A Flood-Frequency Relation Based on Regional Record Maxima

GEOLOGICAL SURVEY PROFESSIONAL PAPER 434-F

Prepared in cooperation with the U.S. Atomic Energy Commission

A Flood-Frequency Relation Based on Regional Record Maxima

By P. H. CARRIGAN, JR.

STATISTICAL STUDIES IN HYDROLOGY

GEOLOGICAL SURVEY PROFESSIONAL PAPER 434-F

Prepared in cooperation with the
U.S. Atomic Energy Commission

UNITED STATES GOVERNMENT PRINTING OFFICE, WASHINGTON : 1971

UNITED STATES DEPARTMENT OF THE INTERIOR

ROGERS C. B. MORTON, *Secretary*

GEOLOGICAL SURVEY

W. A. Radlinski, ***Acting Director***

Library of Congress catalog-card No. 75-611783

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 - Price 40 cents (paper cover)

CONTENTS

ILLUSTRATIONS

TABLES

A FLOOD-FREQUENCY RELATION BASED ON REGIONAL RECORD MAXIMA

By P. H. Carrigan, Jr.

ABSTRACT

Flood intensity varies randomly in both time and space over a region. In flood-frequency analysis the regional variation is commonly related, through regression analysis, to variations in basin and climatic characteristics. Thus, the sample size is reduced to one time sample applicable to the entire region. This type of analysis virtually limits the recurrence-interval estimate for the maximum annual flood to the length of time for which records are available.

Techniques of analysis introduced in this paper increase the sample size by taking account of the random variations of flood intensity in both time and space. The magnitude of increase in the sample size is limited by the statistical dependence between annual flood records for the region; that is, the less the correlation between records the greater the effective sample size. As a result of increasing the sample size, the recurrence interval for the maximum event is increased; the increase may conceivably be as great as the number of station-years of record, if the records are independent.

The recurrence interval for the maximum annual flood in the region is estimated by a computer-simulation model. In order to apply the model, the assumption is made that the sets of concurrent records for a region are reducible to identical distributions bv use of an appropriate scaling factor.

The dependence among records is preserved by including in the model the sample correlation coefficients for all pairs of record sets. By means of a principal-component transform, the matrix of correlation coefficients is used in connection with a normal random-number generator to generate synthetic sets of normally distributed flood intensities equal in number and duration to observed sets of records. The maximum event from each synthetic record is found, and the maxima are ranked in descending order. The exceedance probabilities (reciprocals of recurrence intervals) of the ordered maxima are repeatedly determined to estimate the average normal exceedance probability for each order. The scheme of computing exceedance probabilities of known distribution for ordered maxima selected from synthetic records of known distribution provides an estimate of probability that is independent of the statistical distribution of the observed floods.

The construction of a flood-frequency relation for any station in the region can be accomplished by graphically relating the observed ordered maxima, reduced to a common base, to the corresponding estimated recurrence interval. The procedure of constructing a flood-frequency relation is illustrated by applying it to floods in the Big Lost River basin near the National Reactor Testing Station in southeastern Idaho.

INTRODUCTION

Current practices of regional flood-frequency analysis define regions in which the records of annual hydrologic extremes have, for the most part, identical distributions. The method of analysis is that of multiple regression, whereby floods of specific recurrence intervals or the statistical parameters of flood distributions are related to basin and climatic characteristics. The method composites the experience over a region but, except by extension based on the assumed distribution of events, it produces results that are limited virtually to the length of the period of sampling. Thus, no advantage is taken of the facts that spatial sampling to some degree is equivalent to time sampling and that some of the experience within the region may furnish information about the expectancy of events for a period much longer than the actual period of record. Work by Conover and Benson (1963) suggests that this information can be realized by appropriate treatment of combined records from several streamflow stations in the defined area. In the Conover and Benson method of analysis, the process is assumed to be ergodic, so that spatial sampling is equivalent to time sampling of extreme events. The result is to increase the size of the sample to nk events in a region (k=years of record, n=number of sets of record), provided the records are independently and identically distributed.

If the Conover and Benson method of regional flood-frequency analysis were applied to records of annual extremes, the records could be reduced to identical distributions by normalizing them to the magnitudes of the mean annual events at each streamflow station. However, the records for a region usually are not independent but are correlated. Hence, another method of analysis must be developed for treating the combined information from several identically distributed sets of records. This report presents a procedure to estimate exceedance probabilities of floods from records that are dependent and identically distributed. The maximum recurrence interval for the region, estimated by this procedure, lies between $k+1$ and $f(\rho)k+1$ where $f(\rho)$ is a function of the correlation between record sets ($1 \leq f(\rho) \leq n$) for the range in correlation coefficients ($0 \leq \rho \leq 1$).

The scheme of Conover and Benson for treating the combined information from independent and identically distributed records is to compute expected exceedance probabilities (independent of the distribution) for the ordered record maxima. This scheme is the key to developing regional frequency relations for dependent records.

Once a region having identically distributed sets of concurrent annual flood records is defined, the regional flood-frequency relation is developed as follows:

1. The records are reduced to a common base.
2. The maximum events from each reduced record are ordered.
3. The correlation matrix for the records is computed.
4. Computer routines developed in this study are used to provide, through simulation alogorithms that utilize the correlation matrix, estimates of exceedance probabilities associated with ordered maxima from n, number of k-year records.
5. The frequency relation is defined graphically by relating exceedance probabilities to the ordered maxima. Reversing the transformation process that reduces records to an identical distribution produces regionalized frequency relations for annual floods at individual stations.

After briefly illustrating the Conover and Benson method of obtaining exceedance probabilities for independent flood records, this report introduces the mathematical and computer-simulation foundations for estimating exceedance probabilities for dependent flood records. Then follow discussions on the application of the procedure to simple models which help to describe the influence of the number of records, correlation coefficients, and record length on exceedance probabilities; the development of regional flood-frequency relations in the Big Lost River basin, Idaho; and the extension of the procedure to incomplete and to serially correlated records. The complete computer program is presented at the end of the report.

ACKNOWLEDGMENTS

This study of flood-frequency relations was part of a research project on the statistical occurrence of extreme hydrologic events in relation to the design and selection of sites for nuclear reactor facilities. The work was done on behalf of the Division of Reactor Development and Technology, U.S. Atomic Energy Commission.

Early considerations for estimating the probability of occurrence of rare events favored analytical approaches. However, after discussions with W. J. Conover of Kansas State University, it was decided to discard methods relying on analytical approaches. U.S. Geological Survey personnel made the following contributions to the study: M. A. Benson recommended use of Monte Carlo techniques; discussions with N. C. Matalas resulted in much of the structuring in the Monte Carlo model; E. J. Gilroy drew attention to the use of a moving-average scheme to simulate multivariate distributions; and M. S. Hellman kindly supplied an improved computer algorithm for finding the eigenvectors of a correlation matrix.

CONOVER AND BENSON METHOD OF DETERMINING PROBABILITIES

Assume that n concurrent sets of k annual extremes are obtained from n identical and independent distributions. The maximum events, one from each record, are ordered from highest to lowest. The probability that another event f exceeds these ordered maxima

$\{f_i;\ i=1, 2, \ldots n\}$ is

$$P[f>f_i]=\sum_{m=0}^{i-1}\frac{n!}{(n-m)!k\prod_{j=0}^{m}\left(n+\frac{1}{k}-j\right)}$$

for $i=1, 2, \ldots n$ (Conover and Benson, 1963, p. E159).

Example.—The following three sets of four random events $\{x_{ij};\ i=1, 2, 3 \text{ and } j=1, 2, 3, 4\}$ are from independent and identical distributions:

x_{ij}
[$n=3$, $k=4$]

j	i = 1	2	3
1	3	69	3
2	38	24	48
3	17	61	60
4	32	30	83

The ordered record maxima $\{f_i;\ i=1,\ 2,\ 3\}$ are, respectively, 83, 69, and 38. The probability that another event f exceeds the second highest maxima is

$$P[f>69]=\sum_{m=0}^{2-1}\frac{3!}{(3-m)!k\prod_{j=0}^{m}\left(3+\frac{1}{4}-j\right)}$$

$$=\frac{3!}{3!4\left(3+\frac{1}{4}\right)}+\frac{3!}{2!4\left(3+\frac{1}{4}\right)\left(3+\frac{1}{4}-1\right)}$$

$$=\frac{1}{13}+\frac{3}{29.25}=0.180.$$

METHOD FOR DETERMINING PROBABILITIES FROM DEPENDENT RECORDS

The probability of another random event f exceeding one of the ordered record maxima f_i where the records are from dependent distributions cannot be analytically determined. The probability can be determined, however, through use of Monte Carlo simulation techniques. In this application of simulation for dependent records, it is assumed that the probability of another random event exceeding an ordered record maximum is independent of the identical distributions for which the records of random events are a sample. Such an assumption has been verified by Conover and Benson (1963) and proven by Conover (1965) for independent random events.

The simulation algorithm, which incorporates a normal random-number component having zero mean and unit variance, generates n sets of k events that simulate the dependence among n records of k hydrologic extremes. The maximum event y_i from each set of normally distributed variables $\{x_{i1},\ x_{i2},\ \ldots\ x_{ik};\ i=1,\ 2,\ \ldots\ n\}$ is found, and the maxima are ordered from highest to lowest, so that $y_1>y_2>\ \ldots >y_n$. All y_i are from normal distributions having zero mean and unit variance, so their exceedance probabilities are given by

$$P[y>y_i]=\frac{1}{\sqrt{2\pi}}\int_{y_i}^{\infty}e^{\frac{-y^2}{2}}\,dy,\ \text{for } i=1,\ 2,\ \ldots n.$$

By repeated generation of this simulated set of records, estimates of the average exceedance probabilities for the ordered maxima are obtained.

SIMULATION MODEL

The simulated sets of records are generated by the normal multivariate model

$$X=B\epsilon,$$

where X=matrix of n groups of k events,

$\epsilon=n\times k$ matrix of independent normal random numbers with zero mean and unit variance, and

$B=E\lambda=n\times n$ principal-component transform matrix for the records of annual extremes, in which

E=eigenvector matrix associated with the correlation matrix R obtained from n records of k hydrologic events, and

λ=diagonal matrix whose n diagonal elements are the square roots of the eigenvalues for E.

(See Morrison, 1967, p. 221–247.) This model is based on techniques of synthetic hydrology developed by Matalas (1967, p. 940).

After generation of X, the n maxima are selected and ordered and the exceedance probabilities are computed. The algorithm for computing normal exceedance probabilities is that of Zelen and Severs (1965, p. 932, eq. 26.2.19).

FLOW CHART FOR SIMULATION

The average exceedance probability

$$p_i=P[y>y_i;\ y_1>y_2>\ldots>y_n]$$

$$=\frac{1}{1{,}000}\sum_{m=1}^{1{,}000}\frac{1}{\sqrt{2\pi}}\int_{y_{i,m}}^{\infty}e^{\frac{-y^2}{2}}\,dy$$

was estimated through use of a digital computer. The main program flow chart is shown in figure 1. Routines for computer input-output listings of matrices, matrix operations, and generation of random numbers are adaptations of subroutines developed by IBM (1968) for use in their 360-series computer systems. Formal declarations of most subroutines have been eliminated from the program to increase efficiency. A complete listing of the Fortran language program for the computations is included in the computer program at the end of this report. The parameters n and k may each be as large as 100 using the present configuration of the IBM 360/65 computing system (380,000 bytes available storage) operated by the U.S. Geological Survey.

INFLUENCE OF INTERRECORD CORRELATION ON EXCEEDANCE PROBABILITY

An n-record correlation matrix contains $n(n-1)/2$ correlation coefficients. If $n>3$, problems in management of matrix elements begin to obscure the study of the influence of interrecord correlation on exceedance probabilities. Hence, most simulations of exceedance probabilities were derived from the simple, yet informative, 3-record correlation matrices.

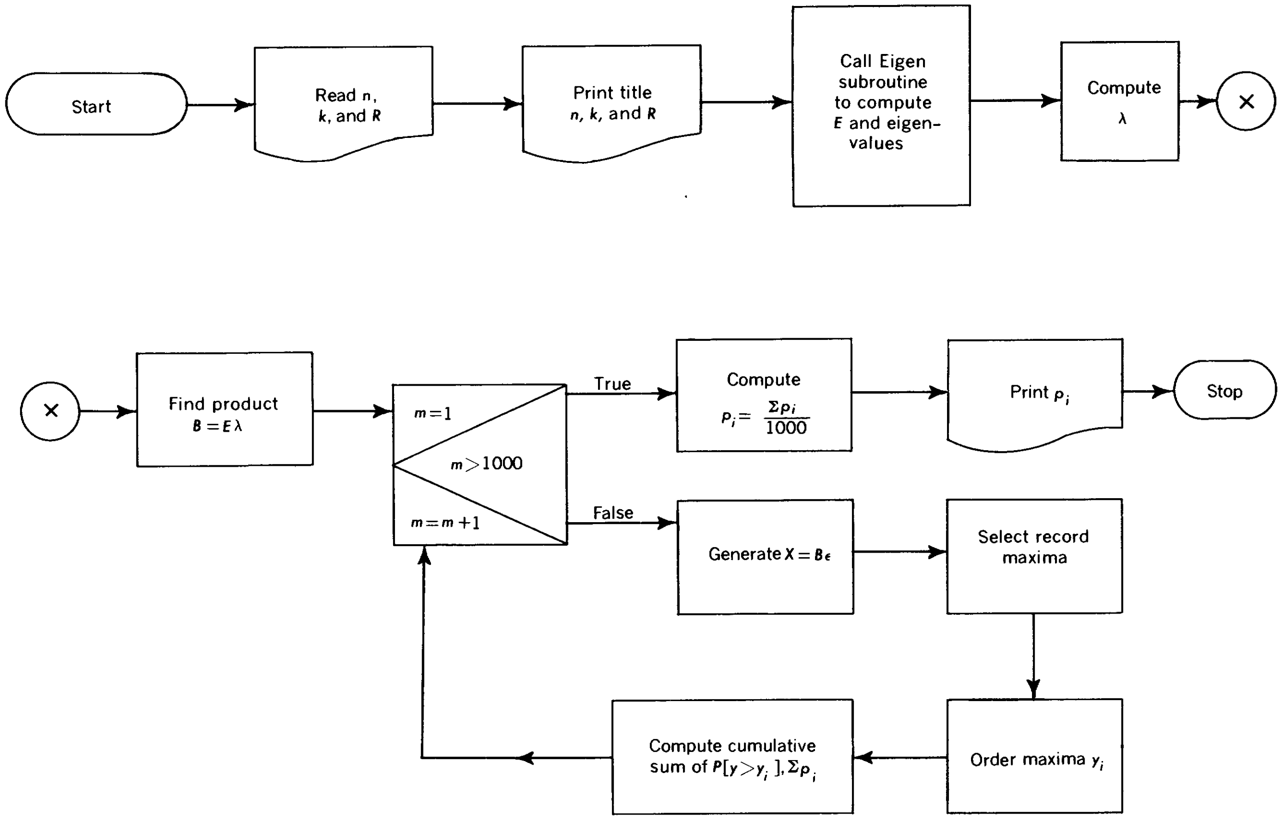

FIGURE 1.—Flow chart for program to simulate $\{p_i(k); i=1, 2, \ldots n\}$.

EQUICORRELATION MATRIX

Using the equicorrelation matrix R (all correlation coefficients equal)

$$\begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$$

for $0 \geq \rho \geq 1$, the expected probabilities $\{p_i;\ i=1, 2, 3\}$ were simulated for $k=10, 20, \ldots 50$.

As shown in figure 2, p_i for any order i changes monotonically, as had been anticipated—very gradually from the theoretical probability at $\rho=0$ (three independent records) to about $\rho=0.8$, and then rapidly to the common theoretical probability at $\rho=1$ (three identical records). Comparison of the simulated probability to the theoretical probability at $\rho=0$ and at $\rho=1$, shown in figure 2, indicates that 1,000 iterations is sufficient for estimates of p_i. Similar monotonic variations in $\{p_i;\ i=1, \ldots n\}$ between $\rho=0$ and $\rho=1$ occurred for $n=2$ and $n=5$ where $k=10, 20, \ldots 50$.

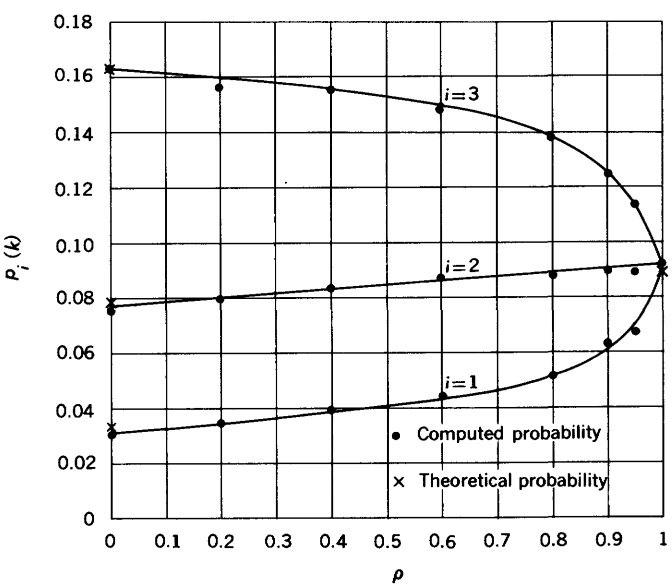

FIGURE 2.—Variation of $\{p_i(k);\ i=1, 2, 3\}$ with ρ for $k=10$.

VERIFICATION THAT p_i IS DISTRIBUTION FREE

The probabilities p_i for $n=2$, $k=10$, and $\rho=0, 0.1, 0.2, \ldots 1.0$ were simulated by the previously described normal variate generating model. The data for these simulation runs are plotted with their related curves in figure 3.

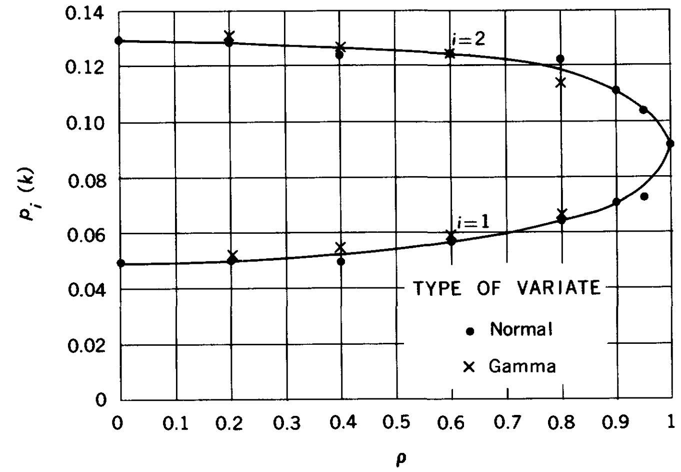

FIGURE 3.—Variation of $p_1(k)$ and $p_2(k)$ with ρ for bivariate gamma and normal distributions of annual extremes; $k=10$.

Another model for simulating a multivariate distribution $\{z_{i,j};\ i=1, 2, \ldots n \text{ and } j=1, 2, \ldots k\}$ is to let

$$z_{i,j}=\sum_{l=1}^{m} u_{(i-1)d+l,\,j},$$

in which $u_{(i-1)d+l,\,j}$=exponentially distributed independent random numbers having unit mean and unit variance, and

$z_{i,j}$=random numbers following a marginal gamma distribution having mean m and variance m, for which the correlation coefficient is

$$\rho(z_i, z_{i+1})=\frac{m-d}{m}.$$

Using this simulation model the mean exceedance probabilities $\{p_i;\ i=1, 2\}$ for the ordered maxima from $z_{i,j}$ (for $m=5$, $n=2$, and $k=10$) were estimated for $\rho=0.2$, 0.4, 0.6, and 0.8 ($d=4$, 3, 2, 1); the coefficient of skew for $z_{i,j}$ was 0.89 (the mean coefficient for a 50-record sample of annual floods was 0.9).

Results for this bivariate gamma method of estimating p_i are shown in figure 3. Within the experimental error experienced, the estimates appear to confirm the assumption that p_i is distribution free.

VARIATION OF p_i WITH k AND n

The probability p_i varies smoothly and monotonically with either k or n. (See figs. 4 and 5 for examples.) The variation with k of p_i for $0<\rho<1$ and n constant is similar to the variation of the theoretical probability ($\rho=0$ and 1) with k. The relation between p_i and n for constant k shifts from being highly curvilenar for ρ near zero to being nearly horizontal for ρ near 1.

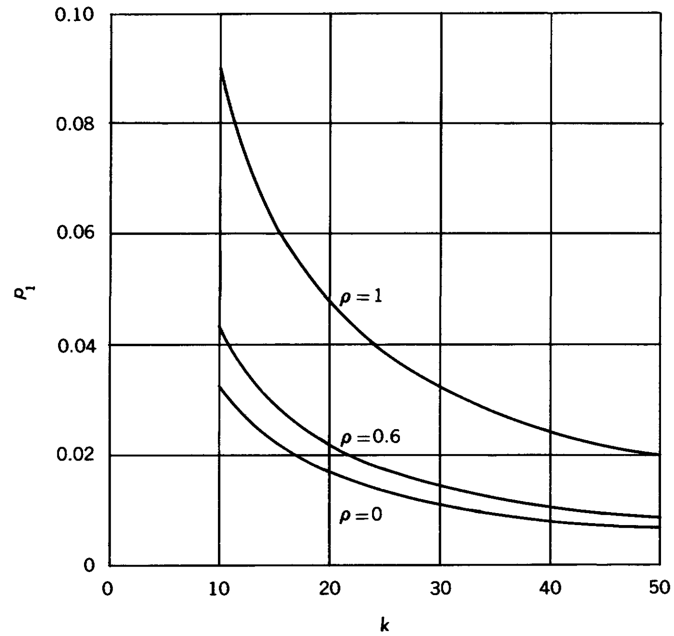

FIGURE 4.—Variation of p_1 with k for $n=3$ and $\rho=0$, 0.6, and 1.

GENERAL CORRELATION MATRIX

The relative insensitivity of p_i to changes in ρ (fig. 2) suggests the use of a simple parameter

$$\bar{\rho}_n=\frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\rho_{ij}.$$

This parameter weights the central tendency of correlation coefficients in the general correlation matrix (ρ_{ij} are not all equal) and replaces ρ in estimating p_i' (the prime referring to a general matrix) from relations such as those in figure 2. Results of tests showed that all the variation in p_i' could not be accounted for by variations in $\bar{\rho}_n$. A parameter that described the dispersion of the coefficients in R about $\bar{\rho}_n$, such as a mean deviation

$$\alpha_n=\frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}|\rho_{ij}-\bar{\rho}_n|$$

or a mean square deviation

$$\delta_n^2=\frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(\rho_{ij}-\bar{\rho}_n)^2,$$

seemed necessary to explain variations in p_i' more fully.

Experiments using third-order correlation matrices were performed to define the variation of $\{p_i';\ i=1, 2, 3\}$ with increases in α_3 or δ_3^2 for constant $\bar{\rho}_3$ (see fig. 6). Under these conditions, p_i' was found to increase monotonically for $i=1$ and $i=2$ and to decrease monotonically for $i=3$.

Under the constraints that R be positive, semidefinite (see Morrison, 1967, p. 60), $\bar{\rho}_n$ be constant, and $0 \leq \rho_{ij} \leq 1$, a maximum α_n exists. For $\bar{\rho}_3 \geq 1/3$, α_3 is maximized if one of the $\rho_{ij}=1$, say $\rho_{12}=1$, and the other coefficients are as follows:

$$\rho_{13}=\rho_{23}=\frac{3\bar{\rho}_3-1}{2};$$

hence, $\rho_{13}=\rho_{23}=\rho$ for $n=2$. For $\bar{\rho}_3<1/3$, α_3 is maximized if $\rho_{12}=3\bar{\rho}_3$ and $\rho_{13}=\rho_{23}=0$. The maximum is

$$\alpha_{3\,\max}=2(\rho_{12}-\bar{\rho}_3).$$

These conditions for which α_3 is a maximum obtain the maximum for δ_3^2. The upper limit for p_i' at a maximum δ_3^2 is shown in figure 7.

The variation of p_i' with $\delta_3{}^2$, shown in figure 6, brings out a point about the maxima. For $i=1$, p_1' increases to the limit p_1 for $n=2$; for $i=3$, p_3' decreases to the limit p_2 for $n=2$; but p_2' does not reach either limit. If $\rho_{12}=1$, then only two ordered maxima occur, one in records 1 and 2 and one in record 3. One of the maxima may be first order and the other may be third order; the second-order maximum does not really exist. It would appear in figure 6 that p_2' is approaching the limit for $\rho=1$ ($p_i=0.0909$ for $k=10$). Results from two other sets of tests—for $k=50$ and $n=3$ and for $k=10$ and $n=5$—indicate that the variations of p_i' with $\bar{\rho}_n$ and $\delta_3{}^2$ are much like those shown in figures 6 and 7.

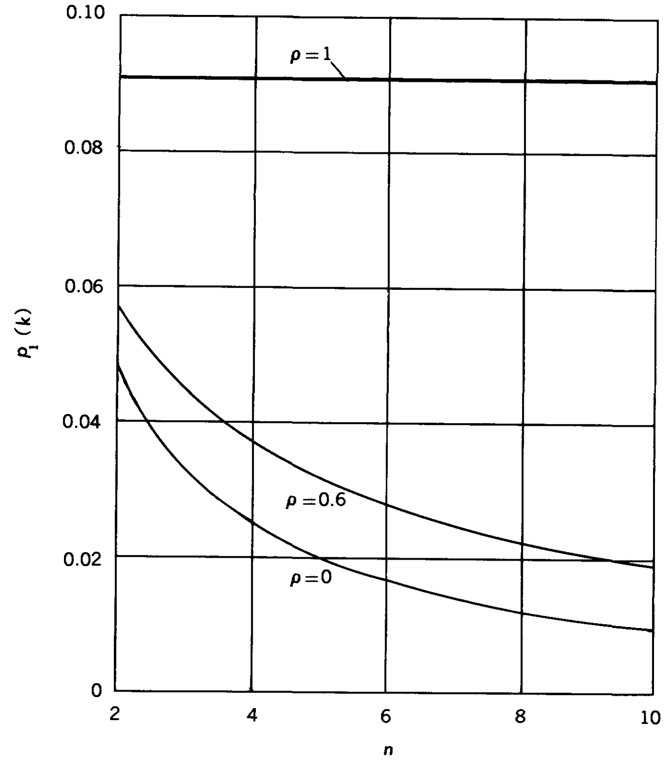

FIGURE 5.—Variation of $p_1(k)$ with n for $k=10$ and $\rho=0$, 0.6, and 1.

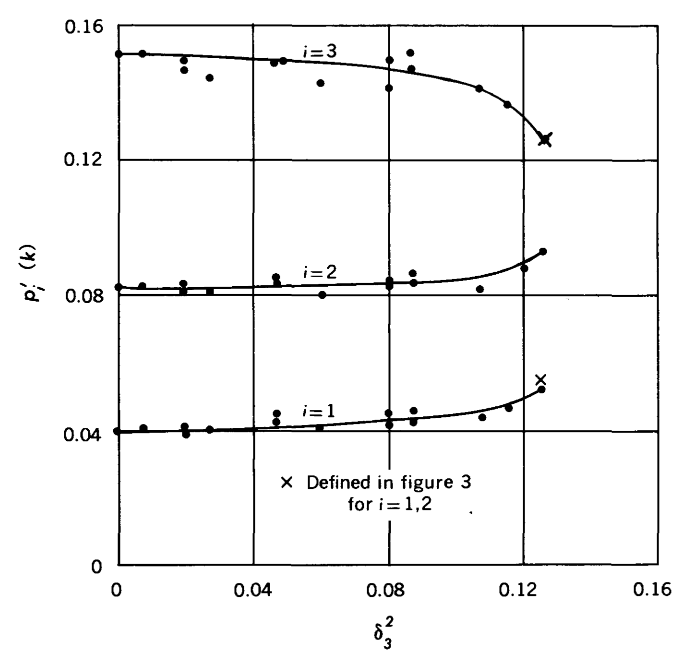

FIGURE 6.—Effect of mean square deviation $\delta_3{}^2$ on $\{p_i'(k);\ i=1,\ 2,\ 3\}$ for $k=10$ and $\rho=0.5$.

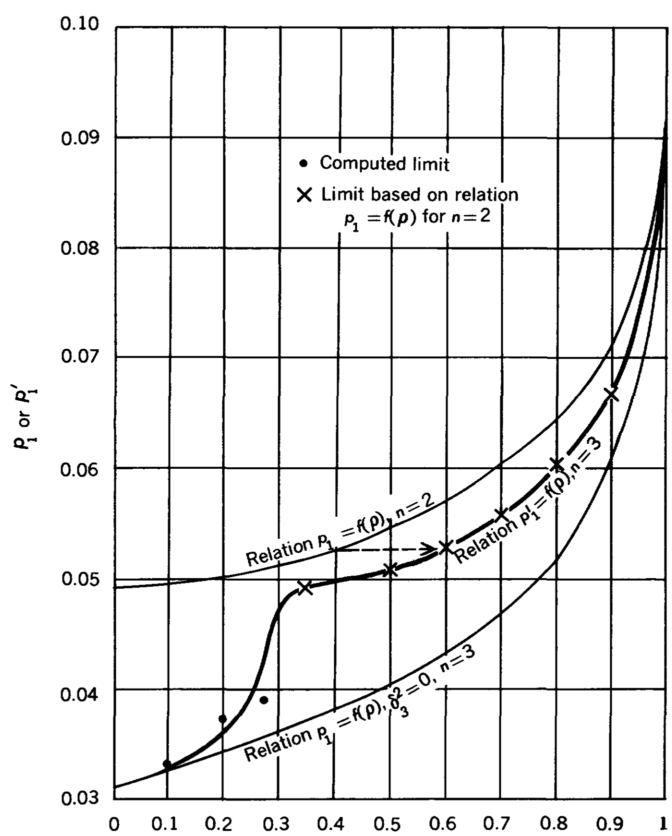

FIGURE 7.—Upper limit of p_1' for $\delta^2_{3\max}$. Arrow indicates shift in plotting position.

An attempt to relate p_i' fully to a practical range in ρ_n, δ_n^2, and k appears to be economically prohibitive because of computer cost. Rather, these results suggest that p_i' should be estimated for each specific correlation matrix and period of record.

APPLICATION OF THE METHOD IN THE BIG LOST RIVER BASIN, IDAHO

Frequencies of floods on the Big Lost River near the National Reactor Testing Station in southeastern Idaho are required in connection with an investigation for the Atomic Energy Commission of the potential hazard from extreme floods. The first step of the flood-frequency analysis is to correlate annual flood records in this part of Idaho with distributions that are identical save for some scaling factor.

From a report on regional flood-frequency relations (Thomas and others, 1963), 48 candidate records were selected for analysis. The coefficient of variation, square of the coefficient of skew, kurtosis, and the standard error of these moments (see Kendall and Stuart, 1963, p. 228–233) were computed for each candidate record. The records of annual floods were assumed to have distributions identical with the selected base record (Big Lost River below Mackay Reservoir) if they satisfied the following criteria:

1. The confidence limits of each moment for the record, defined by its standard error, were within the limits of the same moment for the base record.
2. The median test (Ostle, 1963, p. 473) indicated the acceptability of the hypothesis of identical distribution with the base record at the 5-percent level of significance.

These two criteria were met by 15 records (see table 1).

The next step in the analysis is to reduce the selected records to identical distributions. This is accomplished by normalizing them to the computed mean annual flood.

TABLE 1.—*Name of station and period of record for floods in central Idaho having identical distributions*

Name	*Period of record*
Beaver Creek at Spencer	1950–63
Lime Creek near Bennett	1946–56
Challis Creek near Challis	1944–63
Combined discharge, Big Wood River and Slough at Hailey	1936–57
Big Lost River below Mackay Reservoir near Mackay	1919–67
Surface inflow to Mackay Reservoir near Mackay	1919–67
Big Lost River at Howell Ranch near Chilly	1920–67
Salmon River at Salmon	1920–67
Valley Creek at Stanley	1921–67
Salmon River below Yankee Fork near Clayton	1922–67
Salmon River near Challis	1929–67
Big Lost River at Wild Horse near Chilly	1944–67
South Fork Boise River at Anderson Ranch Dam	1946–67
Big Wood River near Ketchum	1948–67
Birch Creek near Reno	1950–63

At this point in the analysis, the objective is to select those of the 15 concurrent records whose combination maximizes nk to realize the dual advantage of obtaining large recurrence intervals and relating outstanding events in the region to these intervals. The records for which the station years are maximized are listed in table 2.

The correlation matrix R for the six station records

$$\begin{array}{c} \\ 1\\2\\3\\4\\5\\6 \end{array}\begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{bmatrix} 1 & 0.379 & 0.418 & 0.239 & 0.256 & 0.415 \\ & 1 & .941 & .711 & .728 & .802 \\ & & 1 & .677 & .711 & .823 \\ & & & 1 & .971 & .914 \\ & & & & 1 & .934 \\ & & & & & 1 \end{bmatrix} \end{array}$$

and a k value of 46 were inputs to the computer program for which the maxima $\{p'_i;\ i=1, \ldots 6\}$ were computed.

The ordered reduced variate maxima, p'_i, and the associated flood magnitude in the Big Lost River below Mackay Reservoir are listed in table 3. The flood magnitude is equal to the product of the ordered reduced variate maximum and the mean annual flood at this station. The recurrence interval, T, is $1/p_i'$.

The rare-flood reduced variates are plotted against their recurrence intervals in figure 8. The annual floods, in reduced variate form, are also plotted in this figure to define the relation between flood magnitude and flood frequency for the Big Lost River below Mackay Reservoir.

DISCUSSION

The procedure for determining the probability of occurrence of rare flood events may be extended to other kinds of dependent records, only a few alterations being required in the computer program. If the records follow a first-order Markov process as well as being cross correlated, then the simulation model is

$$X_{j+1}=AX_j+B\epsilon_{j+1},$$

where $X_j=n\times 1$ matrix of the jth events ($j=1, 2, \ldots k-1$) in n records,

$A=R_1R_0^{-1}$,

$BB^T=R_0-R_1R_0^{-1}R_1^T$,

$B=E\lambda$, where E is the eigenvector matrix and λ is the diagonal matrix whose elements are the square roots of eigenvalues,

R_0=lag-zero correlation matrix,

R_1=lag-one correlation matrix, and

$\epsilon_j=n\times 1$ matrix of random normal numbers with zero mean and unit variance.

(See Matalas, 1967, for more detail.)

TABLE 2.—*List of annual flood records for analysis of flood-event frequency in the Big Lost River basin, Idaho, for the concurrent period 1922–67*

Station	Name	Coefficient of variation		Square of skew coefficient		Kurtosis		Maximum flood	
		Value	Standard error	Value	Standard error	Value	Standard error	Cubic feet per second	Reduced variate
1	Big Lost River below Mackey Reservoir near Mackay	0. 417	0. 033	0. 397	0. 289	2. 57	0. 49	2, 640	1. 40
2	Big Lost River at Howell Ranch near Chilly	. 394	. 035	. 170	. 165	2. 49	. 43	4, 420	2. 13
3	Surface inflow to Mackay Reservoir near Mackay	. 491	. 052	. 0493	. 089	2. 36	2. 33	3, 240	2. 23
4	Valley Creek at Stanley	. 474	. 048	1. 16	. 71	3. 46	1. 05	2, 000	2. 07
5	Salmon River below Yankee Fork near Clayton	. 354	. 036	. 217	. 317	2. 80	. 63	10, 300	1. 99
6	Salmon River at Salmon	. 380	. 035	. 153	. 144	2. 64	. 34	16, 500	1. 90

TABLE 3.—*Plotting positions for the relation between the rare-flood magnitude and rare-flood frequency at Big Lost River below Mackay Reservoir*

Order	Exceedance probability	Recurrence interval (years)	Reduced variate maximum	Flood magnitude (cfs)
1	0. 00662	151. 1	2. 23	4, 190
2	. 0115	86. 9	2. 13	4, 000
3	. 0159	62. 7	2. 07	3, 890
4	. 0218	45. 8	1. 99	3, 750
5	. 0298	33. 5	1. 90	3, 580
6	. 0432	23. 2	1. 40	2, 640

Through special investigations for historical floods, the period of record for the maximum event may be extended from k to $k+l$ years, where l is the period of extension. The estimation of p_i' for the $(k+l)$-event record may be undertaken provided the correlation matrix defined by k-event records is considered applicable to the longer period and the n maxima for the $(k+l)$-event records are defined.

A somewhat analogous condition to the problem of analyzing historical floods occurs when some records

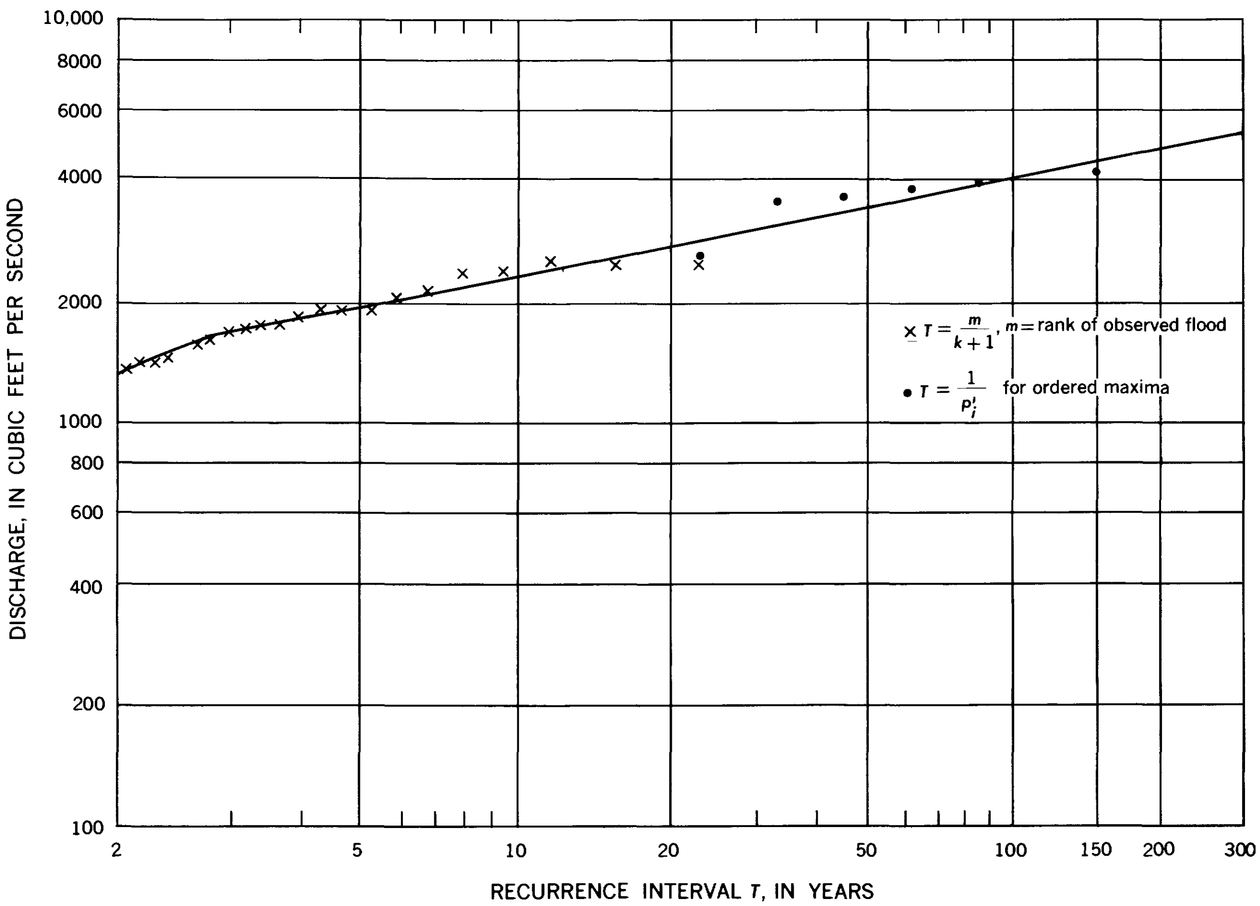

FIGURE 8.—Flood-frequency relation for Big Lost River below Mackay Reservoir based on the probability of occurrence of rare events; $n=6$, $k=46$, and $\bar{\rho}_6=0.661$.

covering a period k have missing events. Assuming known record maxima, some uncertainty about the correlation matrix is introduced for records of this kind. If the matrix is positive semidefinite, a test of the sensitivity of p_i' to changes in individual correlation coefficients would seem appropriate. The variation in the lengths of record may be so large that the correlation matrix is inconsistant. Some suggestions for developing consistent matrices were given by Fiering (1968).

The objective in the analytical procedure for developing a regional flood-frequency relation is to maximize the number of station years (increasing the sample size) which tends to minimize p_i'. Because p_i' varies almost directly with $1/nk$, the objective can be attained by seeking a large number of short records or a few long records. The objection to the use of short records is that the sample may be nonrepresentative in comparison with norms established by much longer sequences of events. Unless statistical comparisons between long and short records indicate nonhomogeneity, no reason exists for rejecting short records. Sampling errors will be larger for short records, but this disadvantage may be outweighed by the substantially larger values of nk that will be provided by the shorter than the longer records.

CONCLUSIONS

The technique of regional flood-frequency analysis that has been introduced takes advantage of the random variation of flood intensities in both time and space. Previous methods of analysis are virtually limited to the consideration of variation in time. A result of applying this new technique is to increase the recurrence interval for the maximum flood event; because of correlation between records, it is unlikely that this recurrence interval would approach the number of station-years of record available.

Using Monte Carlo simulation techniques, the exceedance probability for rare flood events may be estimated by the simultaneous consideration of concurrent records of annual extremes which are dependent and identically distributed. The method for obtaining rare-event probabilities is to generate sets of normal multivariates equal in number and duration to the observed sets of records through the simulation model

$$X=B\epsilon$$

to order the simulated record maxima and to compute the normal exceedance probabilities of the ordered maxima. This scheme of generating synthetic variates preserves the observed dependence in records.

In applying the procedure of estimating exceedance probabilities to regional flood-frequency problems it is assumed that a means of defining the region of identical distributions is at hand, that the estimate of exceedance probability is independent of the underlying statistical distribution of observed floods, and that the sample correlation matrix is a reasonable estimate of the true matrix. Indications are that the probability estimates are independent of the distribution of observed floods.

If the correlation coefficients for the records are all equal, the rare-event probabilities vary in a predictable manner with changes in numbers of records, period of record, and magnitude of the coefficient. If the correlation matrix is a mixture of coefficients, variations in the probabilities are a function not only of an equivalent mean coefficient, the number of records, and the period of record, but also of the deviation of the coefficients about the mean coefficient. Even though the probabilities associated with a general correlation matrix also appear to vary predictably with these variables, any general solution would not seem economically justified because of excessive computer costs (about $2nk$ seconds per run). It would seem to be preferable in the interest of economy to estimate the rare-event exceedance probabilities for each set of data under analysis.

A set of criteria for selecting records having identical distributions has been identified in this report. Other sets of criteria for making this selection are possible, and the most definitive must be sought. The simulation program is equally applicable to many kinds of rare hydrologic and hydrometeorologic events such as precipitation, floods, and temperatures.

REFERENCES

Conover, W. J., 1965, A k-sample model in order statistics: Annals Math. Statistics, v. 36, no. 4, p. 1223–1235.

Conover, W. J., and Benson, M. A., 1963, Long-term flood frequencies based on extremes of short-term records, *in* Geological Survey research, 1962: U.S. Geol. Survey Prof. Paper 450-E, p. E159–E160.

Fiering, Myron B., 1968, Schemes for handling inconsistent matrices: Water Resources Research, v. 4, no. 2, p. 291–297.

International Business Machines, 1968, System/360 scientific subroutine package (360-CM-03X) version III, programmer's manual: White Plains, N.Y., International Business Machines, 454 p.

Kendall, M. G., and Stuart, Alan, 1963, Distribution theory Volume 1 *of* The advanced theory of statistics [2d ed.]: New York, Hafner, 433 p.

Matalas, N.C., 1967, Mathematical assessment of synthetic hydrology: Water Resources Research, v. 3, no. 4, p. 937–945.

421-897 O - 71 - 2

Morrison, D. F., 1967, Multivariate statistical methods: New York, McGraw-Hill, 338 p.

Ostle, Bernard, 1963, Statistics in research [2d ed.]: Ames, Iowa, Iowa, Iowa State Univ. Press, 585 p.

Thomas, C. A., Broom, H. C., and Cummans, J. E., 1963, Magnitude and frequency of floods ln the United States—Part 13, Snake River basin: U.S. Geol. Survey Water-Supply Paper 1688, 250 p., 6 pls.

Zelen, Marvin, and Severs, M. A., 1965, Probability functions, *in* Abramowitz, Milton, and Stegren, I. A., eds., Handbook of mathematical functions: Natl. Bur. Standards Appl. Math. Ser. 55, p. 925–995.

COMPUTER PROGRAM

```
C                                 ********************
C                                 **                **
C                                 ** PROGRAM RAREVENT **
C                                 **                **
C                                 *********************
C
C                                     *           *
C                                     * DISCLAIMER *
C                                     *           *
C
C        ALTHOUGH PROGRAM HAS BEEN TESTED BY ITS CONTRIBUTOR, NO
C        WARRANTY, EXPRESSED OR IMPLIED, IS MADE BY THE CONTRIBUTOR OR
C        THE GOVERNMENT, AS TO THE ACCURACY AND FUNCTIONING OF THE
C        PROGRAM AND RELATED PROGRAM MATERIAL, NOR SHALL THE FACT OF
C        DISTRIBUTION CONSTITUTE ANY SUCH WARRANTY, AND NO
C        RESPONSIBILITY IS ASSUMED BY THE CONTRIBUTOR OR THE
C        GOVERNMENT, IN CONNECTION THEREWITH.
C
C
C        PURPOSE
C           COMPUTES ESTIMATE OF EXPECTED EXCEEDANCE PROBABILITY OF
C           ORDERED COLUMN MAXIMA FROM N CROSS-CORRELATED, CONCURRENT
C           RECORDS OF K EVENTS, WHERE THE RECORDS ARE SIMULATED
C
C        DESCRIPTION OF VARIABLES
C           INPUT
C              N    - NUMBER OF RECORDS. MUST BE 20 OR LESS.
C                     CAPACITY OF THE PROGRAM CAN BE INCREASED FROM N = 20
C                     AND K = 100 BY CHANGING DIMENSION STATEMENTS.
C                     ASSUMING K = 100, N CAN BE AS GREAT AS   100 WITH
C                     PRESENT CONFIGURATION OF THE U.S. GEOLOGICAL
C                     SURVEY COMPUTER SYSTEM.
C              K    - NUMBER OF EVENTS IN A RECORD. MUST BE 100 OR LESS.
C              IX   - SEED FOR NORMAL RANDOM NUMBER GENERATOR. FORMAT
C                       IS I4.
C              RO   - CORRELATION MATRIX (READ IN BY SUBROUTINE MTRXIN)
C              IS   - CODE FOR STORAGE MODE
C                         0 - GENERAL
C                         1 - SYMMETRIC
C                         2 - DIAGONAL
C              FORMATS (PER DATA SET)
C                   N,K - 2I4
C                   RO   - 7F10.0, LISTING BY ROWS, STARTING WITH
C                              DIAGONAL ELEMENT AND CONTINUING WITH
C                              ELEMENTS TO ITS RIGHT.
C                   9    - I1 (INDICATES END OF DATA SET)
C           OUTPUT
C              IER - ERROR CODE FOR INPUT DATA CARDS OF CORRELATION
C                       MATRIX
C                         0 - NO ERROR
C                         1 - INCORRECT NUMBER OF CARDS
C              N    - NUMBER OF RECORDS
C              K    - NUMBER OF EVENTS IN RECORD
C              RO   - CORRELATION MATRIX
C              B1   - DIAGONAL MATRIX CONTAINING SQUAREROOTS OF
C                       EIGENVALUES
C              R    - EIGENVECTOR MATRIX
C              NSD - LEAST NUMBER OF SIGNIFICANT DIGITS IN ELEMENT OF
C                       CHARACTERISTIC EQUATION. IF 0, CHECK STRUCTURE OF
C                       CORRELATION MATRIX
C              P(J)- EXPECTED EXCEEDANCE PROBABILITY FOR JTH-ORDER
C                       COLUMN MAXIMUM
C
```

```
C        SUBROUTINES AND FUNCTION SUBPROGRAMS
C           MTRXIN
C           MTRXOT
C           EGN
C           GMPRD
C           MAX
C           MIN
C           LISD
C
C        METHOD
C           REFER TO P.H.CARRIGAN, JR., 1971, A FLOOD FREQUENCY RELATION
C           BASED ON REGIONAL RECORD MAXIMA, U.S.GEOL.SURVEY PROF.
C           PAPER 434-F
C
C        REMARKS
C           EXECUTION TIME ABOUT 2NK SECONDS PER DATA SET. FORT PLUS
C           LKED TIME LESS THAN 15 SECONDS
C           MAXIMUM LINES OF OUTPUT = 800 + 205*(NO. OF DATA SETS)
C
      DIMENSION X(20,100),Y(20),P(20),JUNK(20)                          #1.0000
      DOUBLE PRECISION D1,D2,D3,D4,D5,D6,U,RO(210),R(400),B(400),B1(400)#1.0010
      DOUBLE PRECISION X1(20),E(20),RBC(210)                            #1.0020
      DATA D1,D2,D3/0.049867347,0.0211410061,0.0032776263/              #1.0030
      DATA D4,D5,D6/0.0000380036,0.0000488906,0.0000053830/             #1.0040
  101 FORMAT(I4)                                                        #1.0050
  102 FORMAT(2I4)                                                       #1.0060
  201 FORMAT('1',12X,'EMPIRICAL EXCEEDANCE PROBABILITIES'/13X,'FOR ORDER#1.0070
     1ED SET OF ',I2,'-RECORD MAXIMA'/18X,'WITH',I4,' EVENTS PER RECORD'#1.0080
     2///)                                                              #1.0090
  202 FORMAT(' ',20X,'CORRELATION MATRIX'//)                            #1.0100
  203 FORMAT('0'/' ',20X,'SQRT EIGENVALUE MATRIX'//)                    #1.0110
  204 FORMAT('0'/' ',20X,'EIGENVECTOR MATRIX'//)                        #1.0120
  205 FORMAT('0',10X,  'THE LEAST NUMBER OF HIGH ORDER SIGNIFICANT DIGIT#1.0130
     1S (NSD) IN A CHECK OF CHARACTERISTIC EQUATION IS',I3,'.'///)      #1.0140
  206 FORMAT('0',9X,'ORDER',8X,8(I2,10X))                               #1.0150
  207 FORMAT(4X,'PROBABILITY',5X,8(F9.6,3X))                            #1.0160
  208 FORMAT('0',10X,'INCORRECT NUMBER OF DATA CARDS FOR CORRELATION MAT#1.0170
     1RIX'//)                                                           #1.0180
  209 FORMAT('0',10X,'MATRIX OF CORRELATION COEFFICIENTS DOES NOT TEST T#1.0190
     1O BE POSITIVE SEMIDEFINITE.(DETERMINANT LESS THAN ZERO.)'/        #1.0200
     2' ',10X,'THIS MAY BE A RESULT OF INACCURACIES IN THE COMPUTER ALGO#1.0210
     3RITHMS.'/' ',10X,'***THIS PROGRAM WILL NOT OPERATE ON THIS DATA SE#1.0220
     4T BEYOND STATEMENT 70***').                                       #1.0230
C
      READ(5,101)IX                                                     #1.0240
    1 READ(5,102,END=500)N,K                                            #1.0250
C
C     READ IN RO
C
      CALL MTRXIN(RO,N,IER)                                             #1.0260
      WRITE(6,201)N,K                                                   #1.0270
C
C     CHECK FOR ERROR IN INPUT
C
      IF(IER)60,61,60                                                   #1.0280
   60 WRITE(6,208)                                                      #1.0290
      GO TO 1                                                           #1.0300
   61 WRITE(6,202)                                                      #1.0310
C
C     PRINT CORRELATION MATRIX
C
      CALL MTRXOT(RO,N,1)                                               #1.0320
C
C     COMPUTE EIGENVECTORS AND EIGENVALUES OF RO
C
      NSD=0                                                             #1.0330
      CALL EGN(RO,R,RBC,N,NSD)                                          #1.0340
C
C     CHECK FOR POSITIVE SEMIDEFINITE CORRELATION MATRIX
C
```

```
      DO 70 I=1,N
      IJ=(I*I+I)/2
      IF(RO(IJ).GE.0.0) GO TO 70
      WRITE(6,209)
      GO TO 71
   70 CONTINUE
C
C     COMPUTE SQUAREROOTS OF EIGENVALUES AND SET OFF-DIAGONAL ELEMENTS
C     OF RO TO ZERO. PUT ELEMENTS IN GENERAL STORAGE MODE.
C
      DO 2 I=1,N
      DO 2 J=1,N
      IF(I-J)5,6,6
    5 IJ=I+(J*J-J)/2
      GO TO 3
    6 IJ=J+(I*I-I)/2
      IF(I.NE.J) GO TO 3
      RO(IJ)=DSQRT(RO(IJ))
      GO TO 4
    3 RO(IJ)=0.0
    4 K1=N*(J-1)+I
    2 B1(K1)=RO(IJ)
C
C     PRINT SQRT OF EIGENVALUES AND EIGENVECTOR MATRICES
C
      WRITE(6,203)
      CALL MTRXOT(B1,N,0)
      WRITE(6,204)
      CALL MTRXOT(R,N,0)
C
C     PRINT SIGNIFICANT DIGITS MESSAGE
C
      IF(NSD.EQ.0) GO TO 73
   72 WRITE(6,205)NSD
C
C     COMPUTE PRINCIPAL COMPONENT MATRIX B
C
   73 CALL GMPRD(R,B1,B,N,N,N)
C
C     INITIALIZE P(J) TO ZERO
C
      DO 7 J=1,N
    7 P(J)=0.0
C
C     IN THE 1000 ITERATION LOOP ENDING IN STATEMENT 10, AN N BY K
C     RECORD OF X(J,KK) IS SIMULATED (USING MATRIX B AND A NORMAL RANDOM
C     NUMBER E), N COLUMN MAXIMA Y(J) ARE SELECTED FROM X(J,KK),
C     Y(J) ARE PLACED IN DESCENDING ORDER (J=1,2,...,N), NORMAL
C     EXCEEDANCE PROBABILITY C FOR EACH Y(J) IS FOUND, AND THE 1000
C     ITEM SUM OF C FOR EACH J, P(J), IS COMPUTED.
C
      DO 10 M=1,1000
C
C     GENERATE X(J,KK) ARRAY
C
      DO 11 KK=1,K
      DO 15 I=1,N
C
C     SUBROUTINE GAUSS IS IMBEDDED IN MAIN PROGRAM BY FOLLOWING
C     10 STATEMENTS
C
      AA=0.0
      DO 12 II=1,12
      IY=IX*65539
      IF(IY)13,14,14
   13 IY=IY+1+2147483647
   14 YFL=IY
      YFL=YFL*0.4656613E-9
      IX=IY
   12 AA=AA+YFL
   15 E(I)=AA-6.0
C
```

```
C     FIND K VECTOR PRODUCTS OF X1 - X(N,K)
C
      CALL GMPRD(B,E,X1,N,N,1)                                          #1.0750
      DO 11 I=1,N                                                       #1.0760
      J=I                                                               #1.0770
   11 X(J,KK)=X1(I)                                                     #1.0780
C
C     FIND COLUMN MAXIMA
C
      DO 20 J=1,N                                                       #1.0790
      D=-1000.0                                                         #1.0800
      DO 21 KK=1,K                                                      #1.0810
      CM=AMAX1(X(J,KK),D)                                               #1.0820
   21 D=CM                                                              #1.0830
   20 Y(J)=D                                                            #1.0840
C
C     ORDER COLUMN MAXIMA
C
      NN=N-1                                                            #1.0850
   30 DO 31 J=1,NN                                                      #1.0860
      IF((Y(J)-Y(J+1)).GE.0.0) GO TO 31                                 #1.0870
      A=Y(J+1)                                                          #1.0880
      Y(J+1)=Y(J)                                                       #1.0890
      Y(J)=A                                                            #1.0900
   31 CONTINUE                                                          #1.0910
      NN=NN-1                                                           #1.0920
      IF(NN.GT.0) GO TO 30                                              #1.0930
C
C     COMPUTE NORMAL EXCEEDANCE PROBABILITY C OF EACH COLUMN MAXIMUM
C     Y(J).  (REF. - EQN. 26.2.19 NBS HNDBK (1965))
C
      DO 10  J=1,N                                                      #1.0940
      IF(Y(J).GE.0.0) GO TO 40                                          #1.0950
      U=-Y(J)                                                           #1.0960
      GO TO 41                                                          #1.0970
   40 U=Y(J)                                                            #1.0980
   41 IF(U.GT.5.0) GO TO 42                                             #1.0990
      IF(U.LT.0.0001) GO TO 43                                          #1.1000
      G=1.0+D1*U+D2*U*U+D3*U*U*U+D4*U**4+D5*U**5+D6*U**6                #1.1010
      C=0.5*G**(-16)                                                    #1.1020
      GO TO 44                                                          #1.1030
   42 C=0.0                                                             #1.1040
      GO TO 44                                                          #1.1050
   43 C=0.5                                                             #1.1060
   44 IF(Y(J).GE.0.0) GO TO 10                                          #1.1070
      C=C-1.0                                                           #1.1080
C
C     SUMMATION OF C = P(J)
C
   10 P(J)=P(J)+C                                                       #1.1090
C
C     COMPUTE MEAN EXCEEDANCE PROBABILITY P(J) FOR JTH-ORDER MAXIMUM
C
      DO 45 J=1,N                                                       #1.1100
   45 P(J)=P(J)/1000.0                                                  #1.1110
C
C     PRINT P(J)
C
      JS=1                                                              #1.1120
      IF(N.LE.8) GO TO 50                                               #1.1130
      JE=8                                                              #1.1140
      GO TO 51                                                          #1.1150
   50 JE=N                                                              #1.1160
   51 DO 52 J=JS,JE                                                     #1.1170
   52 JUNK(J)=J                                                         #1.1180
      WRITE(6,206)(JUNK(J),J=JS,JE)                                     #1.1190
      WRITE(6,207)(P(J),J=JS,JE)                                        #1.1200
      JS=JS+8                                                           #1.1210
      IF(JS.GT.N) GO TO 53                                              #1.1220
      JE=JE+8                                                           #1.1230
```

```
      IF(LE.LT.N) GO TO 51                                              #1.1240
      JE=N                                                              #1.1250
      GO TO 51                                                          #1.1260
   53 CONTINUE                                                          #1.1270
   71 GO TO 1                                                           #1.1280
  500 STOP                                                              #1.1290
      END                                                               #1.1300
      SUBROUTINE MTRXIN(A,NORD,IER)                                     #1.1310
C        PURPOSE
C           READS DATA ELEMENTS OF SYMMETRIC MATRIX FROM CARDS
C
C        DESCRIPTION OF VARIABLES
C           A       - INPUT MATRIX
C           NORD    - ORDER OF INPUT MATRIX. MUST BE 20 OR LESS.
C           IER     - ERROR IN DATA CARDS CODE
C                     IER = 0 NO ERROR
C                         = 1 INCORRECT NUMBER OF CARDS
C
C        LISTING OF DATA CARDS
C           FORMAT IS 7F10.0. COLUMNS 71-80 FOR IDENTIFICATION. DATA
C           ELEMENTS PUNCHED BY ROW STARTING WITH DIAGONAL ELEMENT AND
C           PROCEEDING TO THE RIGHT IN  THAT ROW, WITH LISTING CONTINUED
C           TO NEW CARDS IF NECESSARY. LAST CARD FOR LAST ROW IN MATRIX
C           FOLLOWED BY A CARD WITH 9 PUNCHED IN COLUMN 1.
C
C        REFERENCE
C           ADAPTATION OF SUBROUTINE MATIN. REFER TO INTERNATIONAL
C           BUSINESS MACHINES, 1968, SYSTEM/360 SCIFNTIFIC SUBROUTINE
C           PACKAGE (360-CM-03X) VERSION III, PROGRAMMER'S MANUAL.
C           WHITE PLAINS, NEW YORK, INTERNATIONAL BUSINESS MACHINES,
"           P.453
C
      DOUBLE PRECISION A(1),CARD(8)                                     #1.1320
    1 FORMAT(7F10.0)                                                    #1.1330
      IER=0                                                             #1.1340
      ICOLT=NORD                                                        #1.1350
      IROCR=1                                                           #1.1360
C
C     COMPUTE NUMBFR OF CARDS IN ROW
C
    2 IRCDS=(ICOLT-1)/7+1                                               #1.1370
C
C     SET UP LOOP FOR NUMBER OF CARDS IN ROW
C
      DO 3 K=1,IRCDS                                                    #1.1380
      READ(5,1)(CARD(I),I=1,7)                                          #1.1390
      L=0                                                               #1.1400
C
C     COMPUTE COLUMN NUMBER FOR FIRST FIELD IN CURRENT CARD
C
      JS=(K-1)*7+NORD-ICOLT+1                                           #1.1410
      JE=JS+6                                                           #1.1420
C
C     SET UP LOOP FOR DATA ELEMENTS WITHIN CARD
C
      DO 4 J=JS,JE                                                      #1.1430
      IF(J-NORD)5,5,3                                                   #1.1440
    5 IF(IROCR-J)6,7,7                                                  #1.1450
    6 IJ=IROCR+(J*J-J)/2                                                #1.1460
      GO TO 8                                                           #1.1470
    7 IJ=J+(IROCR*IROCR-IROCR)/2                                        #1.1480
    8 L=L+1                                                             #1.1490
    4 A(IJ)=CARD(L)                                                     #1.1500
    3 CONTINUE                                                          #1.1510
      IROCR=IROCR+1                                                     #1.1520
      IF(NORD-IROCR)9,10,10                                             #1.1530
   10 ICOLT=ICOLT-1                                                     #1.1540
      GO TO 2                                                           #1.1550
```

```
    9 READ(5,1)CARD(1)                                               #1.1560
      IF(CARD(1)-9.D9)11,12,11                                       #1.1570
   11 IER=1                                                          #1.1580
   12 CONTINUE                                                       #1.1590
      RETURN                                                         #1.1600
      END                                                            #1.1610
      SUBROUTINE MTRXOT(A,NORD,IS)                                   #1.1620
C        PURPOSE
C           PRINTOUT LISTING OF ANY SIZED MATRIX
C
C        DESCRIPTION OF VARIABLES
C           A  -   OUTPUT MATRIX
C           NORD - ORDER OF MATRIX
C           IS - STORAGE MODE OF A
C                     IS = 0 GENERAL
C                     IS = 1 SYMMETRIC
C                     IS = 2 DIAGONAL
C
C        SUBROUTINES REQUIRED
C           LOC
C
C        REFERENCE
C           ADAPTATION OF SUBROUTINE MXOUT. REFER TO INTERNATIONAL
C           BUSINESS MACHINES, 1968, SYSTEM/360 SCIENTIFIC SUBROUTINE
C           PACKAGE (360-CM-03X) VERSION III, PROGRAMMER'S MANUAL.
C           WHITE PLAINS, NEW YORK, INTERNATIONAL BUSINESS MACHINES,
C           P.454
      DOUBLE PRECISION A(1),B(8)                                     #1.1630
  201 FORMAT(20X,7(3X,I3,10X))                                       #1.1640
  202 FORMAT('0',11X,I3,7(D16.6))                                    #1.1650
      J=1                                                            #1.1660
    1 LS=1                                                           #1.1670
    2 JNT=J+6                                                        #1.1680
      IF(JNT-NORD)4,4,3                                              #1.1690
    3 JNT=NORD                                                       #1.1700
C
C     NUMBER COLUMNS
C
    4 WRITE(6,201)(JCUR,JCUR=J,JNT)                                  #1.1710
      LE=LS+57                                                       #1.1720
      DO 5 L=LS,LE                                                   #1.1730
C
C     PRINT LINE
C
      DO 6 K=1,7                                                     #1.1740
      KK=K                                                           #1.1750
      JT=J+K-1                                                       #1.1760
      CALL LOC(L,JT,IJ,NORD,NORD,IS)                                 #1.1770
      B(K)=0.0                                                       #1.1780
      IF(IJ)8,8,7                                                    #1.1790
    7 B(K)=A(IJ)                                                     #1.1800
C
C     IF LAST COLUMN, WRITE ROW'S NUMBER AND ELEMENTS
C
    8 IF(JT-NORD)6,9,9                                               #1.1810
    6 CONTINUE                                                       #1.1820
    9 WRITE(6,202)L,(B(JW),JW=1,KK)                                  #1.1830
C
C     IF END OF ROWS, CHECK FOR MORE COLUMNS
C
      IF(NORD-L)10,10,5                                              #1.1840
    5 CONTINUE                                                       #1.1850
C
C     CHECK FOR MORE OUTPUT
C
      LS=LS+58                                                       #1.1860
      GO TO 2                                                        #1.1870
                                                                     #1.1880
```

```
C     END OF COLUMNS, RETURN
                                                                  #1.1890
   10 IF(JT-NORD)11,12,12                                         #1.1900
   11 J=JT+1                                                      #1.1910
      GO TO 1                                                     #1.1920
   12 RETURN                                                      #1.1930
      END                                                         #1.1940
      SUBROUTINE EGN(A,R,B,N,NSD)                                 #1.0000
C        PURPOSE
C           COMPUTE EIGENVALUES, EIGENVECTORS, AND LEAST NUMBER OF
C           SIGNIFICANT DIGITS IN CHARACTERISTIC EQUATION OF A REAL
C           SYMMETRIC MATRIX
C        DESCRIPTION OF VARIABLES
C           A    - MATRIX FOR WHICH EIGENVECTORS AND EIGENVALUES ARE
C                  DESIRED. MATRIX DESTROYED IN COMPUTATION AND RESULTANT
C                  EIGENVALUES STORED IN DESCENDING ORDER IN DIAGONAL.
C           R    - RESULTANT EIGENVECTOR MATRIX
C           B    - ORIGINAL MATRIX TRANSFERRED TO THIS AREA
C           N    - ORDER OF MATRICES
C           NSD  - IF EIGENVALUES ARE TO BE COMPUTED, NSD MUST BE ZERO AT
C                  INPUT, RESULTANT IS LEAST NUMBER OF SIGNIFICANT DIGITS
C                  IN CHARACTERISTIC EQUATION STORED IN NSD. IF NSD = 1,
C                  EIGENVALUES AND LEAST NUMBER OF SIGNIFICANT DIGITS
C                  ARE NOT COMPUTED.
C
C        FUNCTION SUBPROGRAMS REQUIRED
C           LISD
C           MAX
C           MIN
C        REMARKS
C           THIS EXTENSION TO DOUBLE PRECISION FORM OF SUBROUTINE
C           EIGEN DEVELOPED BY MARSHALL HELLMAN, COMPUTER CENTER
C           DIVISION, U.S. GEOLOGICAL SURVEY. REFER TO INTERNATIONAL
C           BUSINESS MACHINES, 1968, SYSTEM/360 SCIENTIFIC SUBROUTINE
C           PACKAGE (360-CM-03X) VERSION III, PROGRAMMER'S MANUAL.
C           WHITE PLAINS, NEW YORK, INTERNATIONAL BUSINESS MACHINES,
C           P.164
C
      DIMENSION A(1), R(1), B(1)                                  #1.0010
      DOUBLE PRECISION A,R,ANORM,ANRMX,THR,X,Y,SINX,SINX2,COSX,COSX2,SIN#1.0020
     1CS,RANGE,B,W,W1,Z,Q                                         #1.0030
  201 FORMAT('0',10X,'NUMBER OF SIGNIFICANT DIGITS IN ELEMENT ',I3,' OF #1.0040
     1CHARACTERISTIC EQUATION'/11X,'MATRIX IS 0. MAGNITUDES OF ELEMENTS #1.0050
     2COMPARED ARE ',D24.16/11X,'AND ', D24.16,'. IF MAGNITUDE OF AN ELE#1.0060
     3MENT IS NOT CLOSE '/11X,' TO ZERO, IT WOULD BE ADVISABLE TO STUDY #1.0070
     4STRUCTURE OF CORRELATION MATRIX'//)                         #1.0080
      MSDR=15                                                     #1.0090
      MSD=15                                                      #1.0100
      MV=NSD                                                      #1.0110
      RANGE=1.0D-14                                               #1.0120
C
C     COPY A INTO B
C
      NX2=N*(N+1)/2                                               #1.0130
      DO 1 I=1,NX2                                                #1.0140
    1 B(I)=A(I)                                                   #1.0150
C
C     GENERATE IDENTITY MATRIX
C
      IF (MV-1) 2,5,2                                             #1.0160
    2 IQ=-N                                                       #1.0170
      DO 4 J=1,N                                                  #1.0180
      IQ=IQ+N                                                     #1.0190
      DO 4 I=1,N                                                  #1.0200
      IJ=IQ+I                                                     #1.0210
      R(IJ)=0.0D0                                                 #1.0220
      IF (I-J) 4,3,4                                              #1.0230
    3 R(IJ)=1.0D0                                                 #1.0240
    4 CONTINUE                                                    #1.0250
C
```

```
C     COMPUTE INITIAL AND FINAL NORMS, ANORM AND ANORMX
C
    5 ANORM=0.0D0                                                    #1.0260
      DO 7 I=1,N                                                     #1.0270
      DO 7 J=I,N                                                     #1.0280
      IF (I-J) 6,7,6                                                 #1.0290
    6 IA=I+(J*J-J)/2                                                 #1.0300
      ANORM=ANORM+A(IA)*A(IA)                                        #1.0310
    7 CONTINUE                                                       #1.0320
      IF (ANORM) 33,33,8                                             #1.0330
    8 ANORM=DSQRT(2*ANORM)                                           #1.0340
      ANRMX=ANORM*RANGE/DFLOAT(N)                                    #1.0350
C
C     INITIALIZE INDICATORS AND COMPUTE THRESHOLD, THR
C
      IND=0                                                          #1.0360
      THR=ANORM                                                      #1.0370
    9 THR=THR/DFLOAT(N)                                              #1.0380
   10 L=1                                                            #1.0390
   11 M=L+1                                                          #1.0400
C
C     COMPUTE SIN AND COS
C
   12 MQ=(M*M-M)/2                                                   #1.0410
      LQ=(L*L-L)/2                                                   #1.0420
      LM=L+MQ                                                        #1.0430
      IF (DABS(A(LM))-THR) 26,13,13                                  #1.0440
   13 IND=1                                                          #1.0450
      LL=L+LQ                                                        #1.0460
      MM=M+MQ                                                        #1.0470
      Z=A(LL)                                                        #1.0480
      Q=A(MM)                                                        #1.0490
      X=(Z-Q)/2                                                      #1.0500
      Y=-A(LM)/DSQRT(A(LM)*A(LM)+X*X)                                #1.0510
      IF (X) 14,15,15                                                #1.0520
   14 Y=-Y                                                           #1.0530
   15 SINX=Y/DSQRT(2*(1+(DSQRT(1-Y*Y))))                             #1.0540
      SINX2=SINX*SINX                                                #1.0550
      COSX2=1-SINX2                                                  #1.0560
      COSX=DSQRT(COSX2)                                              #1.0570
      SINCS=SINX*COSX                                                #1.0580
C
C     ROTATE L AND M COLUMNS
C
      ILQ=N*(L-1)                                                    #1.0590
      IMQ=N*(M-1)                                                    #1.0600
      DO 25 I=1,N                                                    #1.0610
      IQ=(I*I-I)/2                                                   #1.0620
      IF (I-L) 16,23,16                                              #1.0630
   16 IF (I-M) 17,23,18                                              #1.0640
   17 IM=I+MQ                                                        #1.0650
      GO TO 19                                                       #1.0660
   18 IM=M+IQ                                                        #1.0670
   19 IF (I-L) 20,21,21                                              #1.0680
   20 IL=I+LQ                                                        #1.0690
      GO TO 22                                                       #1.0700
   21 IL=L+IQ                                                        #1.0710
   22 X=A(IL)*COSX                                                   #1.0720
      W=A(IM)*SINX                                                   #1.0730
      Y=A(IL)*SINX                                                   #1.0740
      W1=-A(IM)*COSX                                                 #1.0750
      A(IM)=Y-W1                                                     #1.0760
      A(IL)=X-W                                                      #1.0770
   23 IF (MV-1) 24,25,24                                             #1.0780
   24 ILR=ILQ+I                                                      #1.0790
      IMR=IMQ+I                                                      #1.0800
      Z=R(ILR)*COSX                                                  #1.0810
      Q=R(IMR)*SINX                                                  #1.0820
      X=Z-Q                                                          #1.0830
      Z=R(ILR)*SINX                                                  #1.0840
```

```
      Q=-R(IMR)*COSX                                                    #1.0850
      R(IMR)=Z-Q                                                        #1.0860
      R(ILR)=X                                                          #1.0870
   25 CONTINUE                                                          #1.0880
      X=2*A(LM)*SINCS                                                   #1.0890
      Z=A(LL)*COSX2                                                     #1.0900
      Q=-A(MM)*SINX2                                                    #1.0910
      Y=Z-Q                                                             #1.0920
      Z=A(LL)*SINX2                                                     #1.0930
      Q=-A(MM)*COSX2                                                    #1.0940
      W=Z-Q                                                             #1.0950
      W1=-W                                                             #1.0960
      A(LM)=0                                                           #1.0970
      A(LL)=Y-X                                                         #1.0980
      A(MM)=W+X                                                         #1.0990
C
C     TESTS FOR COMPLETION
C
C     TEST FOR M = LAST COLUMN
C
   26 IF (M-N) 27,28,27                                                 #1.1000
   27 M=M+1                                                             #1.1010
      GO TO 12                                                          #1.1020
C
C     TEST FOR L = SECOND FROM LAST COLUMN
C
   28 IF (L-(N-1)) 29,30,29                                             #1.1030
   29 L=L+1                                                             #1.1040
      GO TO 11                                                          #1.1050
   30 IF (IND-1) 32,31,32                                               #1.1060
   31 IND=0                                                             #1.1070
      GO TO 10                                                          #1.1080
C
C     COMPARE THRESHOLD WITH FINAL NORM
C
   32 IF (THR-ANRMX) 33,33,9                                            #1.1090
C
C     SORT EIGENVALUES AND EIGENVECTORS
C
   33 IQ=-N                                                             #1.1100
      DO 37 I=1,N                                                       #1.1110
      IQ=IQ+N                                                           #1.1120
      LL=I+(I*I-I)/2                                                    #1.1130
      JQ=N*(I-2)                                                        #1.1140
      DO 37 J=I,N                                                       #1.1150
      JQ=JQ+N                                                           #1.1160
      MM=J+(J*J-J)/2                                                    #1.1170
      IF (A(LL)-A(MM)) 34,37,37                                         #1.1180
   34 X=A(LL)                                                           #1.1190
      A(LL)=A(MM)                                                       #1.1200
      A(MM)=X                                                           #1.1210
      IF (MV-1) 35,37,35                                                #1.1220
   35 DO 36 K=1,N                                                       #1.1230
      ILR=IQ+K                                                          #1.1240
      IMR=JQ+K                                                          #1.1250
      X=R(ILR)                                                          #1.1260
      R(ILR)=R(IMR)                                                     #1.1270
   36 R(IMR)=X                                                          #1.1280
   37 CONTINUE                                                          #1.1290
C
C     DETERMINE LEAST NUMBER OF SIGNIFICANT DIGITS
C
      IF (MV-1) 38,48,38                                                #1.1300
   38 CONTINUE                                                          #1.1310
C
C     COMPUTE MATRIX ELEMENTS OF PRODUCT OF ORIGINAL MATRIX AND CHARAC-
C     TERISTIC VECTORS, X, AND OF EIGENVALUES AND CHARACTERISTIC
C     VECTORS, Z.
C
```

```
      DO 45 I2=1,N
      Y=A(I2*(I2+1)/2)
      I1=N*(I2-1)
      DO 43 L=1,N
      X=0
      DO 42 M=1,N
      IF (M-L) 39,40,40
   39 I=L*(L-1)/2+M
      GO TO 41
   40 I=M*(M-1)/2+L
   41 K=I1+M
   42 X=X+B(I)*R(K)
      Z=Y*R(I1+L)
C
C     DETERMINE LEAST NUMBER OF SIGNIFICANT DIGITS IN X OR Y FOR EACH
C     ELEMENT OF MATRIX
C
      NSD=LISD(X,Z,15)
      IF(NSD)43,49,43
   49 I1L=I1+L
      WRITE(6,201)I1L,X,Z
   43 MSD=MIN(NSD,MSD)
      MSDR=MIN(MSD,MSDR)
      K1=I2*(I2+1)/2+1
      IF (I2-N) 44,46,46
   44 A(K1)=MSD
   45 MSD=15
   46 DO 47 I=1,N
      K=I*(I+1)/2
      L=K+1
   47 B(K)=A(L)
      B(NX2)=MSD
      NSD=MSDR
   48 CONTINUE
      RETURN
      END
      FUNCTION LISD(X,Y,N)
C        PURPOSE
C           TO FIND SMALLER NUMBER OF SIGNIFICANT DIGITS IN TWO NUMBERS
C           X AND Y, IF NUMBER IS LESS THAN N
C
C        REMARKS
C           PROGRAMMED BY MARSHALL HELLMAN, COMPUTER CENTER DIVISION,
C           U.S. GEOLOGICAL SURVEY
C
      DOUBLE PRECISION X,Y,EPS,T
      LISD=0
      IF(DABS(Y).GE.DABS(X)) GO TO 1
      T=X
      IF(T.EQ.0.0) GO TO 2
      GO TO 3
    1 T=Y
      IF(T.EQ.0.0) GO TO 2
    3 EPS=DLOG10(DABS(T))
      IX=EPS
      IF(EPS.GE.0.0) GO TO 4
      IF(EPS.GE.IX) GO TO 4
      IX=IX-1
    4 IF(Y.NE.X) GO TO 5
    2 LISD=N
      GO TO 6
    5 T=DLOG10(DABS(Y-X))
      IT=T
      IF(T.GE.0.0) GO TO 7
      IF(T.GE.IT) GO TO 7
      IT=IT-1
    7 LISD=IX-IT-1
      LISD=MAX(LISD,0)
    6 RETURN
      END
```

```
      FUNCTION MAX(M,N)                                                 #1.1900
C        PURPOSE
C           TO DETERMINE IF INTEGER M IS LARGER THAN SELECTED INTEGER N
C
C        REMARKS
C           PROGRAMMED BY MARSHALL HELLMAN, COMPUTER CENTER DIVISION,
C           U.S. GEOLOGICAL SURVEY
C
      IF(M.GT.N) GO TO 1                                                #1.1910
      MAX=N                                                             #1.1920
      RETURN                                                            #1.1930
    1 MAX=M                                                             #1.1940
      RETURN                                                            #1.1950
      END                                                               #1.1960
      FUNCTION MIN(M,N)                                                 #1.1970
C        PURPOSE
C           TO DETERMINE IF INTEGER M IS SMALLER THAN SELECTED INTEGER N
C
C        REMARKS
C           PROGRAMMED BY MARSHALL HELLMAN, COMPUTER CENTER DIVISION,
C           U.S. GEOLOGICAL SURVEY
C
      IF(M.LE.N) GO TO 1                                                #1.1980
      MIN=N                                                             #1.1990
      RETURN                                                            #1.2000
    1 MIN=M                                                             #1.2010
      RETURN                                                            #1.2020
      END                                                               #1.2030
      SUBROUTINE GMPRD(A,B,R,N,M,L)                                     #1.2040
C
C        PURPOSE
C           PRODUCT OF TWO GENERAL-MODE-STORAGE MATRICES
C        DESCRIPTION OF VARIABLES
C           A  -   FIRST INPUT MATRIX, N ROWS, M COLUMNS
C           B  -   SECOND INPUT MATRIX, M ROWS, L COLUMNS
C           R  -   OUTPUT MATRIX, N ROWS, L COLUMNS
C
C        REFERENCE
C           INTERNATIONAL BUSINESS MACHINES, 1968, SYSTEM/360 SCIENTIFIC
C           SUBROUTINE PACKAGE (360-CM-03X) VERSION III, PROGRAMMER'S
C           MANUAL. WHITE PLAINS, NEW YORK, INTERNATIONAL BUSINESS
C           MACHINES, P.99
C
      DOUBLE PRECISION A(1),B(1),R(1)                                   #1.2050
      IR=0                                                              #1.2060
      IK=-M                                                             #1.2070
      DO 1 K=1,L                                                        #1.2080
      IK=IK+M                                                           #1.2090
      DO 1 J=1,N                                                        #1.2100
      IR=IR+1                                                           #1.2110
      JI=J-N                                                            #1.2120
      IB=IK                                                             #1.2130
      R(IR)=0.0                                                         #1.2140
      DO 1 I=1,M                                                        #1.2150
      JI=JI+N                                                           #1.2160
      IB=IB+1                                                           #1.2170
    1 R(IR)=R(IR)+A(JI)*B(IB)                                           #1.2180
      RETURN                                                            #1.2190
      END
```

U. S. GOVERNMENT PRINTING OFFICE : 1971 O - 421-897