.0	0.236	0.233	0.235	775.0	774.0	795.0	797.0
13	779.0	0.255	0.253	0.255	783.0	783.0	804.0	806.0
14	779.0	0.275	0.273	0.275	793.0	793.0	813.0	815.0
15	804.0	0.295	0.293	0.294	802.0	801.0	821.0	823.0
16	807.0	0.315	0.313	0.314	811.0	810.0	829.0	832.0
17	807.0	0.335	0.333	0.333	820.0	819.0	838.0	840.0
18	824.0	0.354	0.353	0.353	830.0	828.0	846.0	848.0
19	830.0	0.374	0.372	0.373	838.0	837.0	855.0	856.0
20	840.0	0.394	0.392	0.392	846.0	845.0	863.0	864.0
21	860.0	0.414	0.412	0.412	856.0	855.0	872.0	873.0
22	860.0	0.434	0.432	0.431	866.0	864.0	880.0	881.0
23	880.0	0.454	0.452	0.451	875.0	873.0	889.0	890.0
24	880.0	0.473	0.472	0.471	884.0	882.0	897.0	898.0
25	890.0	0.493	0.492	0.490	893.0	891.0	906.0	907.0
26	900.0	0.513	0.512	0.510	907.0	903.0	915.0	916.0
27	912.0	0.533	0.532	0.529	919.0	916.0	924.0	925.0
28	930.0	0.553	0.551	0.549	933.0	930.0	934.0	934.0
29	943.0	0.572	0.571	0.569	947.0	944.0	943.0	944.0
30	956.0	0.592	0.591	0.588	961.0	958.0	953.0	953.0
31	970.0	0.612	0.611	0.608	975.0	973.0	964.0	964.0
32	985.0	0.632	0.631	0.627	989.0	986.0	974.0	974.0
33	998.0	0.652	0.651	0.647	1003.0	999.0	985.0	985.0
34	1015.0	0.671	0.671	0.667	1016.0	1010.0	997.0	997.0
35	1020.0	0.691	0.691	0.686	1030.0	1026.0	1009.0	1008.0
36	1030.0	0.711	0.711	0.706	1044.0	1041.0	1021.0	1020.0
37	1055.0	0.731	0.730	0.725	1058.0	1054.0	1035.0	1034.0
38	1070.0	0.751	0.750	0.745	1073.0	1068.0	1050.0	1048.0
39	1080.0	0.771	0.770	0.765	1087.0	1083.0	1065.0	1064.0
40	1110.0	0.790	0.790	0.784	1100.0	1096.0	1082.0	1079.0
41	1110.0	0.810	0.810	0.804	1119.0	1107.0	1100.0	1097.0
42	1145.0	0.830	0.830	0.824	1150.0	1141.0	1120.0	1117.0
43	1170.0	0.850	0.850	0.843	1179.0	1169.0	1141.0	1138.0
44	1190.0	0.870	0.870	0.863	1207.0	1197.0	1166.0	1162.0
45	1210.0	0.889	0.890	0.882	1235.0	1225.0	1196.0	1189.0
46	1260.0	0.909	0.910	0.902	1277.0	1250.0	1230.0	1223.0
47	1330.0	0.929	0.929	0.922	1337.0	1316.0	1273.0	1263.0
48	1390.0	0.949	0.949	0.941	1397.0	1373.0	1329.0	1313.0
49	1480.0	0.969	0.969	0.961	1489.0	1437.0	1413.0	1388.0
50	1700.0	0.989	0.989	0.980	1678.0	1562.0	1575.0	1509.0
				RMSE	9.08	22.15	31.07	40.42
	Max$\lvert Q(fitted) - \dot{Q}(formula)\rvert$				22.0	138.0	125.0	191.0

$\dot{Q}(\,)$ = Flood estimated from plotting position formula for the same non-exceedance probability



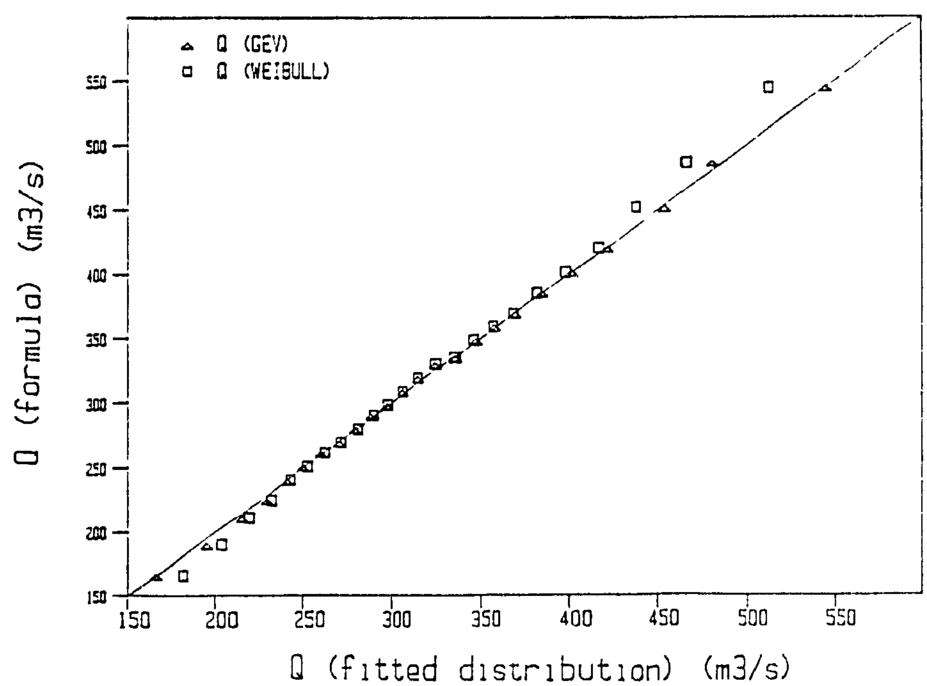


Figure 6.3 Quantile-quantile plot of P3, GEV and Weibull formulas (Dee River, $C_{su} = 0\ 7, N = 24$).

Table 6.3 Numerical comparison of P3, GEV, and Weibull formulas ($C_{su} = 0.7, N = 24$).

Dee River					P3 distribution		GEV distribution	
m	Q_{fitted} (m^3/s)	P3	GEV	Weibull	$\hat{Q}$ (P3)	$\hat{Q}$ (Weibull)	$\hat{Q}$ (GEV)	$\hat{Q}$ (Weibull)
1	165.0	0.028	0.028	0.040	181.0	189.0	165.0	181.0
2	191.0	0.069	0.069	0.080	203.0	207.0	195.0	203.0
3	210.0	0.110	0.110	0.120	217.0	219.0	214.0	218.0
4	225.0	0.152	0.151	0.160	228.0	230.0	228.0	231.0
5	240.0	0.193	0.193	0.200	239.0	241.0	240.0	242.0
6	251.0	0.234	0.234	0.240	248.0	250.0	250.0	252.0
7	260.0	0.275	0.275	0.280	257.0	259.0	260.0	262.0
8	268.0	0.317	0.316	0.320	267.0	268.0	270.0	271.0
9	280.0	0.358	0.357	0.360	276.0	277.0	279.0	280.0
10	290.0	0.399	0.398	0.400	285.0	286.0	288.0	289.0
11	299.0	0.440	0.439	0.440	294.0	294.0	297.0	298.0
12	308.0	0.481	0.480	0.480	303.0	303.0	307.0	307.0
13	320.0	0.523	0.522	0.520	314.0	313.0	316.0	316.0
14	330.0	0.564	0.563	0.560	327.0	326.0	326.0	326.0
15	335.0	0.605	0.604	0.600	339.0	337.0	336.0	336.0
16	349.0	0.646	0.645	0.640	351.0	349.0	347.0	346.0
17	360.0	0.688	0.686	0.680	364.0	361.0	358.0	357.0
18	370.0	0.729	0.727	0.720	376.0	373.0	371.0	369.0
19	385.0	0.770	0.768	0.760	388.0	385.0	385.0	383.0
20	402.0	0.811	0.809	0.800	403.0	397.0	401.0	398.0
21	421.0	0.852	0.850	0.840	427.0	420.0	420.0	416.0
22	453.0	0.894	0.892	0.880	450.0	442.0	454.0	437.0
23	485.0	0.935	0.933	0.920	493.0	476.0	480.0	466.0
24	545.0	0.976	0.974	0.960	561.0	522.0	545.0	513.0
				RMSE	6.7	8.9	2.3	9.7
				Max$\lvert Q(fitted) - \hat{Q}(formula) \rvert$	16.0	24.0	5.0	32.0

$\hat{Q}$( ) = Flood estimated from plotting position formula for the same non-exceedance probability

6.2 Historical Flood Records

This section involves the application of plotting position formulas developed in the present study to actual flood records which contain some extraordinary floods. The application is divided into three parts. First, the P3 and GEV formulas [eqns. (4-16) and (4-17)] which have been developed for systematic flood records are applied to historical flood data, and their performances are assessed. Second, effects of the uncertainty in the identification of flood base level (denoted here as Q_0) are examined. Third, the plotting position formulas for systematic and historical flood records are compared in terms of the bias in flood quantile estimation. In addition, the E-W formula based on the well-known Weibull concept (Hirsch, 1987) is also considered in this numerical application.

Similar to the comparison approach used in the previous section for systematic flood samples, the floods quantiles estimated by plotting position formulas are also compared with those computed by some conventional distribution fitting methods For historical flood records, the method of historically weighted moments, which has been widely used in practice (see, e.g., Condie and Lee, 1982), is employed in this study for estimating the parameters of fitted distributions. In the case of systematic flood records, the method of moments (for P3 distribution) and the PWM method (for GEV distribution) will be used for parameters estimation, as described in the previous section.

The application of the plotting position formulas proposed to actual flood samples with historical information will be illustrated by two examples using flood data at two sites: the Huangbizhuang River at Huangbizhuang, China (UNESCO, 1987), and the Boyne River near Carman, U.K. (Pilon et al., 1985). These two stations give two disparate examples representing two regions with different hydrologic con ditions. The record of annual maximum floods for the Huangbizhuang River is available from 1794 to 1974, and there exist two very large floods that occurred during the period of systematic gauging (1949-1974). In the case of Boyne River, there are three extreme floods which were observed during the period of continuous systematic gauging (1956-1982), and one extraordinary flood which was recorded in 1893. Results of the comparison are illustrated in the following.

6.2.1 Application of P3 and GEV Plotting Formulas to Historical Flood Data

As mentioned above, there exist very large floods in the systematic flood records for the Huangbizhuang and Boyne rivers. The purpose of the present graphical and numerical comparisons is to assess the performance of the P3, GEV and Weibull formulas when they are applied to these two particular sets of flood data. Fig. 6.4 and Table 6.4 show respectively results of the graphical and numerical comparisons for the Huangbizhuang River. As compared to the GEV and Weibull formulas, it is found that the P3 formula gives flood estimates closest to the values computed from the fitted P3 distribution. Similarly, the GEV formula give the best flood estimates as compared to those computed from the GEV distribution which is fitted to Boyne River flood data (Fig. 6.5, Table 6.5).

However, the results shown in Figs. 6.4-6.5 and Tables 6.4-6.5 indicate that, without censoring the flood records to consider some extreme floods above a given base level as historical floods, the P3 and GEV distributions did not give a good fit to the observed data. The bias was found to be largest at the upper end of the plots (Figs. 6.4-6.5). Hence, the P3 and GEV formulas which have been recommended for systematic flood records are not appropriate for cases where there exist historical

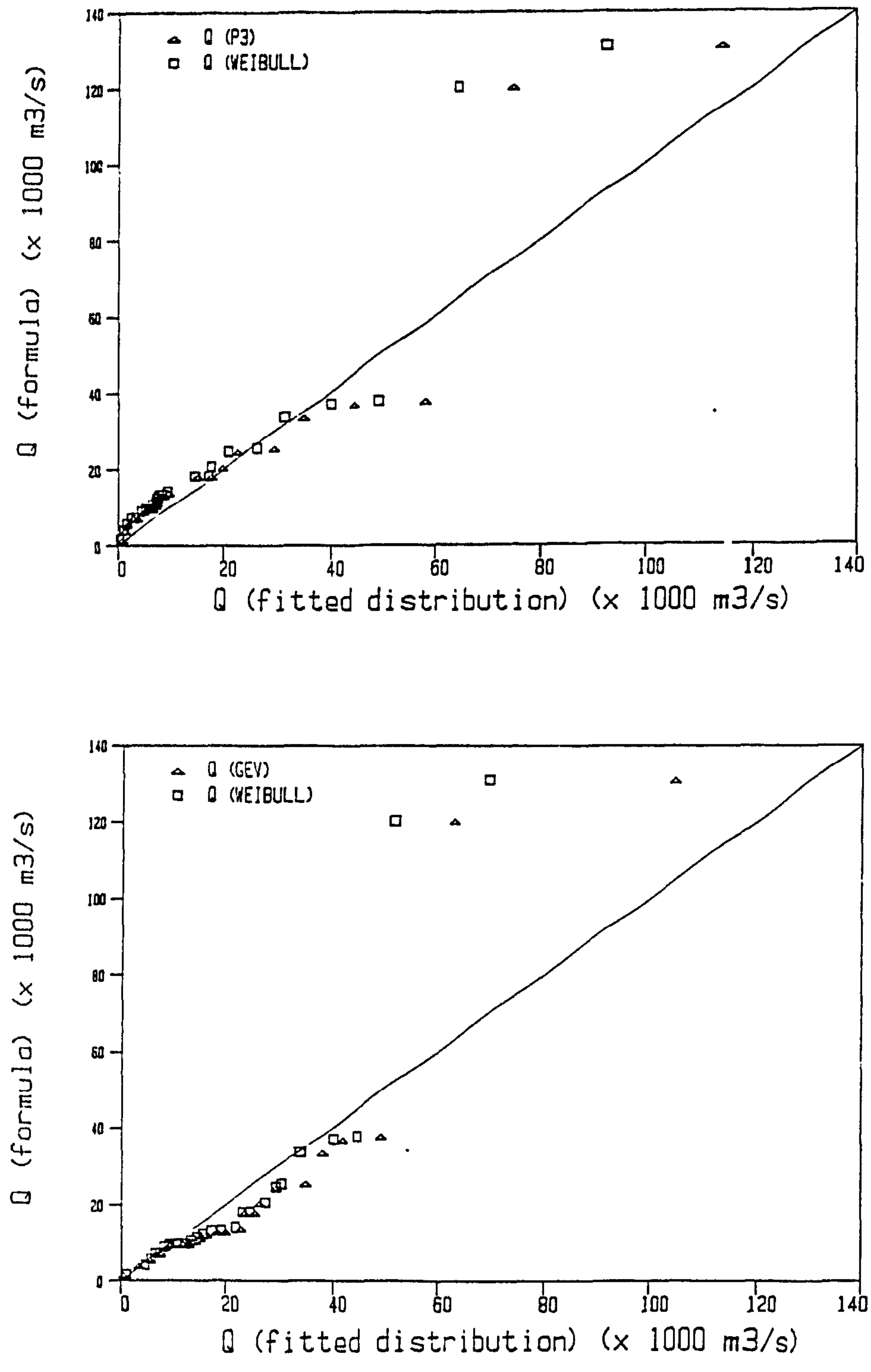
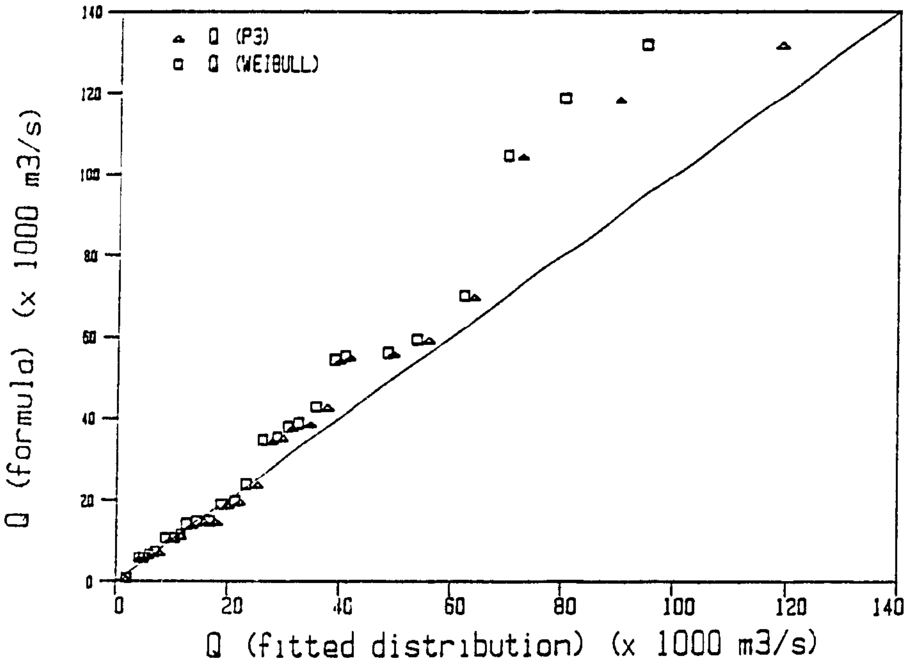


Figure 6.4 Quantile-quantile plot of P3, GEV and Weibull formulas (Huangbizhuang River, $C_{su} = 3.0, N = 25$).

Table 6.4: Numerical comparison of P3, GEV, and Weibull formulas ($C_{su} = 3.0, N = 25$)

Huangbizhuang River					P3 distribution		GEV distribution	
m	Q_{fitted} (m^3/s)	P3	GEV	Weibull	$\hat{Q}$ (P3)	$\hat{Q}$ (Weibull)	$\hat{Q}$ (GEV)	$\hat{Q}$ (Weibull)
1	200	0.061	0.033	0.038	100	50	50	100
2	398	0.100	0.073	0.077	150	120	350	500
3	550	0.138	0.113	0.115	200	180	600	620
4	750	0.176	0.153	0.154	300	250	700	700
5	760	0.214	0.193	0.192	400	350	800	800
6	920	0.252	0.233	0.231	550	480	900	880
7	950	0.291	0.273	0.269	600	580	1000	980
8	950	0.329	0.313	0.308	650	620	1200	1100
9	980	0.367	0.353	0.346	700	680	1300	1250
10	1100	0.405	0.392	0.385	750	720	1400	1350
11	1160	0.443	0.432	0.423	780	760	1500	1450
12	1260	0.481	0.472	0.462	800	790	1600	1550
13	1300	0.520	0.512	0.500	850	830	1800	1700
14	1330	0.558	0.552	0.538	900	880	2000	1900
15	1370	0.596	0.592	0.577	1000	970	2300	2200
16	1800	0.634	0.632	0.615	1500	1450	2400	2350
17	1850	0.672	0.672	0.654	1750	1700	2600	2500
18	2080	0.711	0.711	0.692	2000	1750	2700	2800
19	2450	0.749	0.751	0.731	2300	2100	3000	3000
20	2550	0.787	0.791	0.769	3000	2700	3500	3100
21	3350	0.825	0.831	0.808	3500	3200	3800	3400
22	3700	0.863	0.871	0.846	4500	4000	4200	4000
23	3820	0.901	0.911	0.885	5800	5000	5000	4500
24	12000	0.940	0.951	0.923	7500	6500	6300	5200
25	13100	0.978	0.991	0.962	11500	9300	10500	7000
				RMSE	1090	1398	1351	1870
	Max$\left\|Q(fitted) - \hat{Q}(formula)\right\|$				4500	5500	5700	6800

$\hat{Q}$( ) = Flood estimated from plotting position formula for the same non-exceedance probability



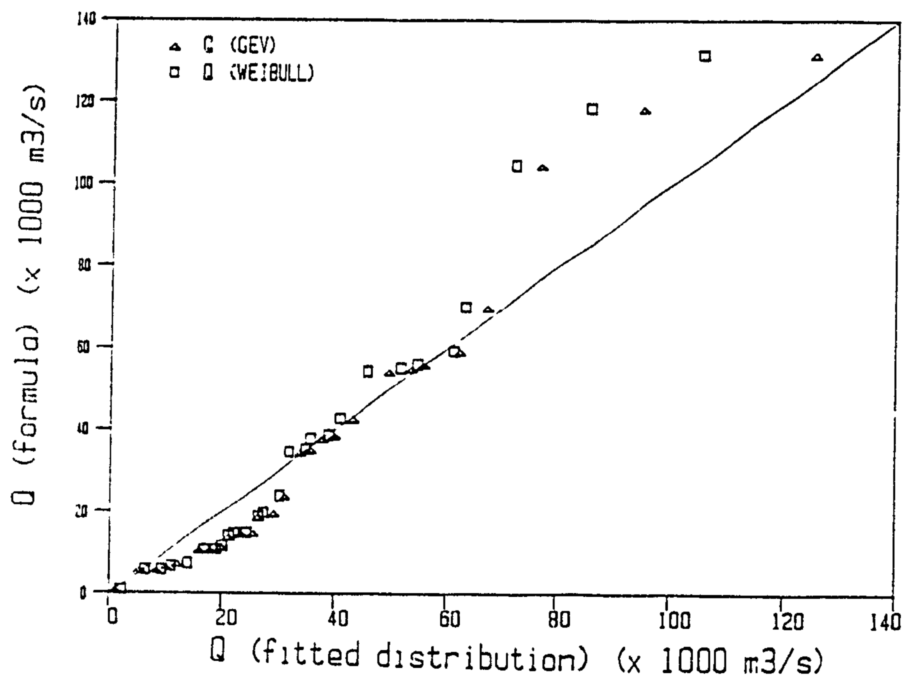


Figure 6.5 Quantile-quantile plot of P3, GEV and Weibull formulas (Boyne River, $C_{su} = 1.9, N = 27$).

Table 6.5: Numerical comparison of P3, GEV, and Weibull formulas ($C_{su} = 1.9, N = 27$)

Boyne River					P3 distribution		GEV distribution	
m	Q_{fitted} (m^3/s)	P3	GEV	Weibull	$\hat{Q}$ (P3)	Q (Weibull)	$\hat{Q}$ (GEV)	Q (Weibull)
1	1.2	0.038	0 027	0 036	2	2	1	2
2	5 5	0.074	0 064	0.071	4	4	5	6
3	6 1	0 110	0 101	0.107	5	5	8	9
4	6.5	0 146	0 138	0.143	6	6	10	11
5	7.4	0 182	0 174	0.179	8	7	12	14
6	10.7	0 219	0.211	0 214	10	9	16	17
7	10.8	0 255	0.248	0 250	11	11	18	19
8	11 4	0 291	0 284	0 286	12	12	20	20
9	13 9	0.327	0 321	0.321	14	13	21	21
10	14 5	0 364	0 358	0.357	16	15	23	22
11	14.9	0 400	0 395	0.393	18	17	25	24
12	19 3	0.436	0.431	0 429	20	19	26	26
13	19.4	0.472	0 468	0 464	22	21	29	27
14	23.8	0.508	0.505	0 500	25	23	31	30
15	34.6	0.545	0 542	0 536	28	26	34	32
16	35 7	0 581	0 578	0.571	30	29	36	35
17	37.9	0 617	0 615	0.607	32	31	38	36
18	38 8	0.655	0 652	0.643	35	33	40	39
19	43.0	0 689	0 688	0.679	38	36	43	41
20	54.1	0 726	0 725	0 714	40	39	50	46
21	55.2	0 762	0 762	0.750	42	41	54	52
22	56 1	0 798	0 799	0 786	50	49	56	55
23	59.5	0 834	0 835	0 821	56	54	62	61
24	69.7	0 870	0 872	0 857	64	62	67	63
25	105.0	0 907	0 909	0 893	73	70	77	72
26	119.0	0 943	0 946	0.929	90	80	95	85
27	132.0	0 979	0 982	0 964	119	95	125	105
				RMSE	10	14	9	12
	Max $\left\|Q(fitted) - \hat{Q}(formula)\right\|$				32	35	28	33

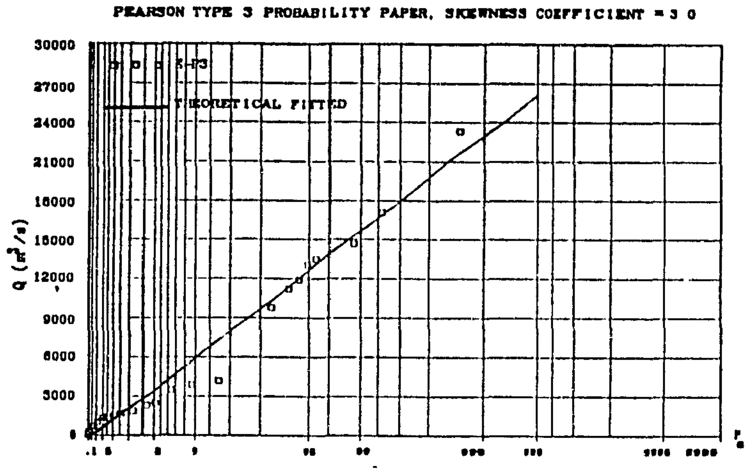
$\hat{Q}(\ )$ = Flood estimated from plotting position formula for the same non-exceedance probability

floods in the data samples. The use of exceedance-based formulas (Section 4.2) would significantly improve the fitting of theoretical distributions to the observed historical flood data, as will be shown in the following sections.

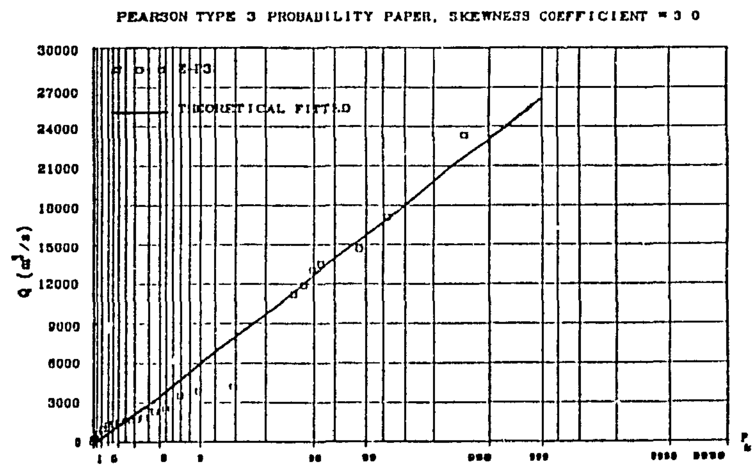
6.2.2 Effects of the Uncertainty in Flood Base Level

One of th difficulties in the analysis of historical flood records involves the accurate determination of a base level Q_0 above which all floods are considered as historical floods. In practice Q_0 is frequently established as being equal to the smallest known historical flood. In the following, an objective method for selecting an optimum value of Q_0 will be proposed using as selection criterion the best-fit of the assumed theoretical distribution to the observed data. More specifically, the best-fit condition is assessed using the graphical comparison between empirical (plotting position) and fitted probabilities, as well as the numerical comparison between flood estimates from plotting position formulas and fitted theoretical distributions (in terms of the minimum RMSE and the smallest absolute difference). To illustrate the use of the proposed method for identifying an optimum value of Q_0, the effects of various base levels on the performance of fitted theoretical distribution are examined. The sensitivity analysis is performed using historical flood data of the two rivers considered in the previous section.

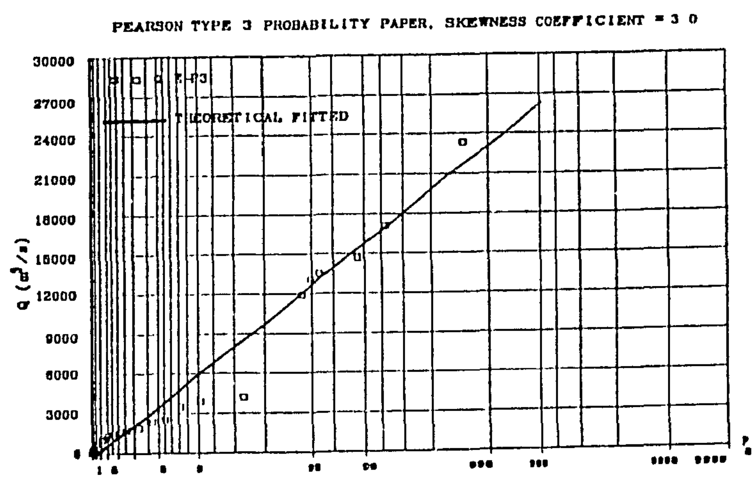
For Huangbizhuang River, a close examination of the flood record suggests some different base levels (9000 m^3/s, 10000 m^3/s and 13000 m^3/s) which could be selected for the sensitivity analysis. Results of graphical and numerical comparisons for the E-P3 formula and the fitted P3 distribution are shown in Fig 6.6 and Table 6.6, respectively; and in Fig. 6.7 and Table 6.7 for the E-GEV formula and the fitted GEV distribution . It can be seen that, with a base level selected



$Q_0 = 9000\ m^3/s$



$Q_0 = 10000\ m^3/s$



$Q_0 = 13000\ m^3/s$

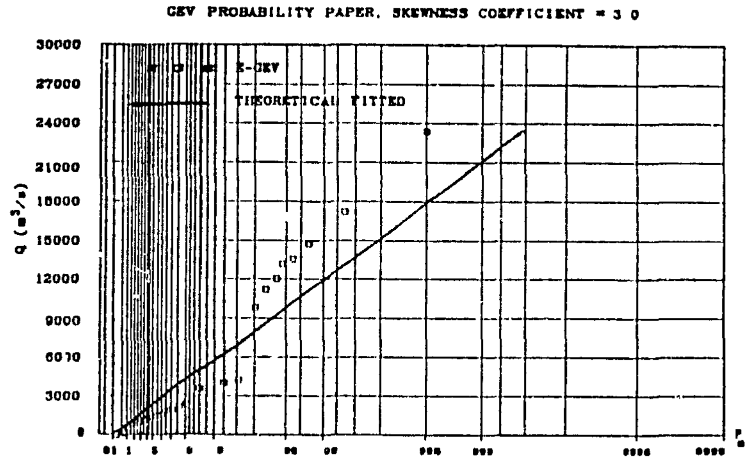
Figure 6.6 Sensitivity of P3 distribution for different base levels (Huangbizhuang River).

Table 6 6 Sensitivity of P3 distribution for different base levels (Huangbizhuang River).

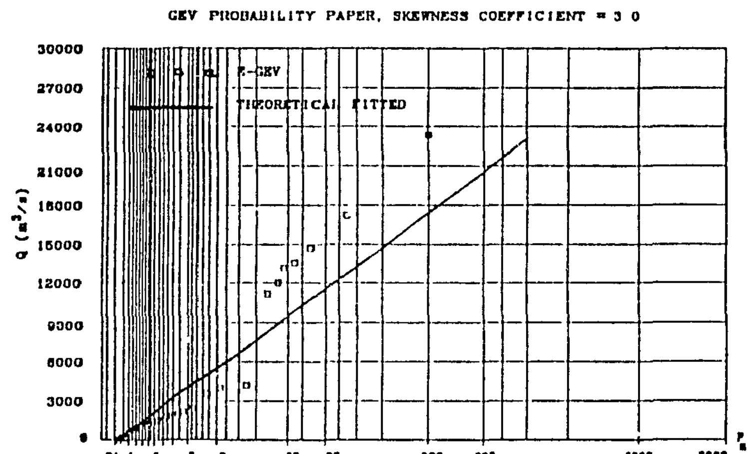
Huangbizhuang River

Q_{fitted}	$Q_0 = 9000\ m^3/s$		$Q_0 = 10000\ m^3/s$		$Q_0 = 13000\ m^3/s$	
(m^3/s)	E-P3	$Q(E-P3)$ (m^3/s)	E-P3	$Q(E-P3)$ (m^3/s)	E-P3	$Q(E-P3)$ (m^3/s)
200	0 077	100	0 077	100	0 078	100
398	0 116	400	0 116	400	0 117	400
550	0 155	500	0 115	500	0 156	500
750	0 193	600	0 195	610	0 196	610
760	0 232	700	0 234	710	0.235	710
920	0 271	800	0 273	810	0.275	810
950	0 310	850	0 312	860	0 314	860
950	0 349	900	0 351	910	0 353	910
980	0 388	950	0 390	960	0 393	960
1100	0 427	980	0 430	990	0.432	990
1160	0 466	1000	0 469	1050	0 471	1060
1260	0 505	1200	0 508	1210	0 511	1220
1300	0 544	1250	0 547	1260	0 550	1260
1330	0 583	1300	0 586	1350	0 590	1360
1370	0 622	1400	0 625	1450	0 629	1500
1800	0 661	18J0	0 665	1700	0 668	1750
1850	0 700	1850	0 704	2000	0 708	2100
2080	0 739	2000	0 743	2500	0 747	2600
2150	0 777	3000	0 782	3100	0 786	3200
2550	0 816	3200	0 821	3900	0 826	4000
3350	0 855	1100	0 860	4900	0 865	5000
3700	0 894	5500	0 899	6000	0 905	6100
3820	0 933	7000	0 938	8000	0.944	8500
9650	0 965	10000	--	-	-	-
11500	0 970	11200	0 970	11200	-	-
12000	0 974	12000	0 975	12000	0 975	12000
13100	0 979	12900	0 979	12900	0 980	12900
13500	0 983	13000	0 984	13000	0 984	13000
14750	0 988	15100	0 988	15100	0 989	15100
17150	0 993	17000	0 993	17000	0 993	17000
23750	0 997	21900	0 997	21!900	0 997	21900
	RMSE	779		1027		1135
Max $\|Q_{fitted} - Q_{formula}\|$		3180		4180		4680

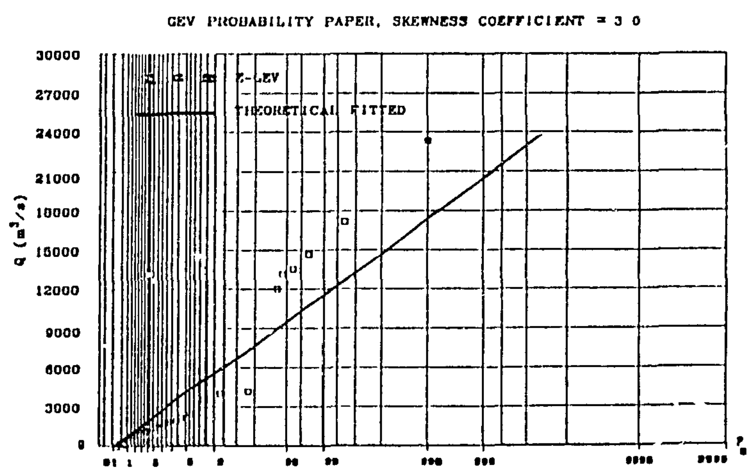
$Q(\,)$ = Flood estimated from plotting position formula for the same non-exceedance probability



$Q_0 = 9000\ m^3/s$



$Q_0 = 10000\ m^3/s$



$Q_0 = 13000\ m^3/s$

Figure 6.7 Sensitivity of GEV distribution for different base levels (Huangbizhuang River).

Table 6 7: Sensitivity of GEV distribution for different base levels (Huangbizhuang River)

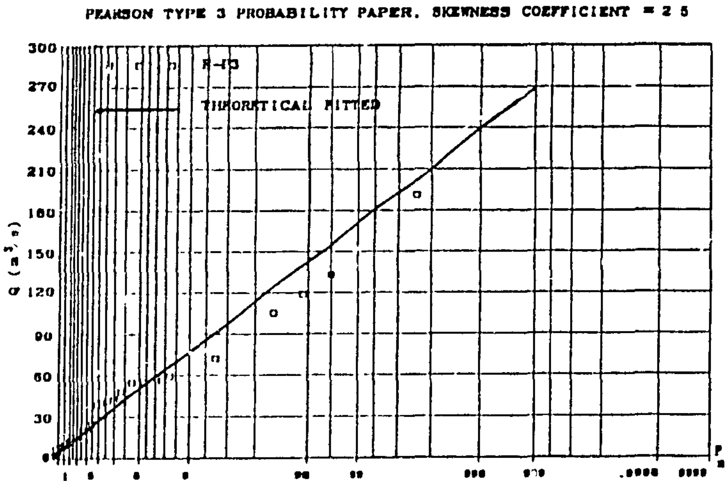
Huangbizhuang River

Q_{fitted}	$Q_0 = 9000\ m^3/s$		$Q_0 = 10000\ m^3/s$		$Q_0 = 13000\ m^3/s$	
(m^3/s)	E-GEV	$\hat{Q}(E-GEV)$ (m^3/s)	E-GEV	$\hat{Q}(E-GEV)$ (m^3/s)	E-GEV	$\hat{Q}(E-GEV)$ (m^3/s)
206	0 033	100	0 033	150	0 033	100
398	0 074	600	0 075	400	0 075	450
550	0 116	800	0 116	500	0 117	700
750	0 157	1000	0 158	800	0.159	850
760	0 198	1100	0 199	900	0 201	990
920	0 240	1200	0 241	1000	0.242	1100
950	0 281	1400	0.282	1100	0 284	1200
950	0 322	1800	0 324	1200	0 326	1300
980	0 363	1900	0 366	1300	0 368	1500
1100	0 405	2000	0 407	1500	0 409	1600
1160	0 446	2100	0 449	1700	0 451	1800
1260	0 487	2200	0 490	1800	0 493	2000
1300	0 529	2500	0 532	2000	0 535	2200
1330	0 570	2800	0 573	2300	0 577	2500
1370	0 611	3000	0 615	2600	0 618	2800
1800	0 653	3100	0 656	2900	0 660	3000
1850	0 691	3400	0 698	3100	0 702	3200
2080	0 735	3600	0 739	3500	0 744	3600
2450	0 777	4000	0 781	3600	0 786	4000
2550	0 818	4300	0 823	4000	0 827	4300
3350	0 859	5000	0 864	5000	0 869	5000
3700	0 900	6200	0 906	5700	0 911	6000
3820	0 942	7000	0 947	7000	0 953	7400
9650	0 960	8000	-	-	-	-
11500	0 966	8600	0 966	8300	-	-
12000	0 971	9200	0 971	9000	0 971	8900
13100	0 976	9500	0 976	9300	0 977	9200
13500	0 982	10000	0 982	10000	0 982	9900
14750	0 987	11000	0 987	10500	0.987	10400
17150	0 993	13000	0 993	12500	0 993	12400
23750	0 998	18000	0 998	17500	0 998	17000
	RMSE	2203		2263		2369
Max $\lvert Q_{fitted} - Q_{formula} \rvert$		5750		6250		6750

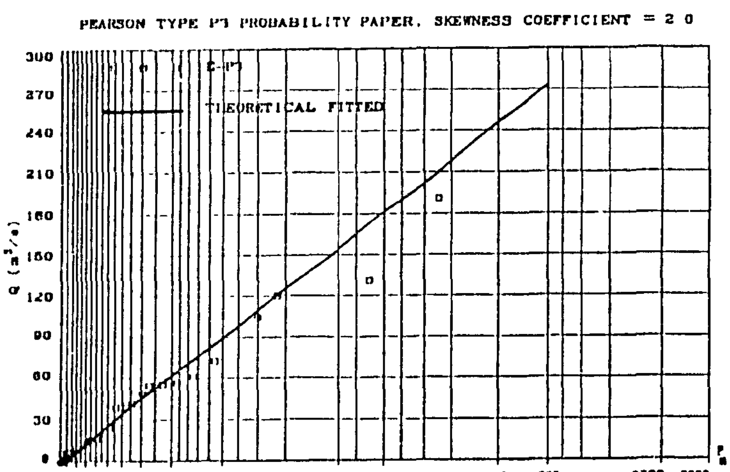
$\hat{Q}(\ )$ = Flood estimated from plotting position formula for the same non-exceedance probability

at 9000 m^3/s, the theoretical P3 distribution shows the best fit to the observed data (Fig. 6.6). Moreover, at this base level, the RMSE and maximum absolute difference of flood quantiles estimated from the E-P3 formula and the fitted P3 distribution were found to be smallest (Table 6.6). Hence, the E-P3 formula performs much better than the E-GEV in this case. Therefore, the P3 distribution is the most appropriate distribution for the Huangbizhuang River, and a flood base level at 9000 m^3/s should be selected to obtain the best fit of this distribution to the observed historical flood data.

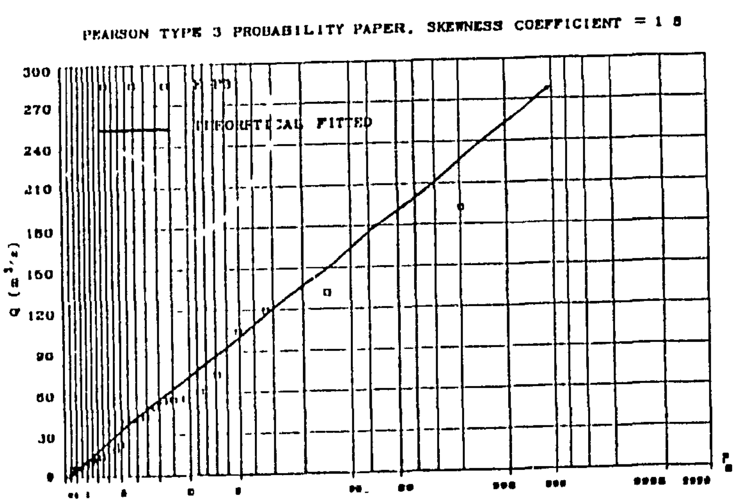
A similar sensitivity analysis procedure can be applied to the historical flood record of Boyne River. After examining the data, base levels at 100 m^3/s, 120 m^3/s, and 150 m^3/s can be chosen. Figure 6.8 and Table 6.8 show results of the sensitivity analysis for the E-P3 formula and the fitted P3 distribution. Results for the E-GEV formula and the fitted GEV distribution are presented in Fig. 6.9 and Table 6.9. It is found that, at a base level of 100 m^3/s, the GEV distribution provides a best fit to the data (Fig. 6.8). Further, at this base level, the RMSE and maximum absolute deviation of flood quantiles estimated from the E-GEV formula and the fitted GEV distribution were found to be minimum (Table 6.8). Therefore, the GEV distribution is more suitable than the P3 for this river; and the base level should be selected at 100 m^3/s to obtain the best fit of the GEV to the observed data.



$Q_0 = 100\ m^3/s$



$Q_0 = 120\ m^3/s$



$Q_0 = 150\ m^3/s$

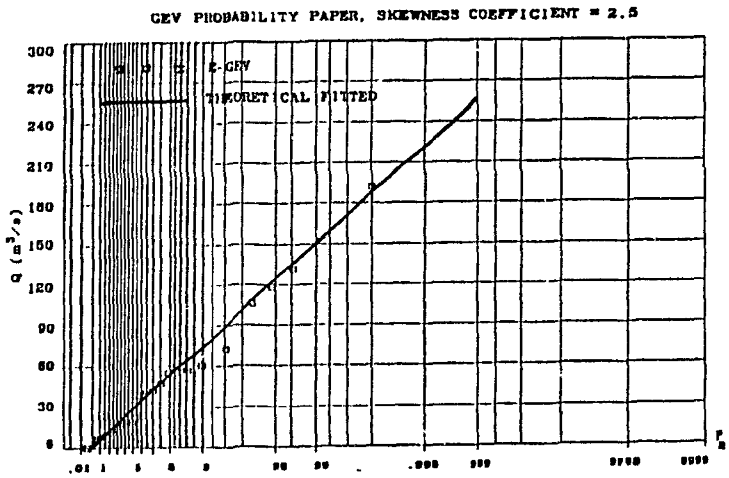
Figure 6.8 Sensitivity of P3 distribution for different base levels (Boyne River).

Table 6.8: Sensitivity of P3 distribution for different base levels (Boyne River).

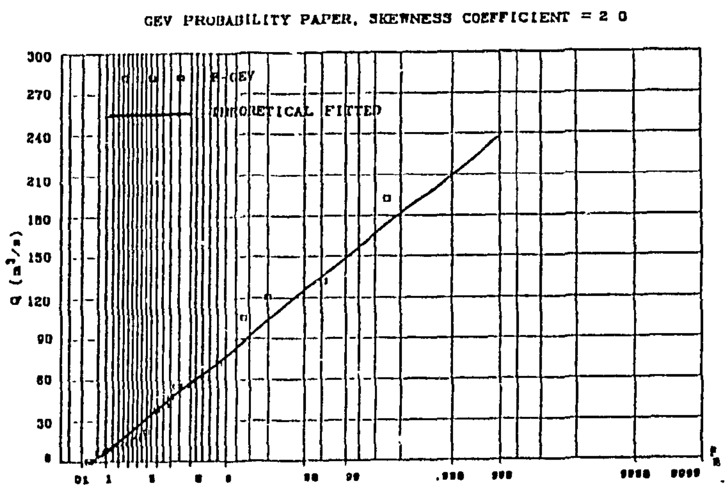
Boyne River

Q_{fitted}	$Q_0 = 100\ m^3/s$		$Q_0 = 120\ m^3/s$		$Q_0 = 150\ m^3/s$	
(m^3/s)	E-P3	$\hat{Q}(E-P3)$ (m^3/s)	E-P3	$\hat{Q}(E-P3)$ (m^3/s)	E-P3	$\hat{Q}(E-P3)$ (m^3/s)
1.2	0 047	1	0 039	0 5	0 036	2
5.5	0.086	2	0.076	1	0 072	5
6.1	0.124	3	0 113	2	0.108	6
6.5	0 163	4	0 149	3	0 144	8
7.4	0.201	5	0.186	5	0 180	10
10.7	0.240	6	0 223	6	0 215	12
10.8	0 278	9	0 259	9	0 251	16
11 4	0.317	10	0.296	10	0.287	18
13 9	0 355	11	0.333	12	0 323	20
14.5	0.394	12	0.369	15	0.359	22
14.9	0.432	15	0 406	18	0 395	24
19.3	0 471	16	0 443	20	0 430	26
19.4	0 509	17	0 480	24	0 466	30
23 8	0.548	18	0 516	26	0 502	32
34.8	0.586	22	0 553	30	0 538	35
35 7	0.625	28	0 590	32	0 574	40
37 9	0.663	32	0 626	36	0.610	43
38.8	0 702	35	0 663	40	0 646	46
43.0	0.741	40	0 700	44	0.681	52
54.1	0 779	45	0.736	50	0 717	56
55 2	0 818	52	0.773	55	0 753	62
56.1	0.856	62	0 810	64	0.789	68
59.5	0.895	65	0 846	73	0.825	78
69.7	0 933	90	0 883	84	0 861	88
105 0	0.967	124	0 920	110	0 896	100
119.0	0 976	140	0.956	120	0 932	116
132.0	0.985	154	0.987	190	0 968	150
187.0	0.995	200	0.995	210	0 996	225
	RMSE	9		13		11
Max$\lvert Q_{fitted} - \hat{Q}_{formula}\rvert$		22 0		58 0		38 0

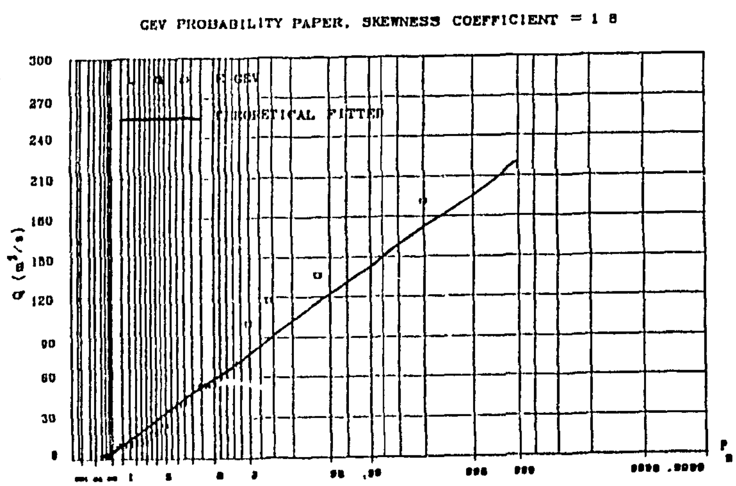
$\hat{Q}(\ )$ = Flood estimated from plotting position formula for the same non-exceedance probability



$Q_0 = 100\ m^3/s$



$Q_0 = 120\ m^3/s$



$Q_0 = 150\ m^3/s$

Figure 6.9 Sensitivity of GEV distribution for different base levels (Boyne River).

Table 6.9: Sensitivity of GEV distribution for different base levels (Boyne River).

Boyne River

Q_{fitted} (m^3/s)	$Q_0 = 100\ m^3/s$ E-GEV	$Q_0 = 100\ m^3/s$ $\hat{Q}(E-GEV)$ (m^3/s)	$Q_0 = 120\ m^3/s$ E-GEV	$Q_0 = 120\ m^3/s$ $\hat{Q}(E-GEV)$ (m^3/s)	$Q_0 = 150\ m^3/s$ E-GEV	$Q_0 = 150\ m^3/s$ $\hat{Q}(E-GEV)$ (m^3/s)
1.2	0.031	1	0.028	1	0.027	2
5.5	0.070	5	0.065	4	0.063	7
6.1	0.110	6	0.103	7	0 099	10
6.5	0.149	7	0.140	9	0.136	15
7.4	0 189	10	0 177	12	0.172	18
10.7	0 228	11	0.214	15	0.208	20
10.8	0.268	13	0 252	17	0 245	21
11.4	0.307	15	0.289	19	0.281	23
13 9	0.347	18	0 326	21	0.317	25
14.5	0.386	20	0.364	22	0.354	26
14 9	0 426	21	0 401	25	0.390	28
19.3	0.465	24	0.438	28	0 426	30
19.4	0.505	27	0 475	30	0 463	32
23.8	0 544	30	0 513	32	0 499	35
34 8	0.584	34	0 550	35	0 535	38
35.7	0.623	36	0 587	38	0 571	40
37 9	0.663	38	0 625	40	0 608	42
38 8	0.702	41	0.662	44	0 644	45
43.0	0.742	46	0 699	48	0.680	48
54.1	0.781	54	0 737	52	0 717	52
55.2	0.821	56	0.774	55	0.753	55
56.1	0 861	62	0.811	60	0 789	58
59.5	0.900	72	0 848	66	0 826	65
69.7	0.940	87	0.886	73	0.862	73
105.0	0.964	106	0.923	86	0.898	79
119.0	0.974	121	0 960	103	0 935	88
132.0	0.985	134	0.985	135	0.971	115
187 0	0.996	187	0 995	172	0.996	170
	RMSE	5		8		12
Max $\lvert Q_{fitted} - \hat{Q}_{formula} \rvert$		16 3		19 0		31 0

$\hat{Q}(\cdot)$ = Flood estimated from plotting position formula for the same non-exceedance probability

6.2.3 Comparison Between Plotting Formulas for Systematic and Historical Flood Records

In Section 6.2.1, it has been shown that it was not appropriate to analyze historical flood records using plotting position formulas (e.g., P3 and GEV formulas) which were derived for systematic data. In the following, graphical and numerical comparisons of flood estimates from formulas recommended for systematic records, and from those suggested for combined historical and systematic data are carried out in order to give a general impression of the magnitude of the differences between results. The performances of P3 and E-P3 formulas are assessed using flood data of the Huangbizhuang River since the P3 distribution has been shown to be the most suitable for this river (Section 6.2.1). Similarly, results of flood estimates from GEV and E-GEV formulas are compared using Boyne River flood record. Further, the optimum base levels determined in the previous section for Huangbizhuang and Boyne Rivers are used in the analysis of historical flood records in the present comparisons. Finally, the E-W formula is also considered in this comparative study because of its popularity in practice, as mentioned above.

For the Huangbizhuang River, results of the graphical and numerical comparisons between P3 and E-P3 formulas are shown in Fig. 6.10 and Table 6.10, respectively. It can be seen that the use of the E-P3 for historical flood data provides the minimum RMSE and the smallest bias in flood estimates as compared to those given by the P3 and the E-W (Table 6.10). Results shown in Fig. 6.10 indicate also that the E-P3 formula appears to be preferable to the E-W. Similarly, for Boyne River, it was found that the E-GEV formula provides the best performance as compared with the GEV and E-W formulas (Fig. 6.11 and Table 6.11).

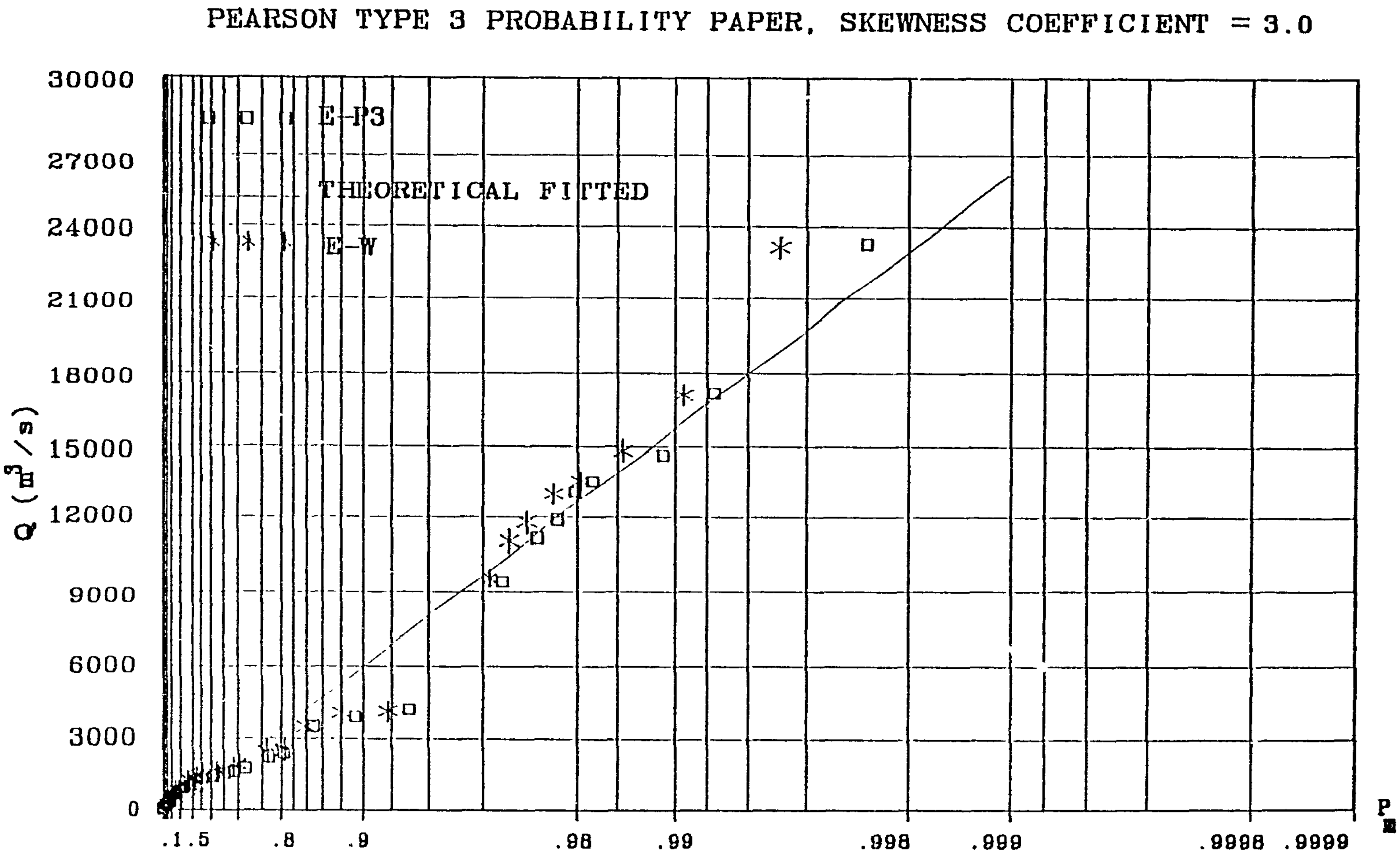


Figure 6.10 Comparison of E-P3 and E-W formulas (Huangbizhuang River, $C_{su} = 3.0, N = 181$).

Table 6 10: Comparison of E-P3 and E-W formulas.

Huangbizhuang River

Q_{fitted}	Systematic		Systematic + Historic $Q_0 = 9000\ m^3/s$			
(m^3/s)	P3	$Q(P3)$ (m^3/s)	E-P3	$Q(E-P3)$ (m^3/s)	E-W	$Q(E-W)$ (m^3/s)
200	0 061	100	0.077	100	0 040	50
398	0.100	150	0.116	400	0 080	150
550	0 138	200	0.155	500	0.119	420
750	0.176	300	0.193	600	0.159	550
760	0.214	400	0 232	700	0.199	650
920	0.252	550	0.271	800	0.239	750
950	0.291	600	0 310	850	0.279	820
950	0.329	650	0.349	900	0.319	880
980	0 367	700	0 388	950	0.358	930
1100	0 405	750	0 427	980	0 398	970
1160	0.443	780	0.466	1000	0.438	990
1260	0 481	800	0 505	1200	0.478	1100
1300	0.520	850	0.544	1250	0 518	1220
1330	0.558	900	0 583	1300	0 558	1280
1370	0.596	1000	0.622	1400	0.597	1350
1800	0.634	1500	0 661	1800	0.637	1500
1850	0.672	1750	0.700	1850	0.677	1820
2080	0.711	2000	0.739	2000	0.717	1900
2450	0 749	2300	0 777	3000	0.757	2900
2550	0.787	3000	0 816	3200	0.797	3100
3350	0 825	3500	0.855	4100	0 836	4000
3700	0.863	4500	0 894	5500	0.876	5000
3820	0.901	5800	0.933	7000	0.916	6600
9650	-	-	0.965	10000	0.961	9800
11500	-	-	0 970	11200	0.966	10000
12000	0.940	7500	0.974	12000	0.971	11200
13100	0.978	11500	0.979	12900	0.975	12100
13500	-	-	0.983	13000	0.980	12900
14750	-	-	0.988	15100	0 985	14000
17150	-	-	0 993	17000	0.990	15600
23750	-	-	0 997	21900	0.995	18500
	RMSE	1090		779		1212
Max	$\lvert Q_{fitted} - \hat{Q}_{formula} \rvert$	4500		3180		5250

$Q(\ )$ = Flood estimated from plotting position formula for the same non-exceedance probability

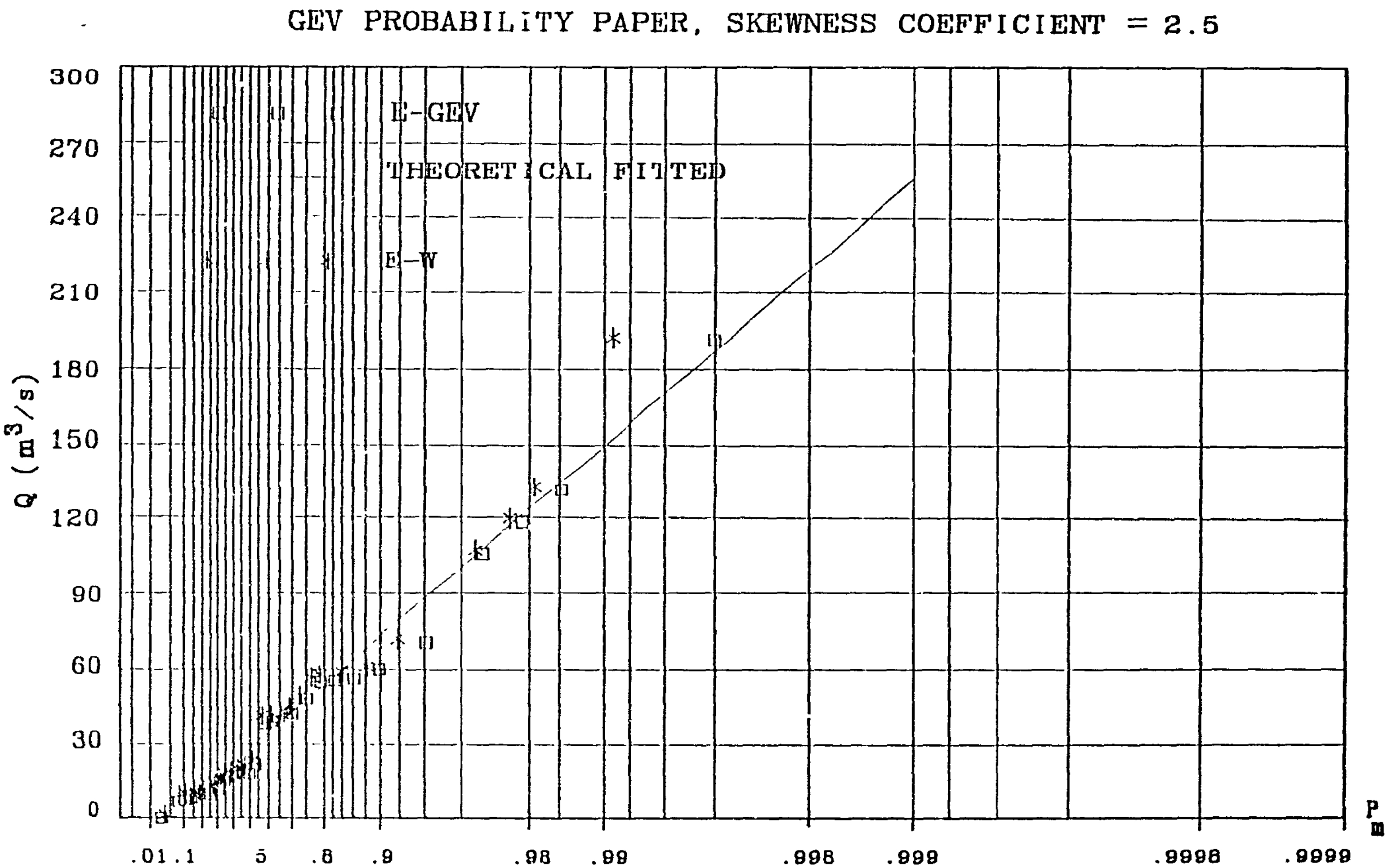


Figure 6.11 Comparison of E-GEV and E-W formulas (Boyne River, $C_{su} = 2.5, N = 90$).

Table 6.11: Comparison of E-GEV and E-W formulas.

Boyne River

Q_{fitted}	Systematic		Systematic + Historic $Q_0 = 100\ m^3/s$			
(m^3/s)	GEV	$Q(GEV)$ (m^3/s)	E-GEV	$Q(E-GEV)$ (m^3/s)	E-W	$\hat{Q}(E-W)$ (m^3/s)
1.2	0.027	1	0.031	1	0.038	2
5.5	0.064	5	0.070	5	0.076	5
6.1	0.101	8	0.110	6	0.115	6
6.5	0.138	10	0.149	7	0.153	8
7.4	0.174	12	0.189	10	0.191	10
10.7	0.211	16	0.228	11	0.229	12
10.8	0.248	18	0.268	13	0.268	13
11.4	0.284	20	0.307	15	0.306	16
13.9	0.321	21	0.347	18	0.344	17
14.5	0.358	23	0.386	20	0.382	19
14.9	0.395	25	0.426	21	0.420	22
19.3	0.431	26	0.465	24	0.459	23
19.4	0.468	29	0.505	27	0.497	26
23.8	0.505	31	0.544	30	0.535	28
34.8	0.542	34	0.584	34	0.573	32
35.7	0.578	36	0.623	36	0.612	35
37.9	0.615	38	0.663	38	0.650	37
38.8	0.652	40	0.702	41	0.688	39
43.0	0.688	43	0.742	46	0.726	42
54.1	0.725	50	0.781	54	0.764	50
55.2	0.762	54	0.821	56	0.803	55
56.1	0.799	56	0.861	62	0.841	60
59.5	0.835	62	0.900	72	0.879	66
69.7	0.872	67	0.940	87	0.917	78
105.0	0.909	77	0.964	106	0.964	106
119.0	0.946	95	0.974	121	0.973	115
132.0	0.982	125	0.985	134	0.982	128
187.0	-	-	0.996	187	0.991	155
	RMSE	9		5		7
Max $\lvert Q_{fitted} - \hat{Q}_{formula} \rvert$		28		16.3		32

$\hat{Q}(\ )$ = Flood estimated from plotting position formula for the same non-exceedance probability

In summary, it is important to recognize the fundamental difference between a flood record with historical and systematic data and a record which is entirely systematic. Depending on the nature of the data sample, the choice of exceedance plotting formulas (E-P3, E-GEV) rather than their traditional counterparts (P3, GEV) could significantly improve the estimation of floods. Further, the proposed objective method for identifying the base level for historical flood data analysis using the special P3 and GEV probability papers developed in the present study demonstrate the convenience in the use of these papers in practice.

CHAPTER 7

CONCLUSIONS

The following conclusions can be drawn from this study :

(1) The PWM method was found to be suitable for the analytical derivation of exact plotting positions for P3 and GEV distributions. Further, it was shown that the estimation of exact plotting positions using the PWM procedure was preferable to both the direct numerical integration method and the Monte Carlo simulation technique.

(2) For practical applications, simple unbiased plotting position formulas have been developed for P3 and GEV distributions, and for both systematic and historic flood records. It was found that the proposed formulas provided a better approximation to the exact plotting positions than several existing formulas. In particular, the formulas developed in this study for historical flood data can provide less bias in discharge estimation than several existing formulas.

(3) As compared with existing plotting positions, the formulas suggested in this study are conceptually more flexible and computational more convenient because they can take explicitly into account the skewness coefficient of the underlying distributions, and thus can be used for both symmetrical and non-symmetrical distributions.

(4) The well-known Weibull and E-W formulas were shown to be biased for both P3 and GEV distributions. Therefore, they should be used only for the uniform

distribution for which they were specifically derived. Further, it was noted that, although Arnell formula was especially recommended for the GEV distribution, results given by this formula were found to be as biased as those obtained from the Weibull formula.

(5) Results of numerical examples using actual flood data have indicated the adequacy of the new plotting position formulas developed in this study. Therefore, it can be concluded that the proposed plotting position formulas are the most appropriate for the P3 and GEV distributions.

(6) Finally, the development of special probability papers for the P3 and GEV distributions as suggested in this study would provide a convenient and practical tool for the application of these distributions in engineering practice.

STATEMENT OF ORIGINALITY

To the best of the author's knowledge, this investigation represents the following original contributions to the development of new plotting position formulas for some well-known distributions in engineering hydrology :

(1) The Probability Weighted Moment (PWM) theory was first used to find the exact plotting positions for P3 distribution.

(2) New Plotting Position formulas for P3 and GEV distributions are developed for use with systematic flood records.

(3) New Plotting Position formulas for P3 and GEV distributions are developed for use with historical flood records.

(4) The probability papers for P3 and GEV distributions are first developed for use in engineering application.

REFERENCES

Adamowski, K., 1981. Plotting Position Formula for Flood Frequency, Water Resources Bulletin, Vol. 17, No. 2, pp. 197-201.

Arnell, N.W., Beran. M., and Hosking, J.R.M., 1986. Unbiased Plotting Position for the General Extreme Value Distribution, Journal of Hydrology, Vol. 86, pp. 59-69

Barnett, V., 1975. Probability Plotting Methods and Order Statistics, Applied Statistics, Vol. 24, pp. 95-108.

Beable, M.E., and McKerchar, A.I., 1982. Regional Flood Estimation in New Zealand, New Zealand National Water and Soil Cons. Org., Water and Soil Tech. Publ. No. 20, 132 pp.

Beard, L.R., 1943. Statistical Analysis in Hydrology, Transactions of the ASCE, Vol. 108, pp. 1110-1160.

Benson, M.A., 1950. Use of Historical Data in Flood-frequency Analysis, Trans Am Geophys. Union, Vol. 31, No. 3, pp. 419-424.

Benson, M.A., 1962. Plotting Positions and Economics of Engineering Planning, Journal of the Hydraulics Division, ASCE, Vol. 88, No. HY6, pp. 57-71

Benard, A. and Bos-Levenbach, E.C., 1953. The Plotting of Observations on Probability Paper, Statistica, No. 7, pp. 163-173.

Bernier, J., Miquel, A., Lebosse, and Griffet, A., 1986. Use of Additional Historical Information for Estimation and Goodness of Fit of Flood Frequency Models, Paper Presented at the International Symposium on Flood Frequency and Risk Analysis, Louisiana State University, May 14-17.

Blom, G., 1958. Statistical Estimates and Transformed Beta-Variable, John Wiley and Sons, New York.

Bobée, B., and Morin, G., 1973. Statistique d'ordre de la Loi Pearson III et de sa forme Dérivée a Asymétrie Négative, Revue de Statistique Appliquée, Vol. XXI, No. 4, pp. 69-80.

Bobée, B., and Robitaille, R., 1975. Correction of Bias in the Estimation of the Coefficient of Skewness, Water Resources Research, Vol. 11, No. 6, pp. 851-854.

Bobée, B., and Robitaille, R., 1977. The Use of the Pearson Type 3 and Log Pearson Type 3 Distributions Revisited, Water Resources Research, Vol. 13, No. 2, pp. 427-443.

California Department of Public Works, 1923. Flow in California Streams, Bulletin No. 5, Division of Engineering and Irrigation, California State Printing Office, Sacramento.

Chegodayev, N.N., 1953. Computation of Runoff on Small Catchments (Russian), Transzhedorizdat, Moscow.

Chegodayev, N.N., 1955. Formulas for the Calculation of the Confidence of Hydrological Quantities. In Chow, V.T. (Editor), Handbook of Applied, Hydrology, McGraw-Hill, New York, 1964.

Chernoff, H., and Lieberman, G.J., 1956. The Use of Generalized Probability Paper, Annuals of Mathematical Statistics, Vol. 27, pp. 806-818.

Chow, V.T., 1964. Plotting of Data, Chapter 8 in Handbook of Applied Hydrology, McGraw-Hill, New York, pp. 8.28-8.29.

Condie, R. and Lee, K., 1982. Flood Frequency Analysis with Historic Information, Journal of Hydrology, Vol. 58, pp. 47-61.

Cunnane, C., 1978. Unbiased Plotting Positions - A Review, Journal of Hydrology, Vol. 37, pp. 205-222.

Environment Canada 1983. Historical Streamflow Summary, Water Resources Branch, Inland Waters Directorate, Ottawa, Ontario, Canada.

Foster, H.A., 1936. Methods for Estimating Floods, U.S. Geological Survey Water-Supply Paper 771: Floods in the United States, U.S Government Printing Office, Washington, D.C.

Fuller, W.S., 1914. Flood Flows. Transactions of the ASCE, Vol. 77, pp. 564-617; Discussion: pp. 618-694.

Galton, F., 1899. A Geometric Determination of the Median of a System of Normal Variants from Two of its Centiles, Nature, No. 61, pp. 102-104.

Gerson, M., 1975. The Techniques and Uses of Probability Plotting, The Statistician, Vol. 24, pp. 235-258.

Gerard, R., and Karpuk, E.W., 1979. Probability Analysis of Historical Flood Data. Journal of the Hydraulics Division, ASCE, Vol. 105, No. HY9, pp. 1153-1165.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C., and Wallis, J.R., 1979. Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form, Water Resources Research, Vol. 15, No. 5, pp. 1049-1054.

Gringorten, I.I., 1963. A Plotting Rule for Extreme Probability Paper, Journal of Geophysical Research, Vol. 68, No. 3, pp. 813-814.

Gumbel, E.J., 1943. On the Plotting of Flood Discharges, Trans. Am. Geophys. Res., Vol. 68, No. 3: pp. 813-814.

Gumbel, E.J., 1947. The Distribution of the Range. Annuals of Mathematical Statistics, No. 18, pp. 384-412.

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

Harter, H.L., 1964. Expected Values of Exponential, Weibull and Gamma Order Statistics, Report No. AD 436-763, U.S. National Technical Information Service, Springfield, Virginia.

Harter, H.L., 1984. Another Look at Plotting Positions, Communications in Statistics, Theory and Methods, Vol. 13, No. 13, pp. 1613-1633.

Hazen, A., 1914. Storage to be Provided in Impounding Reservoirs for Municipal Water Supply, Transactions of the ASCE, Vol. 77, No. 1308, pp. 1547-1550.

Hirsch, R.M. 1987. Probability Plotting Position Formulas for Flood Records with Historical Information, Journal of Hydrology, Vol. 96, pp. 185-199.

Hirsch, R.M., and Stedinger, J.R., 1987. Plotting Positions for Historical Floods and Their Precision, Water Resources Research, Vol. 23, No. 4, pp. 715-727.

Hosking, J.R.M., Wallis, J.R., and Wood, E.F., 1985. Estimation of the Generalised Extreme Value Distribution by the Method of Probability Weighted Moments, Technometrics, Vol. 27, No. 3, pp. 251-261.

In-na, N., and Nguyen, V.T.V., 1989. An Unbiased Plotting Position Formula for General Extreme Value Distribution, Journal of Hydrology, Vol. 106 (in press).

Interagency Advisory Committee on Water Data, 1982. Guidelines for Determining Flood Flow Frequency, Washington, D.C.

Jenkinson, A.F., 1955. The Frequency Distribution of the Annual Maximum (or Minimun) Values of Meteorological Elements, Q.J.R. Meteorol.Soc., No. 81, pp. 158-171.

Johnson, L.G., 1951. The Median Ranks of Sample Values in their Population with an Application to Certain Fatigue Studies, Industrial Mathematics, Vol. 2, pp. 1-9.

Kimball, B.F., 1946. Assignment of Frequencies to a Completely Ordered Set of Sample Data. Trans. Am. Geophys. Union, Vol. 27, pp. 843-846; Discussion: Vol. 28 (1947), pp. 951-953.

Kimball, B.F., 1947. Reply to E.J. Gumbel's Discussion of Kimball, 1946. (q.v.) Trans. Am. Geophys. Union, Vol. 28, No. 6 : pp. 952-953.

Kimball, F., 1960. Plotting Positions on Probability Paper, American Statistical Association Journal, pp. 546-560.

King, J.R., 1981. Probability Charts for Decision Making (Revised edition), Team Publications, Tamworth, New Hampshire (First edition, 1971, Industrial Press, New York).

Lebedev, V.V., 1952. Gidrologiya i Gidrometriya w Zadačah (Hydrology and Hydrometry Problem), Gidrometeorologiceskoe Izdatel'stvo, Leningrad.

Mage, D., 1982. An Objective Graphical Method for Testing Normal Distributional Assumptions Using Probability Plots, The American Statistician, Vol. 36, No. 2, pp. 116-120.

Matalas, N.C., Slack, J.R., and Wallis, J.R., 1975. Regional Skew in Search for a Parent, Water Resources Research, Vol. 11, No. 6, pp. 815-826.

NERC, 1975a. Flood Studies Report, Vol. I, U.K. Natural Environment Research Council, London.

NERC, 1975b. Flood Studies Report, Vol. IV, U.K. Natural Environment Research Council, London.

Nguyen, V.T.V., In-na, N., and Bobee, B., 1989. A New Plotting Position Formula for Pearson Type III Distribution, Journal of Hydraulics Division, American Society of Civil Engineering (ASCE), (in press).

Pilon, P.J., Condie, R., and Harvey, K. D., 1985. Consolidated Frequency Analysis Package (CFA) User Manual Version 1 - DEC PRO Series, Water Resources Branch, Inland Waters Directorate, Environment Canada, Ottawa, Ontario, Canada.

Reich, B.M., and Renard, K.G., 1981. Application of Advances in Flood Flow Frequency Analysis, Water Resources Bulletin, Vol. 17, No. 1, pp. 67-74.

Sinclair, C.D., and Ahmad, M.I., 1988. Location-invariant Plotting Positions for PWM Estimation of the Parameters of the GEV Distribution, Journal of Hydrology, Vol. 99, pp. 271-279.

Singh, K.P., and Sinclair, R.A., 1972. Two Distribution Method for Flood Frequency Analysis, Journal of Hydraulics Division, ASCE, Vol. 98, No. HY1, pp. 29-44.

Srikanthan, R., and McMahon, T.A., 1981. Log Pearson Type III Distribution - An Empirically Derived Plotting Position (Short Communications), Journal of Hydrology, Vol. 52, pp. 161-163.

Stacy, E.W., 1962. A Generalization of the Gamma Distribution, Annals of Mathematical Statistics. Vol. 33, pp. 1187-1192.

Stacy, E.W. and Mihram, G.A., 1965. Parameter Estimation for a Generalized Gamma Distribution, Technometrics, Vol. 7, pp. 349-358.

Stedinger, J.R., and Cohn, T., 1986. Flood Frequency Analysis with Historical and Paleoflood Information, Water Resources Research, Vol. 22, No. 5, pp. 785-793.

Stipp, J. R., and Young, G. K., 1971. Plotting Positions for Hydrologic Frequencies, Journal of Hydraulics Division, ASCE, Vol. 97, No. HY1, pp. 219-222.

Sutcliffe, J.V., 1987. The Use of Historical Records in Flood Frequency Analysis, Journal of Hydrology, Vol. 96, pp. 159-171.

Tukey, J.W., 1962. The Future of Data Analysis, Annuals of Mathematical Statistics, Vol. 33, No. 1, pp. 1-67.

UNESCO, 1987. Casebook of Methods for Computing Hydrological Parameters for Water, Projects, Paris, pp. 286.

U.S. Federal Emergency Management Agency, 1982. Flood Insurance Study Guidelines and Specifications for Study Contractors, Washington, D.C.

U.S. National Research Council, 1985. Safety of Dams–Flood and Earthquake Criteria, National Academy Press, Washington, D.C.

U.S. Water Resources Council, 1967. A Uniform Technique for Determining Flood Flow Frequencies, Bulletin 15, Washington, D.C.

U.S. Water Resources Council, 1977. Guidelines for Determinating Flood Flow Frequency, Vol. 6, Bull. 17A, pp. 1-3, revised.

Webster, R., 1985. PRODESIGN II: the Easy to Use CAD System, User's Manual, American Small Business Computers, Pryor, Oklahoma, 125 pages.

Weibull, W., 1939. A Statistical Theory of Strength of Materials, Ing. Vet. Ak. Handl., No. 151, Generalstabens Litografiska Anstals Forlag, Stockholm.

Wilk, M.B., Gnanadesikan, R., and Huyett, M.J., 1962. Probability Plots for the Gamma Distribution, Technometrics, Vol. 4, No. 1, pp. 1-20.

World Meteorological Organization, 1969. Estimation of Maximum Floods, Tech. Note 98, Geneva.

Xuewu, J., Jing, D., Shen, H.W., and Salas, J.D., 1984. Probability Plots for Pearson Type III Distribution, Journal of Hydrology, Vol. 74, pp. 1-29.

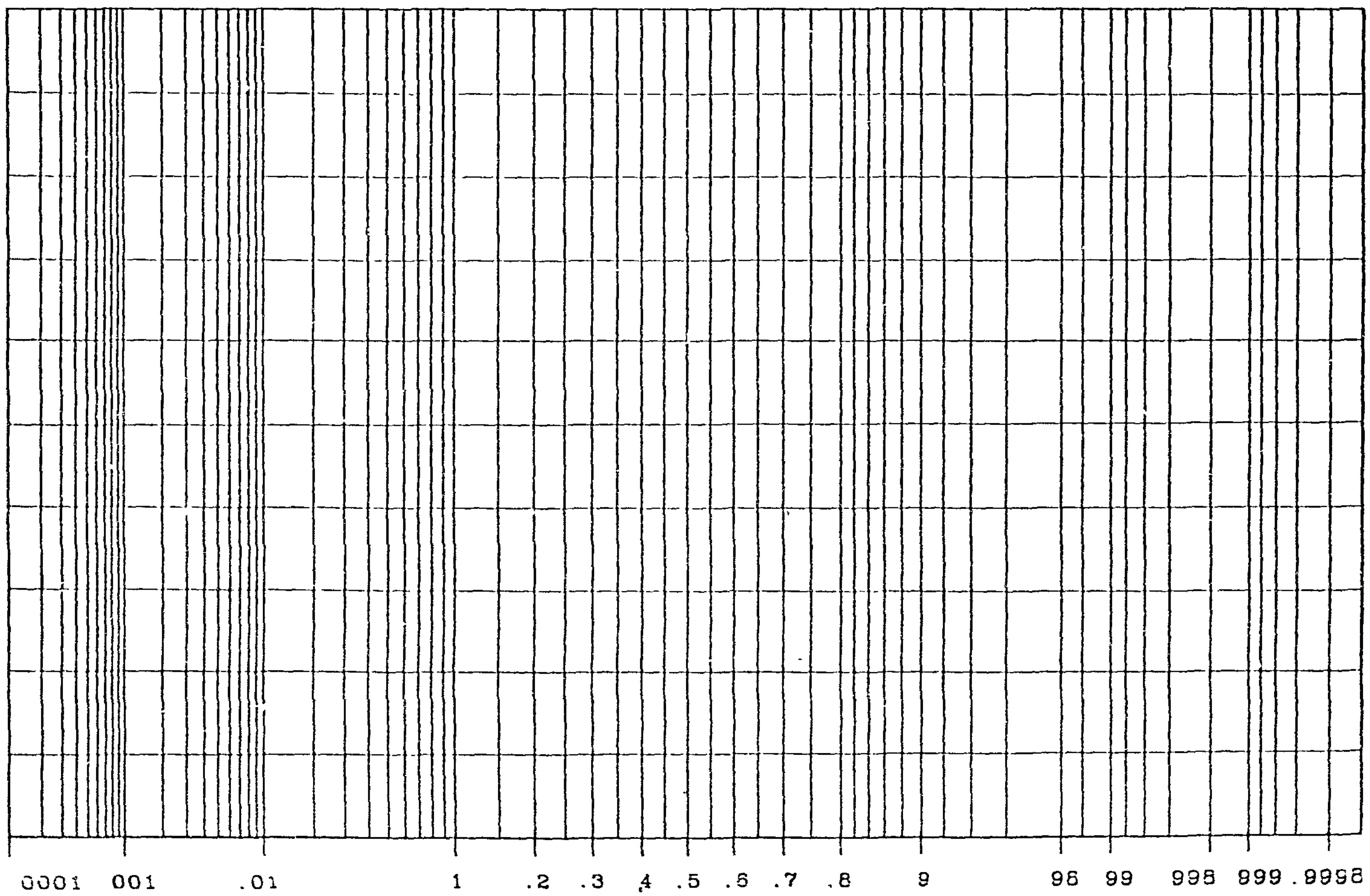
Yevjevich, V., 1972. Probability and Statistics in Hydrology, Water Resour. Publ., Fort Collins, Colo., pp. 89-92.

Zhang, Y., 1982. Plotting Positions of Annual Flood Extremes Considering Extraordinary Values, Water Resources Research, Vol. 18, No. 4, pp. 859-864.

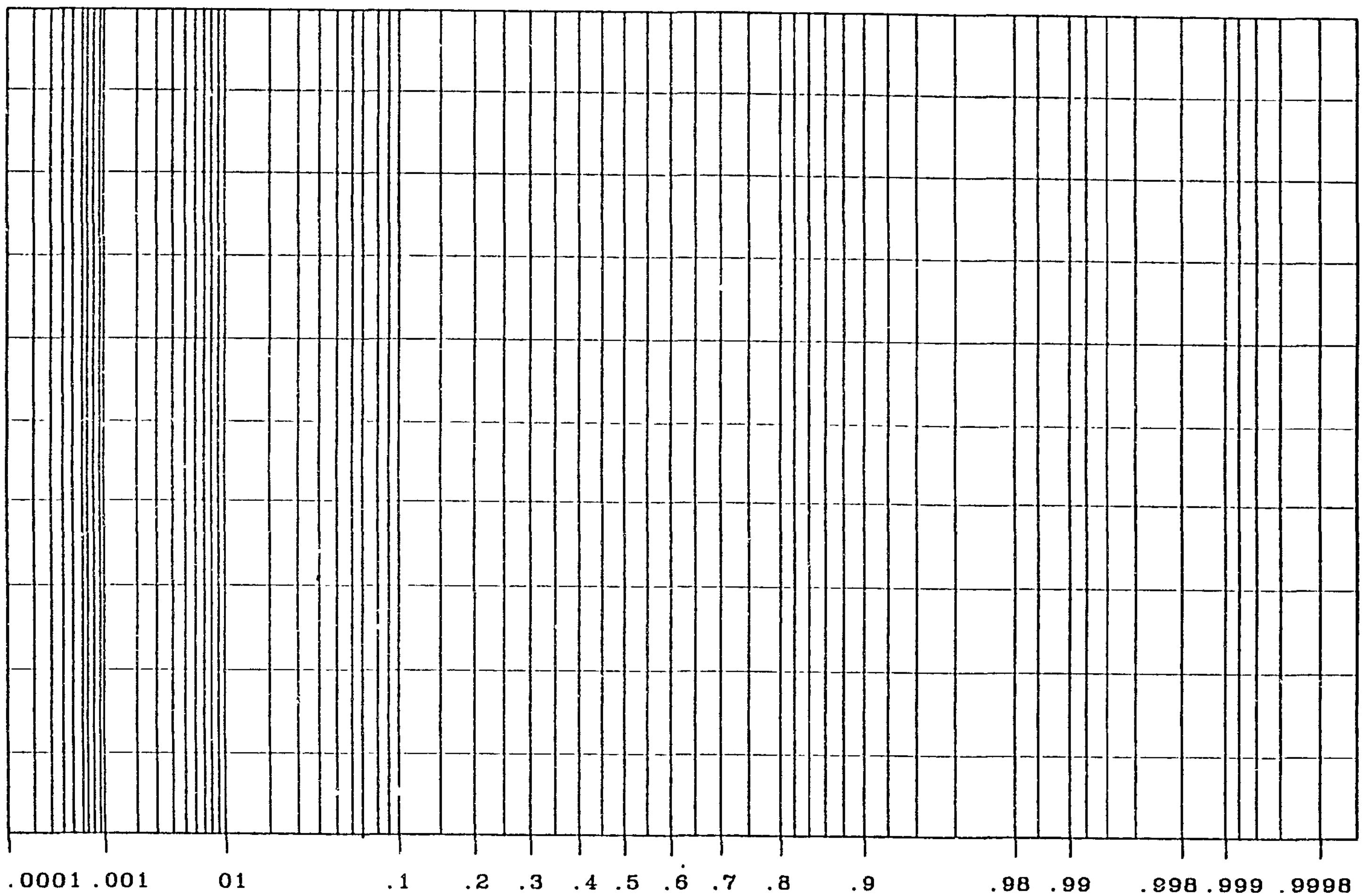
APPENDIX A

Probability Papers for Pearson Type III Distribution.

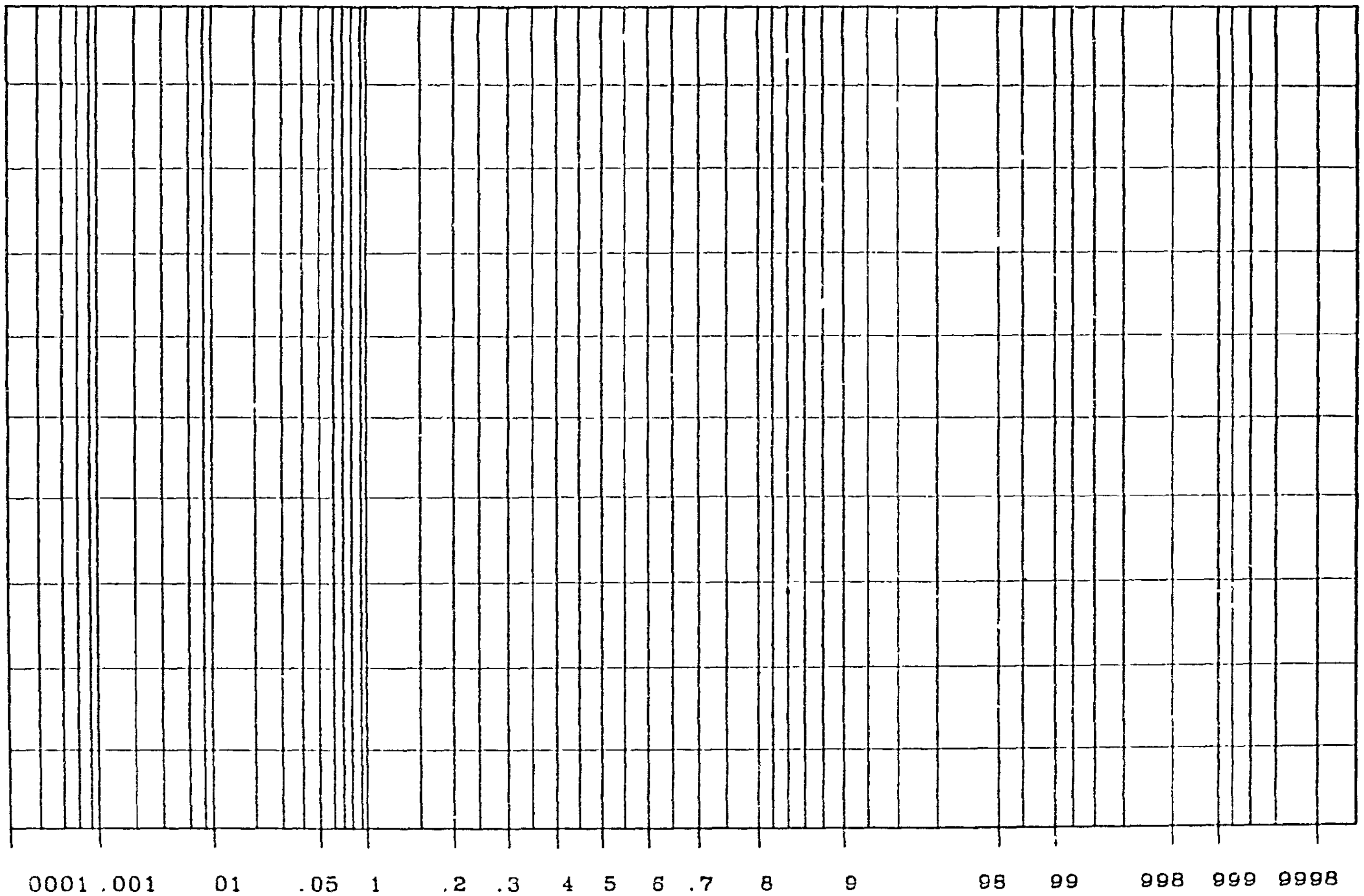
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.0



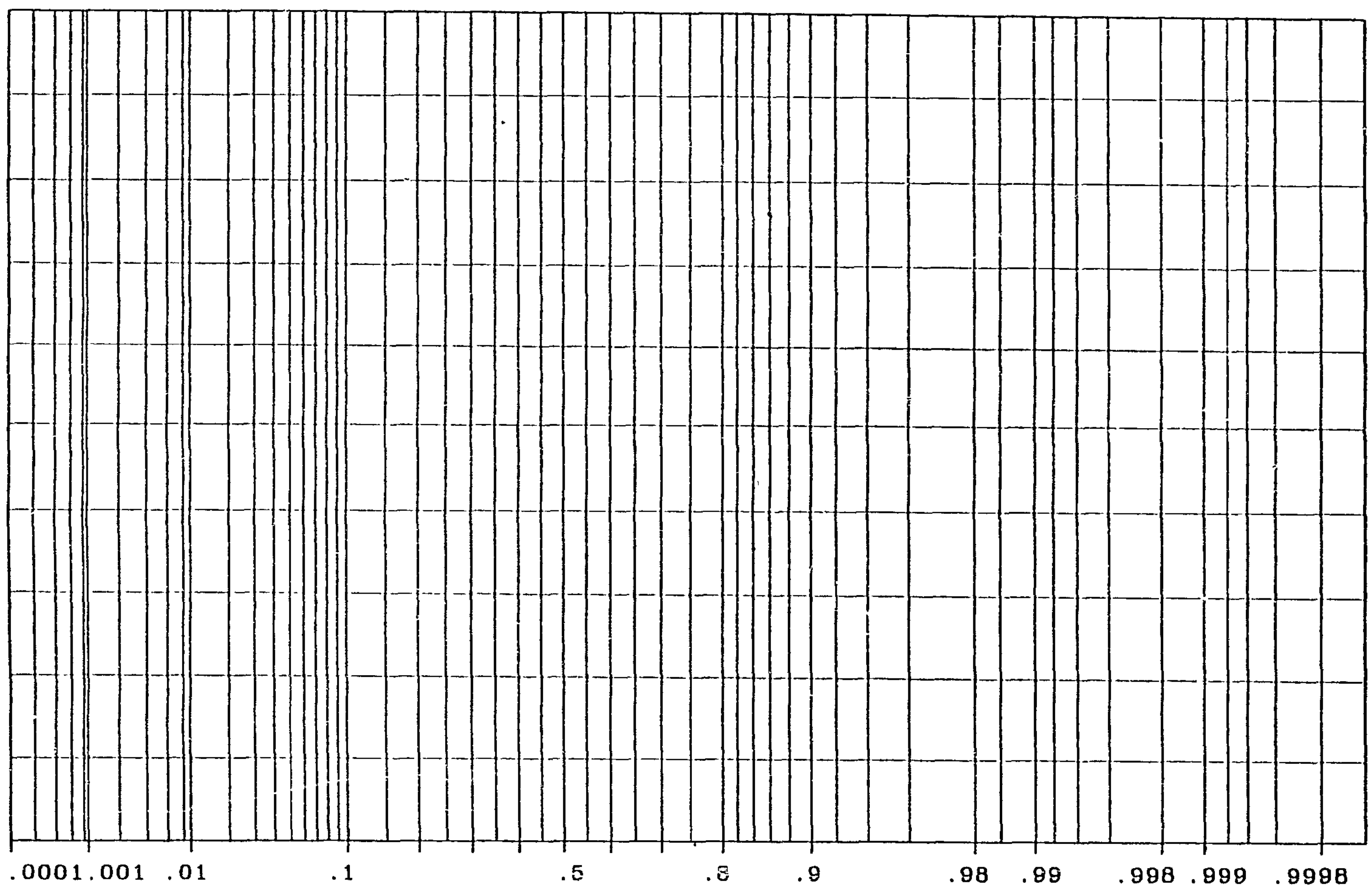
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.1



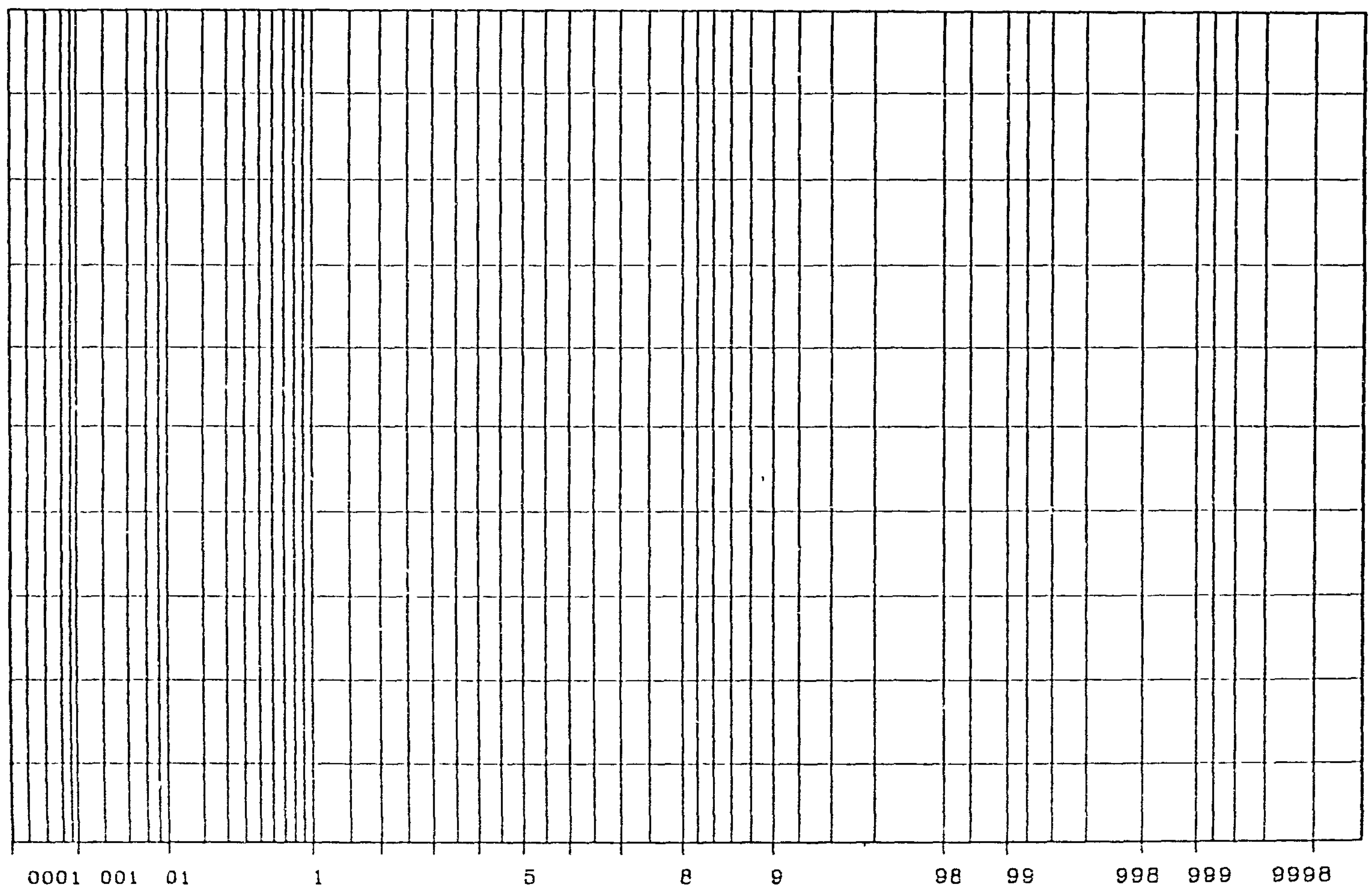
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.2



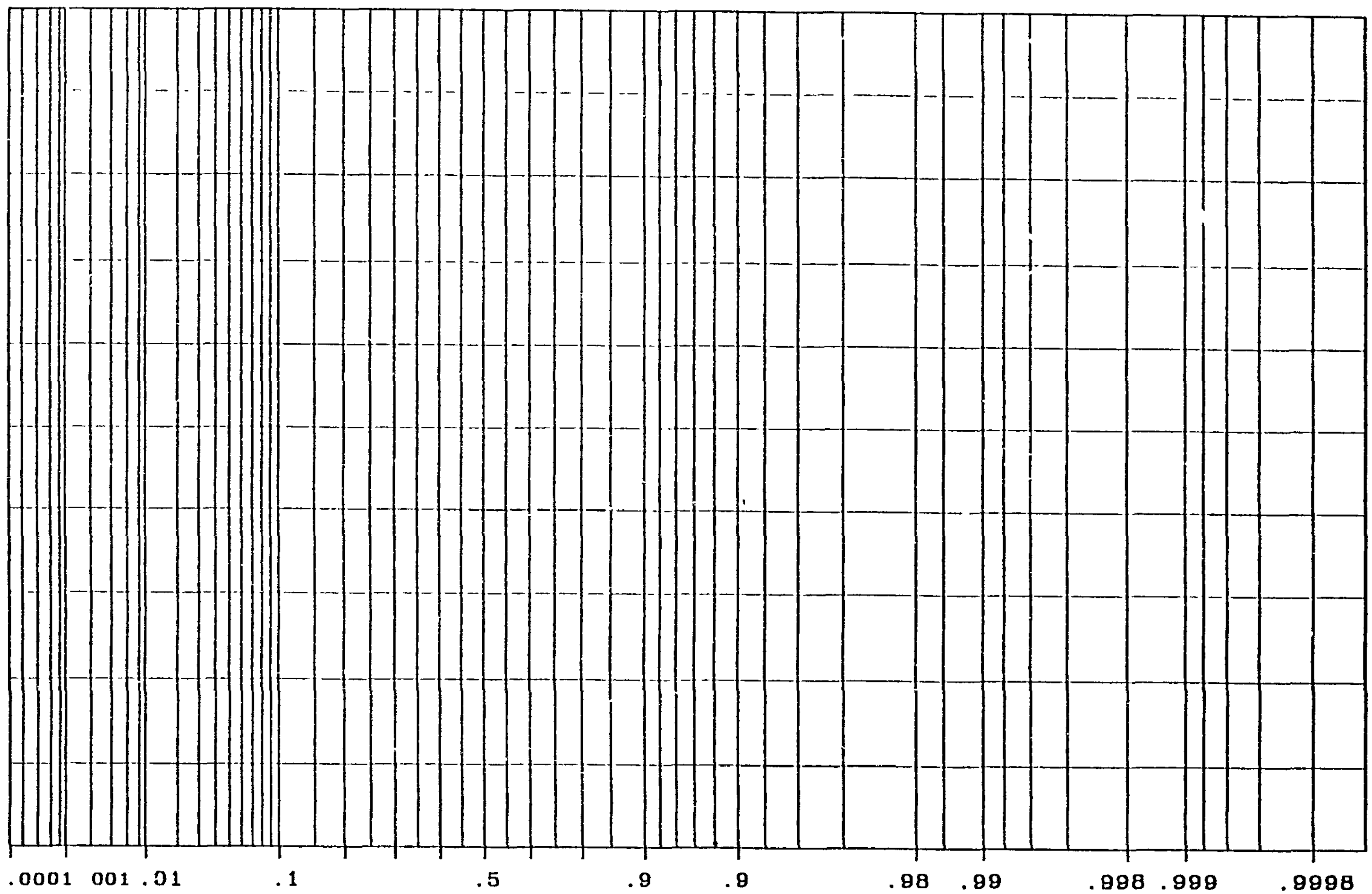
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.3



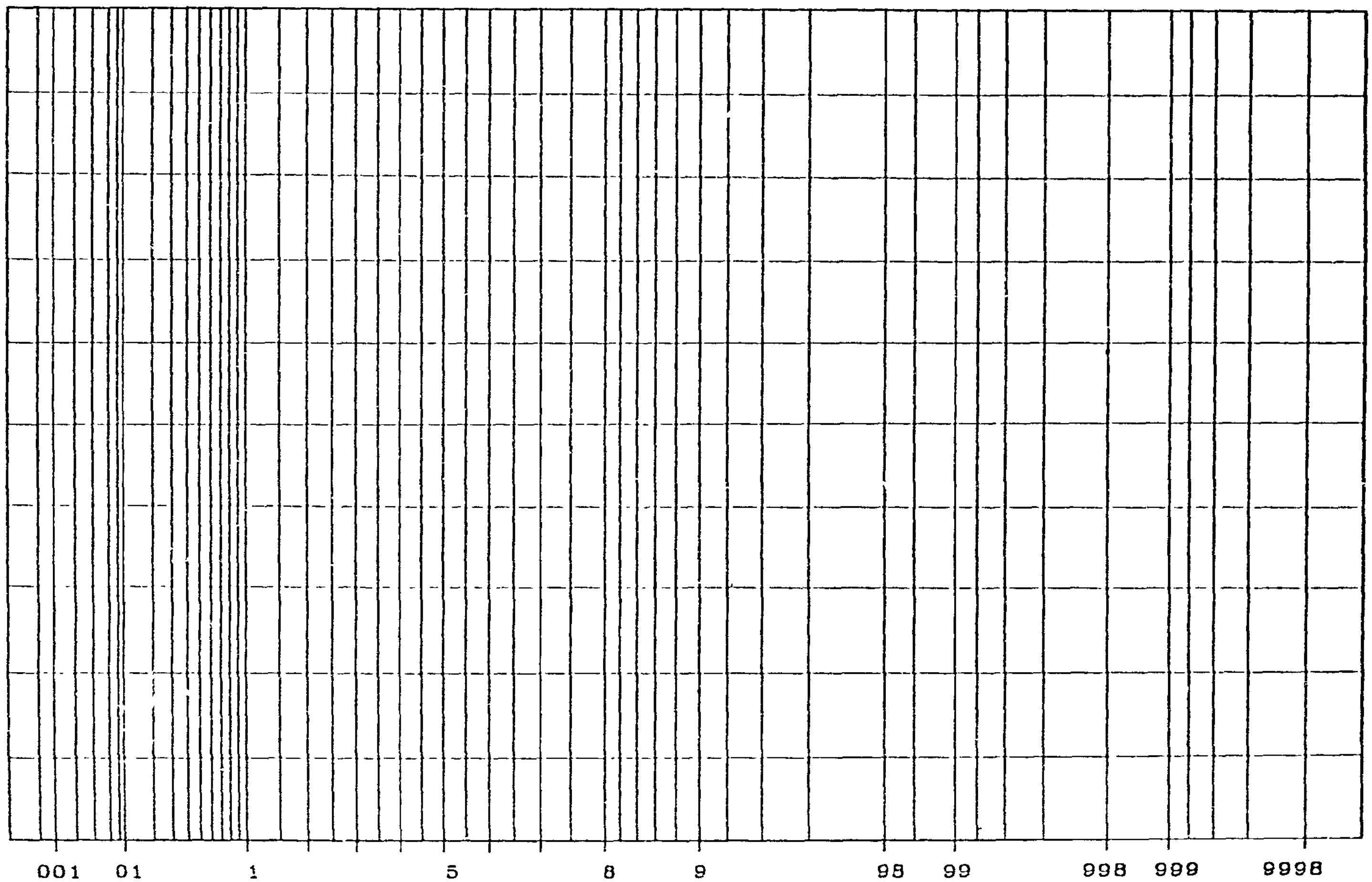
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.4



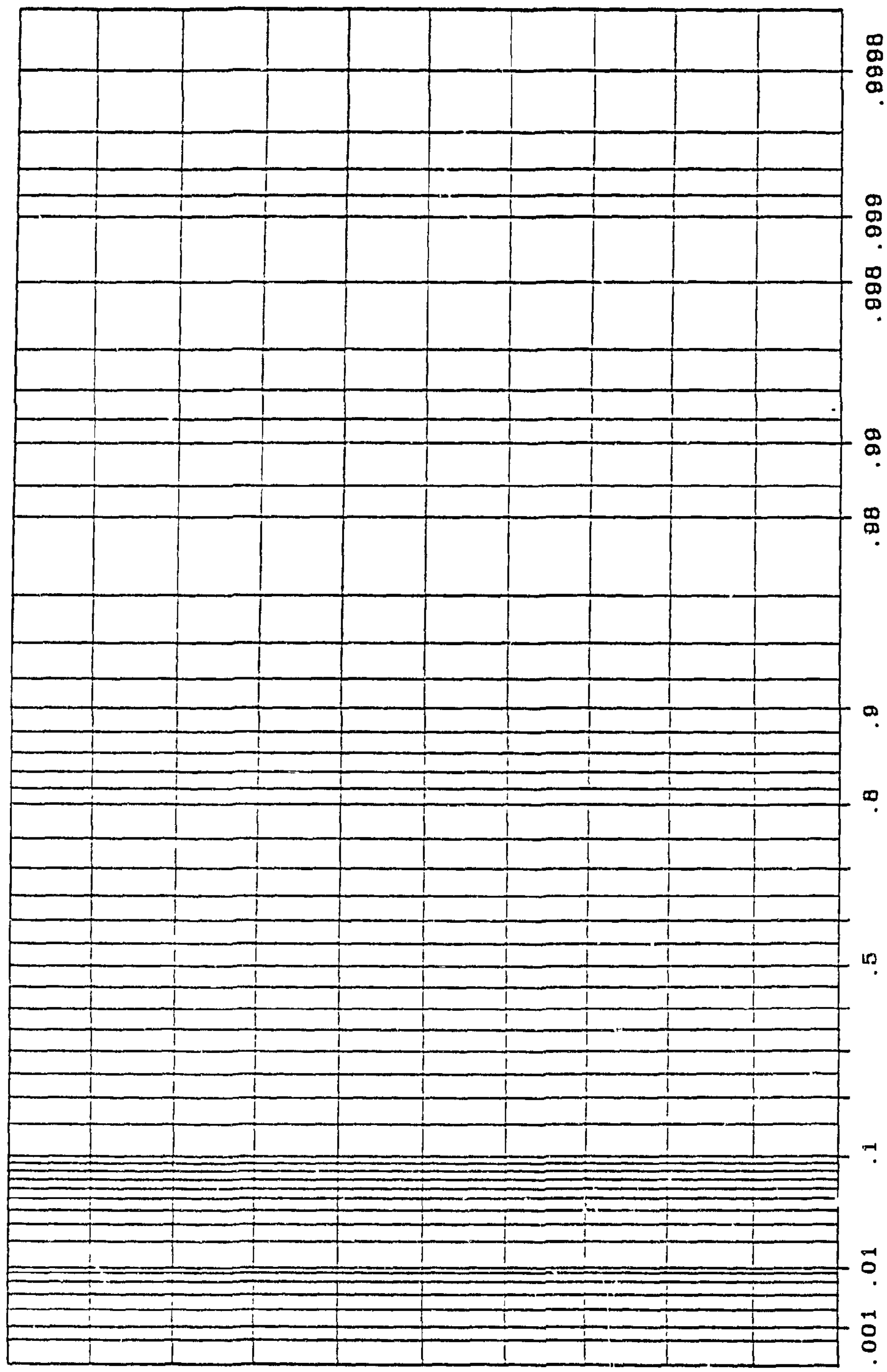
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.5



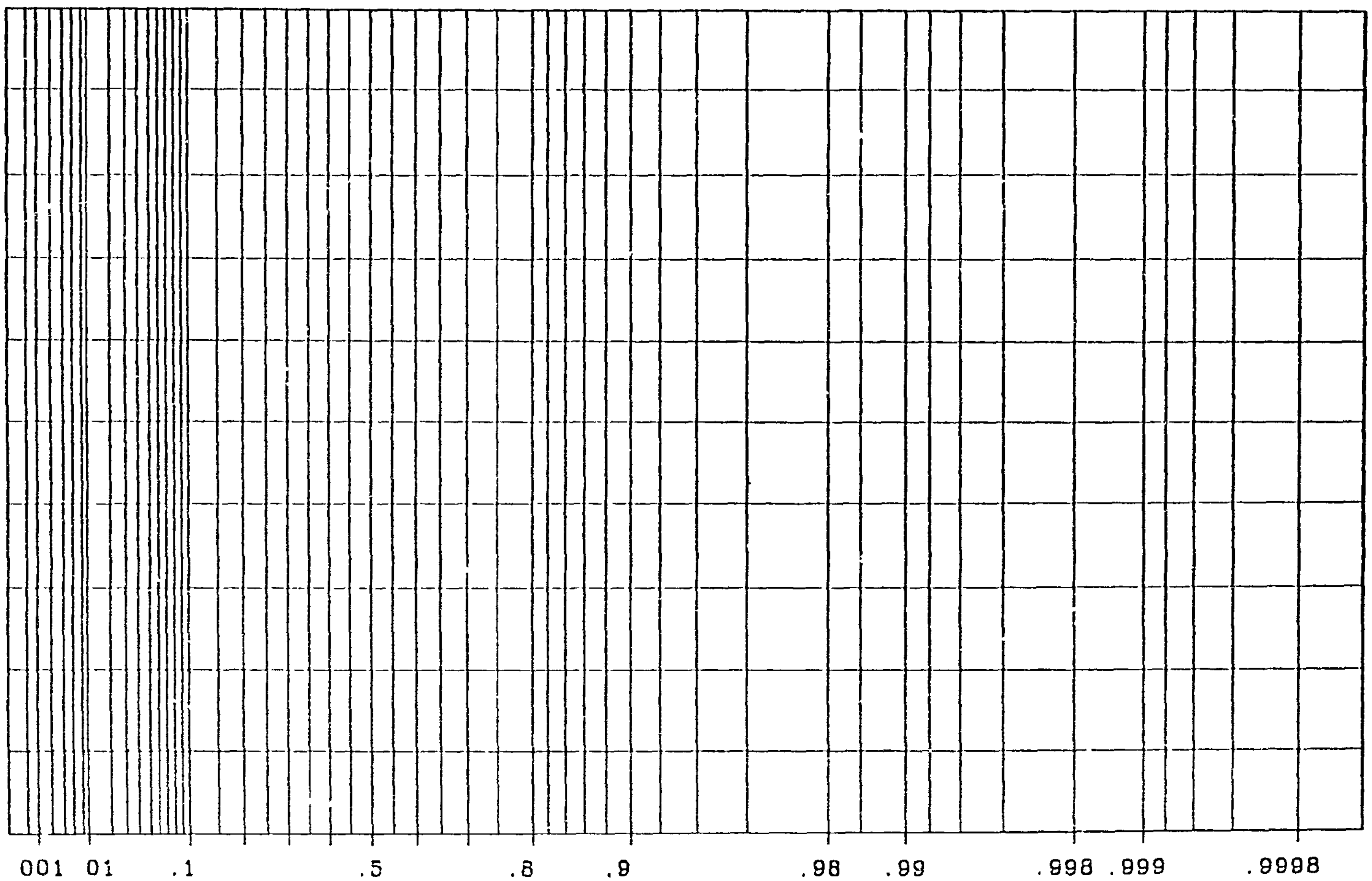
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.6



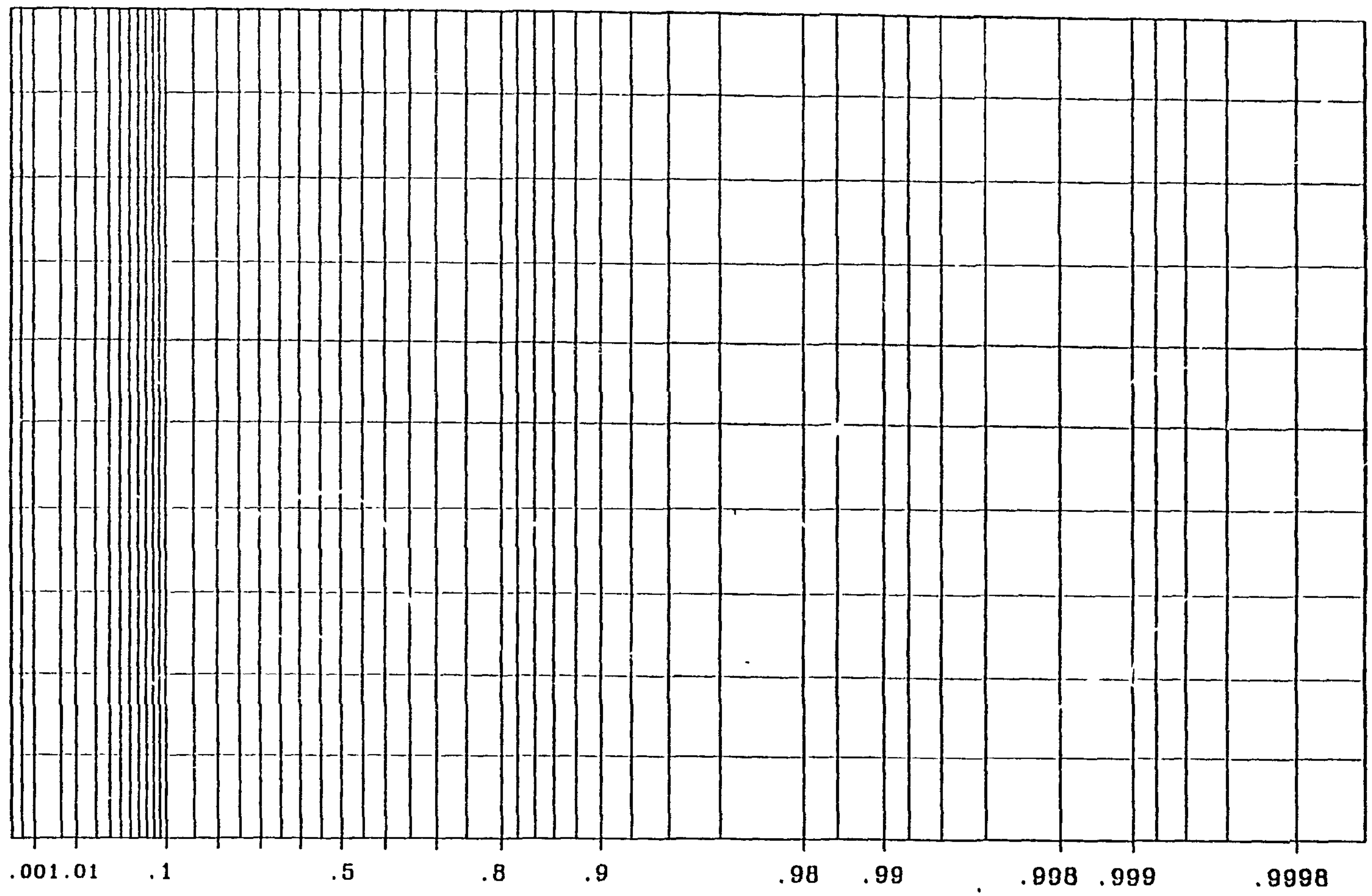
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.7



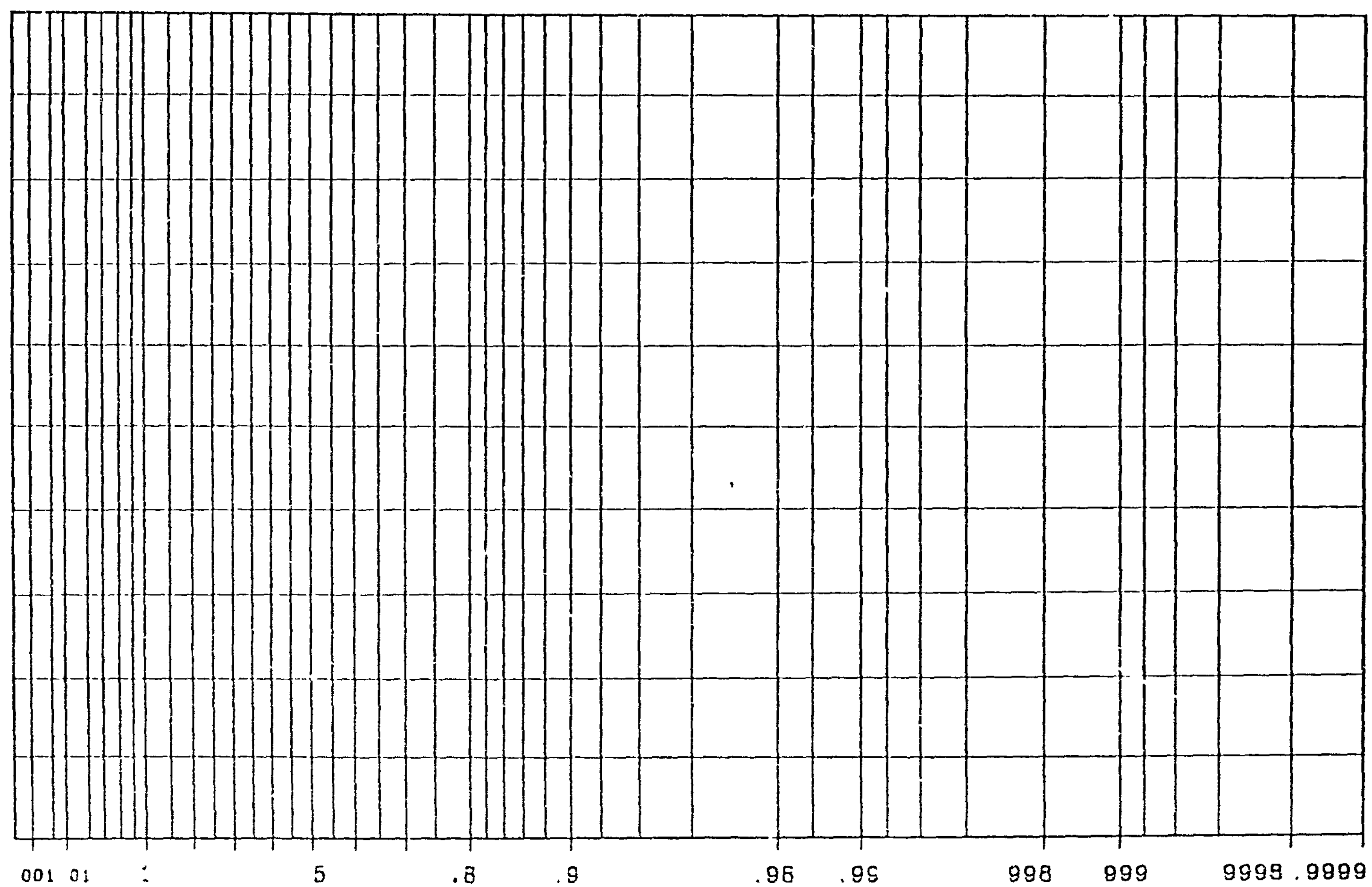
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.8



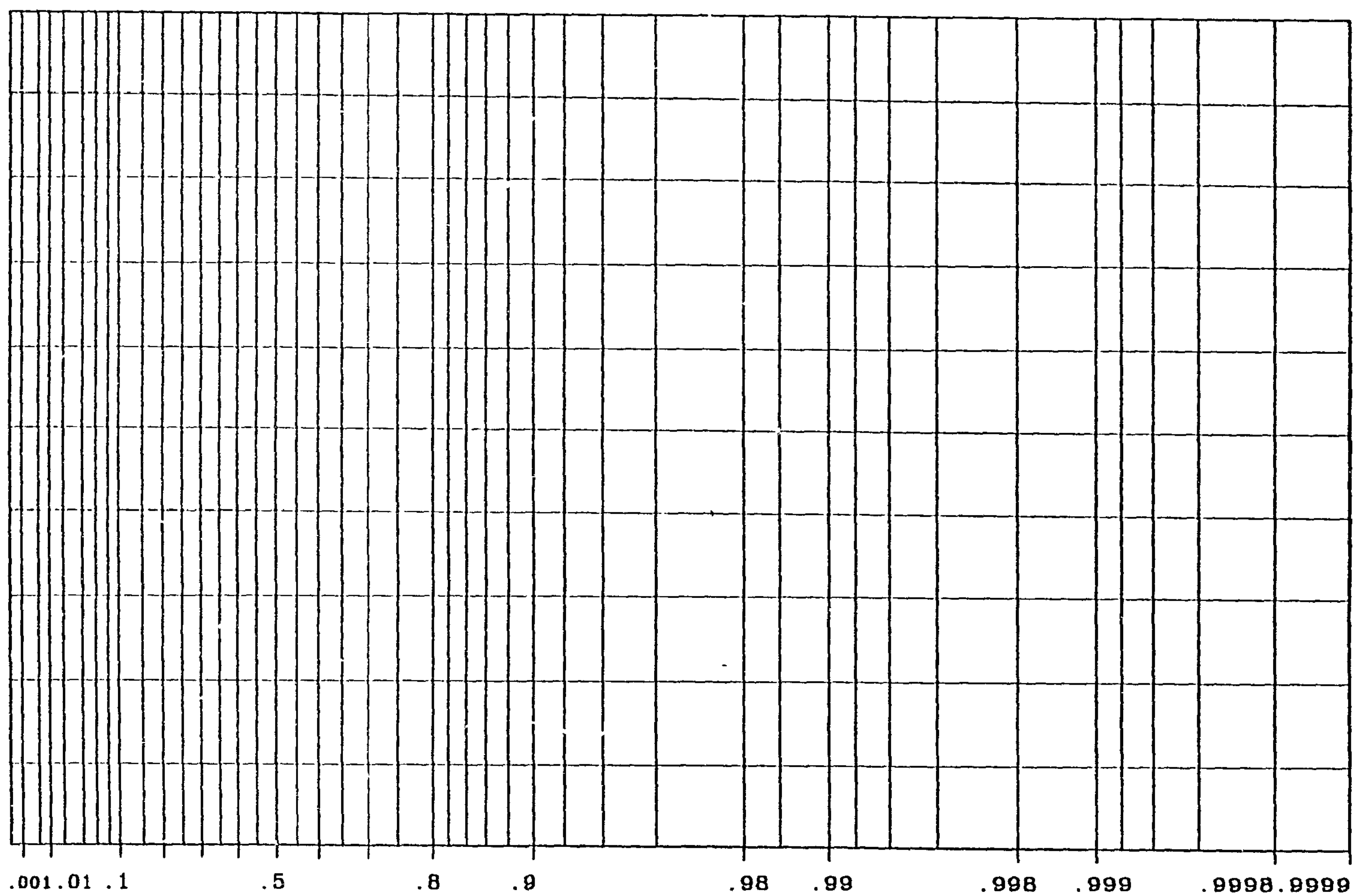
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.9



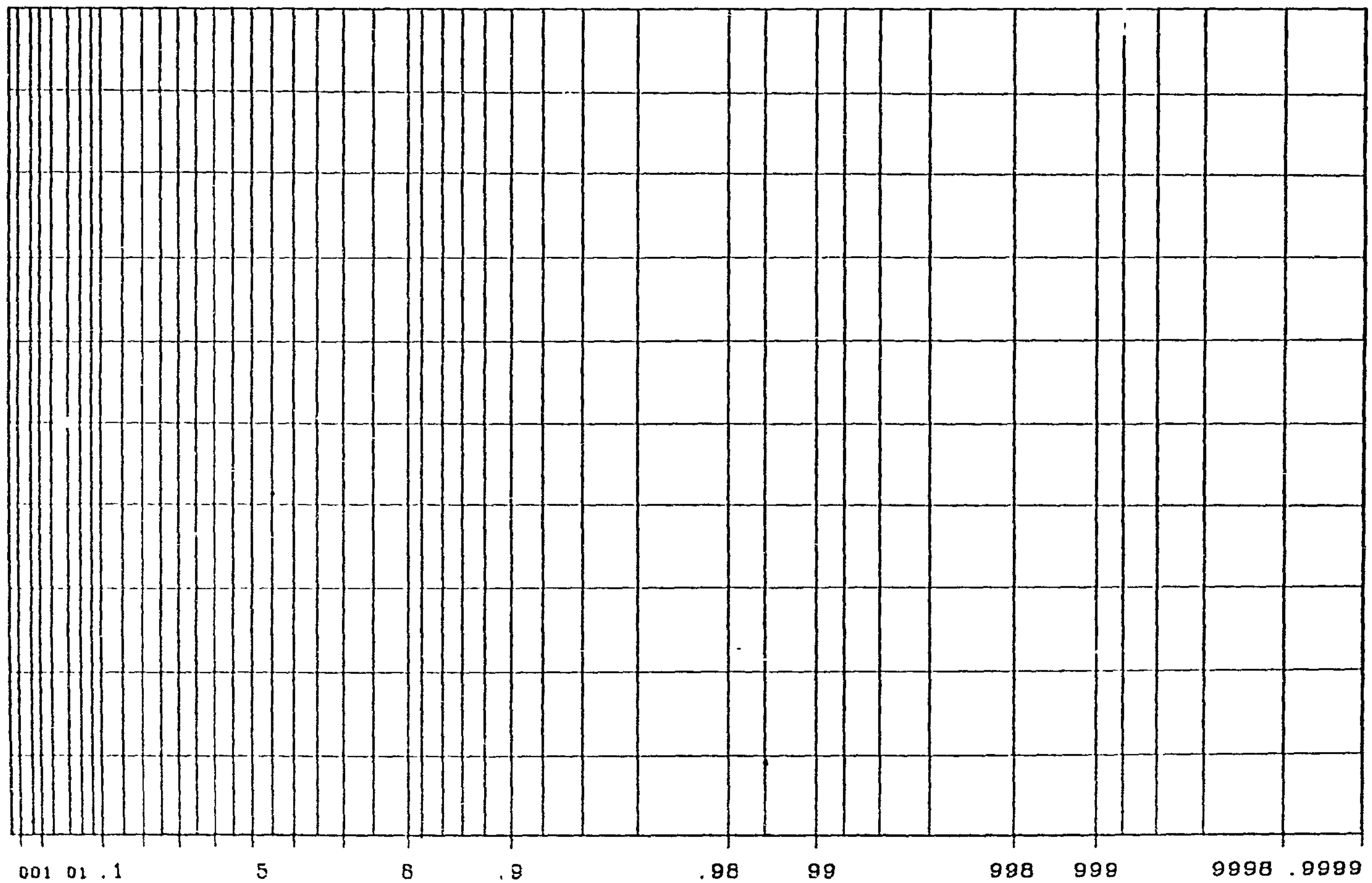
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.0



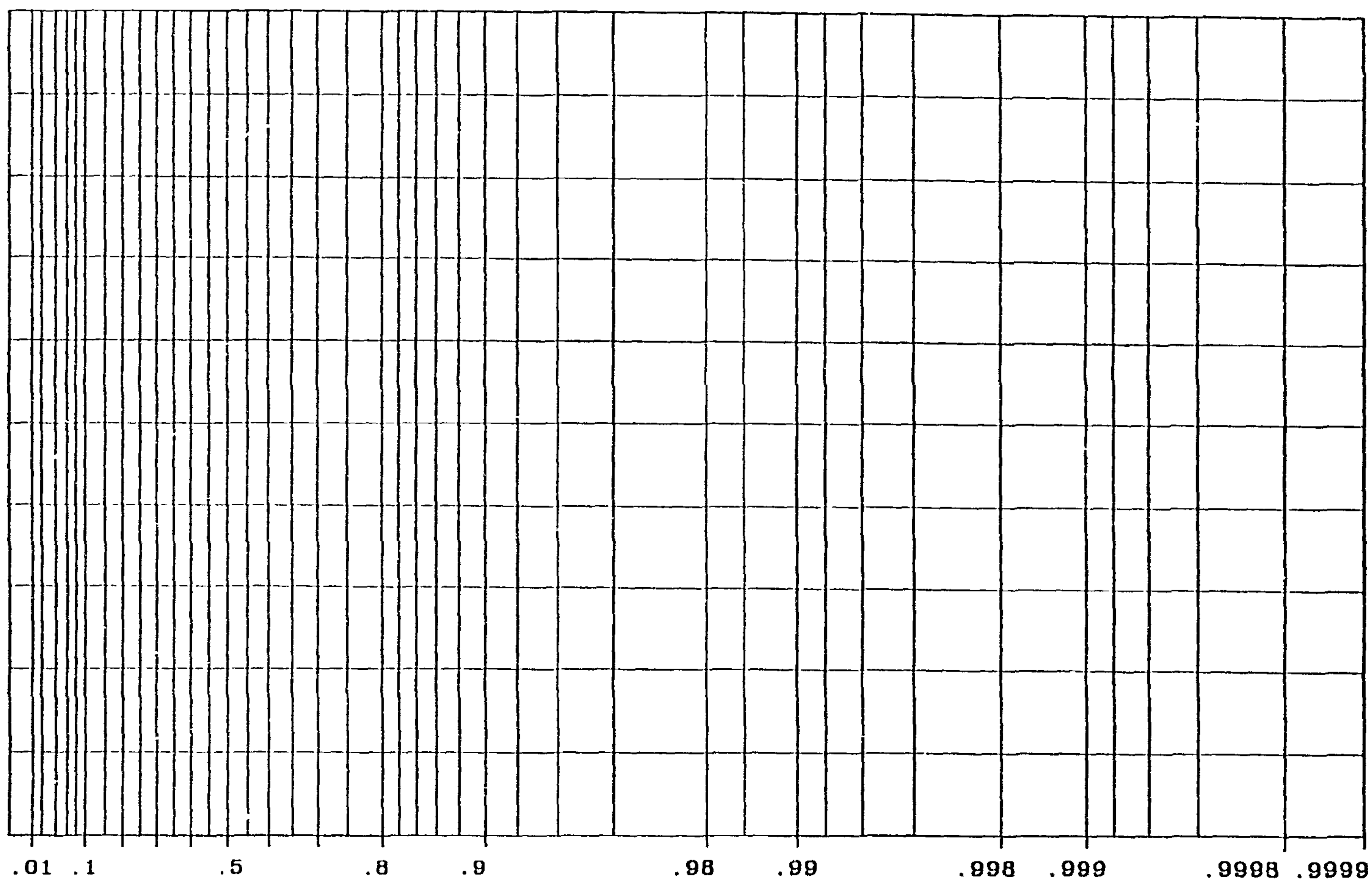
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.1



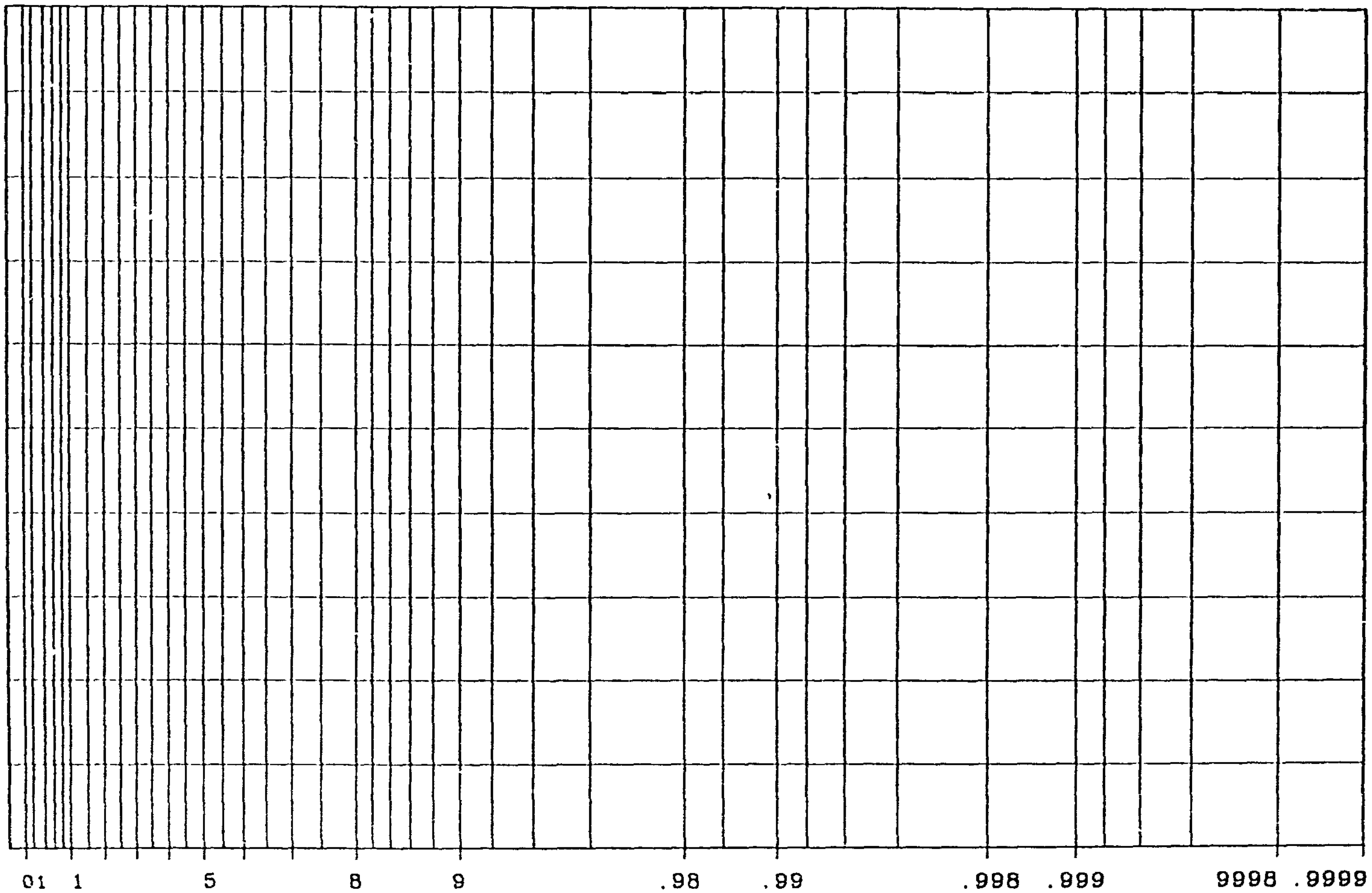
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.2



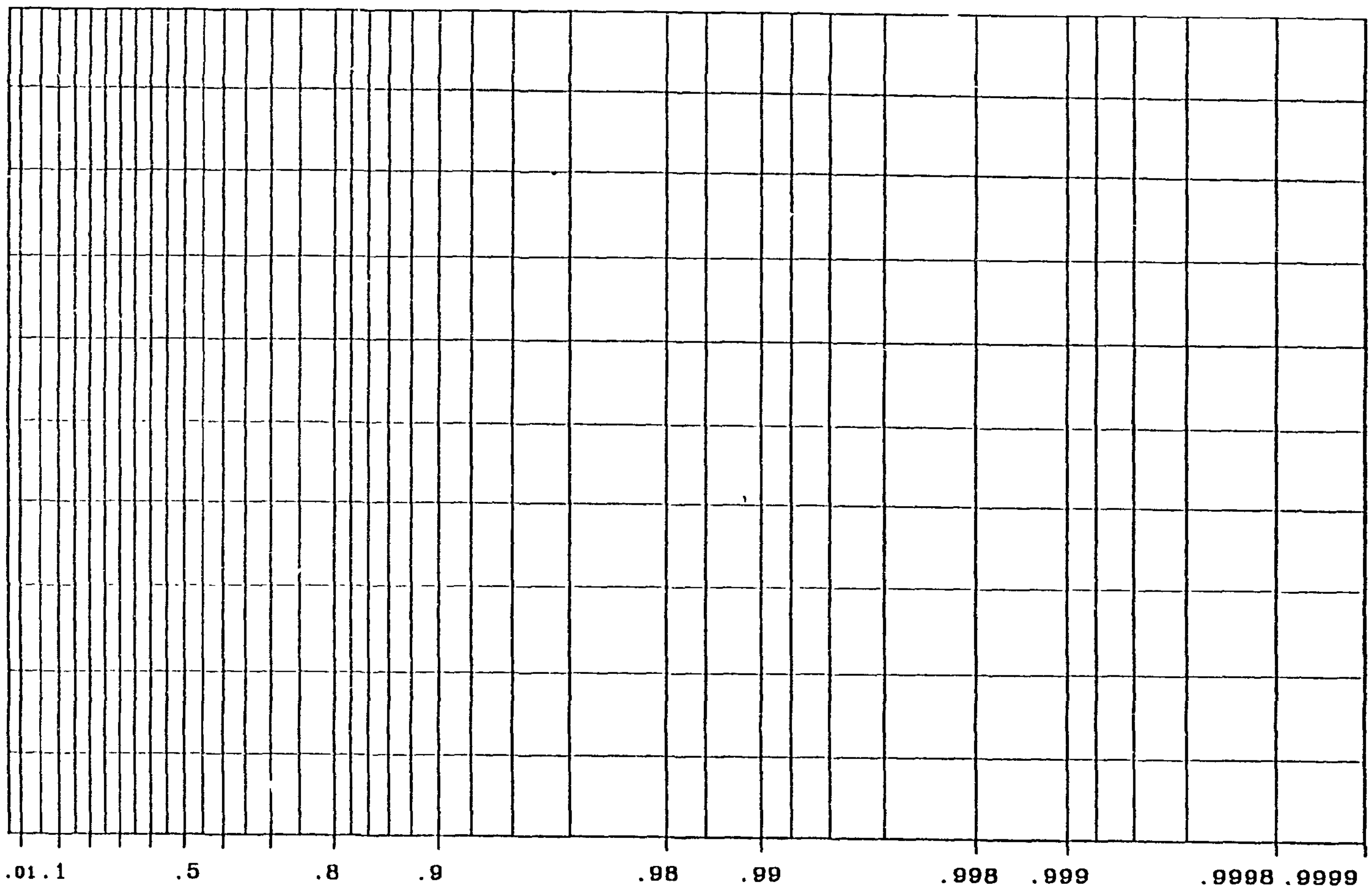
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.3



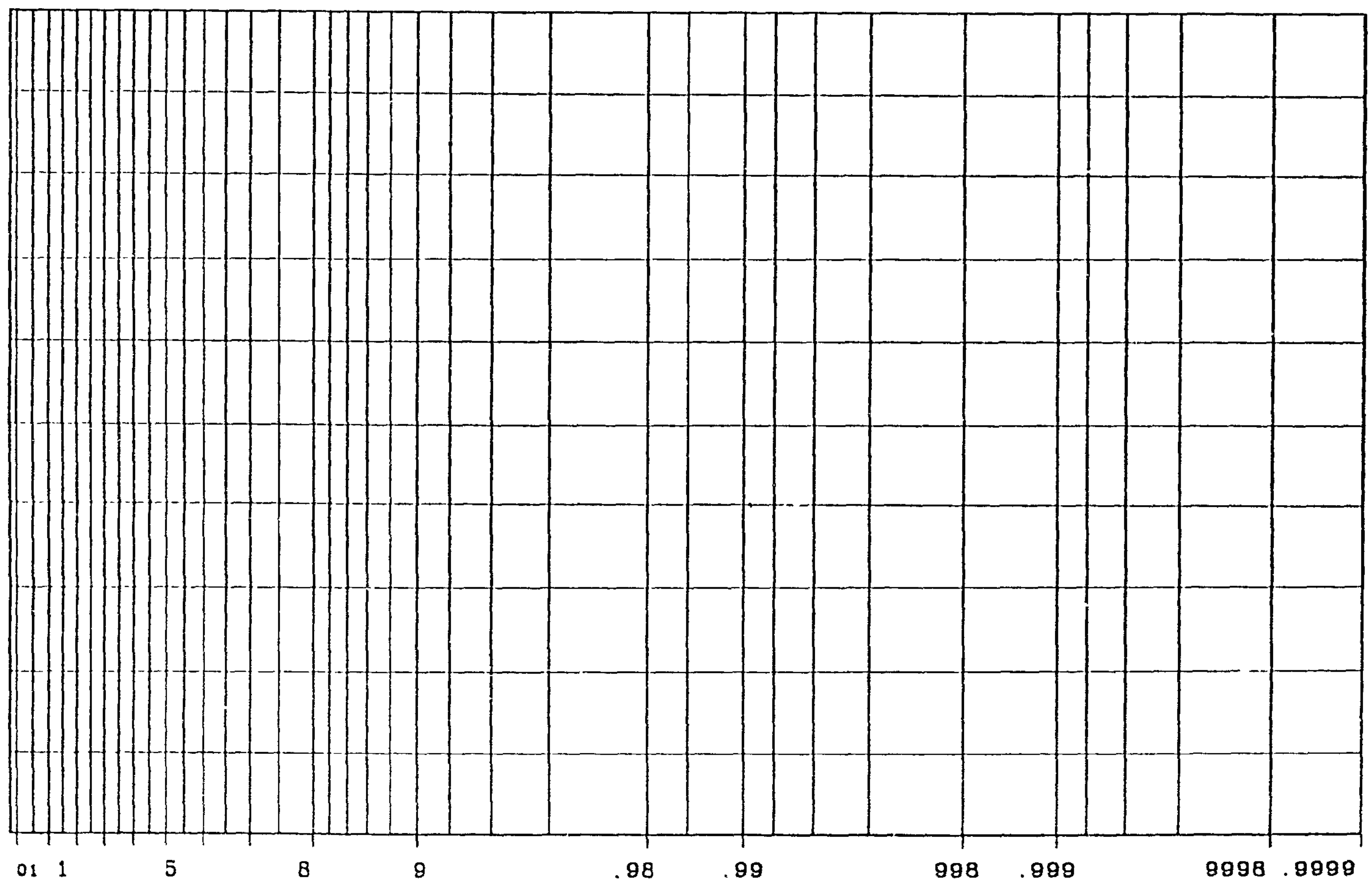
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.4



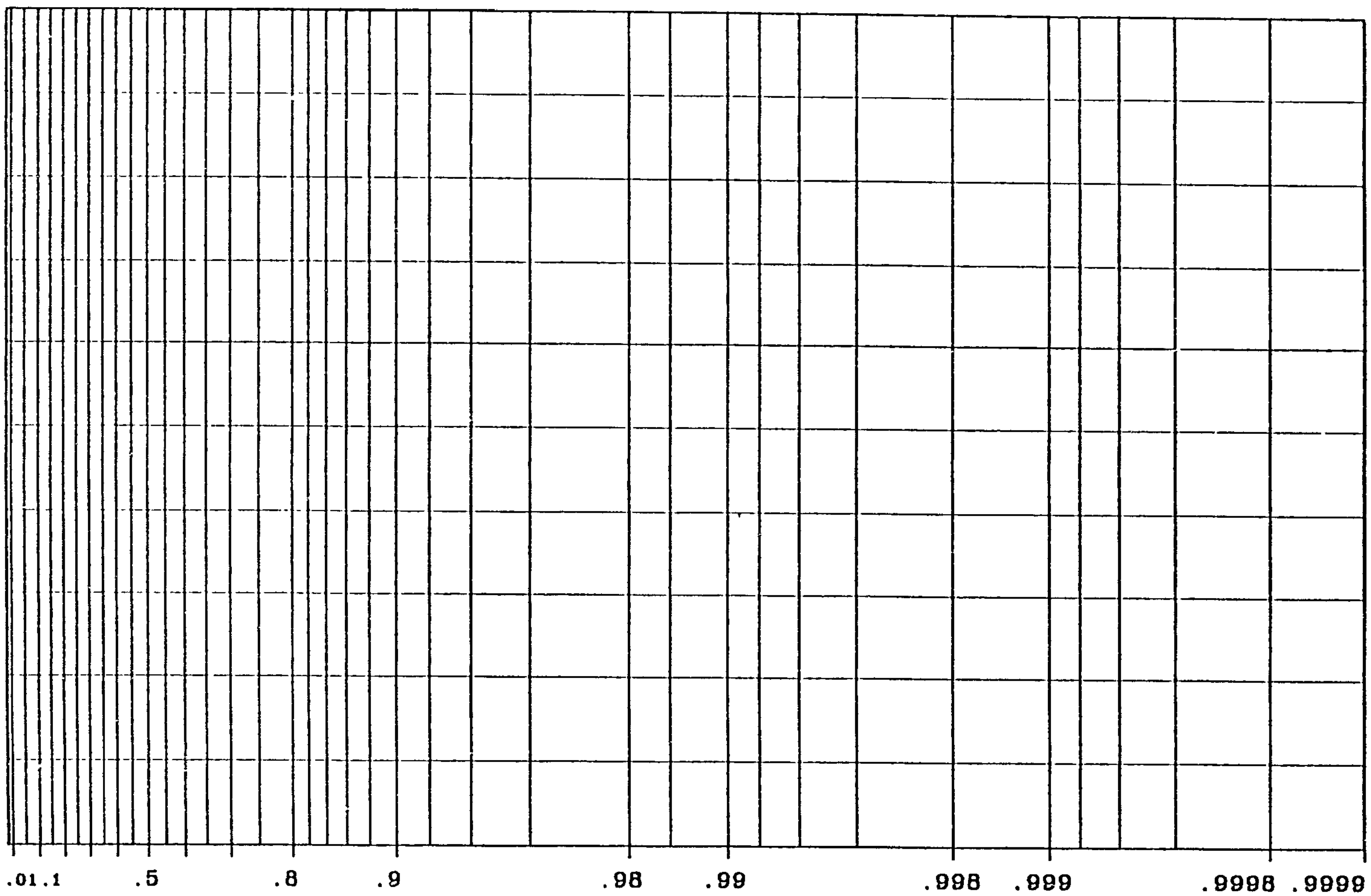
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.5



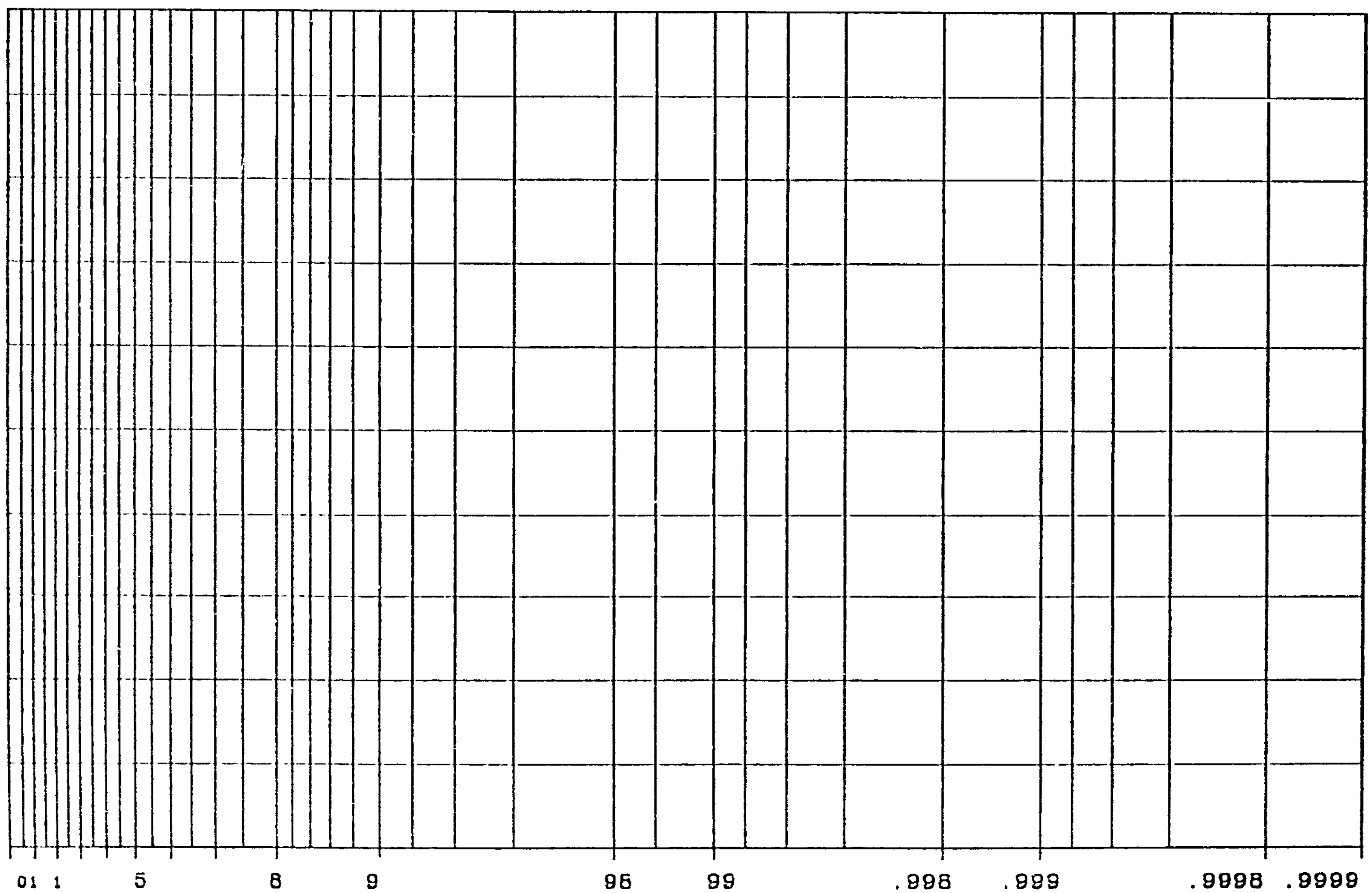
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.6



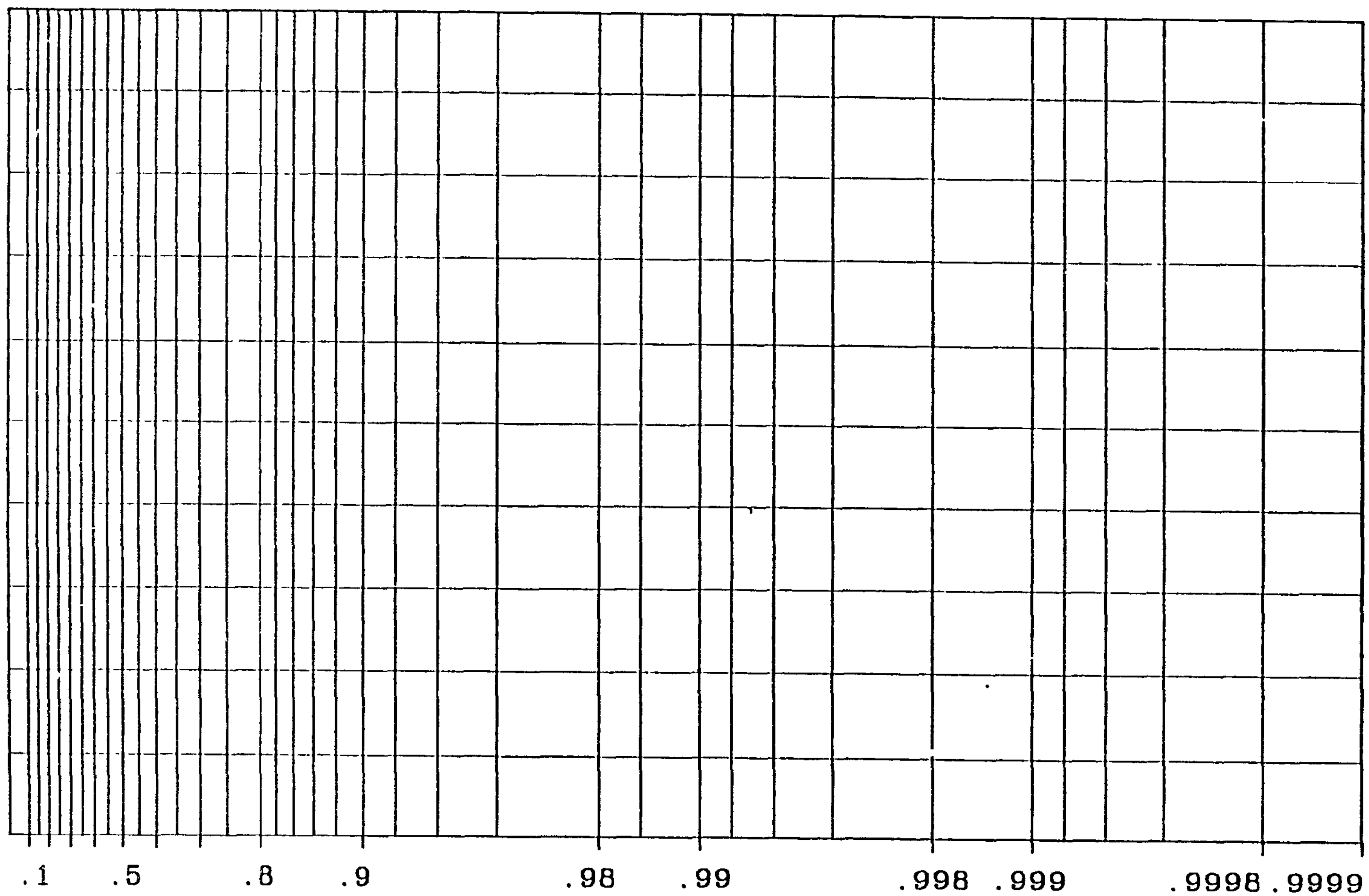
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.7



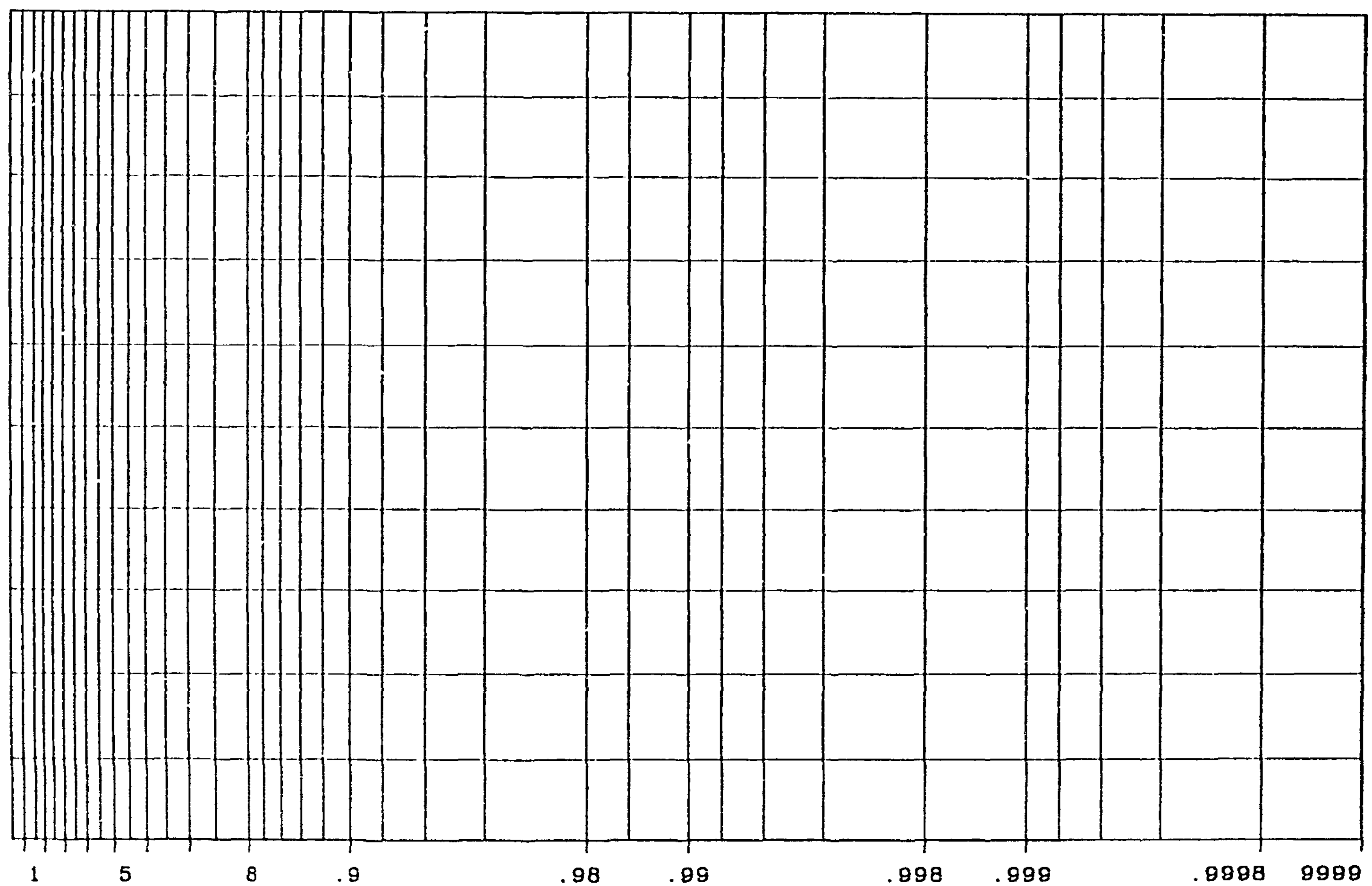
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.8



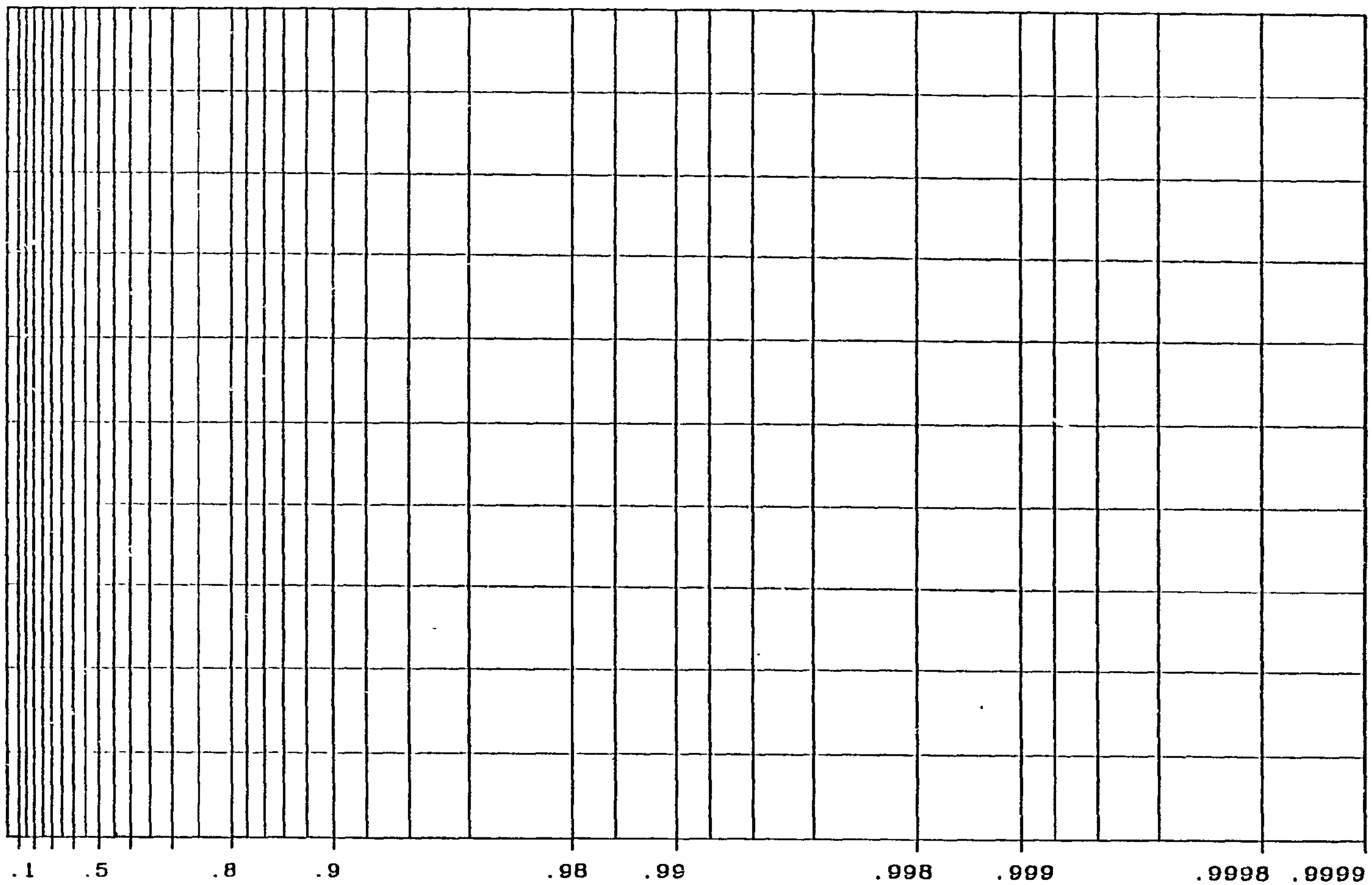
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.9



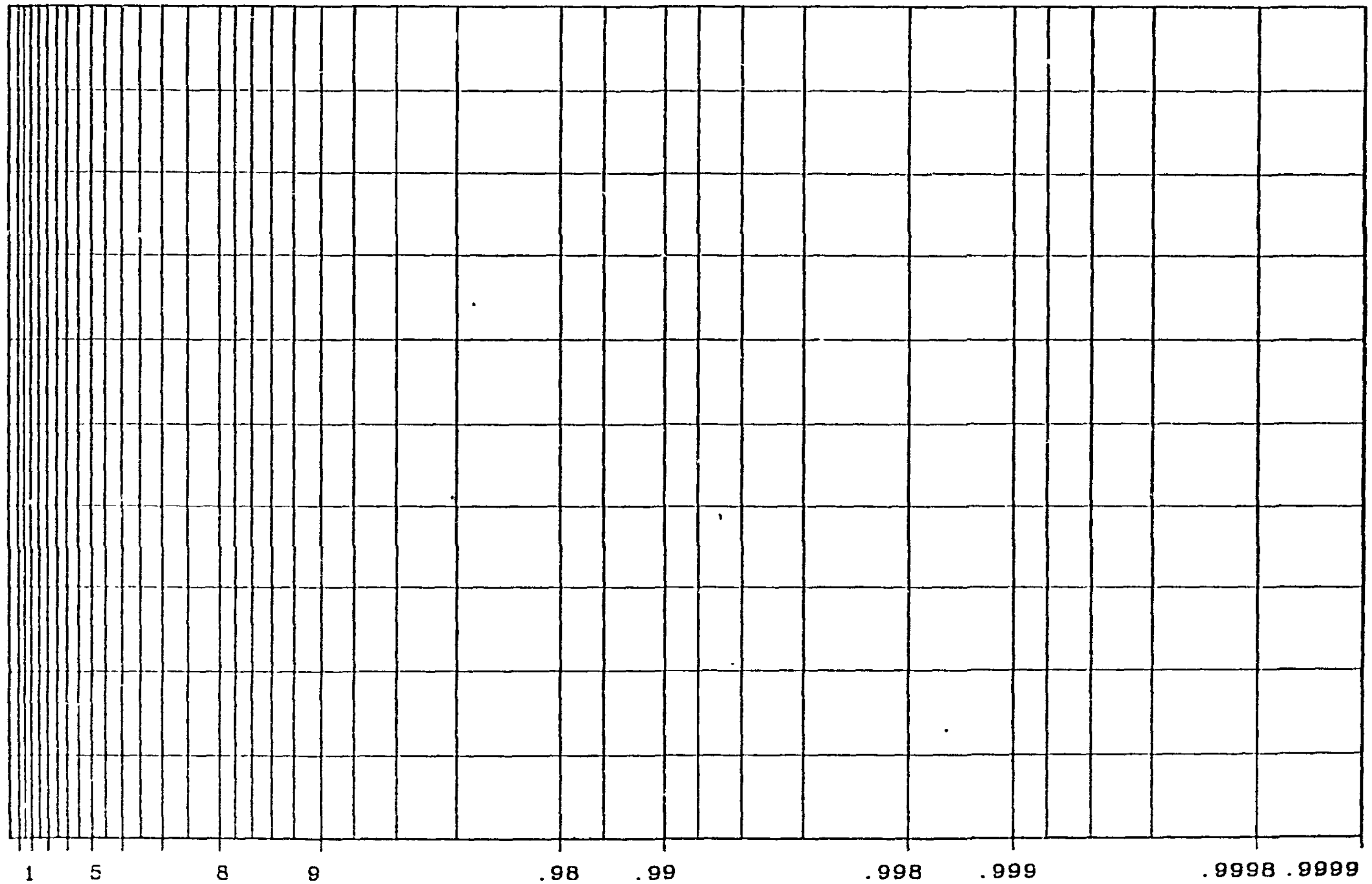
PEARSON TYPE P3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.0



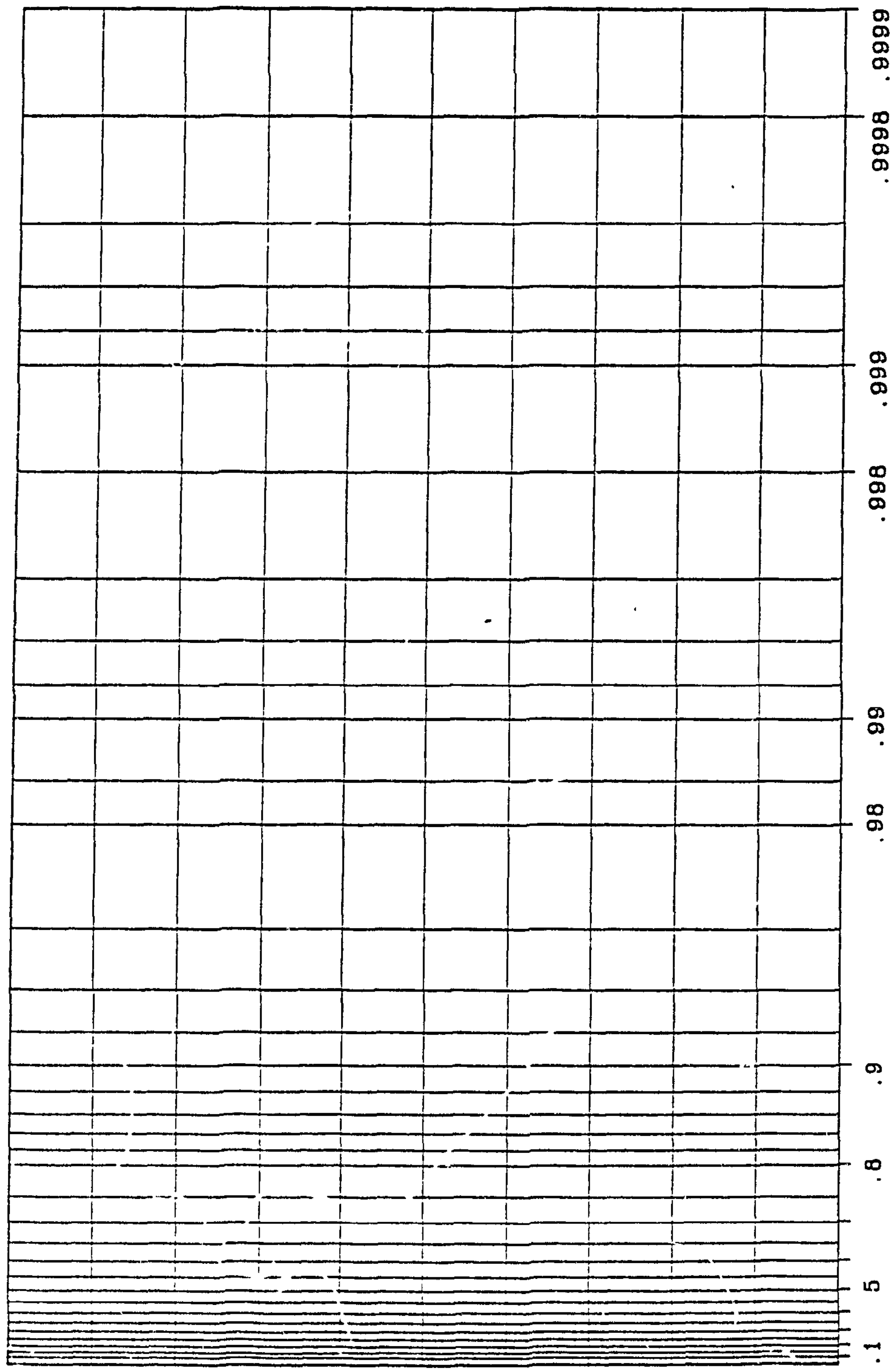
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.1



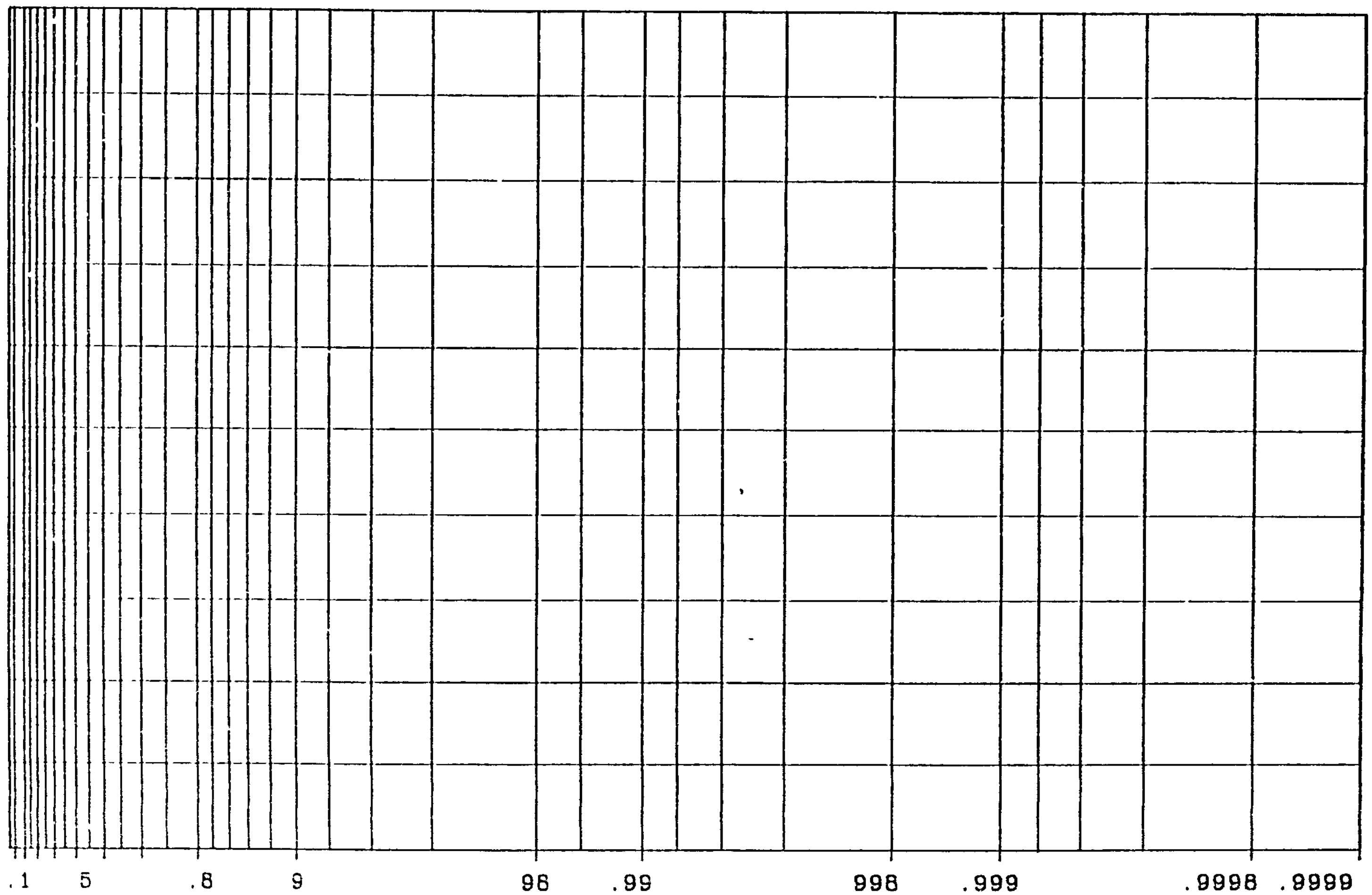
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.2



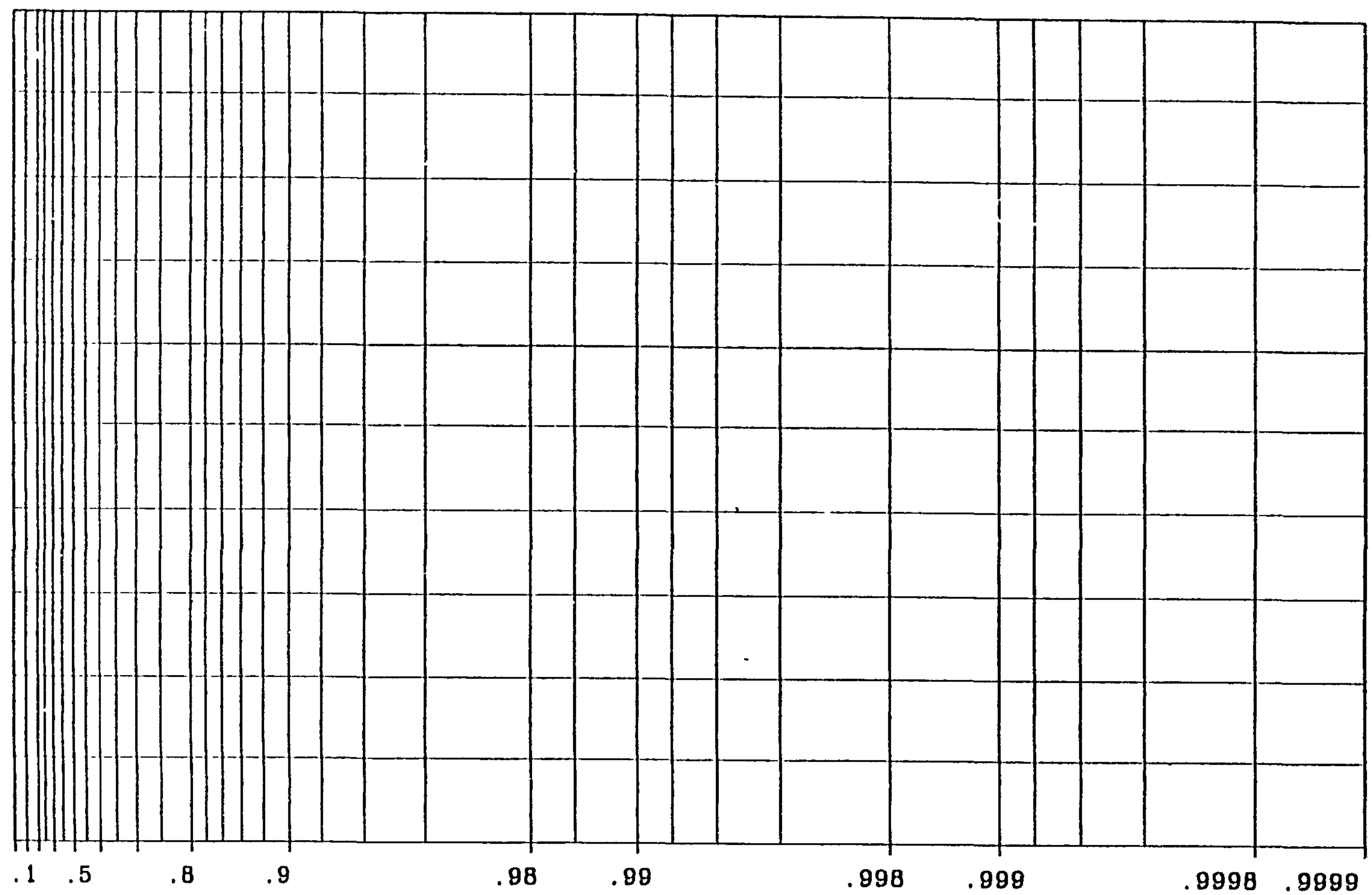
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.3



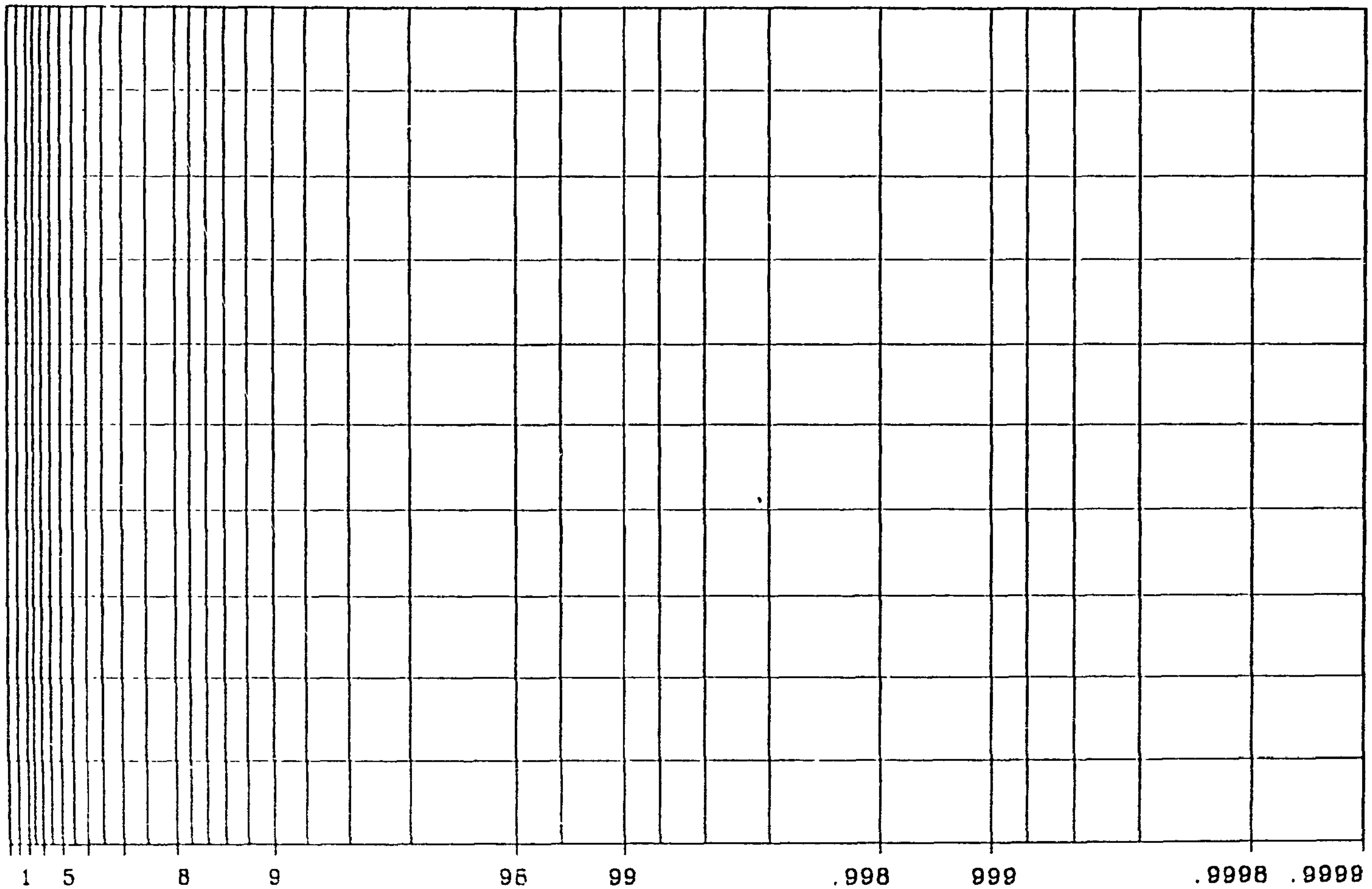
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.4



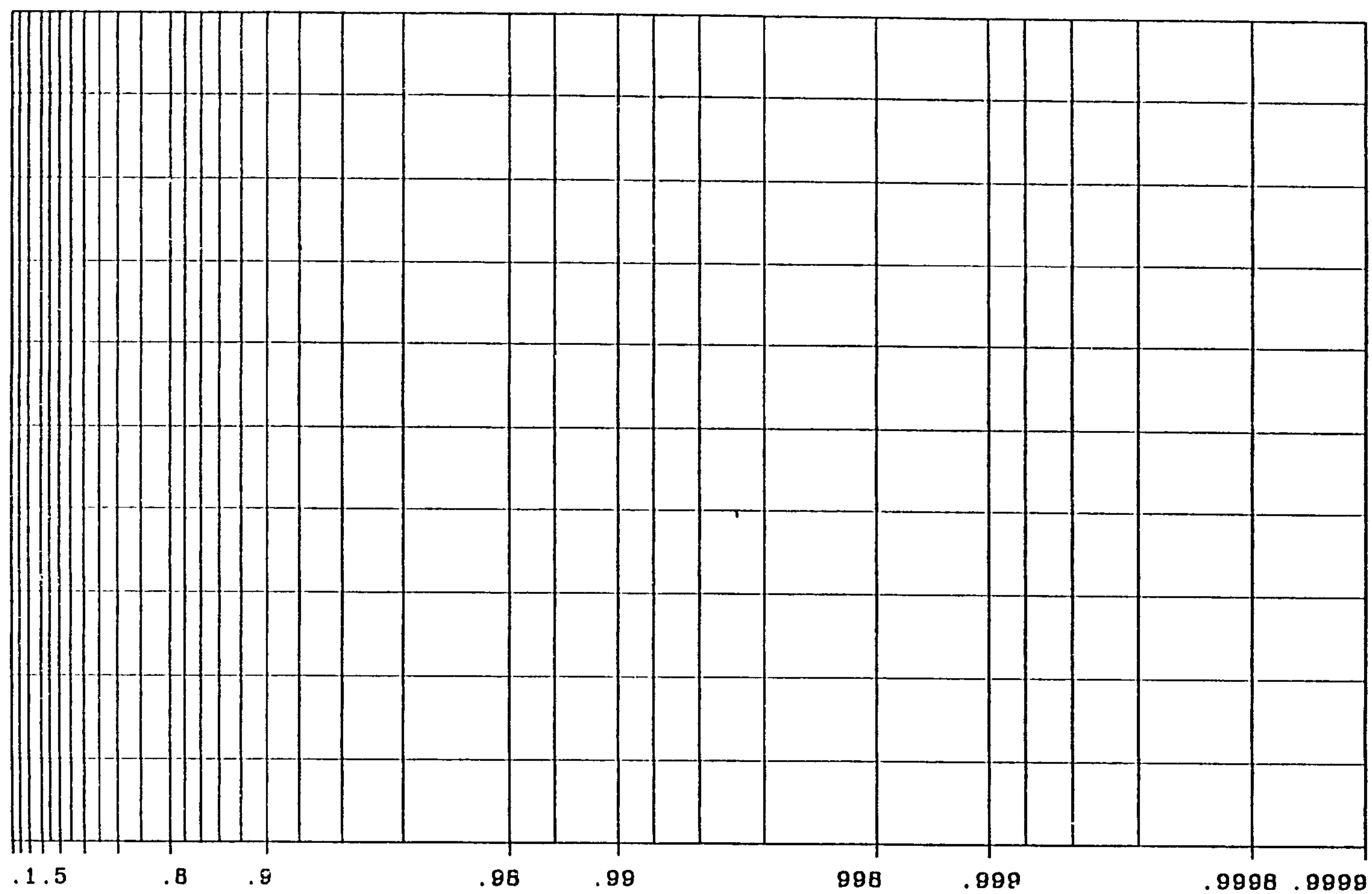
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.5



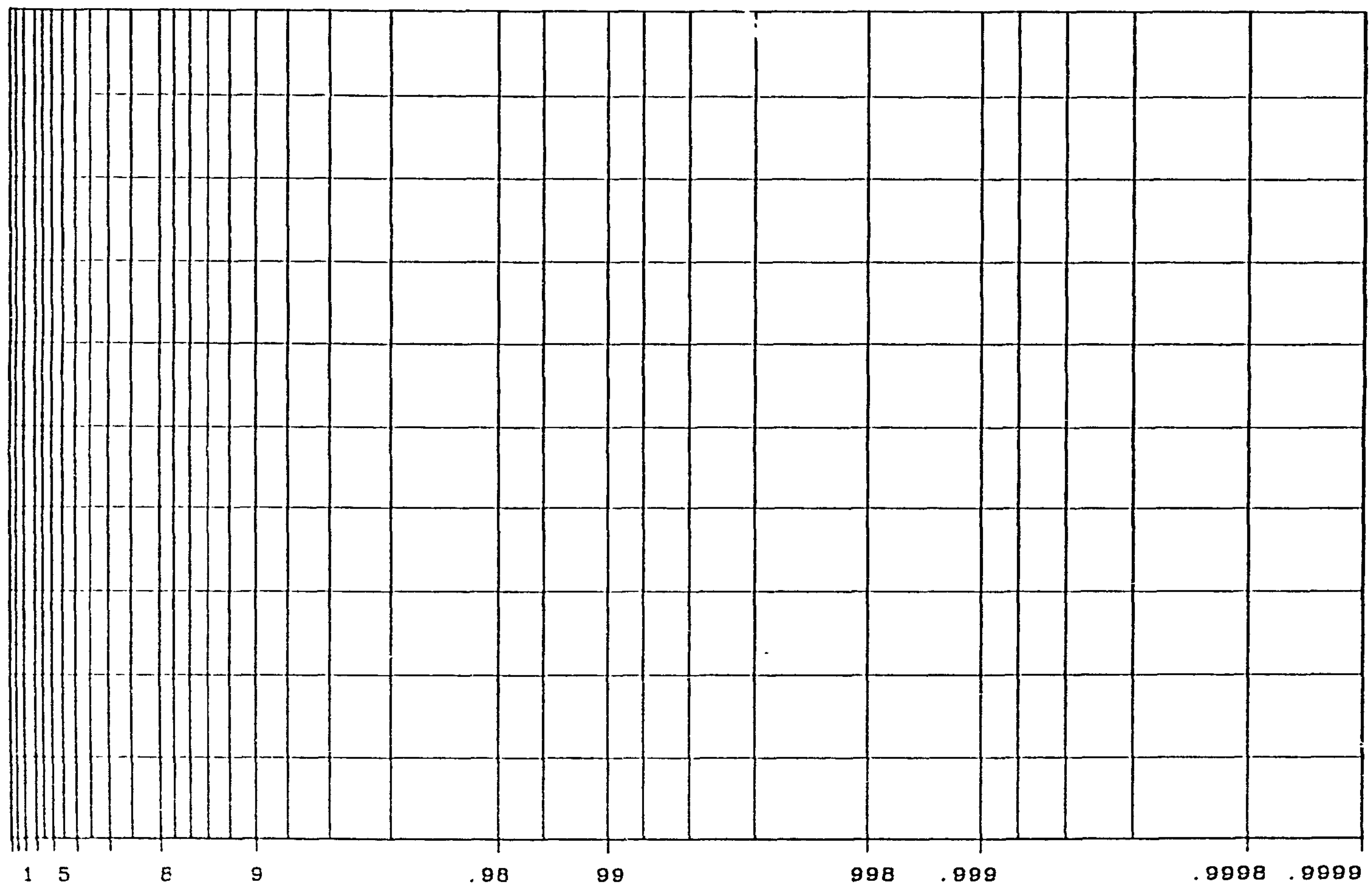
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.6



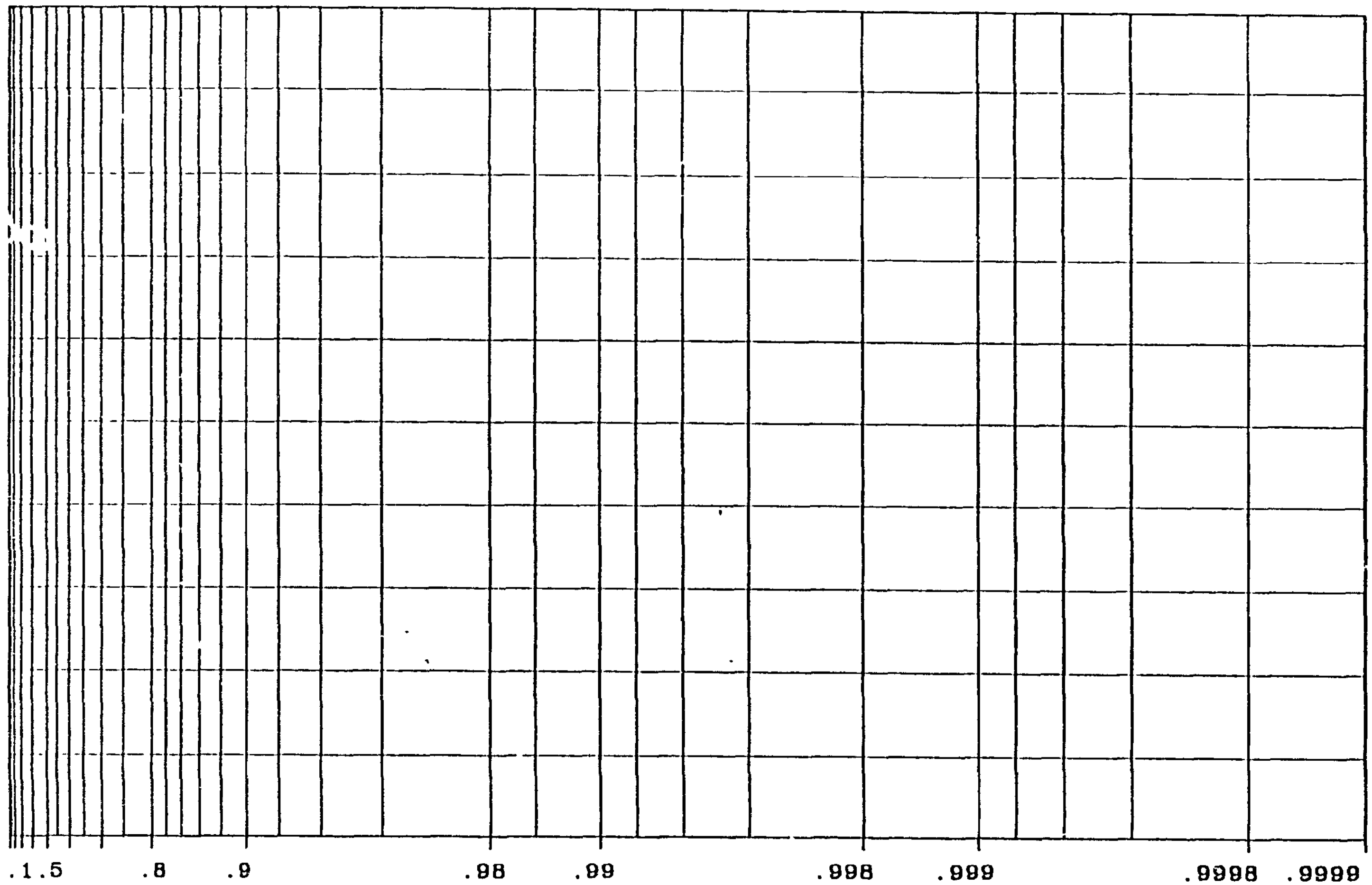
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.7



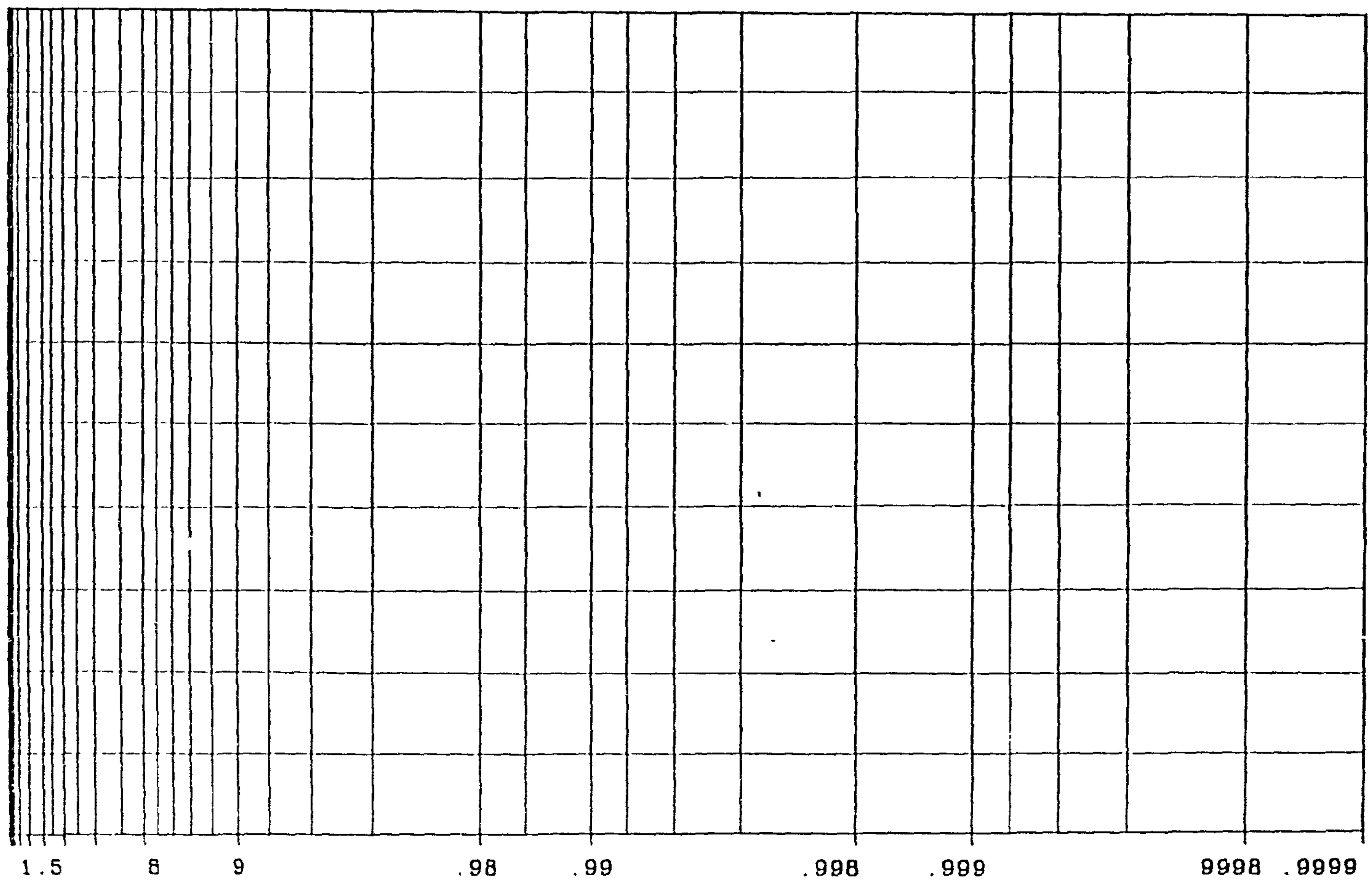
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.8



PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.9



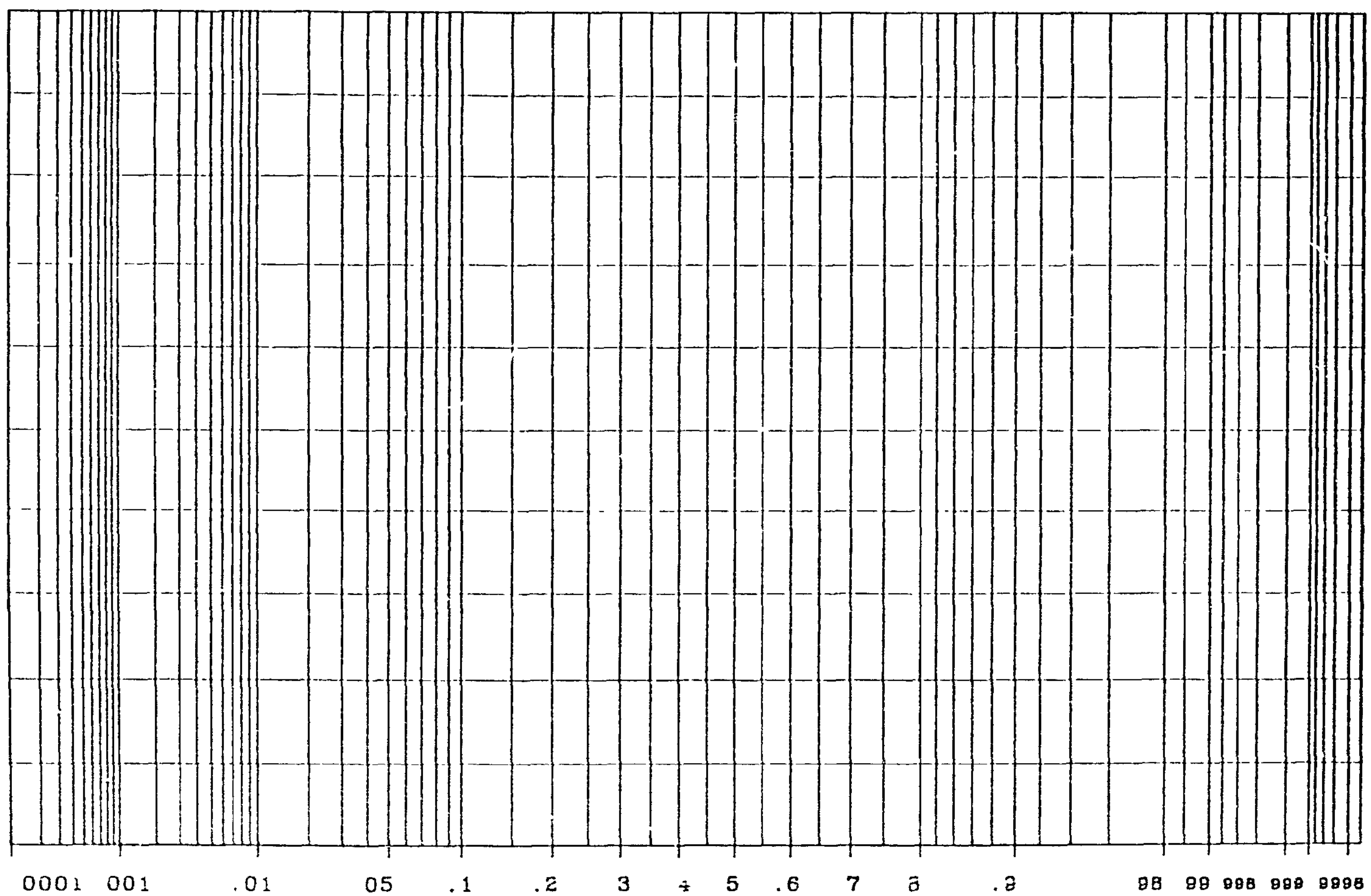
PEARSON TYPE 3 PROBABILITY PAPER, SKEWNESS COEFFICIENT = 3.0



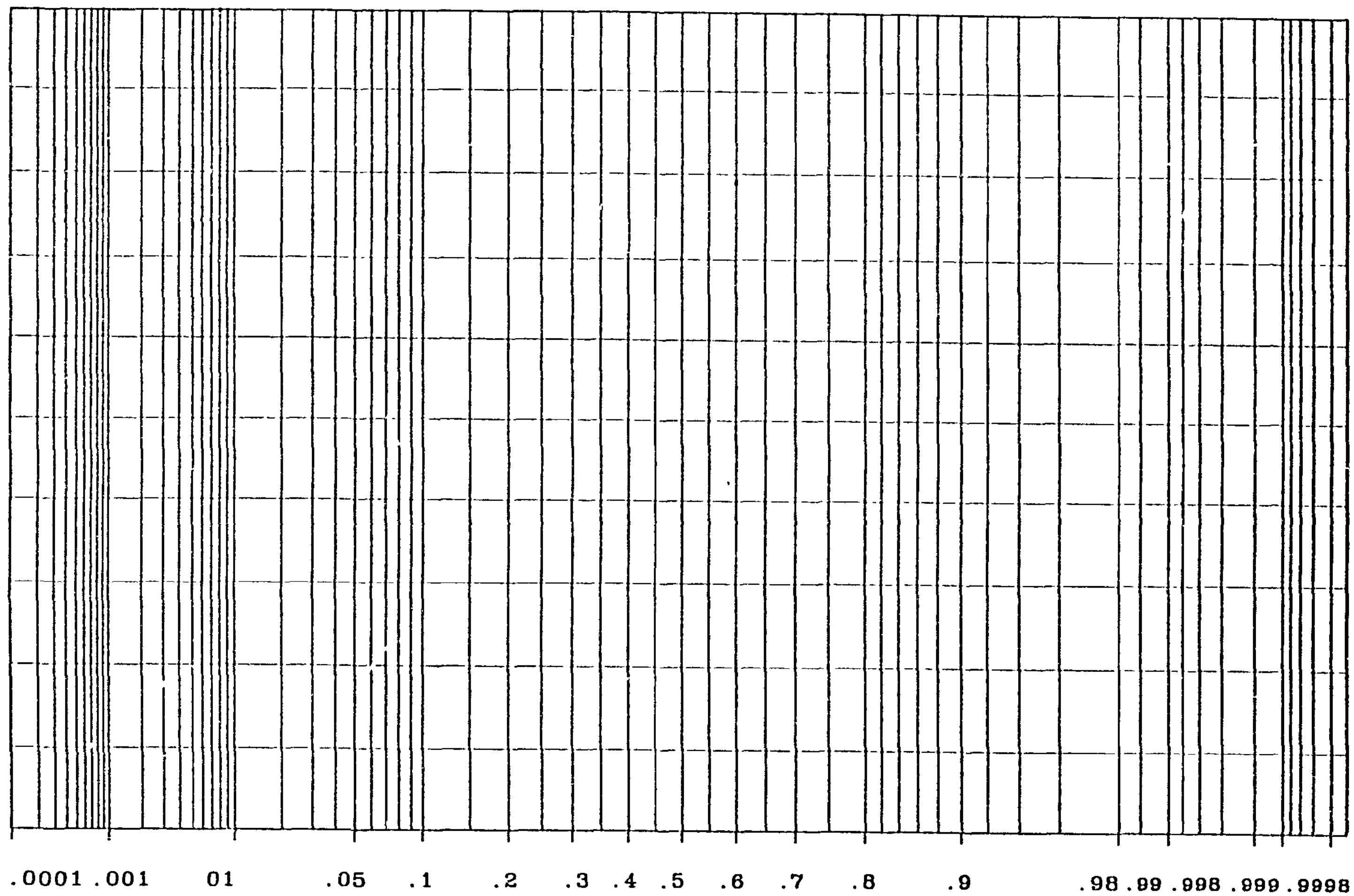
APPENDIX B

Probability Papers for General Extreme Values Distribution.

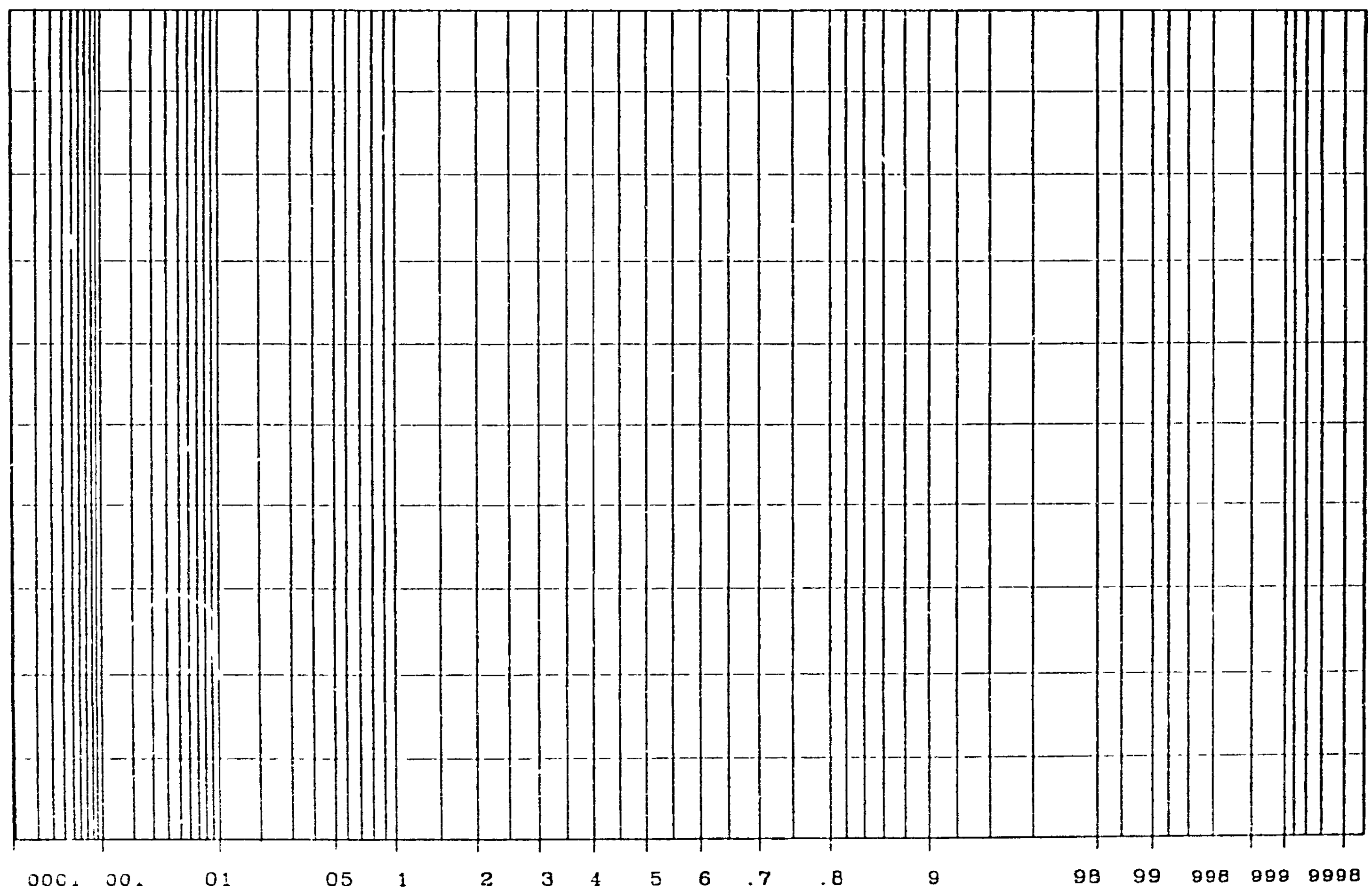
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.0



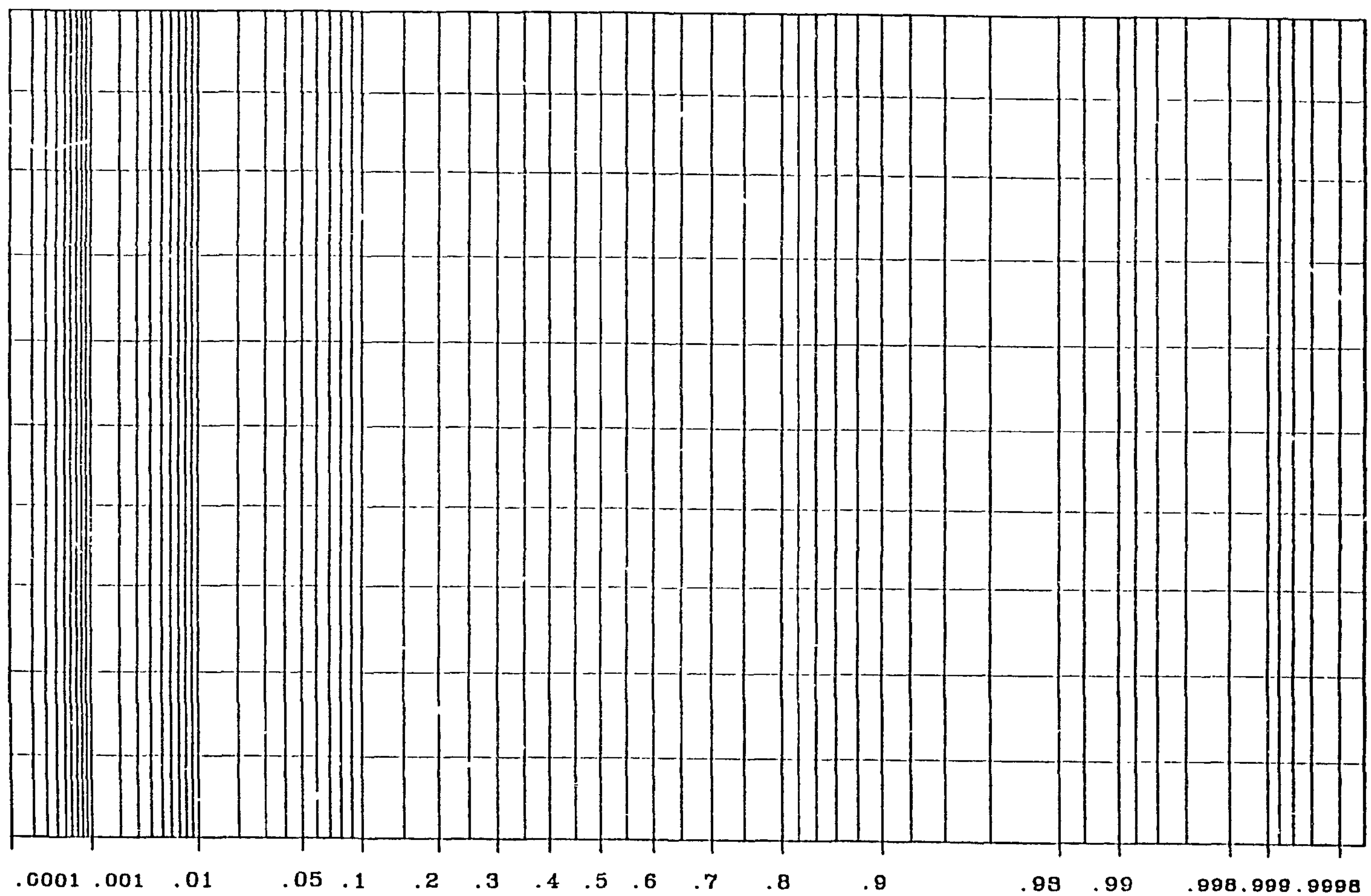
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.1



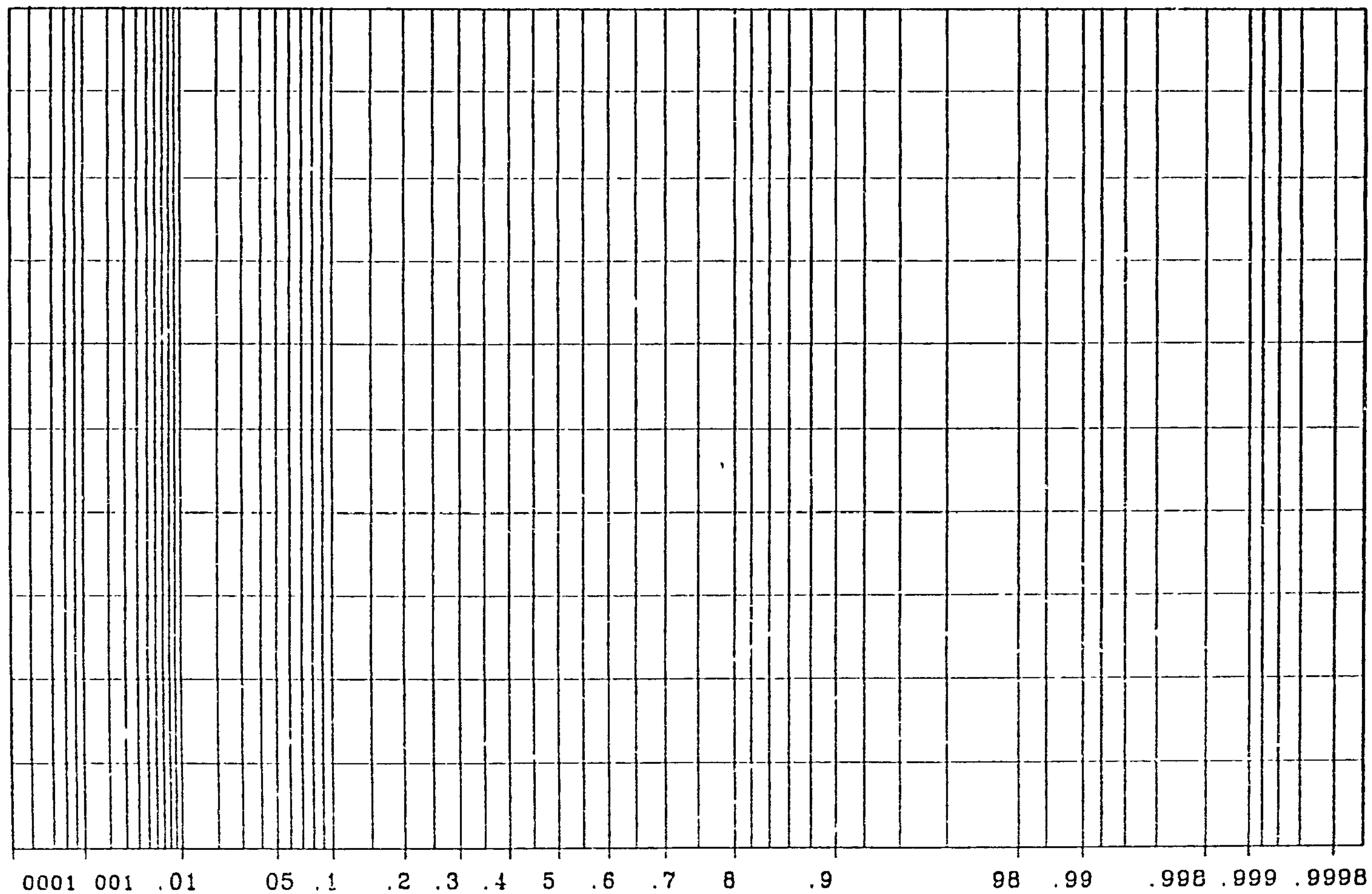
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.2



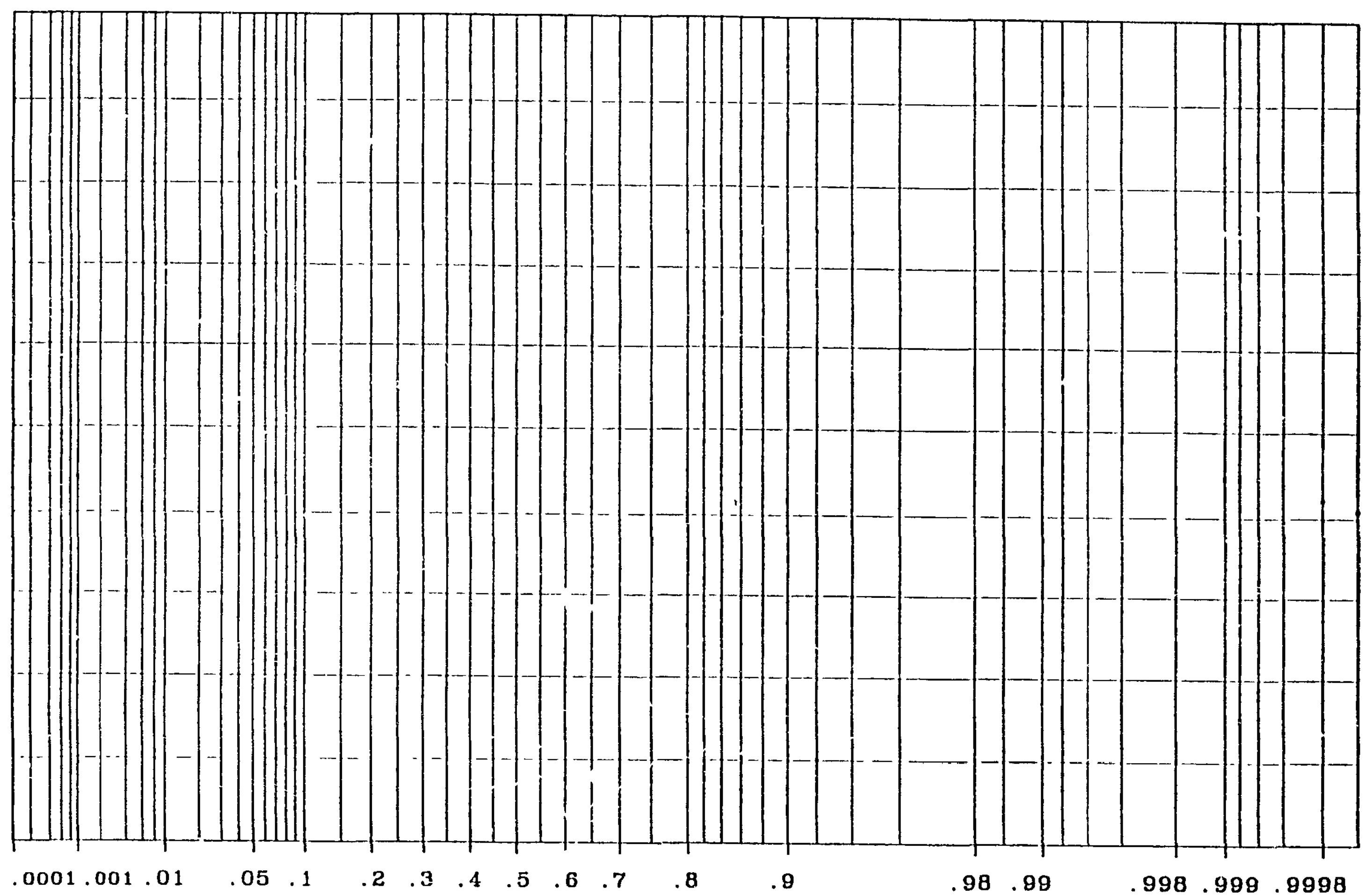
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.3



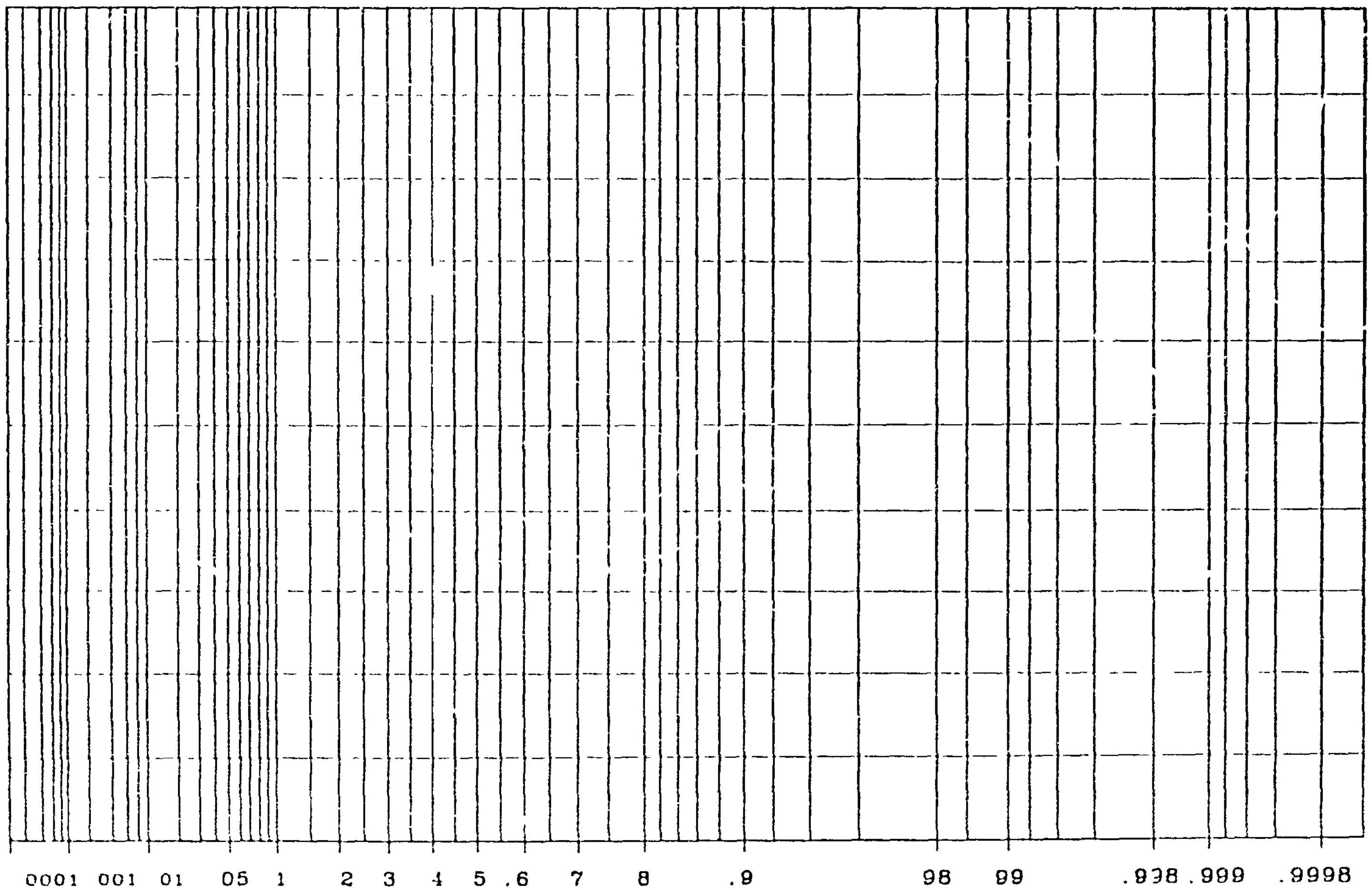
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.4



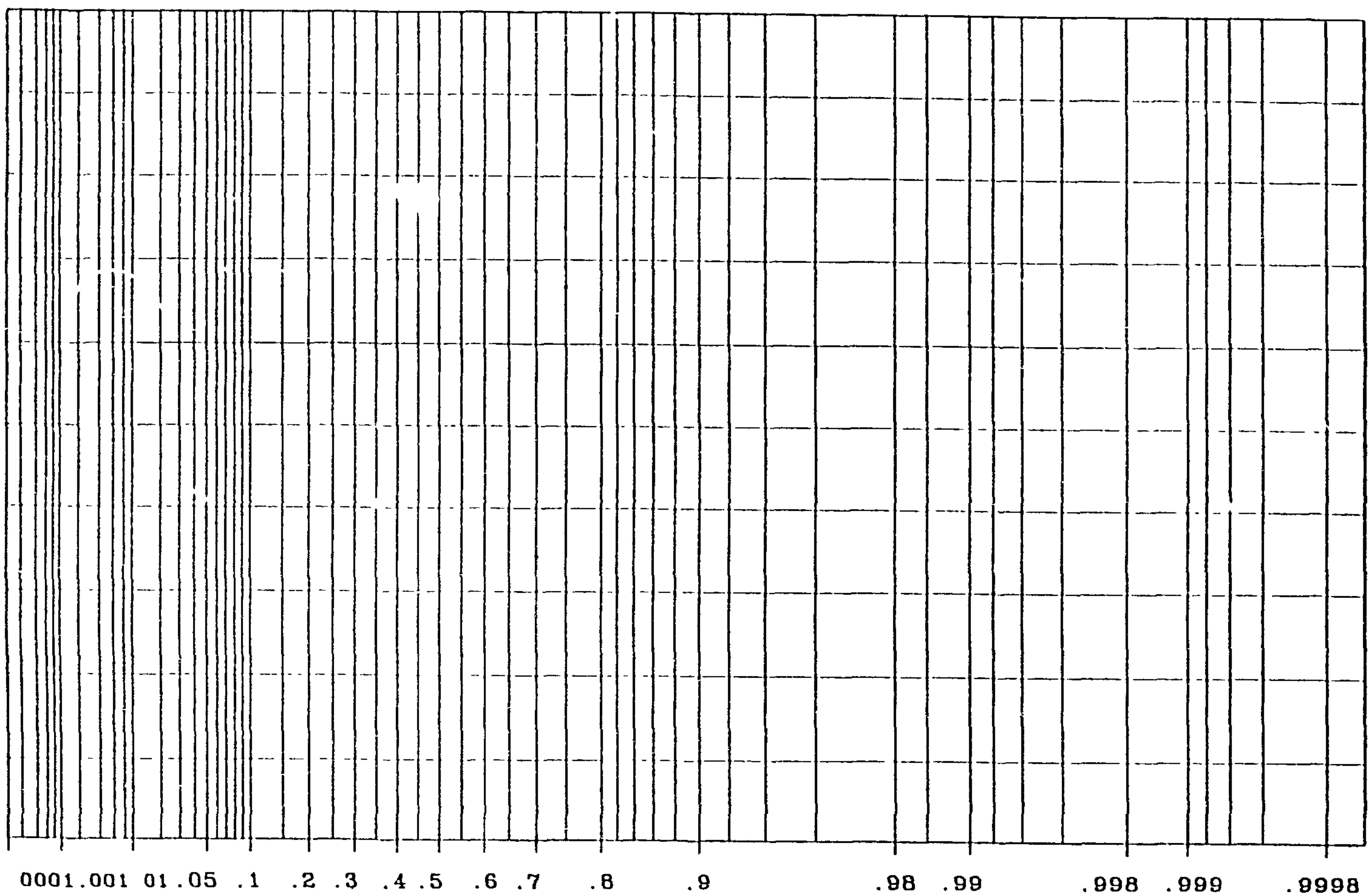
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.5



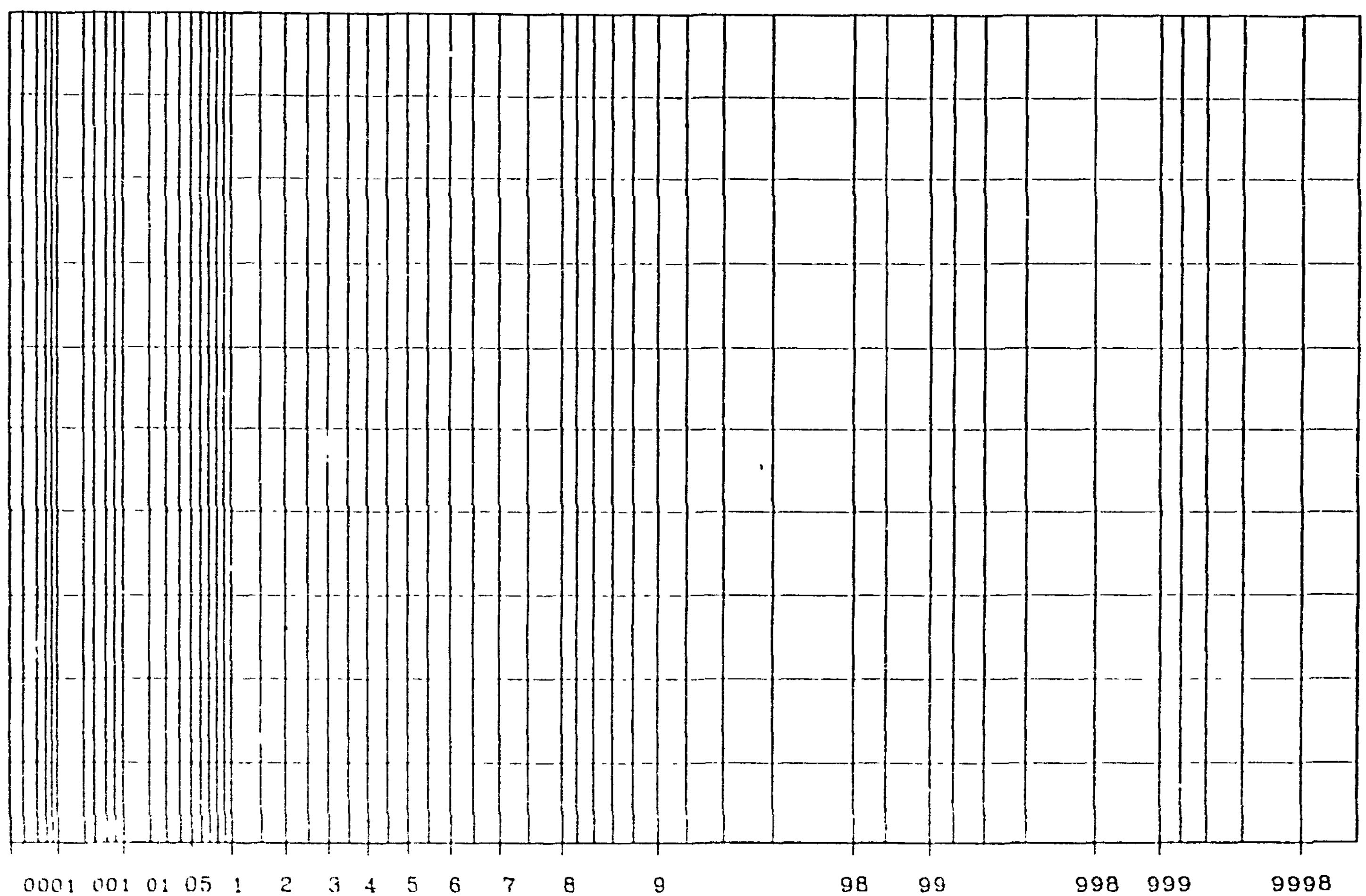
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.6



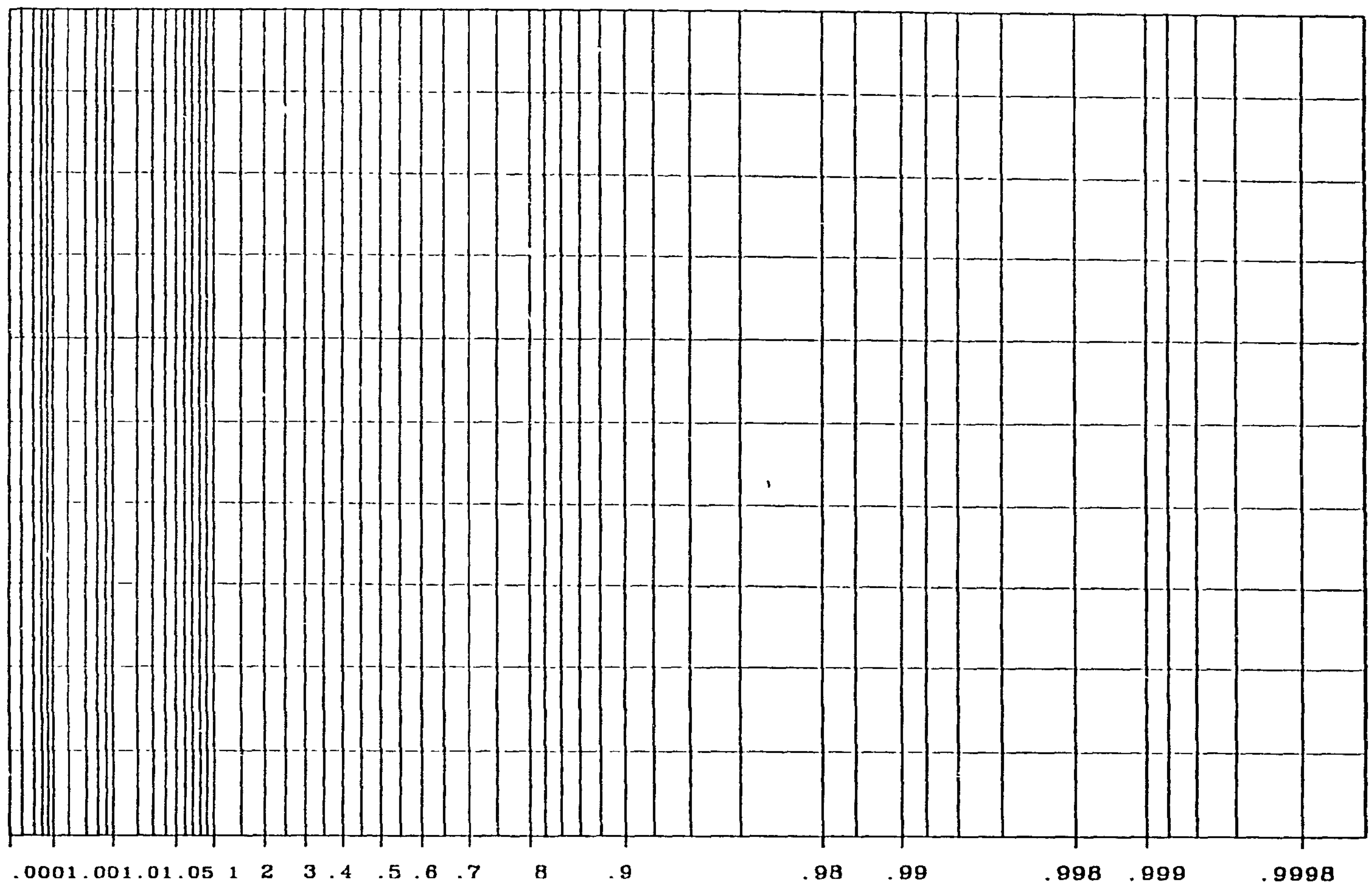
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.7



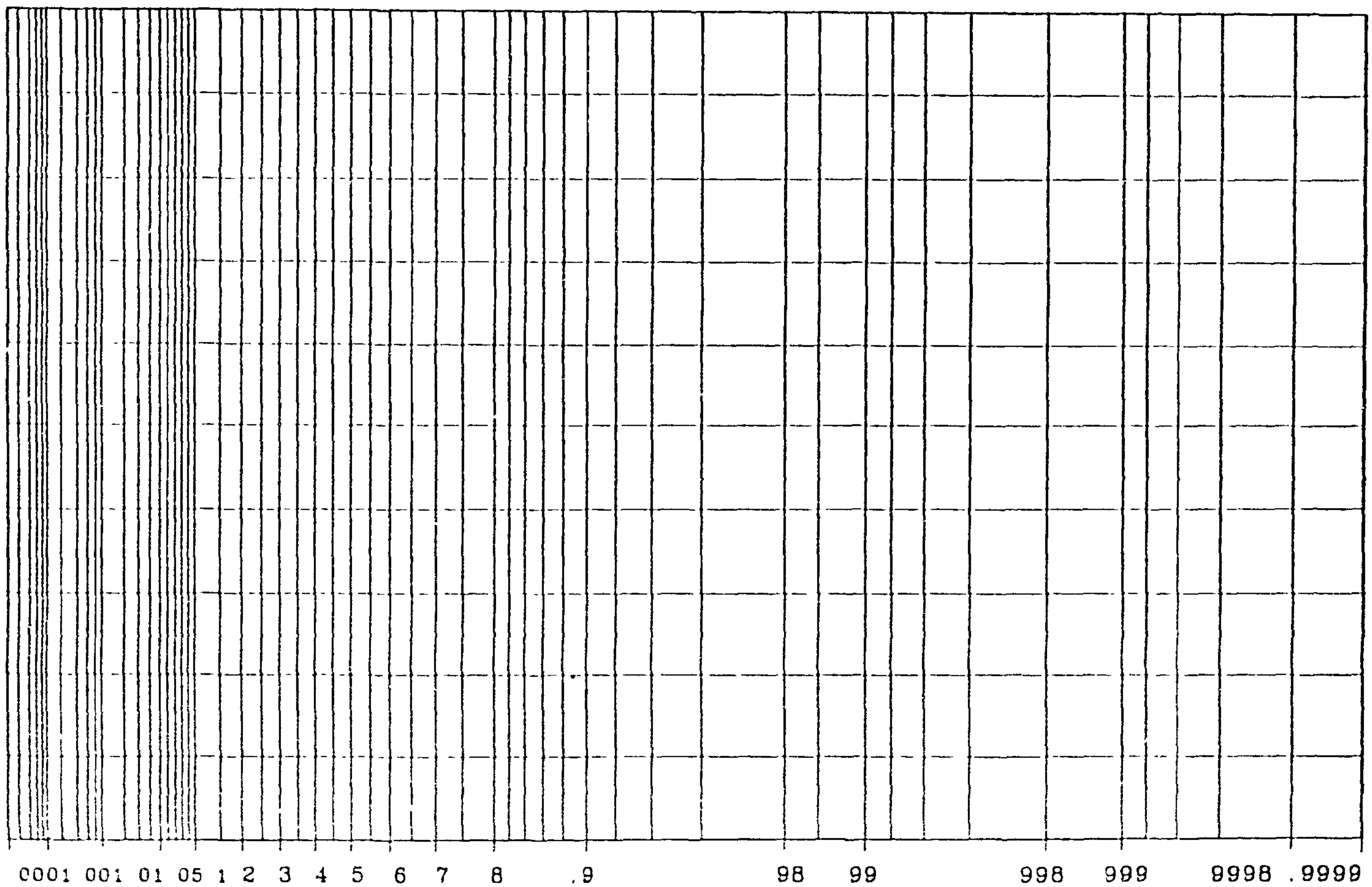
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.8



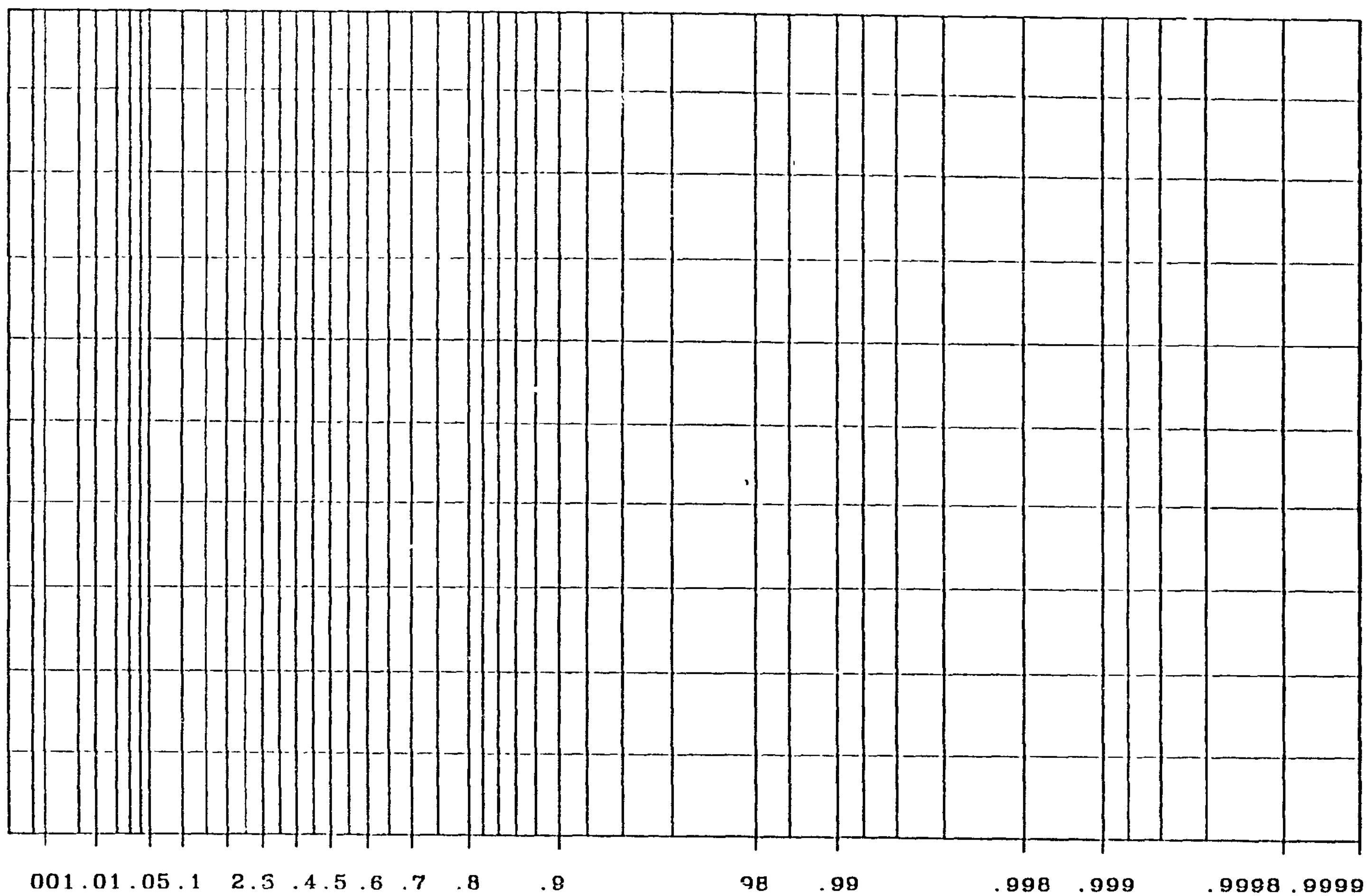
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 0.9



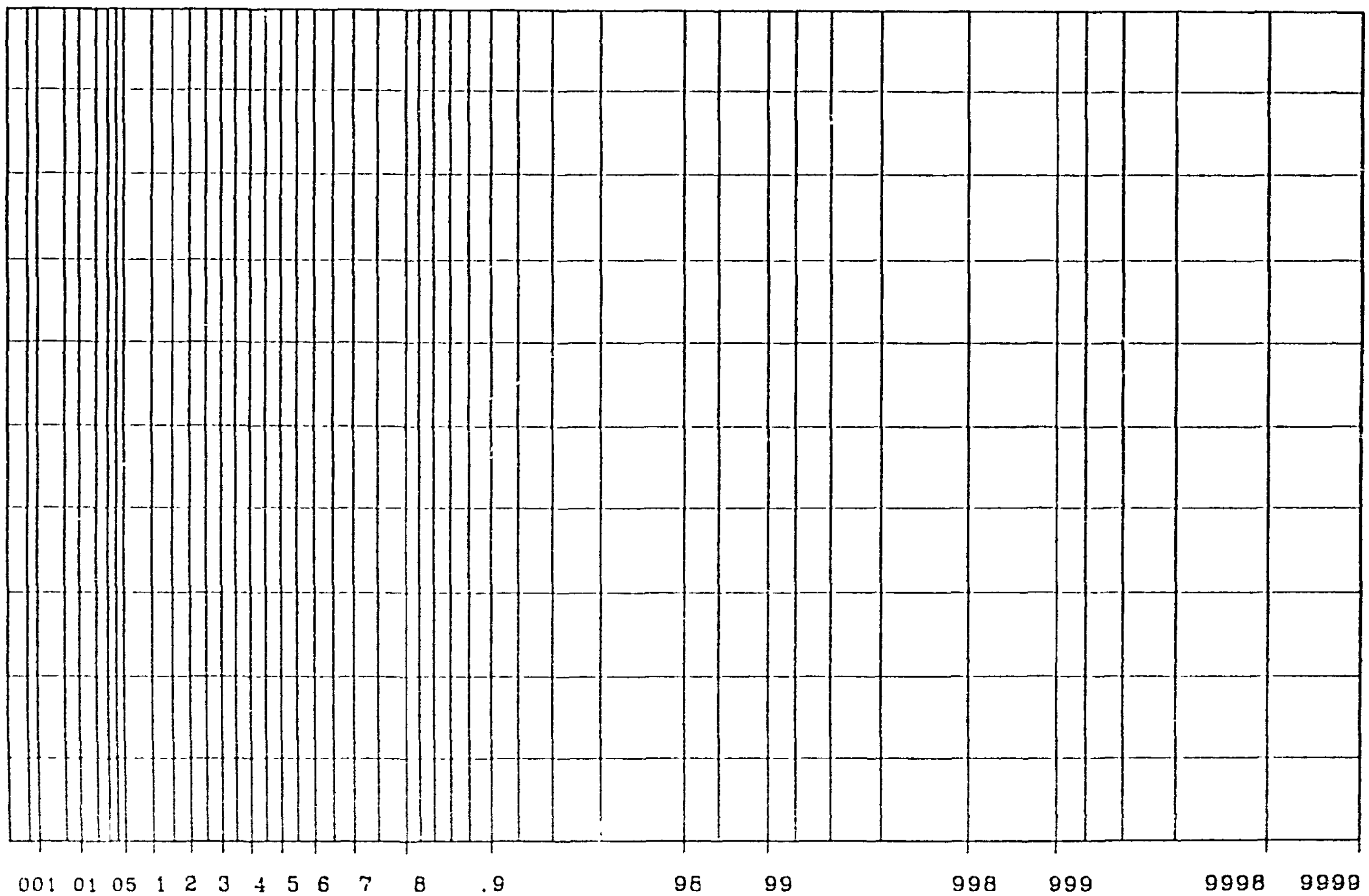
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.0



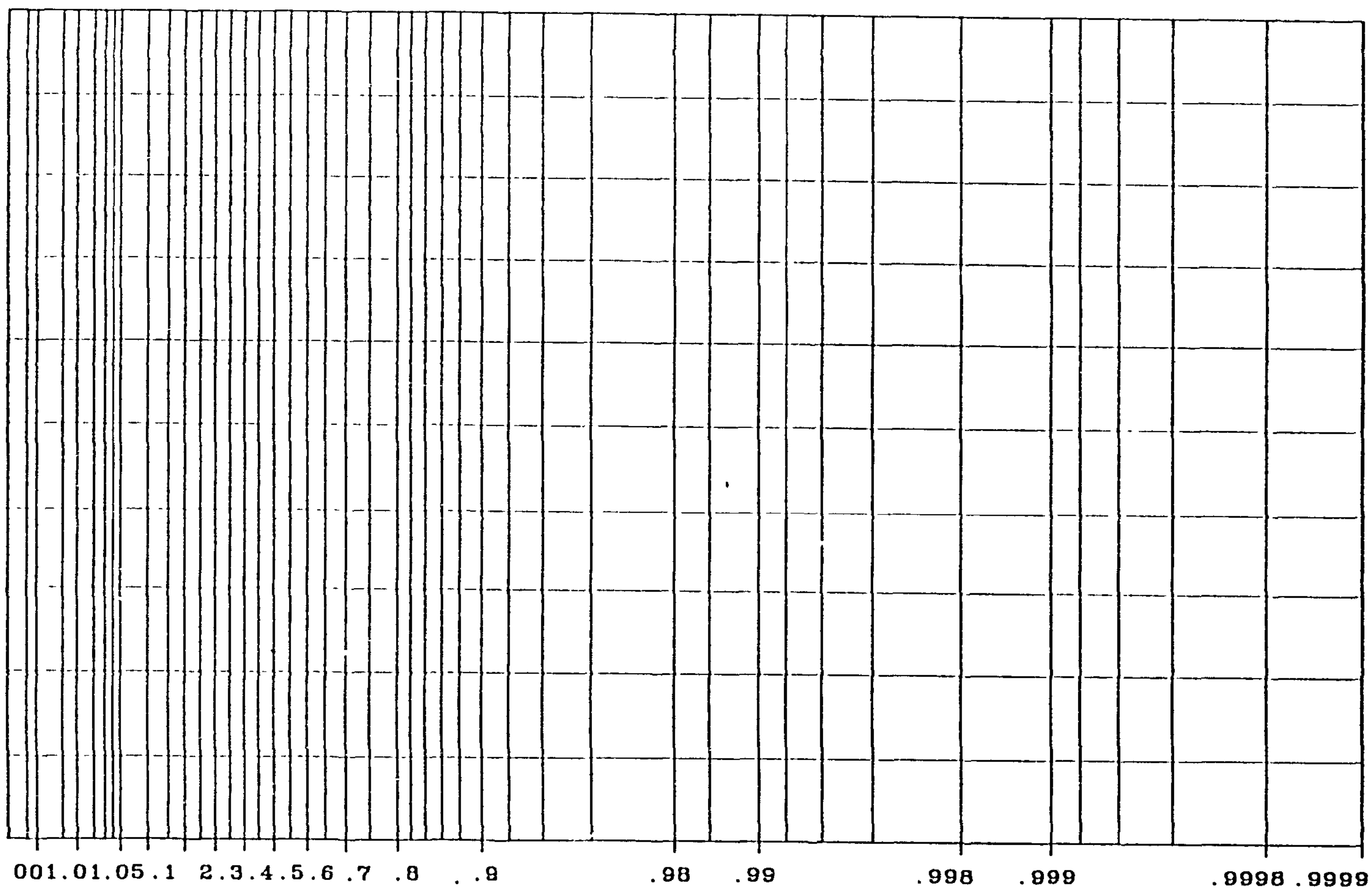
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.1



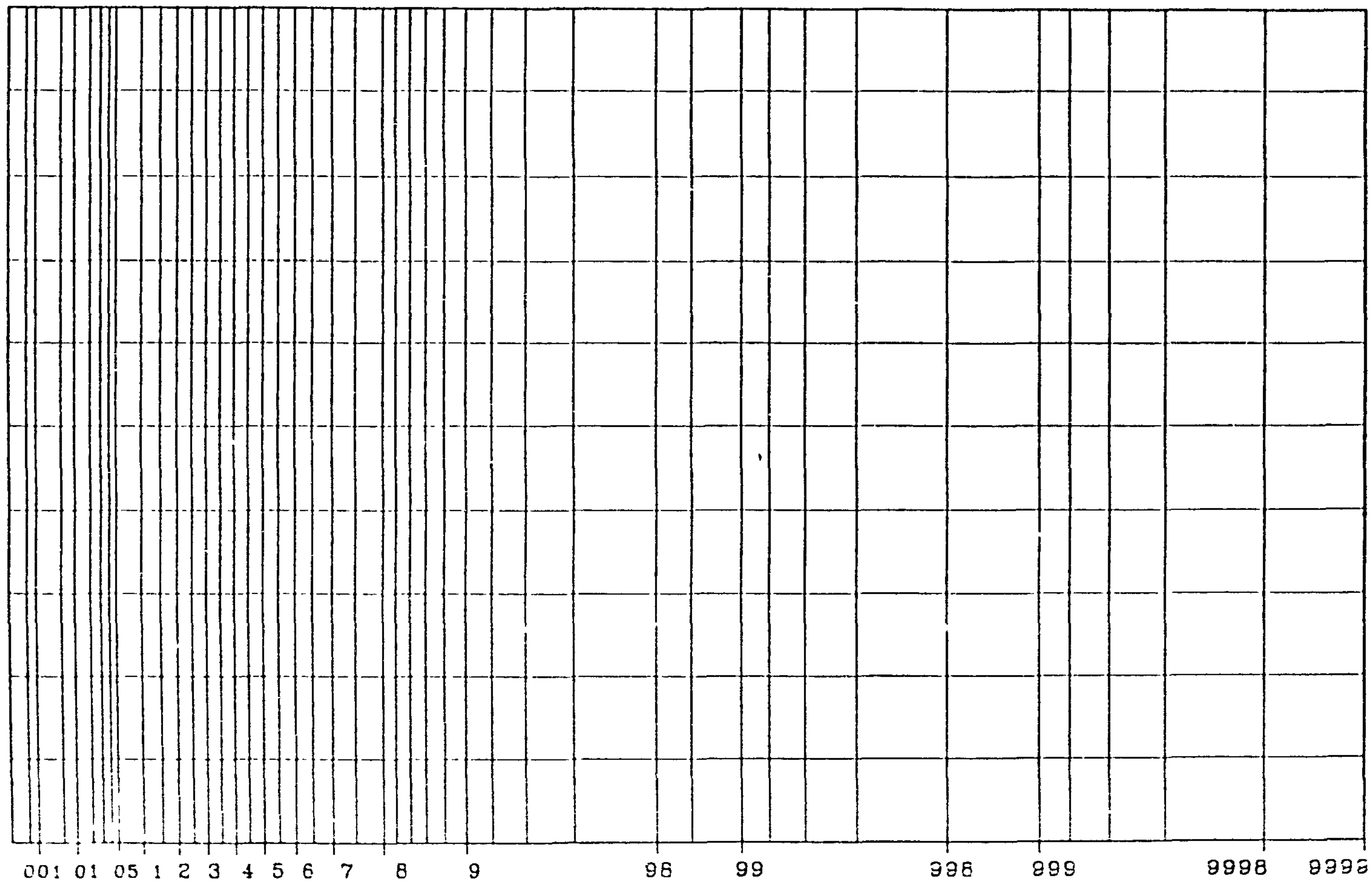
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.2



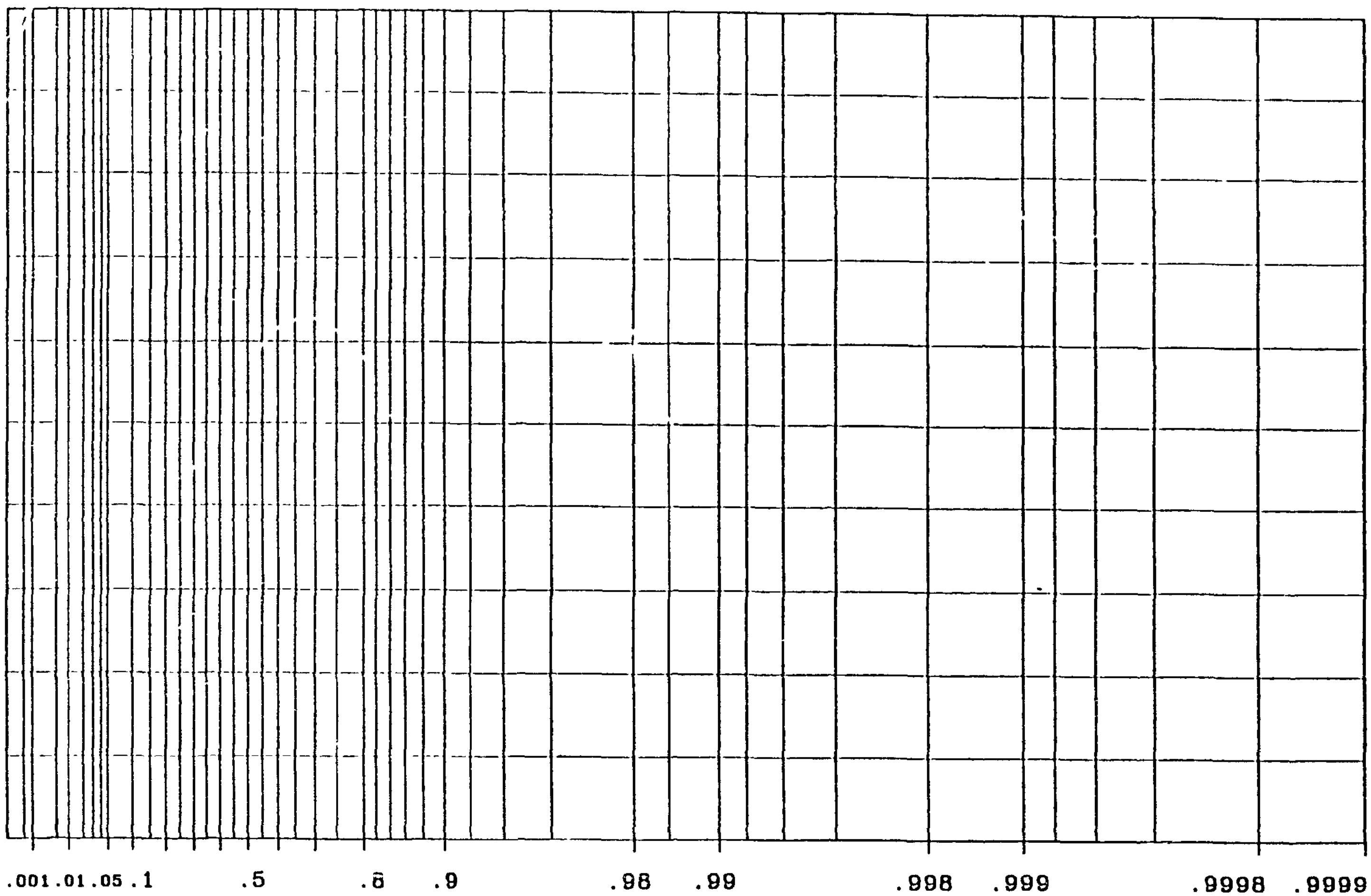
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.3



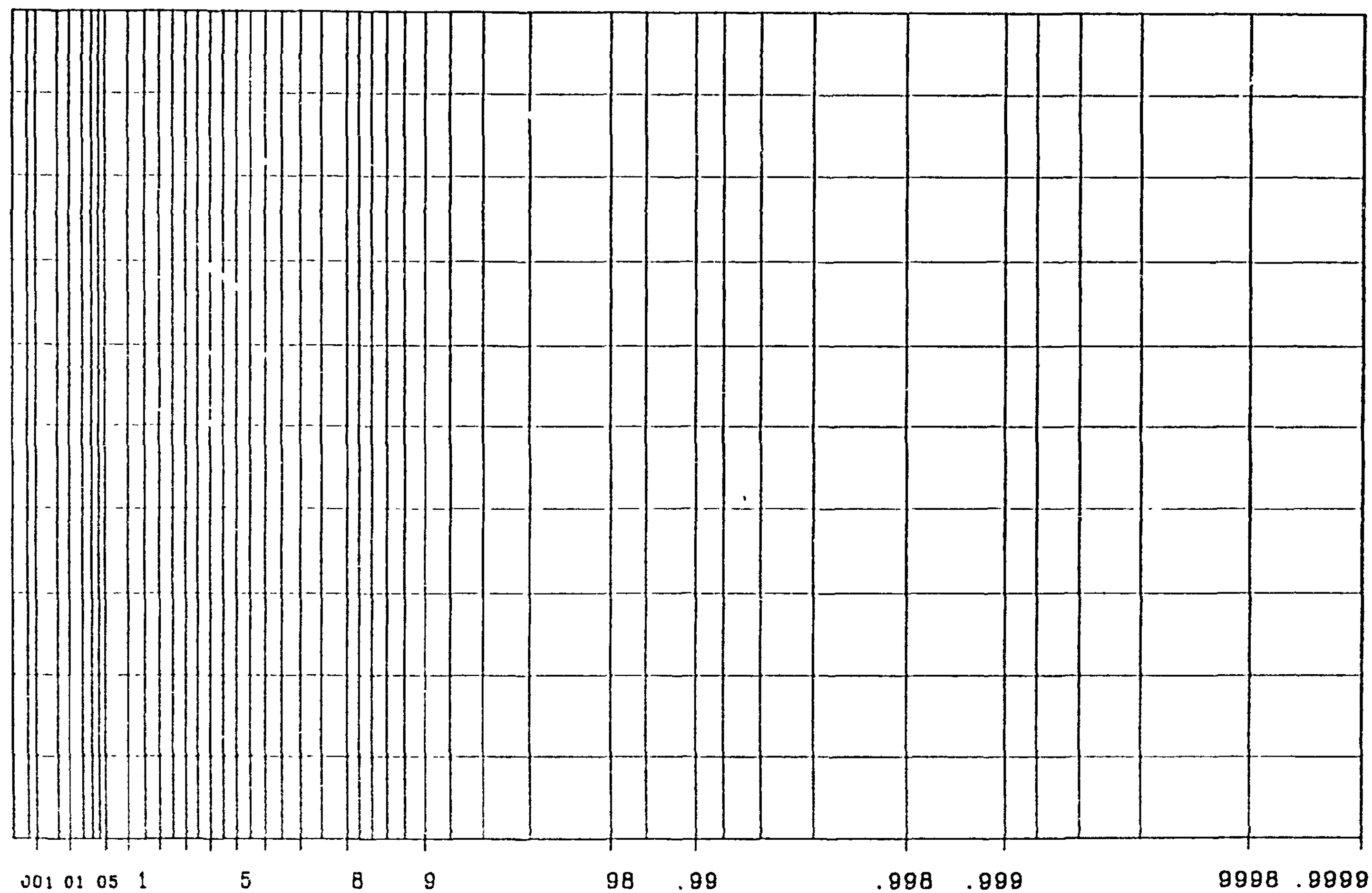
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.4



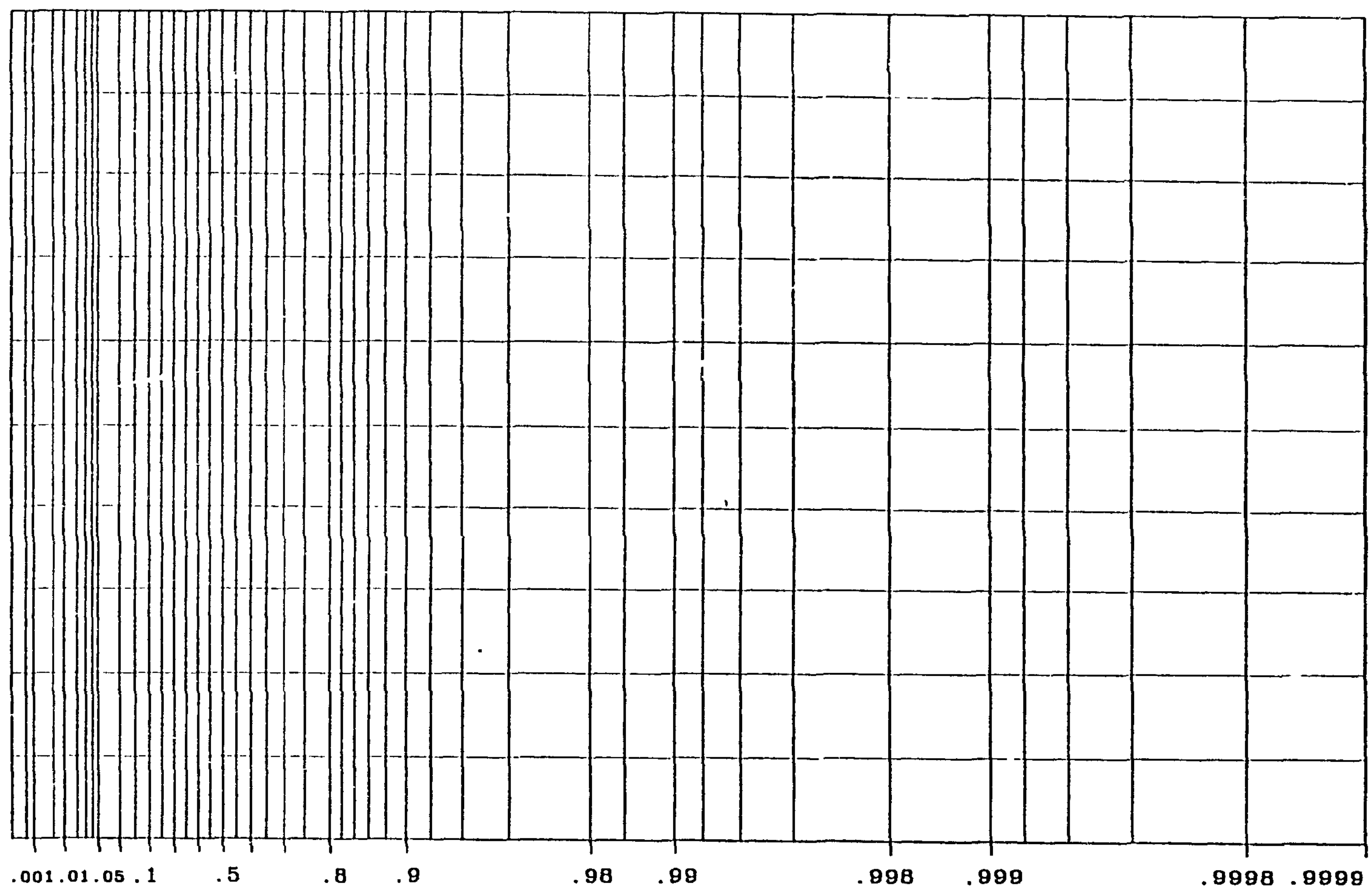
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.5



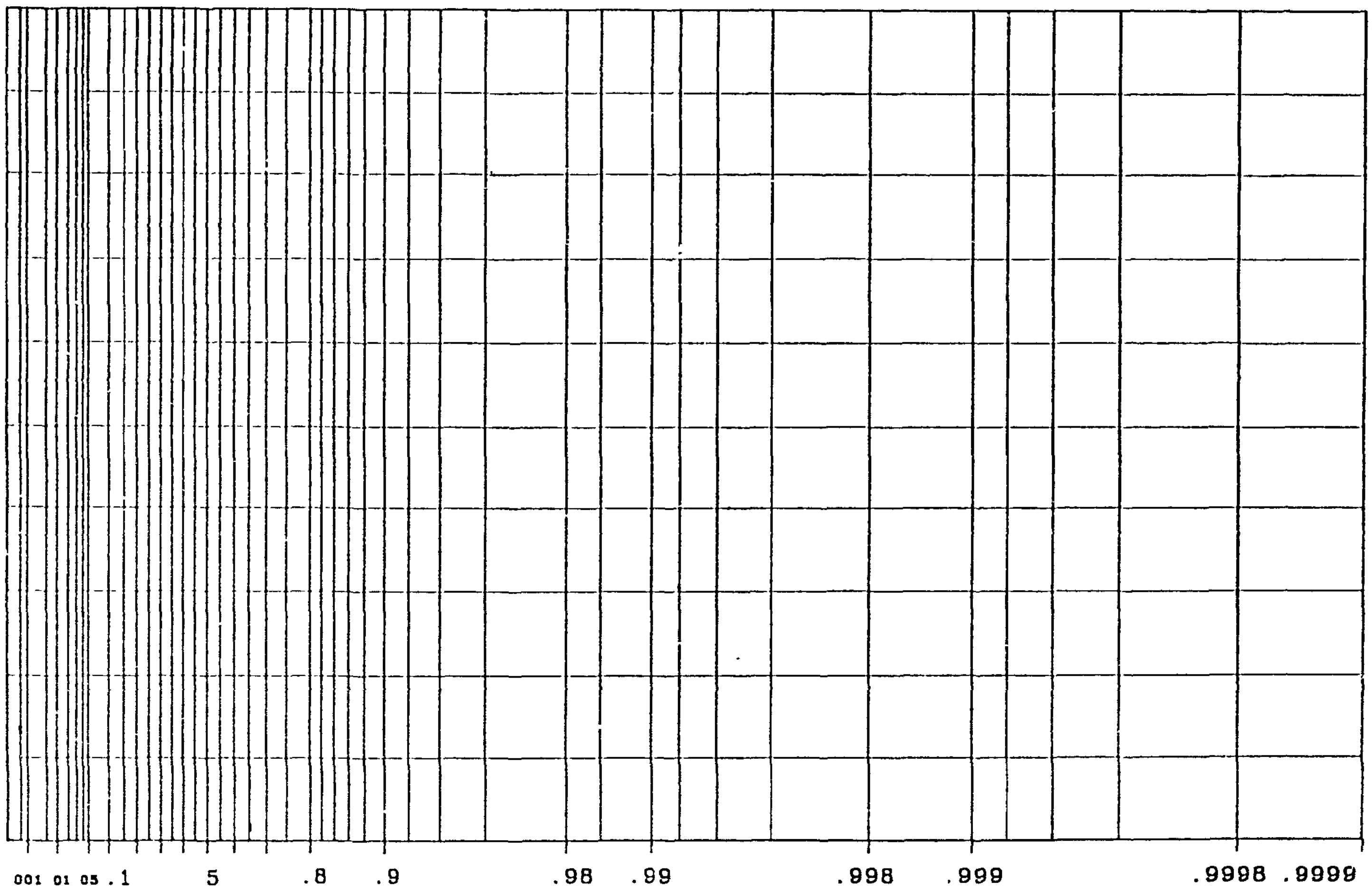
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.6



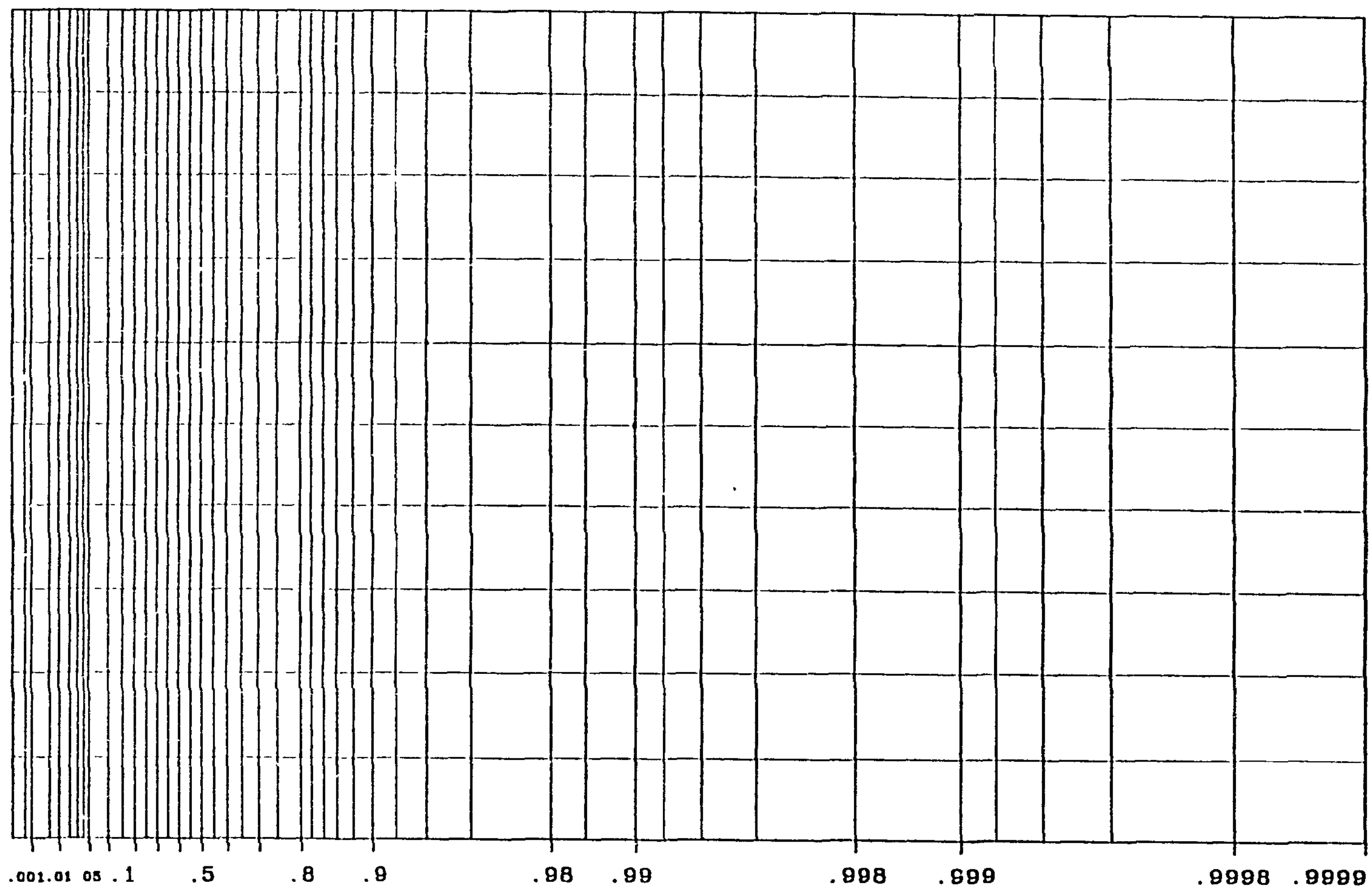
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.7



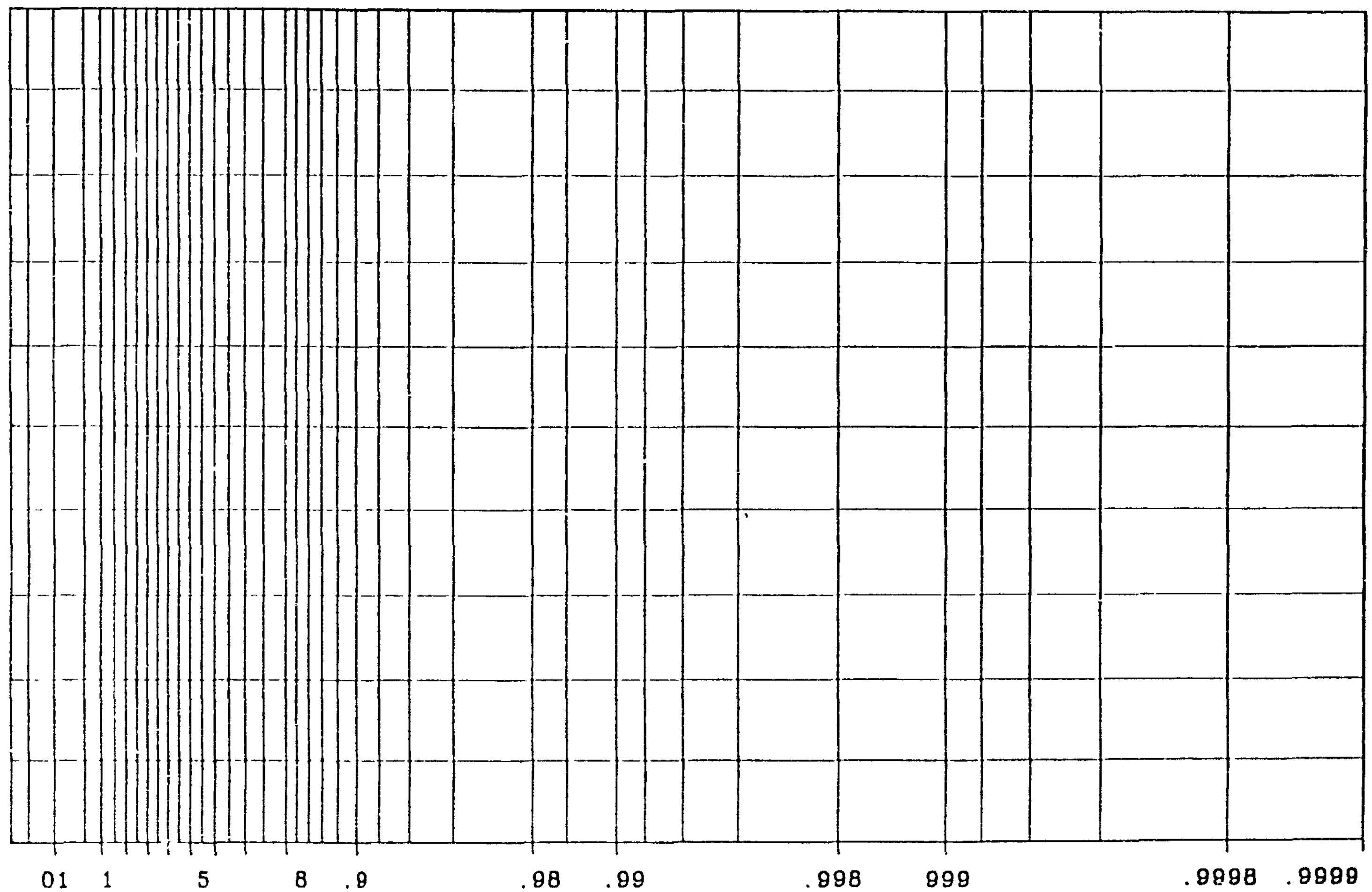
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.8



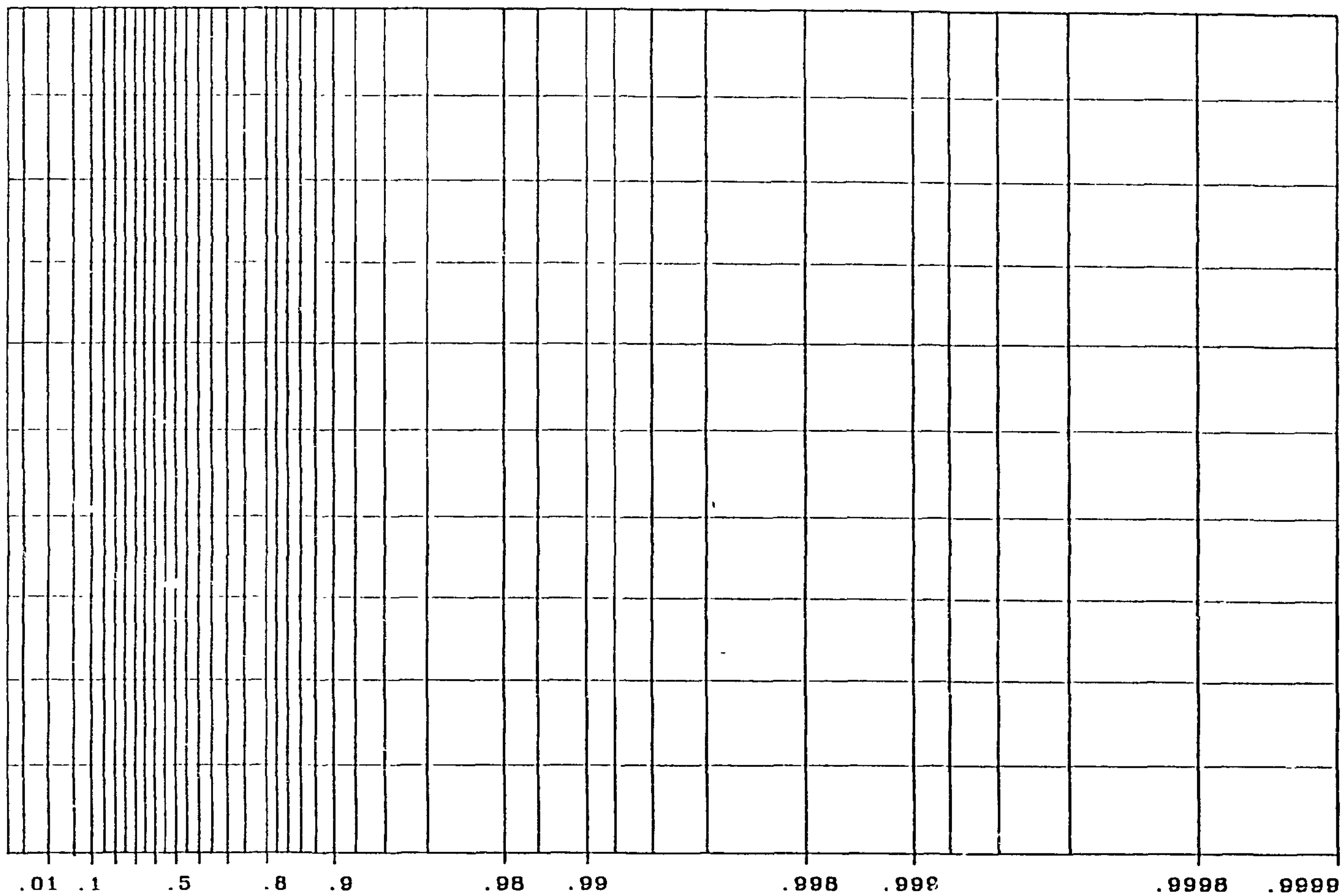
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 1.9



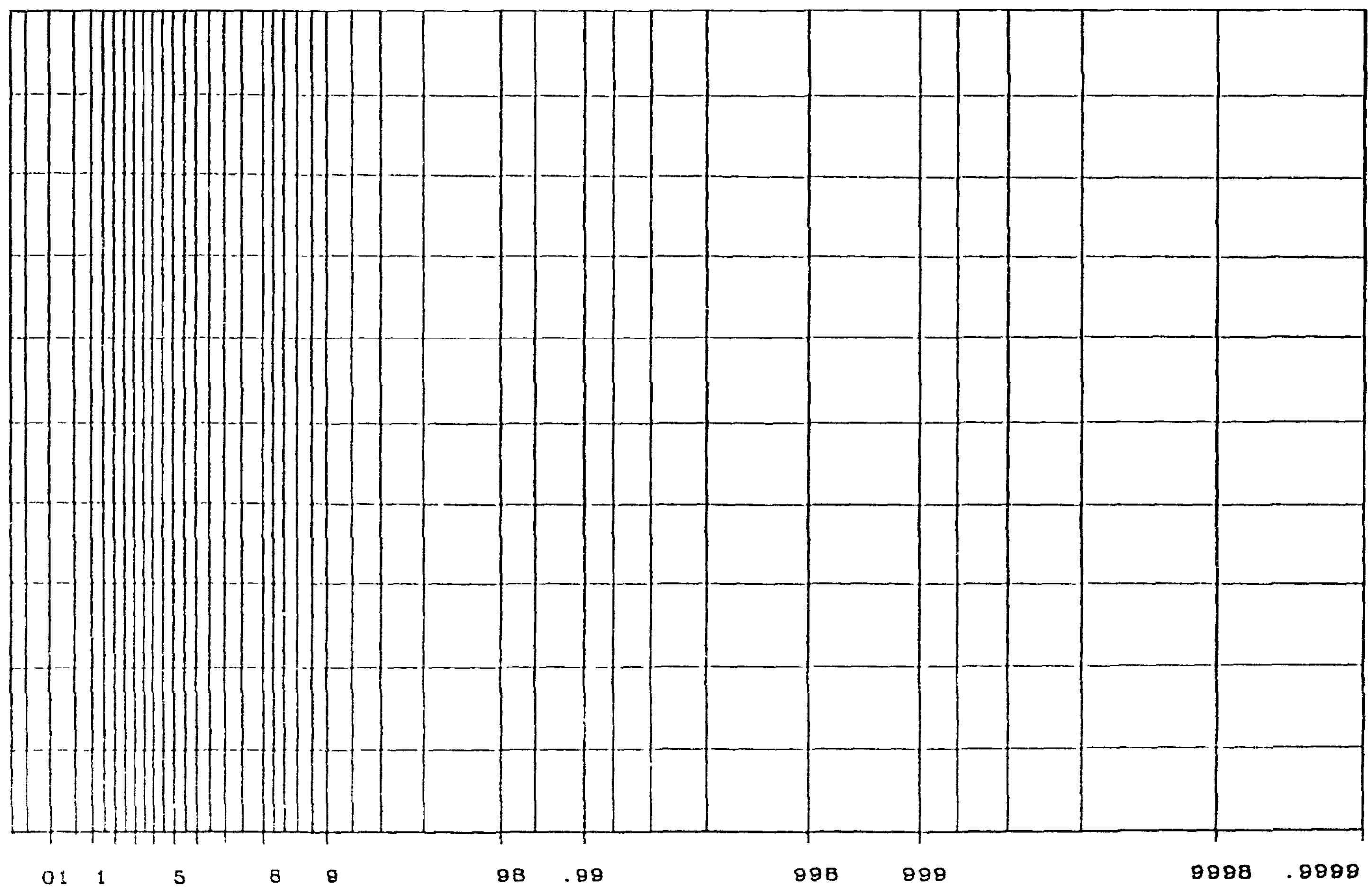
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.0



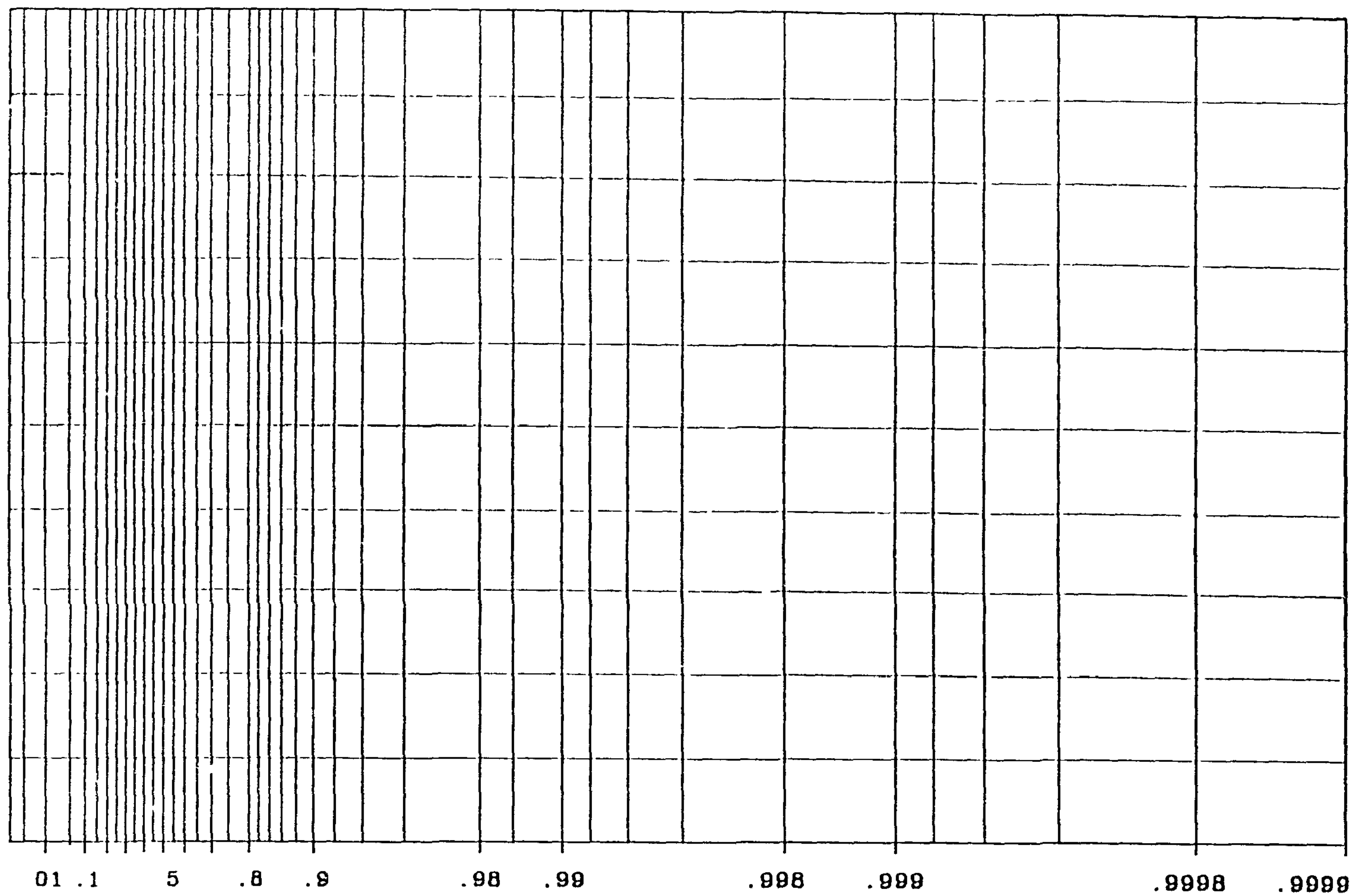
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.1



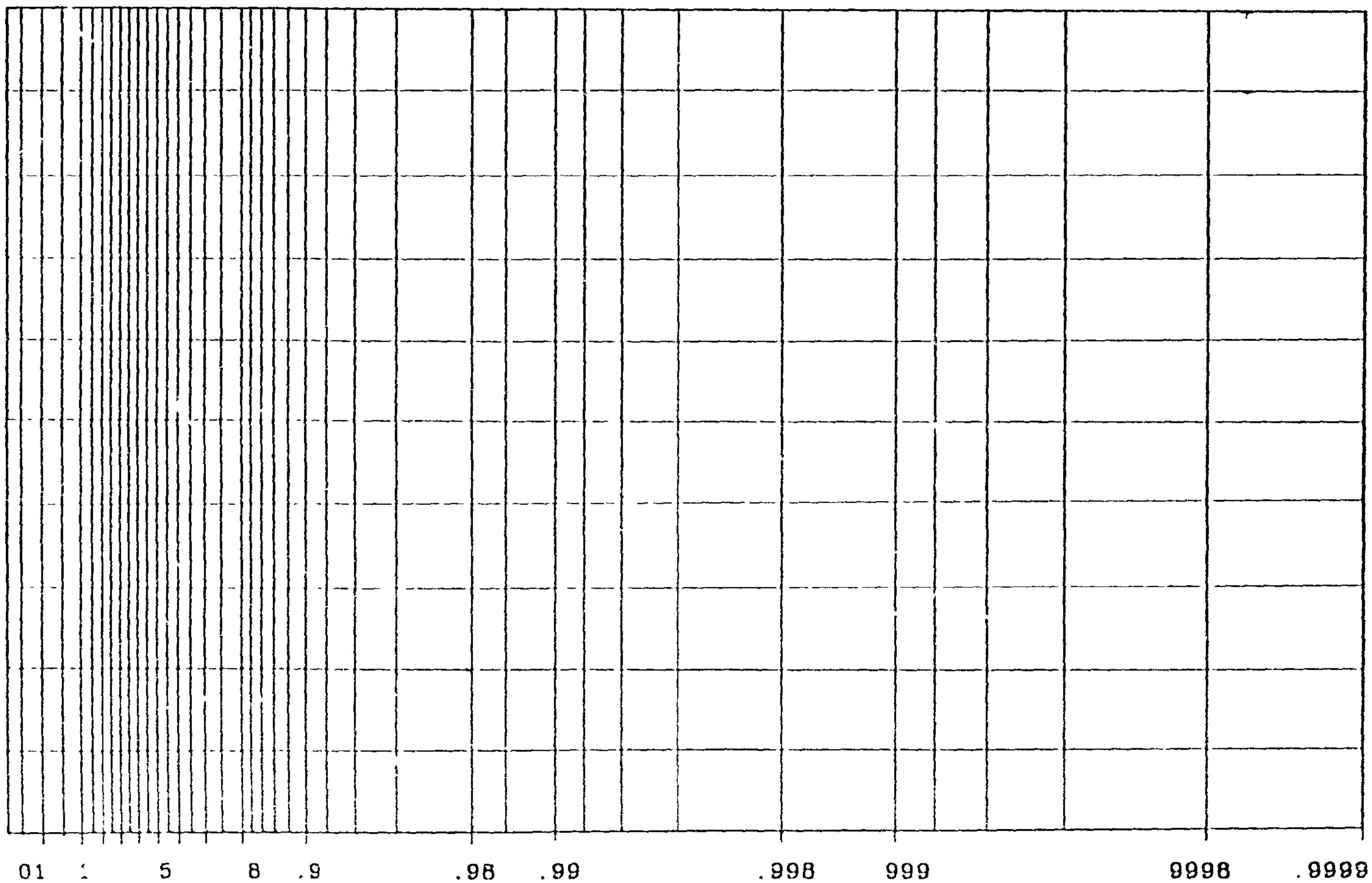
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.2



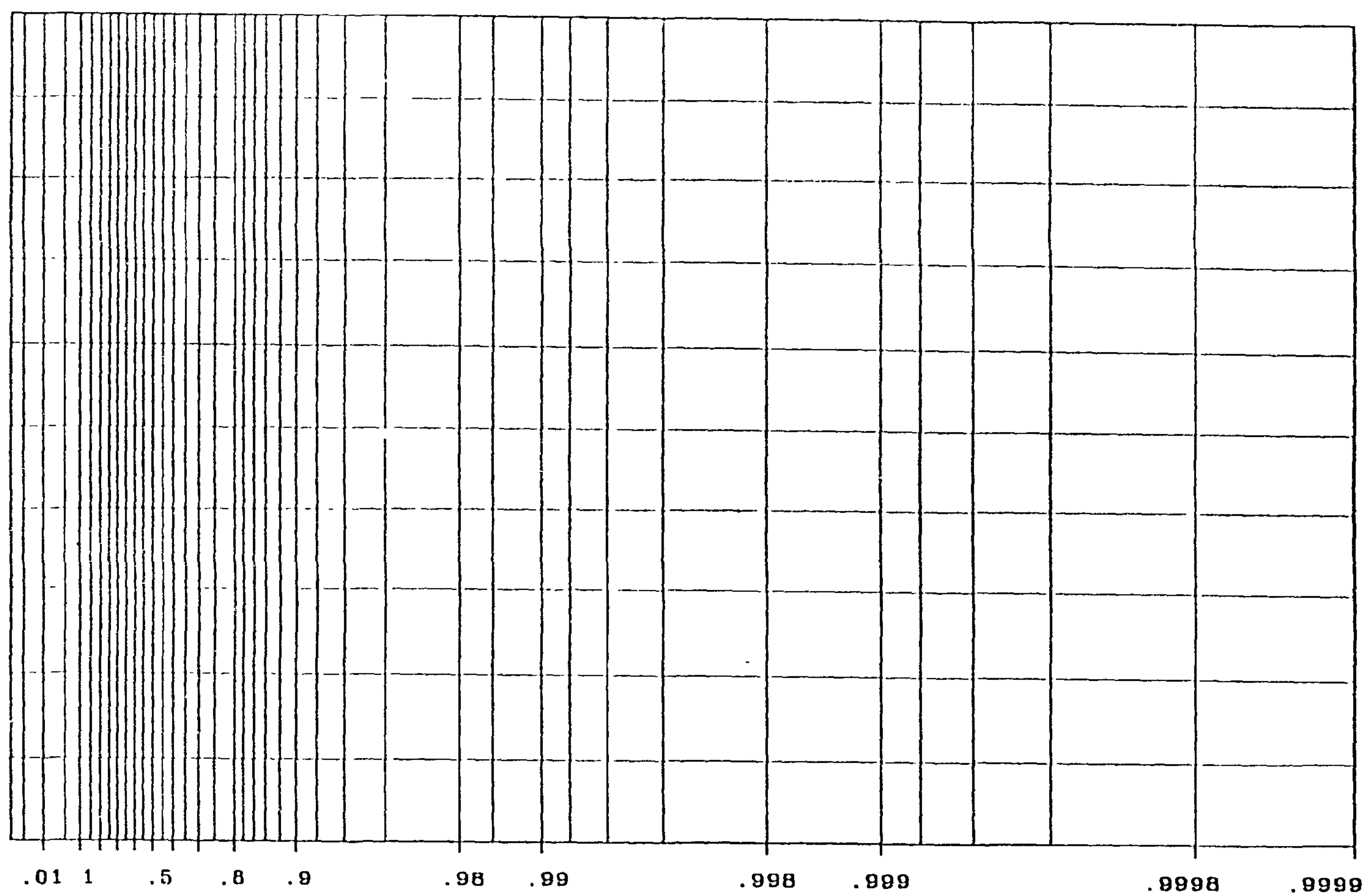
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.3



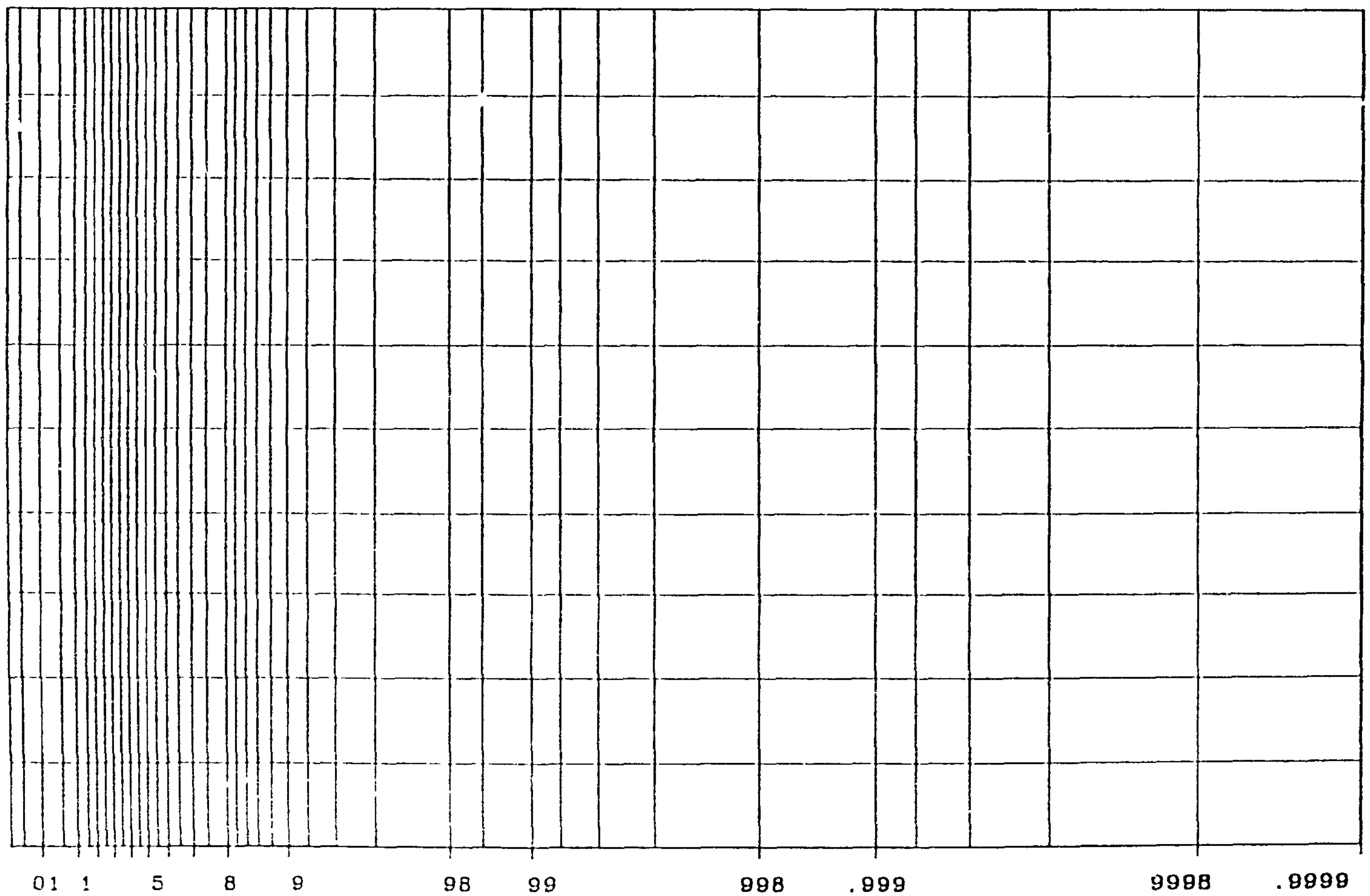
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.4



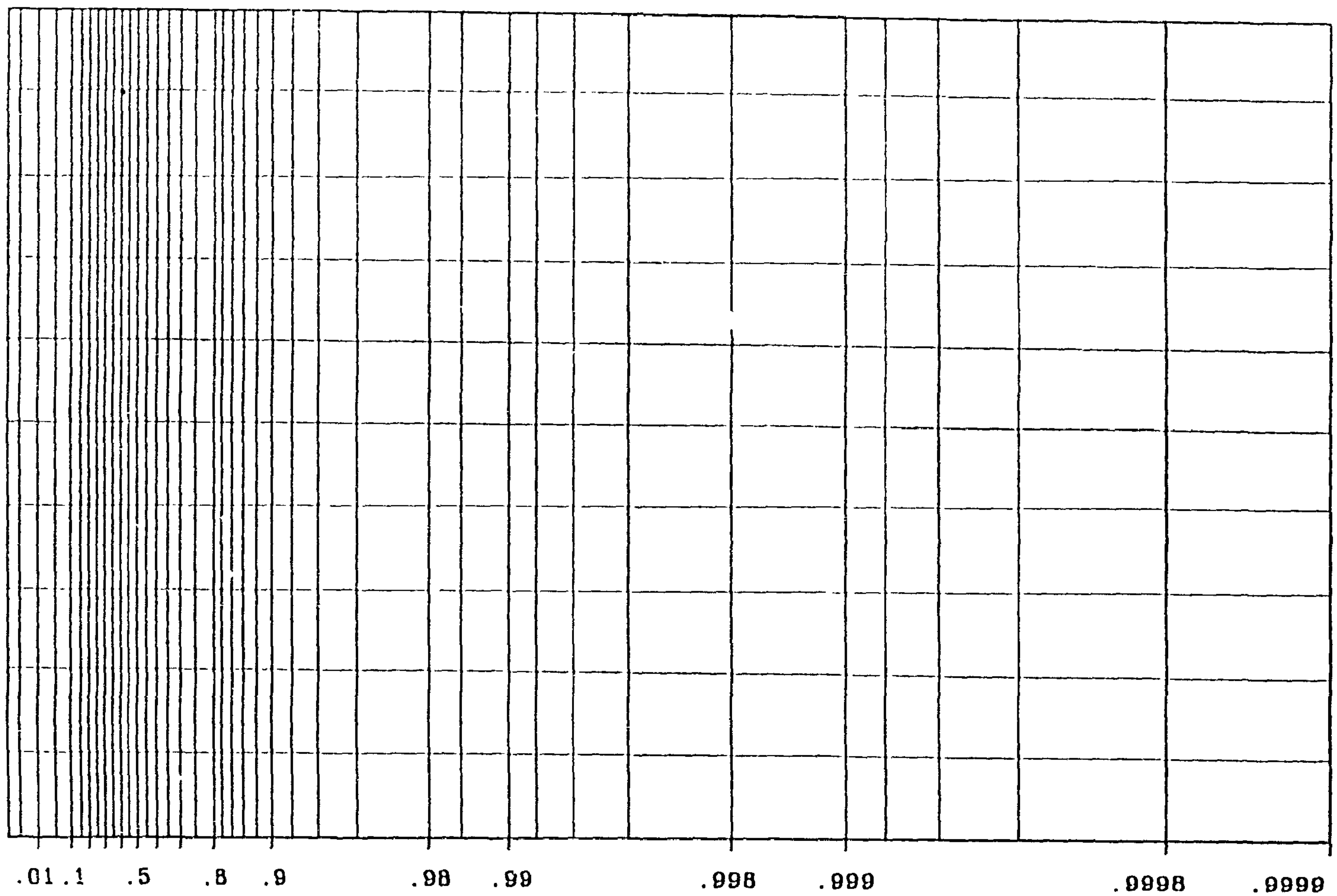
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.5



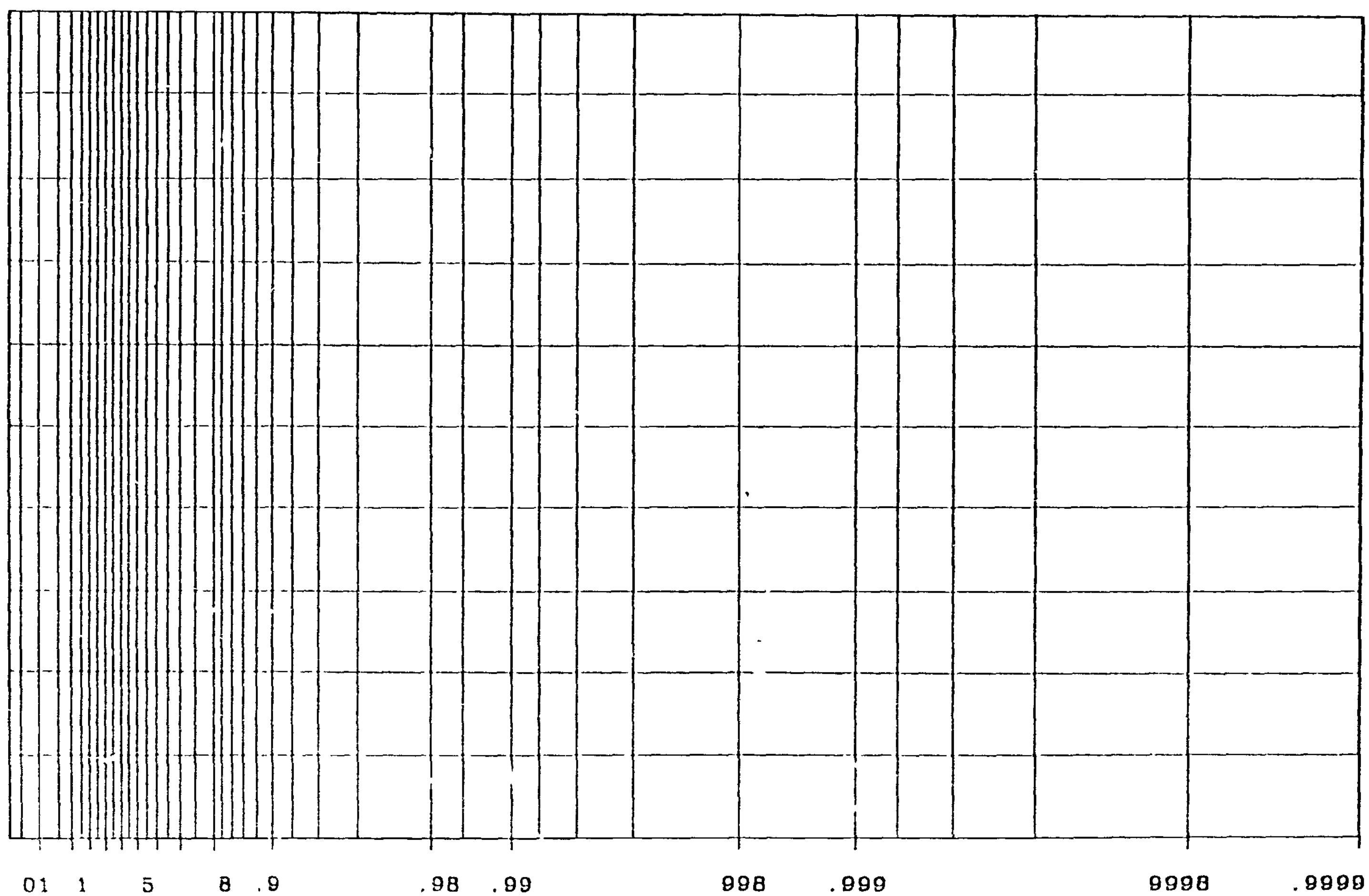
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.6



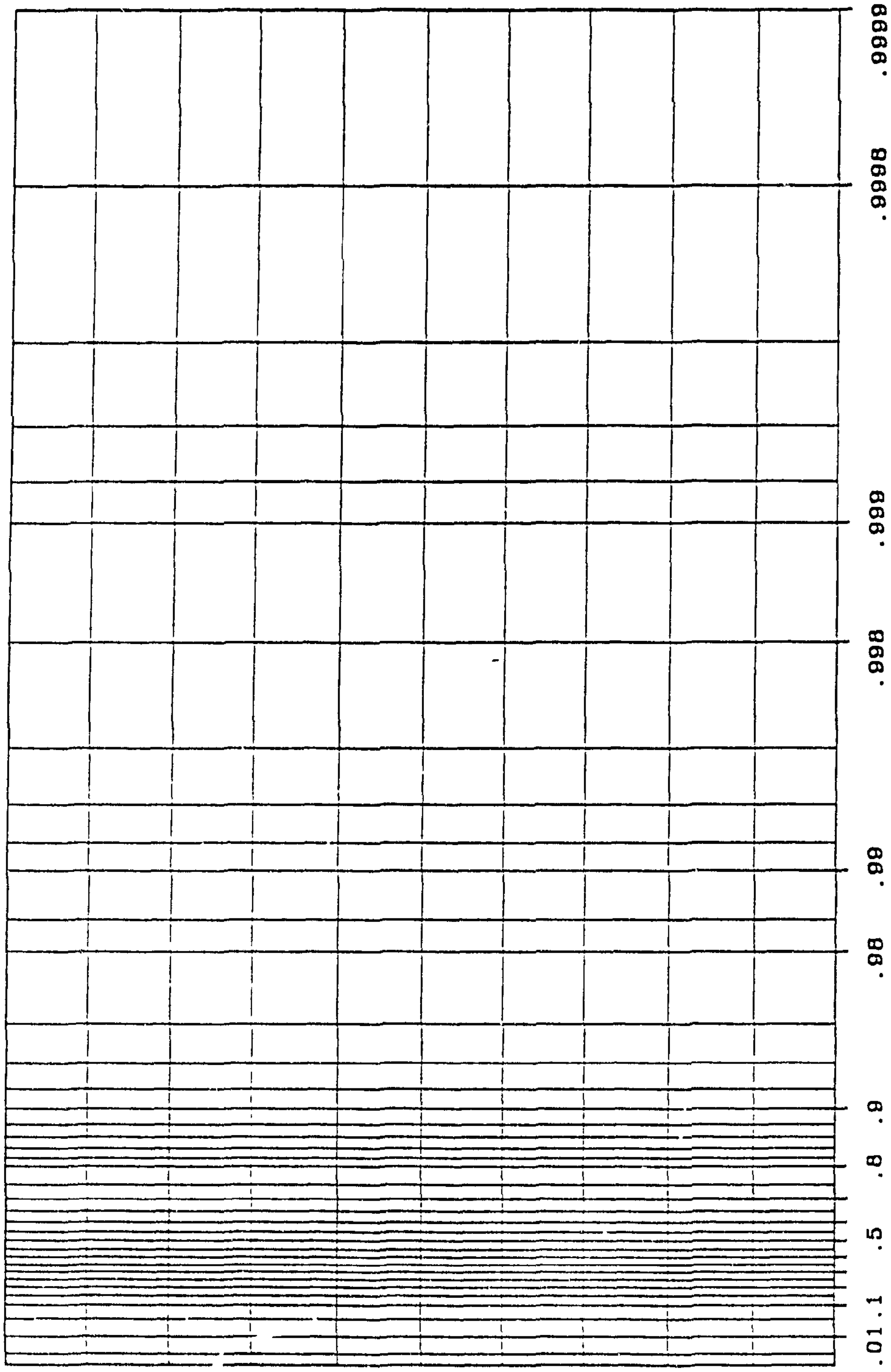
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.7



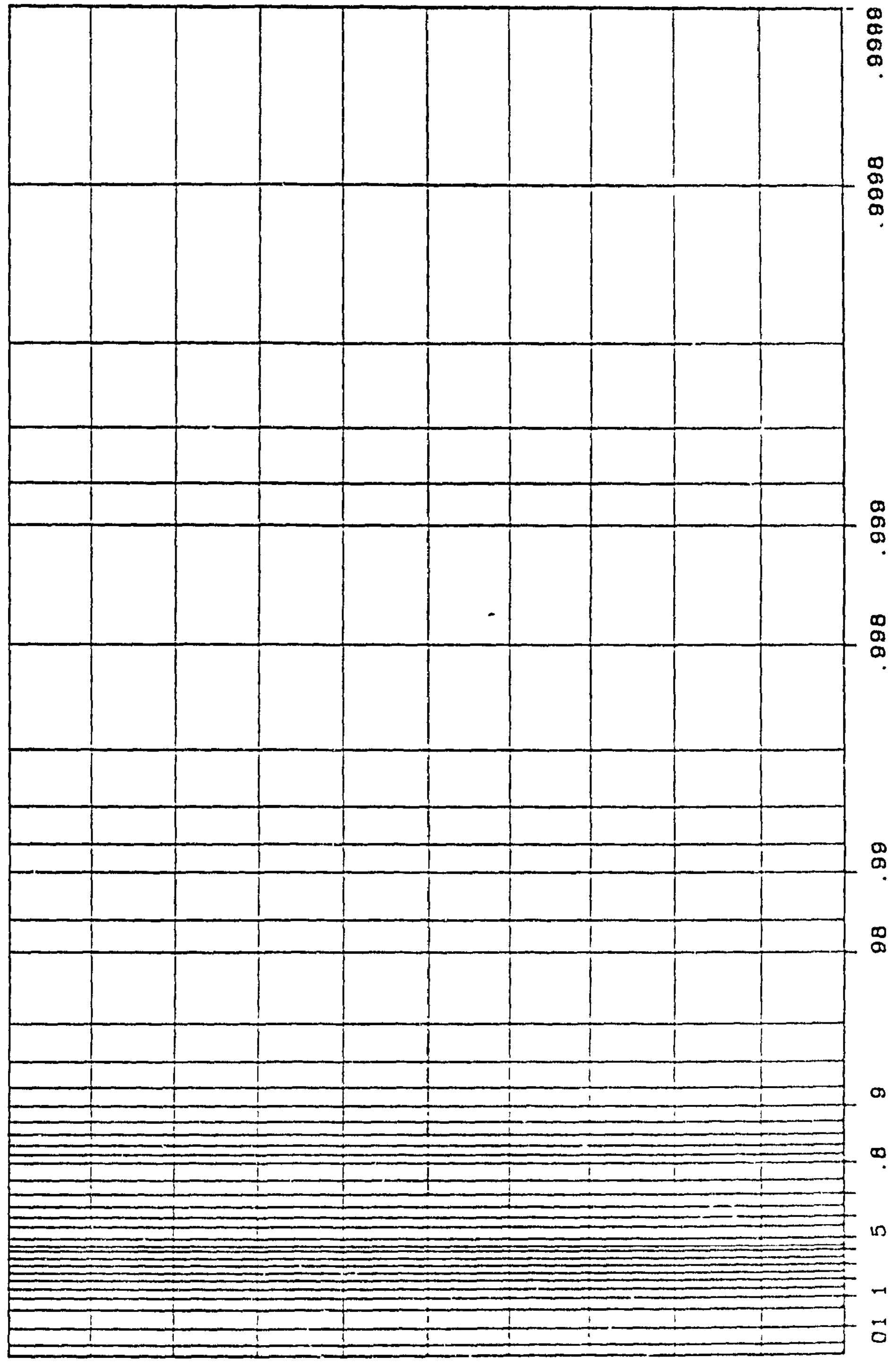
GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.8



GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 2.9



GEV PROBABILITY PAPER, SKEWNESS COEFFICIENT = 3.0



EV1 PROBABILITY PAPER