A Program Suite for the Statistical Analysis and Comparison of Historical and Synthetic Hydrologic Records. July 1978.

Author:
Lindner, M. A.

Publication details:
Commissioning Body: I. B. M. Australia
Report No. UNSW Water Research Laboratory Report No. 153
0858242656 (ISBN)

Publication Date:
1978

DOI:
https://doi.org/10.4225/53/57994a0a2f177

THE UNIVERSITY OF NEW SOUTH WALES
water research laboratory
Manly Vale N.S.W. Australia

A PROGRAM SUITE FOR THE STATISTICAL ANALYSIS AND COMPARISON OF HISTORICAL AND SYNTHETIC HYDROLOGIC RECORDS

by

Michael A. Lindner

Report No. 153

July 1978.

A PROGRAM SUITE FOR THE STATISTICAL ANALYSIS AND COMPARISON

OF HISTORICAL AND SYNTHETIC HYDROLOGIC RECORDS

Errata

The following corrections should be made to the text:

Page ii,	line 8:	Should read "made available by" instead of "included by"
Page ii,	line 12:	Should read "made available by" instead of "included by"
Page 5,	line 5:	Should read "Table 1.1" instead of "Table 1"
Page 94,	line 31:	Should read "1979" instead of "1978"
Page 165,	figure title:	Should read "SAMPLE OUTPUT" instead of "OUTPUT"

M. A. LINDNER,
17th April, 1979.

The University of New South Wales
School of Civil Engineering

A PROGRAM SUITE FOR THE STATISTICAL ANALYSIS AND COMPARISON OF HISTORICAL AND SYNTHETIC HYDROLOGIC RECORDS

by

Michael A. Lindner

Report No. 153
July, 1978

https://doi.org/10.4225/53/57994a0a2f177

BIBLIOGRAPHIC DATA SHEET	1. REPORT No. 153	2. I.S.B.N. 0/85824/265/6

3. TITLE AND SUBTITLE	4. REPORT DATE
A Program Suite for the Statistical Analysis and Comparison of Historical and Synthetic Hydrologic Records	July 1978

5. AUTHOR(S)

Michael A. Lindner

6. SPONSORING ORGANIZATION

School of Civil Engineering, The University of New South Wales

7. SUPPLEMENTARY NOTES

8. ABSTRACT Synthetically generated records of streamflow, rainfall and evaporation may be used in river valley simulation models as an aid to decision-making. Before the records may be used in the simulation model, however, they need to be validated by a comparison of their statistical properties with those of the historical records. A suite of statistical programs has been developed to assist in this comparison, firstly by evaluating a number of statistical properties of a set of historical records, secondly by evaluating the same properties for sets of synthetic records, and finally, by calculating a number of measures of the "difference" between the historical and synthetic record statistics. The programs have been designed so that they can be put to use with very little difficulty and have been dimensioned for up to twenty-five records in a set. All the programs work with monthly records. They have been written in Fortran IV and are compatible with American National Standard Fortran. The programs have been documented in detail, including descriptions of the theory used, the computational procedures employed, the program logic, options, inputs and outputs. A problem chosen as an example is described, and selected illustrative output is given to illustrate one application of the programs, to demonstrate how the suite may be implemented on a computer system, and to provide a test case for checking implementation.

9. DISTRIBUTION STATEMENT

10. KEY WORDS Simulation, Stochastic Processes, Synthetic Hydrology, Statistical Analysis, Computer Programs, Streamflow, Rainfall, River Basin Planning

11. CLASSIFICATION	12. NUMBER OF PAGES	13. PRICE
	216	

PREFACE

This report gives a detailed description of a suite of computer programs developed by the author and used in the selection of a multi-site synthetic data generation technique for the synthesis of 16 streamflow and rainfall series for the Murray River Valley, and for the statistical validation of the generated synthetic hydrologic data. Because this type of problem is of general interest and importance, and because the programs provide a powerful practical tool, this report has been prepared as a guide to their use.

Subprograms for the calculation of statistical properties not included in the statistical analysis program can be written by the user to suit his particular purposes and be readily incorporated into the statistical analysis program. Appropriate comparison programs for these statistics can be added.

D. T. Howell,
Senior Lecturer,
School of Civil Engineering.
July 1978

ACKNOWLEDGEMENTS

Acknowledgements are made to IBM Australia Ltd for providing finance and facilities in support of this work and to my colleagues David Doran and Geoff Wright for constructive comments and suggestions at the beginning of the work.

Subroutines RANK, TIE, and HIST are taken from the System/360 Scientific Subroutine Package [*International Business Machines* 1970]. Subroutine TAB4 is an extensively modified version of subroutine TAB1 of the same package. Listings of these subroutines are included by permission of IBM.

Subroutines PLOTR, SCALE2 and FORM2 are taken from the Biomedical Computer Programs [*Dixon* 1973]. Listings of these subroutines are included by permission of Professor W.J. Dixon, Health Sciences Computing Facility, University of California, Los Angeles, California.

ABSTRACT

Synthetically generated records of streamflow, rainfall and evaporation may be used in river valley simulation models as an aid to decision-making. Before the records may be used in the simulation model, however, they need to be validated by a comparison of their statistical properties with those of the historical records. A suite of statistical programs has been developed to assist in this comparison, firstly by evaluating a number of statistical properties of a set of historical records, secondly by evaluating the same properties for sets of synthetic records, and finally, by calculating a number of measures of the "difference" between the historical and synthetic record statistics.

The programs have been designed so that they can be put to use with very little difficulty and have been dimensioned for up to twenty-five records in a set. All the programs work with monthly records. They have been written in Fortran IV and are compatible with American National Standard Fortran.

The programs have been documented in detail, including descriptions of the theory used, the computational procedures employed, the program logic, options, inputs and outputs. A problem chosen as an example is described, and selected illustrative output is given to illustrate one application of the programs, to demonstrate how the suite may be implemented on a computer system, and to provide a test case for checking implementation.

TABLE OF CONTENTS

TABLE OF CONTENTS cont.

TABLE OF CONTENTS cont.

TABLE OF CONTENTS cont.

LIST OF FIGURES

LIST OF FIGURES cont.

LIST OF TABLES

1. INTRODUCTION

Simulation models are commonly used as an aid to decision-making in the design and operation of water resource systems [*Hufschmidt and Fiering* 1966, *Blainey* 1970, *Texas Water Development Board* 1974]. Traditionally, historically recorded sequences of hydrologic variables, such as, streamflow, rainfall and evaporation have been used as inputs to these models. In so doing, there is an implicit assumption that the historical patterns of streamflow, rainfall and evaporation will recur in the future—a quite unrealistic assumption [*Fiering and Jackson* 1971]. "The use of historical sequences to represent future sequences subjects the design to risk and provides no basis for assessing the risk or evaluating the losses associated with a system that is under- or overdesigned to an unknown extent" [*Matalas and Wallis* 1976].

For these reasons, statistical models have been developed for synthesising large numbers of different possible future sequences of hydrologic variables, each of which is equally likely to occur in the future [*Jackson* 1975, *Matalas* 1975]. For these sequences to be considered realistic realizations of future sequences, their statistical characteristics should be the same as the historical sequences.

In general, it is impossible to devise models which will preserve all the statistical characteristics of the historical sequences. For any one model, the values of certain statistical characteristics of the synthesised sequences may be very little different from the corresponding values of the original historical sequences, but the values of some other statistical characteristics may be quite different. Another model may produce good agreement for a different set of characteristics and poor agreement for some of those for which there was good agreement in the first model. The particular features of any decision-making situation determine which statistical characteristics of the historical sequences should be preserved in the synthetic sequences and which ones do not matter.

As an aid to water resource system design, the suite of computer programs described in this report has been developed, first to evaluate a number of statistical characteristics of a set of historical sequences, then to evaluate the same characteristics for sets of synthesised sequences and, lastly, to compare them. The comparison enables a decision-maker or decision process modeller to choose the synthesising model which best suits his purposes.

The analysis procedure which might be followed by a designer is shown diagrammatically in Figure 1.1*. The area of the analysis in which these programs assist is enclosed by the broken line on this figure. A set of variables to be subjected to this analysis would be selected from among the hydrological and meteorological variables of a region in which there is a system about which decisions are to be made. For example, if a dam is to be sized at a particular site in a river valley, the set of variables chosen for analysis could be the streamflow at the dam site, the streamflow in a tributary entering the main river downstream of the dam site but upstream of the area intended to be irrigated, the rainfall at a representative point in the area to be irrigated, the free water surface evaporation at the same point (from which evapotranspiration is derived), and the free water surface evaporation and rainfall at the dam site. It would be expected that both the streamflows and the rainfalls would be positively correlated with each other, both the evaporations positively correlated with each other but negatively correlated with the streamflows and rainfalls. The historical record of each variable, forming a time series, could be expected to exhibit some persistence measured by serial or auto correlation.

All of the programs work with monthly data, that is, the streamflow records used are volumetric flows in each month, rainfall records are the total precipitation in each month, etc.

For any such record or set of records, an analysis program called "STATS" can evaluate the following:

(a) The first four moments, the coefficient of variation, the minimum and maximum of the variate of each record for each of the twelve months of the year, for all values of the variate of each record without any distinction being made about which month of the year they are taken from, and for the twelve month aggregates, that is, annual values, of the variate of each record,

(b) The monthly and annual serial correlation coefficients for a specified number of lags for the variate of each record to give corre-

---

*All figures (and tables) are grouped at the end of this report. See the List of Figures in the Table of Contents for their page numbers.

lation functions or correlograms,

(c) The correlation coefficients between pairs of variates for a specified number of lags with either variate leading, that is, the "cross correlations" or "spatial correlations" between, for example, streamflows in adjacent or distant catchments,

(d) Histograms for all the monthly values of the variate of each record without regard to the month of the year,

(e) Histograms of run lengths of values of the variate of each record above, below and about the variate median or a specified value,

(f) The surplus, deficit and range in the residual mass curve of the variate of each record,

(g) The "Rippl storage" and storage deficit distribution, together with the drought, drawdown and fill duration distributions of the variate of each record, these evaluations being made for up to five different specified yield levels,

(h) For time periods ranging from one month up to a specified number of months, the maximum and minimum values of aggregate values over the time period of the variate for each record,

(i) The standardised Kendall tau statistic.

Provision is made for line printer plotting of most of these statistics, and for a choice of all or only some of them to be evaluated.

The same analysis program ("STATS") can evaluate any or all of the same statistical properties for any synthetically generated record or sets of records.

A set of comparison programs can compare a number of these statistical properties of the historical record or set of records with those of a synthetically generated record or sets of records. The statistical properties which may be compared are:

(a) the first three moments and minimum and maximum of the variate of each record for each of the twelve months of the year, for all values of the variate of each record without any distinction as to month, and for the annual values of the variate of each record,

(b) the histograms of all values of the variate of each record without regard to the month of the year,

(c) the histograms of run lengths of values of the variate of each record above, below and about the variate median or a specified value,

(d) the "Rippl storage" and storage deficit distribution, the drought, drawdown and fill duration distributions of the variate of each record for each specified yield level, and

(e) the serial correlation coefficients for a specified number of lags of the variate of each record, and the cross correlation coefficients between pairs of variates for a specified number of lags with either variate leading.

These statistical properties are compared by programs called "MOMENT", "FREQ", "RUNS", "YIELD" and "CORREL" respectively, by the evaluation of a number of measures of the "difference" between the historical and synthetic values of the properties, including statistical sampling theory, where applicable.

These particular properties were selected for detailed analysis because of their dominant influence on the design of water resource systems. For example, inflows to a reservoir with a high value of lag one serial (or auto) correlation coefficient will give rise to a different reservoir behaviour than inflows with a low value of lag one serial correlation coefficient, but with the same values of other statistical properties [*Perrens and Howell* 1972].

The programs have been written to conform to American National Standards Institute Fortran IV and can be run on any medium size computer system with only minor changes. They have been designed so that they can be put to use with very little difficulty and can handle quite large systems, being dimensioned for up to twenty-five variates in a set.

The programs have been used to compare two alternative multi-site synthetic generation models [*U.S. Army Corps of Engineers* 1971, *Finzini et al.* 1977] for the synthesis of 16 rainfall and streamflow records for a complex simulation model of the Murray River Valley [*Blainey* 1970]. The programs performed all the necessary computations for the comparison, which would otherwise have involved many man-hours of work, and were found to provide a practical tool for the selection of the "best" generation model for the Murray River Valley [*Lindner* 1978].

Sample execution times and storage requirements for the above comparison on an IBM 360/67 computer system are given in Tables 1.1 and 1.2 respectively*. The historical and synthetic data sets for this example consisted of one set of 16 variates of 77 years of record each and 50 sets of 16 variates of 75 years of record each. Table 1 shows the statistical properties calculated in column 1, the execution times (in minutes) of program STATS for the analysis of the historical and synthetic data sets in columns 2 and 3 respectively, and the execution times (in minutes) of the statistic comparison programs in column 4. The correlation analysis and comparison times given are unrepresentative in that the auto correlation and cross correlation functions were calculated to +24 lags and to +12 and -12 lags respectively in this example. Estimated correlation analysis and comparison times for calculation of the auto correlation and cross correlation functions to +1 lag and to +1 and -1 lags respectively are given in brackets in the table.

The storage requirements of each program are given in Table 2 as the maximum of the "link-edit" and "go" step requirements.

Source listings and card or tape copies of the programs can be obtained at nominal cost by writing to:

Head, Department of Water Engineering,
School of Civil Engineering,
The University of New South Wales,
P.O. Box 1,
KENSINGTON. N.S.W. 2033.
AUSTRALIA.

The programs are furnished, accepted and used by the recipient upon the express understanding that The University of New South Wales makes no warranties concerning the accuracy, completeness, reliability, usability, or suitability for any particular purpose of the programs, and that neither the University nor the author shall be under any liability whatsoever to any person by reason of any use made of the programs.

It would be appreciated if any errors found in the programs were reported to the Head, Department of Water Engineering at the above

---

*All tables (and figures) are grouped at the end of this report. See the List of Tables in the Table of Contents for their page numbers.

address.

A standard format has been used for the documentation of each program in this report and this is described in the following section.

2. PROGRAM SUITE DOCUMENTATION

Program documentation may be described as the "systematic and orderly description of a computer program for the engineer, the programmer, or the operator who is to use a completed program" [*Reti* 1973]. Recent program documentation standards [*Reti* 1973, *Association for Computer Aided Design* 1974 and 1977] have recommended that program documentation should include descriptions of:

(a) the purpose of the program,
(b) the theory used,
(c) the computational procedures employed,
(d) the library subroutines and procedures used,
(e) the program inputs and options,
(f) the program outputs, and
(g) the methods used to check the program.

The standards further recommend that the documentation include a source listing of the program and an example problem and sample run of the program.

These recommendations have been followed in the documentation of the program suite except that source listings have been omitted from the documentation because of printing costs. Source listings may be obtained as described in Section 1. Each program and subroutine listing includes a set of comment statements at the beginning of the listing which briefly describe the program purpose, version, author, latest modifications, and key variables. Comment cards throughout the listing identify major program steps and describe calculations or input/output operations which follow them.

Each program in the suite is described in subsequent sections of this report, each with major subsections entitled—

(a) Introduction,
(b) Fundamental Calculations,
(c) Programming Features,
(d) Input Data Preparation, and
(e) Output Interpretation.

The "Introduction" describes the purpose of the program, the types of calculations performed, and the major program options. Links with other programs through punched card, tape or disk file output are also indicated.

Each program calculates various statistics for each hydrologic variable analysed. In "Fundamental Calculations", these statistics are described and defined and the computational procedures are broadly discussed.

"Programming Features" provides details of the programming of the calculations described in Fundamental Calculations. The program structure and major program loops are discussed. Flow charts of all main programs and of selected subroutines are given. All input, output and scratch files are described. Common blocks are described, and all scalars and variables initialised by data statement or Block Data subroutine are listed. All required Fortran IV Library subroutines are listed. Finally, specifics of computational procedures are provided where considered warranted.

"Input Data Preparation" describes the required input data and how to prepare the input data deck specifying program options, file reference numbers and problem data. Input variables are defined, and data formats and valid ranges of variable values are given.

"Output Interpretation" briefly describes the major sections of the printed program output which is clearly labelled.

The calculations performed by the programs have been checked by independent computation by hand and by alternative programs.

An example problem and sample program output are included in the documentation to illustrate one application of the program suite, to demonstrate how the suite may be implemented on a computer system, and to provide a documented test case for checking implementation. A brief description of the problem is given and the historical and synthetic streamflow and rainfall data is listed. The job control language used to run the programs on an IBM 370/158 computer system is listed. This offers a guide for the use of other systems. The input data deck and selected example output for each program are given.

3. PROGRAM *STATS* DESCRIPTION

3.1 INTRODUCTION

Program STATS is used to calculate selected statistics of monthly hydrologic data. Some statistics are evaluated for each water year month, some for the sequence of monthly values (without regard to water year month) forming a time series, and some for the 12 water year month aggregates forming an annual series. The statistics which may be calculated include:

(a) Kendall's tau coefficient,

(b) monthly moments, extrema and frequency distributions,

(c) annual moments, extrema and frequency distribution,

(d) annual auto correlation function,

(e) time series moments, extrema and frequency distribution,

(f) run length frequency distributions above and below a specified value,

(g) aggregate time series minima and maxima,

(h) time series surplus, deficit, range and theoretical range,

(i) Rippl storage and drought duration; the drought, draw and fill duration frequency distributions and moments; and the storage deficit distribution and moments,

(j) time series auto correlation function, and

(k) time series cross correlation functions.

The calculated values of the user selected statistics are printed out in a tabular form and, optionally, may be also plotted on the line printer.

Program STATS may also be used in conjunction with the statistic comparison programs described in Sections 4, 5, 6, 7 and 8 to compare the numerical values of some of the above statistics in historical and synthetically generated data. To do this STATS is used to calculate the numerical values of the user selected statistics in the historical and synthetically generated data, and to punch and to write these values to disk files in appropriate formats for subsequent input to, and analysis by, the appropriate statistic comparison programs.

3.2 FUNDAMENTAL CALCULATIONS

3.2.1 Introduction

The fundamental purpose of this program is to calculate selected statistics of hydrologic time series data. These statistics, and the methods and equations used in their calculations are described and defined in the following paragraphs.

3.2.2 Trend Analysis

Kendall's tau statistic is a nonparametric measure of the correlation between the values of two variables [*Siegel* 1956, pp. 213-223]. In testing a hydrologic time series for trend, one variable is the time of observation, and the other, the magnitude of the stochastic hydrologic variable. If there is a significant correlation between the time of observation and the magnitude of the stochastic hydrologic variable, then there is a trend in the observed values of the stochastic hydrologic variable.

The Kendall tau statistic for a discrete stochastic hydrologic variable, measured over constant consecutive time periods, such as monthly, weekly or annual streamflow and rainfall, is calculated by the program as follows. Let X_j be the value of the hydrologic variable at the jth time period of N time periods, and let j be the value of the time variable at the jth time period. The value of the time variable at each period is then identical to its rank, and the time variable may thus be considered to be ranked, and ordered in its natural order of occurrence. The rank of each hydrologic variable value is calculated next and substituted for each variable value.

The number of pairs of ranks of the hydrologic variable which are in natural order, that is, the left rank of the pair less than the right rank, is now calculated as follows. The rank of the first hydrologic variable value (X_1) is compared in turn with each subsequent hydrologic variable value rank. When the rank pair is in natural order, one is added to the sum of rank pairs in natural order; when the ranks are tied, zero is added to the sum of rank pairs in natural order; and when the ranks are not in natural order, one is subtracted from the sum of rank pairs in natural order. The rank of the second hydrologic variable value (X_2) is then compared with each subsequent hydrologic variable value rank, and the sum of rank pairs in natural order is incremented or decremented as described above, and so on.

If there are no tied ranks, then the maximum number of rank pairs in natural order, P_1, is for the case of perfect agreement between the hydrologic and time variables, and is:

$$P_1 \quad = \quad 0.5N(N - 1) \tag{3.1}$$

Kendall's tau statistic is then calculated as:

$$TAU \quad = \quad P/P_1 \tag{3.2}$$

where P is the actual number of hydrologic variable rank pairs in natural order.

When there are tied ranks, Kendall's tau statistic is calculated as:

$$TAU \quad = \quad P/((\sqrt{P_1 - CF})\sqrt{P_1}) \tag{3.3}$$

where CF is a correction factor for tied ranks and is equal to:

$$CF \quad = \quad 0.5 \; \Sigma \; m(m - 1) \tag{3.4}$$

where m is the number of hydrologic variable values tied for a particular rank, and the summation is over all tied ranks.

Kendall determined that under the null hypothesis that the two variables are uncorrelated, tau was normally distributed for large N(N > 10) with a mean of zero and a standard deviation of:

$$SD \quad = \quad \sqrt{\frac{2(2N + 5)}{9N(N - 1)}} \tag{3.5}$$

The standardised Kendall tau statistic, Z, is thus:

$$Z \quad = \quad TAU/SD \tag{3.6}$$

The probability of occurrence of any value as extreme as an observed standardised Kendall tau statistic (Z) under the null hypothesis, may be determined by reference to standard normal probability tables. For example, the probability of as extreme a standardised Kendall tau statistic as 3.03 is:

$$p(Z \geq 3.03) + p(Z \leq -3.03) \quad = \quad 0.24\% \tag{3.7}$$

Hence at significance levels of 10%, 5% or 1%, the null hypothesis that the two variables are uncorrelated, is rejected.

3.2.3 Distribution Moments and Extrema

If X_j is the jth value of N values of a discrete stochastic hydrologic variable, then the program calculates the first moment or mean of the distribution of the variable as:

$$M_1 = \overline{X} = \frac{1}{N} \sum_{j=1}^{N} X_j \tag{3.8}$$

The second, third and fourth moments of the distribution are calculated as:

$$M_2 = \frac{1}{N-1} \sum_{j=1}^{N} (X_j - \overline{X})^2 \tag{3.9}$$

$$M_3 = \frac{N}{(N-1)(N-2)} \sum_{j=1}^{N} (X_j - \overline{X})^3 \tag{3.10}$$

$$M_4 = \frac{N^2}{(N-1)(N-2)(N-3)} \sum_{j=1}^{N} (X_j - \overline{X})^4 \tag{3.11}$$

respectively [*Yevjevich* 1972a, pp. 102-111].

The standard deviation and coefficient of variation of the variable's distribution are calculated as:

$$S = \sqrt{M_2} \tag{3.12}$$

$$C_v = S/\overline{X} \tag{3.13}$$

respectively.

The skew and kurtosis coefficients of the variable's distribution are calculated as:

$$C_s = M_3/S^3 \tag{3.14}$$

$$C_k = M_4/S^4 \tag{3.15}$$

respectively.

The extrema of the distribution are the minimum and maximum values of the N values of the variable.

3.2.4 Frequency Distributions

Relative and cumulative frequency distributions of a number of variables may be calculated by the program using a user specified number of class intervals (or bins) and upper and lower bounds, or using a program determined number of class intervals and upper and lower bounds. These frequency distributions are defined as follows.

Let the number of values from a sample of size N of a discrete stochastic hydrologic variable X which lie in the jth class interval be n_j. Then the number n_j is the absolute frequency, or more simply, the frequency of values within the jth class interval. (The sum of the n_j values for all class intervals is of course N.)

The relative frequency or relative probability of the jth class interval is defined as:

$$f_j = n_j/N \tag{3.16}$$

and is equal to the probability of a value of the hydrologic variable X being within the jth class interval. (The sum of the f_j values of all the class intervals is one.) A plot of relative probability against class interval is a relative probability histogram.

The absolute cumulative frequency, or more simply, cumulative frequency, of the jth class interval, is defined as:

$$N_j = \sum_{i=1}^{j} n_i \tag{3.17}$$

and is the number of values of the hydrologic variable X which are less than or equal to the upper bound of the jth class interval. The relative cumulative frequency or cumulative probability is defined as:

$$F_j = \frac{1}{N} \sum_{i=1}^{j} n_i \tag{3.18}$$

and is the probability of a value of the hydrologic variable X being less than or equal to the upper bound of the jth class interval. A plot of cumulative probability against class interval is a cumulative probability histogram.

3.2.5 Run Length Frequency Distributions

The program calculates positive, negative and total run length frequency distributions about a specified value or about the time series median value [*Yevjevich* 1972b, pp. 174-176]. Additionally for each of these distributions, the program calculates the total number of runs and average run length. These distributions are defined as follows.

Let X_j be the value of a discrete stochastic hydrologic variable at the jth time period of N time periods, and X_o be the specified value or time series median value, about which the positive, negative and total run length frequency distributions are to be calculated.

Define $(X_j - X_o) > 0$ as a positive deviation, and $(X_j - X_o) \leq 0$

as a negative deviation. Then a consecutive sequence of M positive deviations, preceded and succeeded by a negative deviation, is a positive run of length M. Similarly, a consecutive sequence of M negative deviations, preceded and succeeded by a positive deviation, is a negative run of length M.

The positive and negative run length frequency distributions show the frequency of occurrence of each positive and negative run length. The total run length frequency distribution is formed by adding the frequencies of positive and negative runs for each run length, and shows the frequency of occurrence of each run length without regard to sign.

The total number of runs of each run length frequency distribution is calculated by summing the frequency of occurrence of each run length. The average run length of each distribution is calculated by dividing the sum of the products of run length and frequency of occurrence for the distribution by the corresponding total number of runs.

In this report, positive, negative and total run length frequency distributions are also referred to as run length frequency distributions above, below and about the specified value or time series median value (X_o).

3.2.6 Aggregate Minima and Maxima

Aggregate minima and maxima values of a hydrologic variable may be calculated for time periods ranging from one month up to a user specified number of consecutive months, and may be scaled by a user specified value or by the variable's time series mean value. Aggregate minima and maxima values are defined as follows.

Let X_j be the jth value of N values of a discrete stochastic hydrologic variable, and let m be the number of consecutive months for which the aggregate minimum and maximum values of the hydrologic variable are to be calculated. Then from the N values of X, N - m + 1 sums or aggregate values of m consecutive values of X may be calculated. The jth sum or aggregate value of m consecutive values of X is defined as:

$$S_{m,j} = \sum_{k=1}^{m} X_{j+k-1} \qquad (3.19)$$

for all $j \leq N - m + 1$. The minimum and maximum values of all the $S_{m,j}$ values are defined as the aggregate minimum and maximum values respectively of m consecutive values of the hydrologic variable X.

The program calculates aggregate minimum and maximum values of all time periods from one month up to the user specified number of months.

To facilitate a comparison of the aggregate minimum and maximum values of each time period of one hydrologic variable with those of another, the program divides the calculated aggregate minimum and maximum values of each time period of the hydrologic variable by the variable's time series mean value or by a user specified scaling factor.

3.2.7 Range Analysis

The program calculates the range of a stochastic hydrologic variable as the difference of the surplus and deficit values of the variable, where the surplus and deficit values are defined as follows [*Yevjevich* 1972b, pp. 131-132].

Let X_j be the inflow to a reservoir during the jth period of N time periods, and let $\overline{X}$ be the average inflow to the reservoir over these N time periods. Assuming that a draw of size $\overline{X}$ is made on the reservoir during each time period, then

$$X'_j = X_j - \overline{X} \tag{3.20}$$

is either the surplus to be stored in the reservoir, or the deficit to be supplied from the reservoir, during the jth time period.

The sum of the X'_j values over i consecutive time periods,

$$S_i = \sum_{j=1}^{i} X'_j \tag{3.21}$$

is a stochastic variable equal to the sum of the deviations of the X_j values from a constant reservoir release of $\overline{X}$ over the i time periods. The maximum S_i value from all the N values of S_i, S_{max}, is defined as the maximum surplus, or simply the surplus, and the minimum S_i value from all the N values of S_i, S_{min}, is defined as the maximum deficit, or simply the deficit.

The difference

$$R = S_{max} - S_{min} \tag{3.22}$$

is defined as the range, and represents the reservoir capacity required to provide a release of $\overline{X}$ each period, based on the one sample of N inflow values from which it was derived.

If it is assumed that the stochastic hydrologic variable is an independent, normally distributed variable, the program calculates an approximation of the theoretical value of the range as

$$E(R) = 1.25S(\sqrt{N} - 1) \tag{3.23}$$

where S is the standard deviation of the stochastic hydrologic variable [*Yevjevich* 1972b, pp. 148-152].

3.2.8 Storage-Yield Analysis

The program performs an extended mass curve or Rippl analysis to determine the storage-yield relationship for a stochastic hydrologic variable. The extensions made to the conventional mass curve or Rippl analysis are discussed below. A monthly time period has been assumed in the discussion.

The required yield is assumed to be constant each month and is specified as a proportion of the hydrologic variable's time series mean value. If X_j is the jth value of N values of the hydrologic variable, then the sum of the hydrologic variable values to the ith month is:

$$S_i = \sum_{j=1}^{i} X_j \tag{3.24}$$

A plot of these sums against time for all months is described as a mass curve.

Figure 3.1 shows part of a hypothetical mass curve, which can be used to define the storage-yield statistics calculated by the program. A constant monthly yield plots as a sloping line on this figure, whose ordinate value increases from one month to the next by the constant monthly yield. The slope of this line represents the required release rate from the hypothetical storage, and the slope of the mass curve, the inflow rate to the hypothetical storage.

The constant monthly yield line is drawn tangential to the mass curve at time period i where the inflow rate to the hypothetical storage is less than the required release rate. The difference in ordinate values of the constant monthly yield line and the mass curve at subsequent time periods, represents the cumulative or total volume of water which must be drawn from the hypothetical storage to meet the required constant monthly release. The maximum difference in ordinate values is referred to as the storage deficit by the program and is equal to the minimum hypothetical storage size to meet the required constant monthly release over the part of the mass curve shown in Figure 3.1.

There is a storage deficit value for every other part of the mass curve where the inflow rate to the hypothetical storage is less than the constant monthly yield or required release rate. The program calculates and prints out the frequency distribution of these storage deficits, together with the mean storage deficit and the standard deviation of the storage deficit distribution. Additionally, for each interval or bin of the frequency distribution, the program prints out the mean and standard deviation of the storage deficit values, and the minimum and maximum storage deficit values within that interval or bin of the frequency distribution.

The maximum storage deficit value of the entire mass curve is the traditional Rippl storage, and is the minimum hypothetical storage size to supply the required constant monthly yield or release rate over the period of analysis.

To enable a comparison of the storage deficit values of one stochastic hydrologic variable with another, the program divides the calculated storage deficit values by the hydrologic variable's time series mean value or by a user specified scaling factor.

For the hypothetical situation depicted in Figure 3.1, from time period i, when the hypothetical storage is full, until time period i + 6 when the maximum ordinate difference occurs, the hypothetical storage is being generally drawndown because the inflow rate is less than the required release rate. This time period is referred to as the drought duration in the program, and for the part of the mass curve shown in Figure 3.1, is equal to 6 months.

The drought duration corresponding to the maximum storage deficit or Rippl storage is referred to as the Rippl drought duration in the program.

The time period from the maximum ordinate difference or maximum drawdown of the hypothetical storage until when the mass curve crosses the constant yield line, at which time the hypothetical storage has refilled, is referred to as the fill (refill) duration in the program and, for the part of the mass curve shown in Figure 3.1, is equal to

4 months.

The time period from the start of drawdown of the hypothetical storage (at time period i) until the refill of the hypothetical storage (at time period i + 10) is referred to as the draw duration in the program and, for the part of the mass curve shown in Figure 3.1, is equal to 10 months.

There is a drought, fill and draw duration associated with each storage deficit value of every other part of the mass curve where the inflow rate to the hypothetical storage is less than the constant monthly yield or required release rate. The program calculates and prints out the drought, fill and draw duration frequency distributions, together with the mean duration and standard deviation of each duration frequency distribution.

3.2.9 Correlation Analysis

Annual and monthly auto correlation functions and monthly cross correlation functions may be calculated to a user specified number of lags. The cross and auto correlation functions are defined as follows.

Let X_j and Y_j be the jth values of N values of discrete stochastic hydrologic variables X and Y. For example, X and Y may represent monthly streamflows in adjacent catchments. Each may exhibit significant persistence or auto correlation, and they may exhibit significant correlation with each other or cross correlation. Let $\overline{X}$ and $\overline{Y}$ be the time series mean values of X and Y, then the cross correlation coefficient of variables X and Y at a lag k, where k is a positive integer, is defined as:

$$r_{xy}(k) = \frac{\frac{1}{N-k}\sum_{j=1}^{N-k}(X_j - \overline{X})(Y_{j+k} - \overline{Y})}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}(X_j - \overline{X})^2}\sqrt{\frac{1}{N}\sum_{j=1}^{N}(Y_j - \overline{Y})^2}} \tag{3.25}$$

The numerator of this equation is the covariance of X and Y at lag k, and the denominator is the product of the square roots of the variance of X and Y respectively. It can be shown that the cross correlation coefficient of X and Y at a lag -k is equal to the cross correlation coefficient of Y and X at a lag k, which may be written as:

$$r_{xy}(-k) = r_{yx}(+k) \tag{3.26}$$

This result is used to calculate cross correlation coefficients of X and Y for negative lags.

The program calculates the cross correlation function between two variables X and Y as the set of correlation coefficients from a lag of -K to a lag of +K, where K is a user specified maximum number of lags.

The auto correlation coefficient of a variable X at a lag k, where k is a positive integer, is defined as:

$$r_{xx}(k) = \frac{\frac{1}{N-k}\sum_{j=1}^{N-k}(X_j - \overline{X})(X_{j+k} - \overline{X})}{\frac{1}{N}\sum_{j=1}^{N}(X_j - \overline{X})^2} \qquad (3.27)$$

It follows that the auto correlation coefficient of a variable at lag 0 is one, that is,

$$r_{xx}(0) = 1.0 \qquad (3.28)$$

and that the auto correlation function is an even function, that is,

$$r_{xx}(-k) = r_{xx}(+k) \qquad (3.29)$$

Hence, the program calculates the auto correlation function of a variable as the set of correlation coefficients from a lag of +1 to a lag of +K, where K is a user specified maximum number of lags.

Monthly cross and auto correlation functions are calculated according to equations 3.25 and 3.27 respectively. A multiplier of 1/N was used in the variance terms of the denominator of these equations because of the large sample sizes being analysed (N approximately 900). However, annual auto correlation functions are calculated according to equation 3.27 except that a multiplier of 1/(N - 1) is used in the denominator because of the smaller sample sizes.

3.3 PROGRAMMING FEATURES

3.3.1 Introduction

The various statistics which may be calculated by the program have been described in 3.1 and 3.2. How the calculation of these statistics is organised, and details of the calculations are given in following paragraphs.

3.3.2 Program Structure

The program is structured around the analysis of a set of hydrologic data consisting of a common number of concurrent years of monthly records (NYEAR) at a number of stations or sites (NSTN). The program has been dimensioned for a maximum of 25 stations each with a maximum common number of 100 concurrent years of monthly records in each hydrologic data set. The program can analyse any number of hydrologic data sets in a single run.

The inner program loop is over the number of stations in the set of hydrologic data to be analysed. The data of the station to be analysed is read in, and may be echo checked, if requested by the user. The program then tests in turn whether a specific statistic is to be calculated, and if so, calculates the statistic and prints the results. If the results are also to be punched or written to disk files, this is done. If the statistic is not to be calculated, the program skips to the next statistic and tests whether it is to be calculated, and so on, to the end of the inner loop.

However, cross correlation functions are not calculated within the inner program loop. In this loop, the program tests whether cross correlation analysis has been specified, and if so, modifies and writes the station's data to a unique disk file for the subsequent calculation of cross correlation functions. (The calculation of correlation functions is more fully described in 3.3.6.)

The outer program loop is over the number of sets of hydrologic data to be analysed. The problem data, such as, the number of stations, the number of years of record, and the program options, such as, the statistics to be calculated and the output required, are read in before the start of the inner program loop. (The problem data and program options that must be specified are fully described in 3.4.)

If the problem data and the program options are the same for each hydrologic data set to be analysed, then this can be specified as described in 3.4 and the problem data and the program options need only then be specified once, for the first hydrologic data set analysed.

If cross correlation analysis was specified for the hydrologic data set being analysed, then this is performed at the end of the inner loop when all station data has been read in, modified and written to unique disk files and all other required statistics calculated, and just before the end of the outer program loop.

At the end of the outer program loop, when all hydrologic data sets have been analysed, all the user specified punch card output, which was temporarily written to user defined disk scratch files in card image format, is transferred to the system card punch file for subsequent punching. (The organisation of card punching is more fully described in 3.3.7.)

It can be seen from the foregoing discussion that it is a relatively simple matter to modify the program to calculate any additional statistics the user may require.

3.3.3 Distribution Moments

The first, second, third and fourth moments of a discrete variables distribution are defined by equations 3.8 to 3.11 respectively. The second, third and fourth moment equations may be expanded and rewritten as:

$$M_2 = \frac{1}{(N-1)} \cdot \left\{ \sum_{i=1}^{N} X_i^2 - N\overline{X}^2 \right\} \tag{3.30}$$

$$M_3 = \frac{N}{(N-1)(N-2)} \cdot \left\{ \sum_{i=1}^{N} X_i^2 - (3\overline{X} \sum_{i=1}^{N} X_i^2) + 2N\overline{X}^3 \right\} \tag{3.31}$$

$$M_4 = \frac{N^2}{(N-1)(N-2)(N-3)} \cdot \left\{ \sum_{i=1}^{N} X_i^4 - 4\overline{X} \sum_{i=1}^{N} X_i^3 + 6\overline{X}^2 \sum_{i=1}^{N} X_i^2 - 3N\overline{X}^4 \right\} \tag{3.32}$$

For computational purposes, define S_1, S_2, S_3 and S_4 as:

$$S_1 = \sum_{i=1}^{N} X_i \tag{3.33}$$

$$S_2 = \sum_{i=1}^{N} X_i^2 \tag{3.34}$$

$$S_3 = \sum_{i=1}^{N} X_i^3 \tag{3.35}$$

$$S_4 = \sum_{i=1}^{N} X_i^4 \tag{3.36}$$

With these definitions, the four moment equations may be re-written as:

$$M_1 = S_1/N \tag{3.37}$$

$$M_2 = \frac{1}{N-1}\,(S_2 - \frac{1}{N} \cdot S_1^2) \tag{3.38}$$

$$M_3 = \frac{N}{(N-1)(N-2)}\,(S_3 - 3M_1S_2 + 2NM_1^3) \tag{3.39}$$

$$M_4 = \frac{N^2}{(N-1)(N-2)(N-3)}\,(S_4 - 4M_1S_3 + 6M_1^2S_2 - 3NM_1^4) \tag{3.40}$$

Equations 3.37 to 3.40 are used in subroutines TALY3, TAB4, TSFREQ and RIPPL5 to calculate the first four moments of a discrete variable's distribution.

To calculate the individual monthly moments or frequency distributions, the monthly station data stored in chronologic order in vector QTS is reordered by water year month in vector QMON by subroutine MATTS, with each water year month's values in chronologic order. The individual monthly moments may then be calculated by subroutine TALY3, or both the individual monthly moments and frequency distributions by subroutine TAB4.

To calculate the annual moments or frequency distribution, the annual time series is first formed in vector QMON from the monthly time series in vector QTS by subroutine FANN. The annual moments may then be calculated by subroutine TALY3, or both the annual moments and frequency distribution by subroutine TAB4.

The mean and standard deviation of a station's time series distribution are used in the calculation of a number of statistics, for example, monthly auto and cross correlation coefficients. Provision has therefore been made in the program to store these statistics of each station. The station time series means are stored in vector VMEAN, and standard deviations in vector VSD. The corresponding elements of vectors IAVER and ISD are set to '1' to indicate that the corresponding statistic has been stored. (The elements of vectors IAVER and ISD are initialised to '0' between the outer and inner program loops.)

3.3.4 Frequency Distributions

The class interval bounds for a frequency distribution analysis are defined by the upper and lower bounds and the number of class intervals or bins specified for the analysis. In this program suite, one class interval or bin is set aside for values less than the specified lower bound, and another for values greater than or equal to the specified upper bound. The net number of class intervals for the analysis is thus the specified number of class intervals minus two. The class interval size is thus calculated as:

$$\text{SINT} = \frac{(\text{BDUP} - \text{BDLOW})}{(\text{NOBIN} - 2)} \tag{3.41}$$

where SINT is the class interval size,
BDUP is the specified upper bound,
BDLOW is the specified lower bound,
and NOBIN is the specified number of class intervals or bins.

The first class interval or bin is for variable values less than BDLOW. The second class interval for values greater than BDLOW but less than or equal to (BDLOW + SINT), and so on. The last or NOBINth class interval is for values greater than or equal to BDUP.

The class interval into which a particular variable value falls is calculated as:

$$\text{INDEX} = \text{MIN}\emptyset((\text{MAX}\emptyset(\text{INT}((\text{VAL} - \text{BDLOW})/\text{SINT} + 1.0),0) + 1),\text{NOBIN}) \tag{3.42}$$

where INDEX is the number of the class interval into which the value falls,
MINØ is a Fortran inbuilt function which returns the minimum of two integer arguments,
MAXØ is a Fortran inbuilt function which returns the maximum of two integer arguments,
INT is a Fortran inbuilt function which truncates a floating point number to an integer,
VAL is the variable value
and BDLOW, SINT and NOBIN are as defined above.

Having determined the class interval into which the variable value falls, the count of values in the interval is incremented by 1 by:

$$\text{FREQ(INDEX)} = \text{FREQ(INDEX)} + 1.0 \tag{3.43}$$

where FREQ is a vector of frequencies of values in each class interval,

and INDEX is as defined above.

Equations 4.41 to 4.43 are used in subroutines TAB4, TSFREQ and RIPPL5 to calculate variable frequency distributions. TAB4 may be used to calculate monthly, annual and time series frequency distributions. For these analyses, TAB4 determines appropriate bounds and uses 20 class intervals.

TSFREQ may be used to calculate time series frequency distributions with 20 class intervals and user specified bounds.

Subroutine RIPPL5 calculates drought, draw and fill duration, and storage deficit frequency distributions using user specified upper bounds and number of class intervals.

3.3.5 Storage-Yield Analysis

The storage-yield analysis described in 3.2.8 may be performed on each hydrologic variable for up to five different yield rates. The user specifies the number of yield rates and the size of each in the input data deck as described in Table 3.3. The program user must also specify in the input data deck appropriate upper bounds for the drought, draw and fill duration distributions, for the storage deficit distribution, and an appropriate number of bins for all these frequency distributions for each variable at each yield level.

The selection and specification of an appropriate upper bound and number of bins for each distribution for an analysis of one variable at one yield level is described in following paragraphs. This procedure may be repeated for each variable for each yield rate and the input data deck prepared as described in Table 3.3.

The upper bounds for the drought, draw and fill duration distributions may be selected from among those assigned to the matrix BDCOM by data statement in the main program. BDCOM is dimensioned as (10,2) and common upper bounds for the drought and draw duration distributions are assigned to column 1, and upper bounds for fill duration distributions to column 2. The user selects the upper bounds to be used for drought, draw and fill duration distributions by specifying the row subscript of matrix BDCOM which has the most appropriate bounds.

Storage deficit values are scaled by the variable's time series mean or by a user specified scaling factor. An appropriate upper bound for the scaled storage deficit distribution is selected from among those assigned to the vector BDDEF by data statement in the main program. BDDEF is dimensioned as (20) and the user selects the upper bound to be used by specifying the subscript of the most appropriate bound.

An appropriate number of bins for these distributions is determined by the user's selection of the row of matrix BDCOM which has the most appropriate drought, draw and fill duration distribution upper bounds. Alternative numbers of bins are assigned to the vector IBINS, dimensioned as (10), by data statement in the main program. These numbers of bins were selected to correspond to the upper bounds of the corresponding rows of BDCOM so that the user selected row of BDCOM could also be used as the subscript of the element of IBINS which has an appropriate number of bins for the drought, draw and fill duration distributions and the storage deficit distribution.

It may be necessary to perform a storage-yield analysis a few times to determine suitable upper bounds for the drought, draw and fill duration distributions, and for the storage deficit distribution, so that there is neither a concentration of values in only a few bins nor beyond the upper bound. Two analyses have generally proven to be adequate.

3.3.6 Correlation Analysis

Annual and monthly auto correlation functions and monthly cross correlation functions are calculated using subroutine COVAR3. This subroutine calculates the dot product of two time series from zero lag to the maximum specified lag and divides each dot product by a given value, which is the time series variance for auto correlation analysis, and the product of the time series standard deviations for cross correlation analysis. It is assumed that each time series has had its mean value subtracted from all values.

Annual and monthly auto correlation functions are calculated from zero lag to the user specified maximum lag by a single call to COVAR3. The annual auto correlation function is calculated immediately after the annual time series moments or frequency distribution. If annual auto correlation analysis has been specified, the program tests whether the annual moments were calculated and, if so, subtracts the annual mean from the annual time series in vector QMON, squares the annual standard deviation to get the annual variance and calls COVAR3 to calculate the annual auto correlation function. If the annual moments were not calculated, the program forms the annual time series, calculates the annual mean and variance, subtracts the mean from each term of the annual series and then calls COVAR3 to calculate the annual auto correlation function.

Monthly cross correlation functions are calculated by two calls to subroutine COVAR3. The first call calculates cross correlation coefficients between the two series, say X and Y, from zero lag to the user specified maximum lag for Y leading X. The second call calculates cross correlation coefficients from zero lag to the user specified maximum lag for Y lagging X by using the equivalence relation of Y lagging X to X leading Y shown by equation 3.26.

It follows from equation 3.26 that there are NSTN(NSTN - 1)/2 distinct cross correlation functions between NSTN stations. The program calculates the cross correlation functions of station 1 with each of stations 2 to NSTN first, followed by the cross correlation functions of station 2 with each of stations 3 to NSTN next, and so on, ending with the cross correlation function of stations (NSTN - 1) and NSTN.

Within the inner program loop described in 3.3.2, the program tests whether monthly auto and/or cross correlation analysis has been specified, and if so, subtracts the mean time series value from each monthly value. If cross correlation analysis was specified, the program writes this modified time series for this station to a unique disk scratch file for later use. The auto correlation function is now calculated, if required.

After all the individual station data has been read in and analysed, cross correlation analysis is performed as described above, if required. The organisation of the calculation of monthly auto and cross correlation functions is more fully described in the program flowchart, shown in Figure 3.2.

3.3.7 Punched Card Output

The program has the facility to punch the calculated values of those statistics which may be compared using the statistic comparison programs. These statistics are listed in Table 3.1. This facility may be used in the historical data analysis to obtain punched values of the historical data statistics suitable for inclusion in the input data decks of the statistic comparison programs.

The statistics are punched to an appropriate format for the statistic comparison programs, and each punched card carries a station identification number, a statistic identification code number, and a card count number. The station identification numbers are those specified by the user except for the cross correlation functions. There are NSTN(NSTN - 1)/2 cross correlation functions between NSTN stations. The station identification numbers used for cross correlation functions are 1, 2, 3, ... NSTN(NSTN - 1)/2, where 1 corresponds to the cross correlation function of the 1st and 2nd stations, 2 to the cross correlation function of the 1st and 3rd stations, and so on. The statistic identification code numbers are listed in Table 3.1.

The station identification, statistic code and card count numbers, assist in splitting the punched output in preparation for input to the appropriate statistic comparison program, and for reordering of the punched cards should they come out of order.

The calculated values of each statistic to be punched are not written directly to the system punch file NPUN, but to temporary (scratch) card image disk files. The file program names and assigned Fortran unit numbers for each statistic are listed in Table 3.1. After all hydrologic data sets have been analysed, the contents of each of these files is read and written to the system punch file NPUN. This action is taken to minimise the splitting and sorting of punch output required to separate the appropriate historical data statistics for each statistic comparison program.

Subroutine PUNI3 is used to write a vector of integer values to a specified card image disk file under format 24I3 in record columns 1-72. Subroutine PUNF12 is used to write a vector of floating point numbers to a specified card image disk file under format 6F12.5 in record columns 1-72. Both subroutines write a station identification number in columns 73-75, a statistic identification number in columns 76-78, and a card count number in columns 79-80.

Subroutine ROUT is used to read each record of a temporary (scratch) card image disk file and transfer it to the system punch file for subsequent punching.

3.3.8 Disk File Output

The program uses disk output and scratch files as listed in Table 3.2. The disk output files are for the calculated values of those statistics which may be compared using the statistic comparison programs. These files are named NF20, NF21, ... NF36 in the program. The statistics written to each file and the statistic comparison program which uses each file are given in Table 3.2.

The disk scratch files are used for cross correlation analysis and for the temporary storage of punched card output. The unit numbers of the files used for cross correlation analysis are held in the vector NFILE. When cross correlation analysis has been specified, each station's time series data is written to a unique disk scratch file for later use as described in 3.3.6.

The disk scratch files used for temporary storage of punched card output are named NPU1, NPU2 .... NPU4 and NPU9, NPU10 .... NPU15. The assigned unit numbers and the statistics temporarily written to each file are given in Table 3.1.

A summary of all the files used by the program suite is given in Table 9.1.

3.3.9 Data Initialisation

The following scalar and vector variables are initialised by data statement in the main program:

(a) the cross correlation analysis scratch file unit numbers (NFILE(1), NFILE(2), ... NFILE(25)),
(b) the synthetic data statistic file unit numbers (NF20, NF21, ... NF36),
(c) the historical data statistic card image scratch file unit numbers (NPU1, NPU2, ... NPU4; NPU9, NPU10 ... NPU15),
(d) the calendar months of the water year (MWYEAR),
(e) the upper bound values for drought, draw and fill duration distributions for storage-yield analysis (BDCOM),
(f) the upper bound values for storage deficit distributions for storage-yield analysis (BDDEF), and
(g) the number of bins for storage-yield analysis (IBINS).

The installation Fortran unit numbers for the card reader, line printer and card punch, and the default unit number for the hydrologic data file, are assigned to the program variables NIN, NOUT, NPUN and INFLOW respectively, at the start of the main program.

The symbols used in line printer histograms and plots are assigned by data statements in subroutines HIST, PLOTS, PLOTR, SCALE2 and FORM2.

3.3.10 Library Subprograms

The Fortran IV Library subprograms required by this program are ABS, ALOG1Ø, AMAX1, AMIN1, FLOAT, INT, MAXØ, MINØ, MOD and SQRT.

3.4 INPUT DATA PREPARATION

The input data for program STATS comprises the problem data, the user selected program options and the hydrologic data to be analysed. The problem data and program options are read from the input data file NIN, and the hydrologic data from the input file INFLOW.

The input data deck is prepared according to Table 3.3. The first card of the input data deck specifies the number of sets of hydrologic data to be analysed and whether the problem data and program options for each set of hydrologic data are the same. If the problem data and program options are the same for each set of hydrologic data, then they are specified for the first set of hydrologic data only. The program will then use the same problem data description and program options for the analysis of all other hydrologic data sets. This program facility is particularly useful in analysing multiple sets of synthetically generated hydrologic data.

If the problem data and program options are not the same for every set of hydrologic data, then they must be specified for each set of hydrologic data.

The second card of the input data deck specifies an identifying run number, whether the hydrologic data is to be echo checked, and the unit number of the file from which the hydrologic data is to be read. The facility to read the hydrologic data from a different file to that of the problem data and program options was included to simplify the analysis of multiple sets of synthetically generated hydrologic data, which would normally have been written to a unique output file by the generation program. This generation program output file can be made the hydrologic data input file for program STATS, and has a default unit number of 8. Alternatively, the hydrologic data may be read from the same file as the problem data and program options by specifying the unit number of the hydrologic data file INFLOW the same as for the problem data and program option file NIN.

The remaining cards of the input data deck may be prepared from Table 3.3.

3.5 OUTPUT INTERPRETATION

A brief description of the program output is given as the printout is clearly labelled. Sample printout from the example problem described in Section 9 is given in Figure 9.8.

The printout for each hydrologic data set analysed consists of a title page and the tabulated/plotted values of the calculated statistics. The title page has subheadings "Job Description", "Data Description" and "Analysis Options", and provides an echo check of the problem data and program options as read from the input data deck. Under the subheading "Job Description", in the top right hand corner of the title page, the number of hydrologic data sets to be analysed and the number of the particular set for which tabulated/plotted results follow, are printed.

An optional listing of station data under the heading "DATA FOR TIME SERIES NO XXX", where "XXX" is replaced by a user assigned identifying code number for the station, and tabulated/plotted values of the user selected statistics, follow for each station in turn.

The calculated values of each user selected statistic are printed out for each station in turn in a tabular format with descriptive row and column headings, and a title identifying the calculated statistics and station data. For example, the row headings for the monthly moments tabulation are "MEAN", "STD DEV", "SKEW", "KURTOSIS", "COEF VAR"; and the column headings are the numerical abbreviations for the months of the year (that is, 1 for January, 2 for February and so on). The tabulation is titled "MOMENT ANALYSIS ------ MONTHLY ------ STATION NO XXX", where XXX is replaced by a user assigned identifying code number for the station.

Line printer histograms and plots of user selected statistics are preceded by a descriptive title identifying the plotted statistic and the station data. For example, monthly frequency distribution histograms are titled "FREQUENCY HISTOGRAM ------ MONTH NO YY ------ STATION NO XXX", where YY and XXX are replaced by the numerical abbreviation for the month and the user assigned identifying code number for the station respectively.

4. PROGRAM *MOMENT* DESCRIPTION

4.1 INTRODUCTION

Frequency distribution moments and extrema have been defined in section 3.2.3. Program MOMENT is used to compare historical data frequency distribution moments and extrema with corresponding synthetic data moments and extrema. The comparison is made by evaluating a number of measures of the difference between the historical and the synthetic data moments and extrema. The values of these measures may be used to assess whether these historical data characteristics have been satisfactorily preserved in the synthetic data.

The means, standard deviations and skew coefficients of the twelve months, the annual and the time series frequency distributions may be compared. The measures of difference calculated for each moment of each distribution may include the range of synthetic values, the number of synthetic sample values outside the statistically acceptable sampling interval, the moments of the synthetic sampling distribution, the synthetic population value and its statistical acceptability, and the percentage errors between historical and synthetic population values.

The minima and maxima of the twelve monthly frequency distributions may be compared. The measures of difference calculated for both minimum and maximum extremes are the range of synthetic values, the percentages of synthetic values less than and greater than their historical extremes, and the first three moments of the synthetic sampling distribution.

The calculated values of these measures of difference for each moment of each distribution and for the monthly minimum and maximum extremes are printed out for each station.

The historical and synthetic data frequency distribution moments and extremes are obtained as output from STATS.

Punched card output may be obtained which compares the first three moments of the monthly frequency distributions of the historical record and the synthetic population for input to the user's graphical presentation system.

4.2 FUNDAMENTAL CALCULATIONS

The fundamental purpose of the program is to compare the frequency distribution moments and extrema of a station's historical record with its corresponding synthetic data moments and extrema. The frequency distributions used are the 12 water year months, and optionally, the annual and time series, making a possible total of 14 distributions. Each synthetic sample's means, standard deviations and skew coefficients for each distribution are compared with their historical counterparts.

The program also compares the synthetic minima and maxima of the 12 water year months frequency distributions with their historical counterparts.

Each of the three moments of the above distributions is analysed in turn. The inner program loop is over the 12 or 14 moment values, corresponding to the above distributions, calculating the measures of difference between the synthetic sample's moment values and the historical moment values. The outer loop is over all the synthetic samples.

Some of the measures of difference used are calculated in the same way for each of the three distribution moments. These measures are (1) the range of synthetic moment values, (2) the moments of the synthetic moment sampling distributions, and (3) the percentage errors between historical and synthetic population moment values.

The minimum and maximum synthetic moment values are stored in matrices VD and VE respectively and are updated as each synthetic sample is read. When all samples have been processed, the range of synthetic moment values has been defined.

The NSAM synthetic sample values for a distribution's moment define a sampling distribution for that moment. The mean, standard deviation and skew coefficient of this sampling distribution are calculated by moment formulae defined in 3.2.3.

The synthetic population value of a distribution moment is the moment value of the composite distribution formed by combining all the NSAM synthetic sample distributions into the one large composite distribution or population.

The percentage error between an historical and a synthetic population moment value is defined as 100 times the difference of the synthetic population and historical values divided by the historical value. The absolute value of each water year month's percentage error is summed and the mean monthly percentage error is calculated. Similarly, the absolute values of the monthly percentage errors over the six wettest contiguous months and the six driest contiguous months are summed, and the mean monthly percentage errors over the six wettest and six driest months calculated.

The synthetic population moments may be related to known sample moments and other statistics by relationships which will now be derived. The first three moments of a sample of N values have been defined in 3.2.3 as:

$$M_1 = \overline{X} = \frac{1}{N} \sum_{i=1}^{N} X_i \tag{4.1}$$

$$M_2 = \frac{1}{(N-1)} \sum_{i=1}^{N} (X_i - \overline{X})^2 \tag{4.2}$$

$$M_3 = \frac{N}{(N-1)(N-2)} \sum_{i=1}^{N} (X_i - \overline{X})^3 \tag{4.3}$$

where M_1 is the first moment or mean, M_2 the second moment or variance, and M_3 is the third moment.

This sample of N values could be divided into p sub-samples each of q values where $q = N/p$. The first three moments of the kth sub-sample where $1 \leq k \leq p$ may be calculated as:

$$M_1^k = \overline{X}_k = \frac{1}{q} \sum_{i=(k-1)q+1}^{kq} X_i \tag{4.4}$$

$$M_2^k = \frac{1}{(q-1)} \sum_{i=(k-1)q+1}^{kq} (X_i - \overline{X}_k)^2 \tag{4.5}$$

$$M_3^k = \frac{1}{(q-1)(q-2)} \sum_{i=(k-1)q+1}^{kq} (X_i - \overline{X}_k)^3 \tag{4.6}$$

where M_1^k, M_2^k and M_3^k are the first, second and third moments of the kth sub-sample respectively.

The population moments are to be calculated from the known sample moments by relationships of the form:

$$M_1 = f(M_1^k, M_2^k, M_3^k) \tag{4.7}$$

$$M_2 = f(M_1^k, M_2^k, M_3^k) \tag{4.8}$$

$$M_3 = f(M_1^k, M_2^k, M_3^k) \tag{4.9}$$

The moment equations 4.2 and 4.3 may be slightly simplified by introducing the approximations:

$$\frac{1}{(N-1)} \simeq \frac{1}{N} \tag{4.10}$$

$$\frac{N}{(N-1)(N-2)} \simeq \frac{1}{N} \tag{4.11}$$

For monthly and annual data N will typically be greater than 50 so that minimal error is incurred by this approximation. On substituting these approximations into equations 4.2 and 4.3, multiplying out and simplifying, the second and third moment equations may be rewritten as:

$$M_2 = \frac{1}{N}\sum_{i=1}^{N} X_i^2 - \overline{X}^2 \tag{4.12}$$

$$M_3 = \frac{1}{N}\sum_{i=1}^{N} X_i^3 - \frac{3\overline{X}}{N}\sum_{i=1}^{N} X_i^2 + 2\overline{X}^3 \tag{4.13}$$

Similarly, the kth sample second and third moments become:

$$M_2^k = \frac{1}{q}\sum_{i=(k-1)q+1}^{kq} X_i^2 - \overline{X}_k^2 \tag{4.14}$$

$$M_3^k = \frac{1}{q}\sum_{i=(k-1)q+1}^{kq} X_i^3 - \frac{3\overline{X}_k}{q}\sum_{i=(k-1)q+1}^{kq} X_i^2 + 2\overline{X}_k^3 \tag{4.15}$$

The mean of the p sample first moments is:

$$= \frac{1}{p}\sum_{k=1}^{p} M_1^k \tag{4.16}$$

$$= \frac{1}{p}\sum_{k=1}^{p} \overline{X}_k \tag{4.17}$$

This relationship may be rewritten as:

$$= \frac{1}{p}\sum_{k=1}^{p} \frac{1}{q}\sum_{i=(k-1)q+1}^{kq} X_i \tag{4.18}$$

$$= \frac{1}{pq}\left\{\sum_{i=1}^{q} X_i + \sum_{i=q+1}^{2q} X_i + \sum_{i=2q+1}^{3q} X_i + \ldots + \sum_{i=(p-1)q+1}^{pq} X_i\right\} \tag{4.19}$$

$$= \frac{1}{N}\sum_{i=1}^{N} X_i \tag{4.20}$$

that is, the mean of the p sample first moments is the population first moment. The program sums the sample first moments in matrix VA and saves the synthetic population first moments in matrix SPCPAV.

The mean of the p sample second moments is:

$$= \frac{1}{p} \sum_{k=1}^{p} M_2^k \tag{4.21}$$

$$= \frac{1}{p} \sum_{k=1}^{p} \left\{ \frac{1}{q} \sum_{i=(k-1)q+1}^{kq} X_i^2 - \overline{X}_k^2 \right\} \tag{4.22}$$

Assume that a relationship exists between the mean of the p sample second moments and the population second moment, and write as:

$$M_2 = \frac{1}{p} \sum_{k=1}^{p} M_2^k + \Delta_2 \tag{4.23}$$

where Δ_2 is a function of known sample and population statistics. Rearranging and expanding this relationship:

$$\Delta_2 = M_2 - \frac{1}{p} \sum_{k=1}^{p} M_2^k \tag{4.24}$$

$$= \frac{1}{N} \sum_{i=1}^{N} X_i^2 - \frac{1}{p} \sum_{k=1}^{p} \frac{1}{q} \sum_{i=(k-1)q+1}^{kq} X_i^2 - \overline{X}^2 + \frac{1}{p} \sum_{k=1}^{p} \overline{X}_k^2 \tag{4.25}$$

Noting that pq = N and that:

$$\sum_{k=1}^{p} \sum_{i=(k-1)q+1}^{kq} X_i^2 = \left\{ \sum_{i=1}^{q} X_i^2 + \sum_{i=q+1}^{2q} X_i^2 + \ldots + \sum_{i=(p-1)q+1}^{pq} X_i^2 \right\} \tag{4.26}$$

$$= \sum_{i=1}^{N} X_i^2 \tag{4.27}$$

then:

$$\Delta_2 = \frac{1}{p} \sum_{k=1}^{p} \overline{X}_k^2 - \overline{X}^2 \tag{4.28}$$

The population second moment is thus equal to the mean of the sample second moments plus the difference between the mean of the sample first moments squared and the population first moment squared. The sample second moments are summed in matrix SPOPSD. The sum of the sample first moments squared are saved from the first moment analysis in matrix SAVSQ. The population second moment is calculated and then the population standard deviation which is saved in matrix SPOPSD.

Assume that a relationship exists between the mean of the p sample third moments and the population third moments, and write as:

$$M_3 = \frac{1}{p} \sum_{k=1}^{p} M_3^k + \Delta_3 \tag{4.29}$$

where Δ_3 is a function of known sample and population statistics. Rearranging and expanding this relationship:

$$\Delta_3 = M_3 - \frac{1}{p} \sum_{k=1}^{p} M_3^k \tag{4.30}$$

$$= \frac{1}{N} \sum_{i=1}^{N} X_i^3 - \frac{1}{N} \sum_{k=1}^{p} \sum_{i=(k-1)q+1}^{kq} X_i^3 - \frac{3\overline{X}}{N} \sum_{i=1}^{N} X_i^2$$

$$+ \frac{3}{N} \sum_{k=1}^{p} \overline{X}_k \sum_{i=(k-1)q+1}^{kq} X_i^2 + 2\overline{X}^3 - \frac{2q}{N} \sum_{k=1}^{p} \overline{X}_k^3 \tag{4.31}$$

Recognising that:

$$\sum_{k=1}^{p} \sum_{i=(k-1)q+1}^{kq} X_i^3 = \sum_{i=1}^{q} X_i^3 + \sum_{i=q+1}^{2q} X_i^3 + \sum_{i=2q+1}^{3q} X_i^3 + \ldots$$

$$+ \sum_{i=(p-1)q+1}^{pq} X_i^3 \tag{4.32}$$

$$= \sum_{i=1}^{N} X_i^3 \tag{4.33}$$

and that N = pq and that q/N = 1/p, then:

$$\Delta_3 = -\frac{3\overline{X}}{N} \sum_{i=1}^{N} X_i^2 + \frac{3}{N} \sum_{k=1}^{p} \overline{X}_k \sum_{i=(k-1)q+1}^{kq} X_i^2 + 2\overline{X}^3 - \frac{2}{p} \sum_{k=1}^{p} \overline{X}_k^3 \tag{4.34}$$

Now the population second moment is:

$$M_2 = \frac{1}{N} \sum_{i=1}^{N} X_i^2 - \overline{X}^2 \tag{4.35}$$

so that:

$$\sum_{i=1}^{N} X_i^2 = N(M_2 + \overline{X}^2) \tag{4.36}$$

and similarly for the kth sample second moment:

$$M_2^k = \frac{1}{q} \sum_{i=(k-1)q+1}^{kq} X_i^2 - \overline{X}_k^2 \tag{4.37}$$

so that:

$$\sum_{i=(k-1)q+1}^{kq} X_i^2 = q(M_2^k + \overline{X}_k^2) \tag{4.38}$$

On substitution of these results and rearranging:

$$\Delta_3 = \frac{1}{p} \sum_{k=1}^{p} \overline{X}_k^3 + \frac{3}{p} \sum_{k=1}^{p} \overline{X}_k M_2^k - \overline{X}(3M_2 + \overline{X}^2) \tag{4.39}$$

The population third moment may thus be calculated as the mean of the sample third moments plus Δ_3. The sample third moments are summed in matrix SPOPSK. The sample first moments cubed are saved in matrix SAVCU from the first moment analysis. The sum of the products of sample first moments times sample second moments is saved from the second moment analysis in matrix SAVXSD. The final term of Δ_3 may be calculated from the saved population results in matrices SPOPAV and SPOPSD. The population third moment is calculated and then divided by the population standard deviation cubed to yield the population skew coefficient which is saved in matrix SPOPSK.

A number of statistical tests are available for testing sample means and standard deviations. The appropriate test is determined by the assumptions made about the population statistics. In this instance the historical means and standard deviations are taken as the population statistics.

The Central Limit theorem states that if a random sample of size N is drawn from a population of known mean (μ_H) and variance (σ_H^2), then as N increases, the distribution of the statistic:

$$\frac{\overline{X} - \mu_H}{\sigma_H/\sqrt{N}} \tag{4.40}$$

where $\overline{X}$ is the sample mean, is approximately normal with mean 0.0 and variance 1.0 regardless of the distribution of the variable X. As sample sizes used are large (N > 50), this result may be used to define the sampling interval for the sample means. At a given significance level α, the statistically acceptable sampling interval for $\overline{X}$ is:

$$\mu_H - N_{\alpha/2} \frac{\sigma_H}{\sqrt{N}} < \overline{X} < \mu_H + N_{\alpha/2} \frac{\sigma_H}{\sqrt{N}} \qquad (4.41)$$

where $N_{\alpha/2}$ is the normal distribution statistic corresponding to the significance level $\alpha/2$. The sampling interval for each of the distribution means, at the specified significance level α, may be determined from the distribution's historical moments (μ_H, σ_H^2) and the number of values belonging to the distribution in each synthetic sample (N). The sampling interval of the synthetic population means of each distribution is calculated similarly, but N is now the product of the number of synthetic samples and the number of values belonging to the distribution in a sample.

Each synthetic sample's distribution means are checked against their sampling intervals and the number of sample means below the lower bound and above the upper bound is counted for each distribution. Whenever any distribution mean lies outside of its sampling interval, the synthetic sample in which this occurs is also noted. Each distribution's synthetic population mean is also compared against its sampling interval and any violation is noted.

The χ^2 test is applied to the synthetic sample and synthetic population standard deviations. This test's usual assumption of a normally distributed population may be relaxed when testing large samples. The statistic:

$$(N-1) \frac{\sigma_S^2}{\sigma_H^2} \qquad (4.42)$$

where σ_H is the population standard deviation, σ_S the sample standard deviation and N the number of values, is thus distributed as χ^2 with N-1 degrees of freedom. The statistically acceptable sampling interval of σ_S at a given significance level α, is thus:

$$\frac{\sigma_H^2}{(N-1)\ \chi^2_{\alpha/2,N-1}} < \sigma_S^2 < \frac{\sigma_H^2}{(N-1)\ \chi^2_{1-\alpha/2,N-1}} \qquad (4.43)$$

where $\chi^2_{\alpha/2,N-1}$ and $\chi^2_{1-\alpha/2,N-1}$ are the upper and lower tail χ^2 statistics respectively, corresponding to significance level $\alpha/2$ and N-1 degrees of freedom. The sampling interval for each distribution standard deviation, at the specified significance level α, may be determined from the distribution's historical standard deviation, the number of values belonging to the distribution in each synthetic sample (N), and the appropriate χ^2 statistics. The sampling interval of the synthetic population standard deviation of each distribution is calculated similarly, but N is now the product of the number of synthetic samples and the number of values belonging to the distribution in a sample.

Each synthetic sample's standard deviations are compared with their sampling intervals and the number of sample standard deviations outside either bound is counted for each distribution. Whenever any standard deviation lies outside of its sampling interval, the synthetic sample in which this occurs is also noted. Each synthetic population standard deviation is also compared with its sampling interval and any violation noted.

In the preceding discussion of statistical tests on means and standard deviations only a single significance level was mentioned. The program, however, performs these tests at two significance levels simultaneously. The significance levels are selected from the commonly used values of 1%, 5%,and 10% by means of the input variables IPCEN1 and IPCEN2. The results of the tests at each significance level are printed out.

The same measures of difference are used to compare the twelve monthly historical and synthetic minima, and the twelve monthly historical and synthetic maxima. Some of these measures of difference used to compare each extreme are calculated in an identical manner to their distribution moment counterparts. These measures are (1) the range of synthetic extreme values, and (2) the three moments of the sampling distribution of the synthetic extreme values.

For a minimum extreme, the program counts the number of synthetic extreme values less than or equal to the historical extreme. For a maximum extreme, the number of synthetic extreme values greater than or equal to the historical extreme are counted. From these totals the percentages of synthetic extreme values more extreme than their historical counterparts are calculated.

4.3 PROGRAMMING FEATURES

4.3.1 Introduction

The calculation of the measures of difference used in comparing a station's historical record moments and extrema with its NSAM synthetic record moments and extrema has been described. How these calculations are organised for many stations is described now.

4.3.2 General Program Outline

The main program is divided into four major sections. Each of the first three sections corresponds to one of the first three distribution moment analyses, and the fourth to the extremes analysis.

Similar calculations are required in the analysis of each moment and these are performed by the subroutines BEGIN, SAM1, SAM2 and OUTMOM. The 'Purpose' section of each program's listing describes its particular function. To facilitate communication among these routines and with the main program, COMMON storage was used.

The major common blocks are PARAM, KEEP, SPACE1 and SPACE2. PARAM contains the parameter variables for the analysis. KEEP the historical moments and extremes matrices, the synthetic population moments and special purpose matrices. SPACE1 and SPACE2 contain real and integer type work matrices used in all four sections.

Matrices are typically dimensioned as (25,14). The first subscript is a station subscript allowing up to 25 stations to be analysed. The second subscript is the distribution subscript, values 1 to 12 inclusive corresponding to the twelve monthly distributions, 13 to the annual distribution and 14 to the time series distribution.

The sequence of operations within each of the first three sections is similar. Subroutine BEGIN initialises the work matrices, and any other necessary matrices are initialised in the main program. In the first and second moment sections, the sampling intervals are calculated next for each station (1 to NSTN) and each distribution (1 to MONTHS). The work matrices used to save the bounds are the same in each section and are labelled in the listing of the first section. (Sampling intervals are not calculated for the third moment.)

The synthetic moment data is processed next. The outer loop is over the number of synthetic samples (1 to NSAM) and the inner loop over the number of stations. After a station's synthetic sample distribution moments are read, they are passed to SAM1, which loops over the distribution moments (1 to MONTHS) calculating the measures of difference described in its 'Purpose' section, and updating the results matrices for that station.

The synthetic population distribution moments are calculated and passed with the historical distribution moments to SAM2 which calculates the measures of difference and final results described in its 'Purpose' section. The sample rejection results are finalised for the first and second moments and OUTMOM is called to print the results for each station. Sample rejection results are then printed (in the case of the first and second moments). Punched output comparing the historical and synthetic population monthly moments may now be produced if required.

The fourth section of the program analyses the monthly extremes. Subroutine BEGIN initialises the work matrices and additional results matrices are initialised in the main program. The synthetic extremes data is read in a double loop structure as in previous sections. However, the comparison of a station's synthetic sample extremes with its historical extremes is performed within a third loop over the 12 monthly distributions. Both minimum and maximum extremes are compared and corresponding results matrices updated within this loop. When all the synthetic samples have been processed, the moments of the extremes sampling distribution are calculated, and the results printed for each station.

4.3.3 Sampling Interval Calculations

The value of the χ^2 statistic in the test on standard deviations is a function of both the significance level (α) and the sample size (N). The significance levels used in this program are 1%, 5% and 10%. Sample sizes will be from less than 100 for monthly and annual data to greater than 500 for time series data.

The values of the χ^2 statistics for the monthly and annual data must be explicitly specified by the user in the DATA BLOCK subprogram. Values for two sample sizes may be specified. The upper tail χ^2 values are assigned to matrix CHIUP, and the lower tail χ^2 values to matrix CHID. The columns of these 2 x 3 matrices correspond to the significance levels 1%, 5% and 10%. The two rows correspond to the two possible sample sizes, presently assigned for N = 75 (row 1) and N = 50 (row 2).

The input variables IPCEN1, IPCEN2 and IYR determine the χ^2 statistics used in the monthly and annual standard deviation tests. IPCEN1 and IPCEN2 define the significance level (i.e., the column) and IYR defines the sample size (i.e., the row) in the matrices CHIUP and CHID.

To change the present sample sizes the user should define the appropriate χ^2 values in the DATA BLOCK subprogram and change the "IF" statement before label "14" in the main program for the new sample sizes.

For the larger sample size (i.e., N > 100), as for the time series standard deviation, the upper and lower χ^2 values may be calculated as:

$$\chi^2(\alpha,N) = f\left(1 - \frac{2}{9f} + Z_p\sqrt{\frac{2}{9f}}\right)^3 \qquad (4.44)$$

where f = N-1, and Z_p is a function of the significance level and the tail [*Thompson* 1941]. This formula is evaluated by function CHILAR. The data statement in this function defines values of Z_p for significance levels of 1%, 5% and 10%.

The sampling interval bounds for the synthetic standard deviations (σ_S) described in section 4.2 were for σ_S^2. As values of σ_S are read by the program, the square root of these bounds is taken and σ_S is compared with these modified bounds for more efficient processing.

4.3.4 Data Initialisation

The upper and lower tail χ^2 statistics for significance levels of 1%, 5% and 10% and sample sizes of 75 and 50 years are assigned to matrices CHIUP and CHID respectively, in the DATA BLOCK subprogram, as discussed in the preceding section.

Normal distribution statistics at 1%, 5% and 10% are assigned to vector RNO1 in the DATA BLOCK subprogram.

The calendar months of the water year are assigned to vector MWYEAR in the DATA BLOCK subprogram.

Values of Z_p for use in calculating large sample χ^2 values at significance levels of 1%, 5% and 10% are assigned in the data statement in function CHILAR.

4.3.5 Library Subprograms

The Fortran IV library subprograms required by this program are ABS and SQRT.

4.4 INPUT DATA PREPARATION

4.4.1 Introduction

To use this program to compare historical monthly, annual and time series moments and monthly extremes with their synthetic counterparts, it is necessary to:

(a) analyse the historical data set and determine the historical moments and extremes,

(b) analyse the synthetic data sets and determine the corresponding synthetic moments and extremes,

(c) prepare the input data deck defining program options, file reference numbers and problem data.

The results of these analyses and the input data deck comprise the program input which is described in Tables 4.1 and 4.2.

4.4.2 Historical Data Analysis

The historical monthly moments and extremes are obtained from an analysis of the historical data set by program STATS which will punch these statistics onto cards in the proper format for input to this program. Annual and time series moments are obtained similarly if they are to be compared with their synthetic counterparts.

The historical monthly moments and extremes data immediately follows the 8th card of the input data deck as shown in Table 4.2. Annual and time series moments optionally follow next.

The station order of the historical moments and extremes, and the annual and time series moments, must be the same as for their synthetic counterparts. This may be ensured by maintaining the same station order in the historical and synthetic data sets analysed by STATS.

4.4.3 Synthetic Data Analysis

Each of the NSAM synthetic data sets of NYEAR years of concurrent flows at each of the NSTN stations must be analysed by program STATS to determine the synthetic data moments and extremes. The calculated synthetic monthly means are written to output file NF20, the monthly standard deviations to NF21, the monthly skew coefficients to NF22, the monthly minima and maxima to NF23 and optionally, the annual moments to NF24 and the time series moments to NF25 (Table 4.1). These output files of STATS are the synthetic data input files for this program.

4.4.4 Input Data Deck

The input data file NIN specifies program options, file reference numbers and problem data. These cards may be prepared from Table 4.2 and are followed by the historical moments and extremes data.

4.5 OUTPUT INTERPRETATION

A brief description of the program output is given as the printout is clearly labelled. Sample printout from the example problem described in section 9 is given in Figure 9.11.

The first output page is entitled "SYNTHETIC MOMENT PROCESSING" and provides an echo check of the program options, file reference numbers, problem data, historical moments and extremes as read from the input data deck by the program.

The synthetic first moment data is echo checked under the heading "FIRST MOMENT DATA", and is followed by the results of the comparison of historical and synthetic first moments for each station under the title "MOMENT 1 STATION XXX", where "XXX" is replaced by a user assigned identifying code number for the station. A summary table showing which stations have had at least one first moment statistically rejected, and in which synthetic sample this occurred, is printed for the two specified significance levels under the heading "MOMENT 1 SAMPLE REJECTIONS AT YY PER CENT", where "YY" is replaced by the appropriate numerical value of the two user specified significance levels.

The synthetic standard deviation data is echo checked under the heading "SECOND MOMENT DATA", and is followed by the results of the comparison of historical and synthetic standard deviations for each station under the title "MOMENT 2 STATION XXX", where "XXX" is the station code number. A summary table showing which stations have had at least one standard deviation statistically rejected, and in which synthetic sample this occurred, is printed for the two specified significance levels under the heading "MOMENT 2 SAMPLE REJECTIONS AT YY PER CENT", where "YY" is the appropriate significance level. A combined summary table showing which stations have had at least one first or second moment statistically rejected is printed for the two specified significance levels under the heading "SAMPLE REJECTIONS ON 1ST AND 2ND MOMENTS AT YY PER CENT", where "YY" is the appropriate significance level.

The synthetic skew coefficient data is echo checked under the heading "THIRD MOMENT DATA", and is followed by the results of the comparison of historical and synthetic skew coefficients for each station under the title "MOMENT 3 STATION XXX", where "XXX" is the station code number.

The synthetic minimum and maximum monthly extremes data is echo checked under the heading "EXTREMES DATA", and is followed by the results of the comparison of historical and synthetic minimum and maximum extremes for each station under the title "EXTREMES STATION XXX", where "XXX" is the station code number.

5. PROGRAM *FREQ* DESCRIPTION

5.1 INTRODUCTION

The time series frequency distribution has been defined in section 3.2.4. Program FREQ is used to compare the time series frequency distribution of historical data with that of synthetic data. The historical and synthetic data time series frequency distributions may be calculated by program STATS. These distributions are compared by the evaluation of a number of measures of the difference between them. The values of these measures may be used to assess whether the historical data time series frequency distribution is preserved in the synthetic data.

The measures of difference calculated to compare the historical and synthetic data time series frequency distributions include the range of synthetic sample relative and cumulative probabilities for each distribution bin or class interval, the number of synthetic sample time series frequency distributions which are statistically different from their historical counterpart, and the statistical acceptability of the synthetic population time series frequency distribution.

The results of these calculations, comparing the historical and the synthetic data time series frequency distributions, are printed out for each station.

The historical data time series frequency distribution may be obtained as punched card output from program STATS. The synthetic data time series frequency distributions may be written to a specified disk or tape file by program STATS.

Punched card output may be obtained which compares the relative and cumulative time series probability distributions of the historical and the synthetic data population for input to the user's graphical presentation system.

5.2 FUNDAMENTAL CALCULATIONS

The fundamental purpose of this program is to compare a station's historical data time series frequency distribution with those of a number NSAM of synthetic data samples for the same station. The historical and synthetic data time series frequency distributions must of course have the same upper and lower bounds and the same number of class intervals or bins for a meaningful comparison to be made.

Each of the NSAM synthetic sample time series frequency distributions is read in in turn. The synthetic sample frequency distribution is transformed to a probability distribution by dividing each bin frequency by the number of monthly values.

The range of the synthetic sample relative and cumulative frequencies for each bin is determined by the program. Each synthetic sample's relative and cumulative frequencies are compared for each bin with the minimum and maximum relative and cumulative frequencies of all the synthetic samples processed to date. When a more extreme bin frequency is found, the appropriate relative or cumulative, minimum or maximum, bin frequency is updated. When all synthetic samples have been processed, the range of the synthetic sample relative and cumulative frequencies for each bin of the distribution is determined. The corresponding bin probabilities are calculated by dividing each bin frequency by the number of monthly values of each synthetic sample.

The synthetic population time series frequency distribution is the frequency distribution formed by adding all the NSAM synthetic sample time series frequency distributions. As each synthetic sample frequency distribution is read in, each of its bin frequencies is added to the corresponding element of matrix VA. When all the synthetic samples have been read in, VA contains the synthetic population time series frequency distribution. The synthetic population relative probability distribution is calculated by dividing each bin frequency of the synthetic population frequency distribution by the total number of monthly values considering all NSAM synthetic samples. The synthetic population cumulative probability distribution may then be formed.

The percentage difference at any bin between the cumulative synthetic population and historical probability distributions is calculated as 100 times the cumulative synthetic population bin probability minus the cumulative historical bin probability, divided by the cumulative historical bin probability.

The Kolmogorov-Smirnov one sample test [*Siegel* 1956, pp. 47-52] is used to test for a statistically significant difference between a sample cumulative probability distribution and an empirical population cumulative probability distribution at a specified significance level α. The greatest absolute probability difference between corresponding bins of the sample and population probability distributions is first determined. If this greatest absolute probability difference is greater than $K_{\alpha/2}/\sqrt{N}$, where N is the number of sample observations ($N > 35$), and $K_{\alpha/2}$ the Kolmogorov-Smirnov statistic for a significance level of α, then a statistically significant probability difference exists between the two distributions, and the sample cannot be considered to come from a population with the given empirical cumulative probability distribution.

The historical time series cumulative probability distribution is treated as the empirical cumulative probability distribution against which the synthetic sample time series cumulative probability distributions are tested at the specified significance level α. When a synthetic sample time series frequency distribution is read in, its cumulative probability distribution is formed, and the greatest absolute probability difference from the historical cumulative probability distribution is determined. If this greatest absolute probability difference is greater than $K_{\alpha/2}/\sqrt{N}$, then the bin at which this sample is statistically rejected and the synthetic sample number are noted.

When all the synthetic samples have been tested, the number of sample rejections for each distribution bin or class interval is printed out. A summary table showing which synthetic sample time series probability distributions were statistically significantly different from the historical time series probability distribution is also printed out.

The cumulative synthetic population time series probability distribution is tested in the same way. The greatest absolute probability difference is first determined and then compared to $K_{\alpha/2}/\sqrt{N}$, where N is now the total number of monthly values considering all NSAM synthetic samples. The program prints out for each bin the absolute difference in cumulative probabilities and indicates a statistically significant difference by a '1' opposite the bin.

This discussion of the application of the Kolmogorov-Smirnov test to the testing of synthetic time series probability distributions has considered only a single significance level α. The program, however, performs the test at two significance levels simultaneously. The significance levels are selected from the commonly used values of

1%, 5% and 10% by means of the input variables IPCEN1 and IPCEN2. The results of the test at each significance level are printed out.

5.3 PROGRAMMING FEATURES

5.3.1 Introduction

The calculation of the measures of difference used to compare a station's historical time series frequency distribution with those of its NSAM synthetic data samples has been described in section 5.2. Some details of the calculations are given in following paragraphs.

5.3.2 General Program Outline

The calculations performed in the comparison of the historical data time series frequency distributions with their synthetic data counterparts are organised within the DO 100 do loop in the main program. The results of these calculations are printed out within the DO 200 do loop of the main program. The details of these calculations and the printing out of results are given in the program flowchart shown in Figure 5.1.

Matrices are typically dimensioned as (25,20). The first subscript is a station subscript, dimensioned for up to 25 stations. The second subscript is a frequency distribution bin subscript, dimensioned for up to 20 bins.

The sample-station rejection matrices SOKSS and SOKSL are dimensioned as (52,26). The first subscript is the synthetic sample number subscript, dimensioned for up to 50 samples; values 51 and 52 correspond to the number of samples rejected and to whether the synthetic population was rejected, respectively, for a station. The second subscript is the station subscript, dimensioned for up to 25 stations, and value 26 corresponds to the number of stations which rejected this sample (first subscript).

The program reads in the historical data time series relative frequency distributions, and based on the common number of years of historical data for each station, calculates the historical data time series relative and cumulative probability distributions. From these probability distributions, the program calculates equivalent historical data time series relative and cumulative frequency distributions for the common number of years of synthetic data for each sample, so that the historical and synthetic data relative and cumulative frequency distributions may be directly compared. These equivalent historical data time series relative and cumulative frequency distributions appear in the print-out.

All bin probabilities in this program have been multiplied by 100, that is, the relative probability distributions sum to 100 and not one.

To calculate the mean absolute percentage difference per class interval or bin of a cumulative probability distribution, the program sums the absolute values of the percentage difference for each bin and divides by the number of bins. However, if the lower bound of the cumulative probability distribution was specified as zero, the program divides by the number of bins minus 1 as there can be no values less than zero.

5.3.3 Data Initialisation

The installation Fortran unit numbers for the card reader, line printer and card punch are assigned to the program variables NIN, NOUT and NPUN, respectively, by data statement at the start of the main program. The Kolmogorov-Smirnov two-tail statistics at significance levels of 1%, 5% and 10% are assigned to vector RKOL by the same data statement.

5.3.4 Library Subprograms

The Fortran IV Library subprograms required by this program are ABS, MINØ and SQRT.

5.4 INPUT DATA PREPARATION

5.4.1 Introduction

To use this program to compare historical and synthetic data time series frequency distributions it is necessary to:

(a) analyse the historical data set and determine the time series frequency distributions,

(b) analyse the synthetic data sets and determine their time series frequency distributions, and

(c) prepare the input data deck defining program options and problem data.

The results of these analyses and the input data deck comprise the program input which is described in Tables 5.1 and 5.2.

5.4.2 Historical Data Analysis

The historical data time series frequency distributions are obtained from an analysis of the historical data set by program STATS, which will also punch these frequency distributions onto cards in the correct format and order for input to this program.

These historical data time series frequency distribution cards immediately follow the fourth card of the input data deck as shown in Table 5.2.

The station order of the historical data time series frequency distributions must be the same as for the synthetic data time series frequency distributions. This may be ensured by maintaining the same station order in the historical and synthetic data sets analysed by STATS.

5.4.3 Synthetic Data Analysis

The same upper and lower bounds, and number of bins, must be used for each station in the synthetic data analysis as was used in the historical data analysis. Each of the NSAM synthetic data sets of NYEAR years of concurrent flows at each of the NSTN stations may be analysed by program STATS to determine the synthetic data time series frequency distributions, which are written to output file NF26.

5.4.4 Input Data Deck

The input data file NIN specifies program options, file reference numbers and problem data. These cards may be prepared from Table 5.2 and are followed by the historical data time series frequency distributions.

5.5 OUTPUT INTERPRETATION

A brief description of the program output is given as the printout is clearly labelled. Sample printout from the example problem described in section 9 is given in Figure 9.13.

The first output page is entitled "SYNTHETIC FREQUENCY ANALYSIS" and provides an echo check of the program options, file reference numbers and problem data. The historical relative and cumulative time series probability distributions of each station are printed out. Corresponding relative and cumulative frequency distributions are calculated for each station for the given common number of years of each synthetic sample and are printed out.

The synthetic time series frequency distribution data is echo checked for each station and each synthetic sample under the heading "FREQUENCY ANALYSIS DATA".

The results of the comparison of each station's historical time series frequency distribution with its synthetic counterparts are printed out for each station in turn under the heading "FREQUENCY ANALYSIS STATION XXX", where "XXX" is replaced by a user assigned identifying code number for the station. The results for each station are divided into individual sample results and population results by the headings "INDIVIDUAL SAMPLES" and "POPULATION RESULTS" respectively. The results printed under these two subheadings are the numerical values of the measures of difference discussed in section 5.2.

Sample-station rejection summary tables are printed showing which stations had synthetic sample time series frequency distributions statistically rejected, and for which samples this occurred. These tables are printed for the two user specified significance levels under the heading "SAMPLE REJECTION MATRIX FOR FREQUENCY ANALYSIS AT YY PER CENT", where "YY" is replaced by the appropriate numerical value of the two user specified significance levels.

6. PROGRAM *RUNS* DESCRIPTION

6.1 INTRODUCTION

Run length frequency distributions have been defined in section 3.2.5. Program RUNS is used to compare the runs behaviour of the historical data with that of the synthetic data. Run length frequency distributions above, below and about the historical median (or a user specified value), may be determined for both the historical and the synthetic data by program STATS. These historical and synthetic data run length frequency distributions are compared in this program by the evaluation of a number of measures of the difference between them. The values of these measures may be used to assess whether the historical runs behaviour is preserved in the synthetic data.

The same measures of difference are calculated for the comparison of historical and synthetic data run length frequency distributions above, below and about the historical median (or user specified value). These measures of difference include the range of synthetic sample relative and cumulative probabilities for each run length, the sampling distribution moments of the synthetic sample total number of runs and average run length statistics, the number of synthetic sample run length distributions statistically different from their historical counterpart, and the statistical acceptability of the synthetic population run length distribution.

The total number of "long" run lengths and the average "long" run length of a data sample are calculated by the program because of the significant effect of "long" runs of low flows on the storage-yield behaviour of a stream. The first three sampling distribution moments and the range of each of these "long" run statistics of the synthetic samples may be compared with the historical data total number of "long" runs and average "long" run length statistics to assess the preservation of this important feature of runs behaviour in the synthetic data samples. (This measure of difference is discussed in the following section where a "long" run is also defined.)

The results of these calculations comparing the historical and the synthetic data run length distributions above, below and about the historical median (or user specified value) are printed out for each station by distribution type (i.e., above, below and about).

Although the output page headings of this program state that the run length frequency distributions being compared are about the median, the program may be used to compare run length distributions about any value.

The historical data run length frequency distributions above and below the median or other specified value may be obtained as punched output from STATS. The synthetic data run length frequency distributions above and below the same median or other specified value are written to a specified disk/tape file by STATS.

Punched output comparing the historical and the synthetic population relative and cumulative run length probability distributions above, below and about the median or user specified value may be obtained for input to the user's graphical presentation system.

6.2 FUNDAMENTAL CALCULATIONS

The fundamental purpose of this program is to compare a station's historical data run length frequency distributions above, below and about the median with those of its NSAM synthetic data samples. As the measures of difference calculated in this comparison are the same for runs above, below or about the median, this description of the calculations will consider only one run length frequency distribution, that below the median.

Each of the NSAM synthetic sample run length frequency distributions below the median is read in in turn. The total number of runs observations and the average run length of the sample distribution are calculated by subroutine TOTAV. This total number of runs and average run length, are referred to as the overall total number of runs and average run length, to distinguish them from their "long" run counterparts.

The synthetic sample frequency distribution is transformed to a probability distribution by dividing each run length frequency value by the overall total number of runs for the sample.

The range of the synthetic sample relative and cumulative probabilities for each run length is determined by the program. Each synthetic sample's relative and cumulative probabilities are compared for each run length to the minimum and maximum relative and cumulative probabilities of all the synthetic samples processed to date. When a more extreme run length probability is found, the appropriate relative or cumulative, minimum or maximum, run length probability is updated. When all the synthetic samples have been processed, the range of the synthetic sample relative and cumulative probabilities for each run length of the distribution is determined. These calculations are performed by subroutine RUNS1.

The synthetic population run length frequency distribution below the median is the frequency distribution formed by adding all the NSAM synthetic sample run length frequency distributions below the median. As each synthetic sample frequency distribution is read in, its number of runs for each run length is added to the corresponding run length of matrix NA by subroutine RUNS1. When all the synthetic samples have been read in, NA contains the synthetic population run length frequency distribution below the median. The overall total number of observations and the average run length of the synthetic population distribution is determined, and the synthetic population relative and cumulative probability distributions are calculated by subroutine OUTPUT.

The percentage difference at any run length between the cumulative synthetic population and historical probability distributions is calculated as 100 times the cumulative synthetic population run length probability minus the cumulative historical run length probability, divided by the cumulative historical run length probability.

The Kolmogorov-Smirnov one sample test [*Siegel* 1956, pp. 47-52] is used to test for a statistically significant difference between a sample cumulative probability distribution and an empirical population cumulative probability distribution at a specified significance level α. The greatest absolute probability difference between corresponding bins of the sample and population probability distributions is first determined. If this greatest absolute probability difference is greater than $K_{\alpha/2}/\sqrt{N}$, where N is the number of sample observations ($N > 35$), and $K_{\alpha/2}$ the Kolmogorov-Smirnov statistic for a significance level of $\alpha/2$, then a statistically significant probability difference exists between the two distributions, and the sample cannot be considered to come from a population with the given empirical cumulative probability distribution.

The historical run length cumulative probability distribution is treated as the empirical cumulative probability distribution against which the synthetic sample run length cumulative probability distributions are tested at the specified significance level α. When a synthetic sample run length frequency distribution is read in, its total number of runs is determined and is used to calculate the sample relative probability distribution and the value of the statistic $K_{\alpha/2}/\sqrt{N}$. The sample's cumulative probability distribution is then formed and the greatest absolute probability difference from the historical cumulative probability distribution is determined. If this greatest absolute probability difference is greater than $K_{\alpha/2}/\sqrt{N}$, then the run length at which this sample is statistically rejected and the synthetic sample number are noted. These calculations are performed by subroutine RUNS1.

When all the synthetic samples have been tested, the number of sample rejections for each distribution run length is printed out. A summary table showing which synthetic sample run length probability distributions were statistically significantly different from the historical run length probability distribution is also printed out.

The cumulative synthetic population run length distribution is tested in the same way. The number of runs comprising this distribution is determined and $K_{\alpha/2}/\sqrt{N}$ calculated. The greatest absolute probability difference is determined and compared to $K_{\alpha/2}/\sqrt{N}$. The program prints out for each run length the absolute difference in cumulative probabilities and indicates a statistically significant difference by a '1' opposite the run length. These calculations are performed in subroutine OUTPUT.

This discussion of the application of the Kolmogorov-Smirnov test to the testing of synthetic run length probability distributions has considered only a single significance level α. The program, however, performs the test at two significance levels simultaneously. The significance levels are selected from the commonly used values of 1%, 5% and 10% by means of the input variables IPCEN1 and IPCEN2. The results of the test at each significance level are printed out.

The first three moments of the sampling distributions of the synthetic sample overall total number of runs and average run length statistics are calculated. The sampling distribution statistics are held in columns 1 and 2 respectively of the matrices VA, VB and VC. The range of these two statistics is also determined and their minimum and maximum values are held in columns 1 and 2 respectively of matrices VD and VE. As each synthetic sample frequency distribution is read in, its overall total number of runs and average run length are calculated by subroutine TOTAV, the sampling distribution statistics are updated, and the minima and maxima are checked and updated as necessary. These calculations are performed in subroutine RUNS1. After all the synthetic samples have been processed, the mean, standard deviation and skew coefficient of the two sampling distributions are calculated and printed by subroutine OUTPUT, together with the range of each statistic.

To compare the "long" runs behaviour of the historical data with that of the synthetic data, it is necessary to decide what run lengths are to be considered "long" and to choose suitable measures of "long" run behaviour.

All run lengths equal to, or longer than a "minimum long run" length, will be considered as "long" runs. The "minimum long" run length is defined as that run length with a probability of exceedance of ILRUN% for runs about the median. As the run length frequency distribution about the median is not continuous but discrete, the longest run length with a probability of exceedance of greater than ILRUN% is taken as the minimum long run length. A value of 5% for ILRUN has proven satisfactory in work to date.

The minimum long run length may be thought of as defining the lower limit of a frequency distribution of long run lengths, the upper limit of which is set by the maximum run length (NRUNS) specified for the analysis. The total number of long runs and the average long run length of this distribution, were chosen as measures of long run behaviour, which could be used to compare historical and synthetic data long run behaviours. This total number of runs and average run length are referred to as the long run total number of runs and average run length.

After the historical run length frequency distributions above and below the median are read in, the historical overall total number of runs and average run length are calculated for each distribution by subroutine TOTAV. These distributions are then added to form the historical run length frequency distribution about the median, and its overall total number of runs and average run length are calculated by

TOTAV. ILRUN% of the overall total number of runs about the median is calculated and is used to determine the minimum long run length, which is saved in vector NRUNL. The actual number of runs about the median, equal to and longer than the minimum long run length, are noted. The historical long run total number of runs and average run length are calculated for all three runs distributions by subroutine TOTAV. The results of these calculations are printed on the first output page which is entitled "SYNTHETIC RUNS ANALYSIS".

The first three moments of the sampling distributions of the synthetic sample long run total number of runs and average run length statistics are calculated. The sampling distribution statistics are held in columns 3 and 4 respectively of the matrices VA, VB and VC. The range of these two statistics is also determined and their minimum and maximum values are held in columns 3 and 4 respectively of matrices VD and VE. As each synthetic sample frequency distribution is read in, its long run total number of runs and average run length are calculated by subroutine TOTAV, the sampling distribution statistics are updated and the minima and maxima checked and updated as necessary. These calculations are performed in subroutine RUNS1. After all the synthetic samples have been processed, the mean, standard deviation and skew coefficient of the two sampling distributions are calculated and printed by subroutine OUTPUT, together with the range of each statistic.

6.3 PROGRAMMING FEATURES

6.3.1 Introduction

The preceding section has described the calculation of the measures of difference used to compare a station's historical data runs behaviour with that shown by its NSAM synthetic data samples. The organisation of these calculations for many stations will now be described.

6.3.2 General Program Outline

The calculations performed in the comparison of each of the historical data run length frequency distributions with their synthetic data counterparts are organised within the DO 400 do loop in the main program. The loop variable is ITYPE which takes on values 1 to 3 corresponding to the comparison of historical and synthetic data run length frequency distributions above, below and about the median respectively. The details of these calculations are described in the flowcharts of programs RUNS, OUTPUT and RUNS1 (Figures 6.1, 6.2 and 6.3).

The similar calculations made for each of these run length frequency distributions are performed by subroutines INIT, RUNS1, TOTAV and OUTPUT. The "Purpose" section of each program's listing summarises its particular function. To facilitate communication among these subroutines and with the main program, COMMON storage was used.

The common blocks are ONE and TWO. Common block ONE contains real and integer type work matrices and the sample-station rejection matrices used in the analysis of each distribution. Subroutine RUNS1's listing includes a description of the use of these work matrices. Common block TWO contains some program option variables, problem data variables and the vector of calculated station minimum long run lengths.

Matrices are typically dimensioned as (25,60). The first subscript is a station subscript, dimensioned for up to 25 stations. The second subscript is the run length subscript, dimensioned for run lengths of up to 60 months.

The sample-station rejection matrices, SOKSS, SOKSL, SOKTS and SOKTL, are dimensioned as (52,26). The first subscript is the synthetic sample number subscript, dimensioned for up to 50 samples; values 51 and 52 correspond to the number of samples rejected and to whether the synthetic population was rejected, respectively, for a station. The second subscript is the station subscript, dimensioned for up to 25 stations, and value 26 corresponds to the number of stations which rejected this sample (first subscript).

All run length probabilities in this program have been multiplied by 100, that is, the relative probability distributions sum to 100 and not one.

Runs above, below and about the median are sometimes referred to as runs up, down and total respectively in the comment statements in the program listings.

6.3.3 Data Initialisation

The Kolmogorov-Smirnov two-tail statistics at significance levels of 1%, 5% and 10% are assigned to vector RKOL at the start of the main program.

The installation Fortran unit numbers for the card reader, line printer and card punch are assigned to the program variables NIN, NOUT and NPUN, respectively, by data statement at the start of the main program.

NOUT is also the variable for the line printer unit number in subprograms OUT1 and OUTPUT. Data statements are used in both subprograms to assign the installation Fortran unit number for the line printer to NOUT.

6.3.4 Library Subprograms

The Fortran IV Library subprograms required by this program are ABS, MINØ and SQRT.

6.4 INPUT DATA PREPARATION

6.4.1 Introduction

To use this program to compare the historical and synthetic data runs behaviours it is necessary to:

(a) analyse the historical data set and determine the run length frequency distributions above and below the median,

(b) analyse the synthetic data sets and determine their run length frequency distributions above and below the historical median,

(c) prepare the input data deck defining program options and problem data.

The results of these analyses and the input data deck comprise the program input which is described in Tables 6.1 and 6.2.

6.4.2 Historical Data Analysis

The historical data run length frequency distributions above and below the median are obtained from an analysis of the historical data set by program STATS, which will also punch these frequency distributions onto cards in the correct format and order for input to this program.

These historical data run length frequency distribution cards immediately follow the fifth card of the input data deck as described in Table 6.2.

The station order of the historical data runs frequency distributions must be the same as for the synthetic data runs frequency distributions. This may be ensured by maintaining the same station order in the historical and synthetic data sets analysed by STATS.

6.4.3 Synthetic Data Analysis

Each of the NSAM synthetic data sets of NYEAR years of concurrent flows at each of the NSTN stations is analysed by program STATS to determine the synthetic data run length frequency distributions above and below the historical data medians. The synthetic data run length frequency distributions are written to output file NF27 by program STATS.

6.4.4 Input Data Deck

The input data file NIN specifies program options, file reference numbers and problem data. These cards may be prepared from Table 6.2 and are followed by the historical data run length frequency distributions above and below the median.

6.5 OUTPUT INTERPRETATION

A brief description of the program output is given as the printout is clearly labelled. Sample printout from the example problem described in section 9 is given in Figure 9.15.

The first output page is entitled "SYNTHETIC DATA ANALYSIS" and provides an echo check of the program options, file reference numbers and problem data. The historical data overall and long run total number of runs and average run length statistics are printed out for the run length frequency distributions above, below and about the historical data medians.

The historical data relative and cumulative run length probability distributions for each station, above, below and about the historical data medians are printed out next under appropriate page headings.

The results of the comparison of the historical data run length frequency distributions above, below and about the historical medians, with their synthetic counterparts, make up the remaining output. The format of the printed results for each run length distribution type differs only in the page headings, which identify a particular run length distribution type. For this reason, only the printout of the results for the run length distribution above the median will be described.

The synthetic run length frequency distribution data above the median is echo checked for each station and each synthetic sample under the heading "RUNS DATA ABOVE THE MEDIAN".

The results of the comparison of each station's historical run length frequency distribution above the median with its synthetic counterparts are printed out for each station in turn, under the heading "RUNS ABOVE THE MEDIAN : STATION XXX", where "XXX" is a user assigned station code number. The results for each station are divided into individual sample results and population results by the headings "INDIVIDUAL SAMPLES" and "POPULATION RESULTS" respectively. The results printed under these two subheadings are the numerical values of the measures of difference discussed in the section 6.2.

Sample-station rejection summary tables are printed, showing which stations had synthetic sample run length frequency distributions above the median statistically rejected, and for which samples this occurred. These tables are printed for the two user specified significance levels, under the heading "SAMPLE REJECTIONS FOR RUNS ABOVE THE MEDIAN", with subheading "AT YY PER CENT", where "YY" is the appropriate significance level.

Sample-station rejection summary tables are also printed for runs above, below and about the median combined for the two user specified significance levels, under the heading "SAMPLE REJECTIONS FOR COMBINED RUNS" with subheading "AT YY PER CENT", where "YY" is the appropriate significance level.

7. PROGRAM *YIELD* DESCRIPTION

7.1 INTRODUCTION

Storage-yield analysis has been described in section 3.2.8. Program YIELD is used to compare the storage-yield behaviour of the historical data with that of the synthetic data. The comparison is made by the evaluation of a number of measures of the difference between the historical and synthetic data storage-yield behaviours. The values of these measures may be used to assess the similarity of the historical and synthetic data storage-yield behaviours.

The Rippl analysis technique is applied to a hydrologic series in program STATS to determine some characteristics of its storage-yield behaviour. These include the Rippl storage and the corresponding drought duration, the storage deficit distribution and the drought, draw and fill duration distributions. The analysis may be made for up to five yield levels. The historical storage-yield characteristics at each yield level are contrasted with their synthetic counterparts.

The first three moments of the synthetic Rippl storage and drought duration distributions are calculated. The number of standard deviations the historical Rippl storage is from the mean of the synthetic Rippl storage distribution is then calculated. The synthetic Rippl storages are ordered from the smallest to the largest value and are printed out with their corresponding drought duration and synthetic sample numbers.

The same calculations are made to compare each of the historical deficit, drought, draw and fill duration distributions with their synthetic counterparts at each yield level. The calculated measures of difference include the range of synthetic sample relative and cumulative probabilities for each frequency interval, the sampling distribution moments of the distribution total and average statistics, the number of synthetic sample distributions statistically different from the historical distribution, and the statistical acceptability of the synthetic population distribution.

The results of these calculations comparing the storage-yield characteristics of the historical and synthetic data are printed out for each station and yield level.

The historical storage-yield behaviour data may be obtained as punched output from program STATS. The synthetic storage-yield behaviour data is written to specified disk/tape files by program STATS.

Punched output comparing the historical and the synthetic population deficit, drought, draw and fill probability distributions may be obtained for input to the user's graphical presentation system.

7.2 FUNDAMENTAL CALCULATIONS

The fundamental purpose of this program is to compare the storage-yield behaviour of a station's historical record with that shown by the NSAM synthetic data samples. This comparison will be made for NYIELD levels, and, as the measures of difference calculated for each yield level are the same, this description of the calculations will consider only one yield level.

The Rippl storage as determined by the traditional Rippl technique, is the minimum storage size that can always fully deliver the constant monthly yield on the basis of the historical or synthetic data sample analysed. The Rippl drought duration is the number of time periods (i.e., months) from the start of the drawdown of this initially full Rippl storage till when it just empties.

The first three moments of the sampling distributions of the NSAM Rippl storages and Rippl drought durations are calculated. The mean, standard deviation and skew coefficient of each of these distributions is then calculated and printed with the historical data Rippl storage and drought duration for comparison.

The number of standard deviations the historical Rippl storage is from the mean of the synthetic Rippl storage distribution is calculated as:

$$\frac{S_H - \overline{S}_S}{\sigma_S} \qquad (7.1)$$

where S_H is the historical Rippl storage, and $\overline{S}_S$ and σ_S the synthetic Rippl storage distribution mean and standard deviation respectively.

As each synthetic sample's Rippl storage and drought duration are read, they are saved in matrices RIPPLE and DURAT respectively, in corresponding matrix elements. The sample number associated with the Rippl storage and drought duration is similarly saved in the corresponding matrix element of SAMPLE. When all of the NSAM synthetic sample

Rippl storages and drought durations have been read, the Rippl storages in matrix RIPPLE are ordered from the smallest to the greatest value by a standard ordering technique. Whenever the order of two Rippl storage values is reversed in the ordering process, the corresponding drought durations and sample numbers are also reversed, so that the matrix element correspondence between Rippl storage, drought duration and sample number values is preserved. The probability of any synthetic Rippl storage being less than or equal to a given synthetic Rippl storage value is calculated as the rank of the given storage value divided by the number of samples plus one. The ordered Rippl storages are printed out with their associated drought durations, sample numbers and probabilities.

These calculations on the Rippl storages and drought durations are all performed in subroutine STOAGE.

The Rippl storage is actually the largest storage of a distribution of different storage sizes, each storage corresponding to a particular dry period in the record, when the stream flow rate was less than the yield rate, and each sized so that it just emptied during that dry period, while still fully satisfying the constant monthly yield. This storage distribution has been named the storage deficit distribution or just deficit distribution.

Each storage of this deficit distribution has associated with it its own drought duration. The distribution of these storage drought durations was named the drought duration distribution or simply the drought distribution. Each storage also has associated with it a refill duration, that is, the number of time periods (i.e., months) it takes for the storage to fill from the empty state. The distribution of these values was named the fill duration distribution or fill distribution. The distribution formed by adding each storage's drought and fill duration times was named the draw duration distribution or just draw distribution.

The number of observations in the draw and fill distributions may occasionally be one less than in the deficit and drought distributions. This happens when the most severe dry period in the record occurs near the end of the record so that the storage does not have the opportunity to refill.

These four distributions are a natural consequence of the application of the Rippl technique of storage analysis for a specified yield level. They may be determined from a station's historical record and compared with their synthetic data counterparts at each yield level. The same measures of difference are calculated for each distribution in making this comparison, and so the calculations will be described for one distribution only, the deficit distribution.

There are NSAM synthetic sample deficit frequency distributions at each yield level for each station. Each synthetic sample distribution is read in in turn with its total number of deficit observations and its average deficit value. Each deficit frequency distribution is transformed to a probability distribution by dividing each bin frequency value by the total number of deficit observations for the distribution.

The range of synthetic sample relative and cumulative probabilities for each bin of the deficit distribution is calculated. Each synthetic sample's relative and cumulative probabilities are compared for each bin to the minimum and maximum relative and cumulative probabilities of all of the synthetic samples processed to date. When a more extreme bin probability is found, the appropriate relative or cumulative, minimum or maximum, bin probability is updated. When all synthetic samples have been processed, the range of the synthetic sample relative and cumulative probabilities for each bin of the distribution is determined.

The first three moments of the sampling distributions of the synthetic distribution total number of observations and average deficit value are calculated. Each synthetic sample's total number of observations and average deficit value are read in with the sample's deficit frequency data. The sampling distribution statistics of the total number of observations and the average deficit value held in matrix VA are then updated by subroutine YIELD1. The range in value of these two distribution statistics is also checked and updated when necessary. After all synthetic samples have been processed, the mean, standard deviation and skew coefficient of the two sampling distributions are calculated and printed by subroutine OUTPUT, together with the range of each statistic.

The synthetic population deficit frequency distribution is the frequency distribution formed by adding all the NSAM synthetic sample deficit frequency distributions. After a synthetic sample frequency distribution is read in, the number of observations in each bin are added to the corresponding bin of matrix NA. When all the synthetic samples have been read in, NA contains the synthetic population frequency distribution. The total number of observations in NA is determined and the synthetic population relative and cumulative probability distributions are calculated by subroutine OUTPUT.

The percentage difference at any bin between the cumulative synthetic population and cumulative historical probability distributions is calculated as 100 times the cumulative synthetic population bin probability minus the cumulative historical bin probability divided by the cumulative historical bin probability.

The Kolmogorov-Smirnov one-sample test [*Siegel* 1956, pp. 47-52] tests for a statistically significant difference between a sample cumulative probability distribution and an empirical population cumulative probability distribution at a specified significance level α. The greatest absolute probability difference between corresponding bins of the sample and population probability distributions is first determined. If this greatest absolute probability difference is greater than $K_{\alpha/2}/\sqrt{N}$, where N is the number of sample observations ($N > 35$) and $K_{\alpha/2}$ the Kolmogorov-Smirnov statistic for a significance level of $\alpha/2$, then a statistically significant probability difference exists between the two distributions, and the sample cannot be considered to come from a population with the given empirical cumulative probability distribution.

The historical deficit cumulative probability distribution is treated as the empirical cumulative probability distribution against which the synthetic sample probability distributions are tested at the specified significance level α. When a synthetic sample deficit distribution and its total number of observations and average value are read in, the number of observations is used to calculate the value of the statistic $K_{\alpha/2}/\sqrt{N}$. The sample's cumulative probability distribution is then formed and the greatest absolute probability difference from the historical cumulative probability distribution is determined. If this greatest absolute probability difference is greater than $K_{\alpha/2}/\sqrt{N}$, then the bin in which this sample is statistically rejected and the synthetic sample number are noted. When all the synthetic samples have been tested, the number of sample rejections for each distribution bin is printed out. A summary table showing which synthetic sample probability distributions were statistically significantly different from the historical deficit probability distribution is also printed. These calculations are performed in subroutine YIELD1 as each synthetic sample is read.

The cumulative synthetic population distribution is tested in the same way. The number of observations comprising this distribution is determined and $K_{\alpha/2}/\sqrt{N}$ calculated. The greatest absolute probability difference is determined and compared to $K_{\alpha/2}/\sqrt{N}$. The program prints out for each bin the absolute differences in cumulative probabilities and indicates a statistically significant difference by a '1' opposite the bin number. These calculations are performed in subroutine OUTPUT.

In the preceding discussion of the application of the Kolmogorov-Smirnov test to the testing of synthetic probability distributions, only a single significance level α was mentioned. The program, however, performs the test at two significance levels simultaneously. The significance levels are selected from the commonly used values of 1%, 5% and 10% by means of the input variables IPCEN1 and IPCEN2. The results of the test at each significance level are printed out.

7.3 PROGRAMMING FEATURES

7.3.1 Introduction

The preceding section has described the calculation of the measures of difference used to compare a station's historical data storage-yield behaviour with that shown by NSAM synthetic data samples at a particular yield level and for a single station. The organisation of these calculations for multiple yield levels and stations will now be described.

7.3.2 General Program Outline

The main program may be divided into two major sections. The first section performs the analysis of the Rippl storages and drought durations by calling subroutine STOAGE. The second section compares the historical and synthetic data storage deficit, drought duration, draw and fill duration distributions.

Similar calculations are made for each of these four distributions and these are performed by subroutines INIT, YIELD1 and OUTPUT. Flowcharts of YIELD1 and OUTPUT are given in Figures 7.2 and 7.3.

The "Purpose" section of each program's listing summarises its particular function. To facilitate communication among these subroutines and with the main program, COMMON storage was used.

The major common blocks are ONE, TWO and THREE. Common block ONE contains real and integer type work matrices, and the sample-station rejection matrices used in the analysis of each distribution. Common blocks TWO and THREE contain file reference number variables, program option variables, problem data variables and some storage-yield behaviour analysis parameters and results.

Blank common is used for the communication of the historical Rippl storages and drought durations from the main program to subroutine STOAGE.

Matrices are typically dimensioned as (25,5,20). The first subscript is a station subscript, dimensioned for the analysis of up to 25 stations. The second subscript is the yield level subscript, dimensioned for up to 5 yield levels. The third subscript is the frequency bin subscript, dimensioned for up to 20 frequency bins.

The sample-station rejection matrices, SOKSS and SOKSL, are dimensioned as (52,6,26). The first subscript is the sample number subscript, dimensioned for up to 50 synthetic samples, values 51 and 52 correspond to the number of samples rejected and to whether the synthetic population was rejected, respectively, for a station (third subscript). The second subscript is the yield level subscript, dimensioned for up to 5 yield levels, and value 6 corresponds to the combined sample-station rejection results for all the NYIELD yield levels. The third subscript is the station subscript, dimensioned for up to 25 stations, and value 26 corresponds to the number of stations which rejected this sample (first subscript).

The organisation of the calculations for the analysis of the synthetic Rippl storages and drought durations is shown in subprogram STOAGE's flowchart (Figure 7.4).

The calculations performed in the comparison of each of the four historical data frequency distributions with their synthetic data counterparts are organised within the DO 500 do loop in the main program. The loop variable is JTYPE which takes on values 1 to 4 corresponding to the comparison of historical and synthetic storage deficit, drought duration, draw duration and fill duration distributions respectively. The organisation of the calculations within this loop is shown in program YIELD's flowchart (Figure 7.1).

7.3.3 Echo Check of Synthetic Frequency Data

A special output technique has been written into the main program to minimise the volume of printed output when an echo check of the synthetic sample frequency data is requested.

The synthetic sample frequency data is printed out in blocks of 40 columns across the page by up to 20 rows down the page. Each column corresponds to a synthetic frequency distribution for a particular station, at a particular yield level and from a particular synthetic sample. Each row of each column corresponds to a frequency distribution bin.

When an echo check is requested, each synthetic frequency distribution is read in in turn into a column of matrix IDATA of 40 columns by 20 rows, and its station number, sample number and yield level number are assigned to the corresponding column of IDATA1, IDATA2 and IDATA3 respectively. When all the columns of IDATA have been filled, output of the block begins. The column headings in IDATA1, IDATA2,and IDATA3 are written out first and then the frequency distribution data in IDATA. This cycle is repeated until all the synthetic frequency distribution data has been read and echo checked.

7.3.4 Data Initialisation

The Kolmogorov-Smirnov two-tail statistics at significance levels of 1%, 5% and 10% are assigned to vector RKOL at the start of the main program.

7.3.5 Library Subprograms

The Fortran IV Library subprograms required by this program are ABS, FLOAT, MAXØ, MINØ and SQRT.

7.4 INPUT DATA PREPARATION

7.4.1 Introduction

To use this program to compare the historical and synthetic data storage-yield behaviours it is necessary to:

(a) analyse the historical data set and determine the Rippl storages and drought durations, and the deficit, drought, draw and fill frequency distributions,

(b) analyse the synthetic data sets and determine the corresponding Rippl storages and drought durations, and the deficit, drought, draw and fill frequency distributions,

(c) prepare the input data deck defining program options, file reference numbers and problem data.

The results of these analyses and the input data deck comprise the program input which is described in Tables 7.1 and 7.2.

7.4.2 Historical Data Analysis

The historical data Rippl storages and drought durations, deficit, drought, draw and fill distributions are obtained from an analysis of the historical data set by program STATS, which will also punch these storage-yield characteristics and the analysis parameters onto cards in the correct format and order for input to this program.

These historical storage-yield analysis parameters and characteristics immediately follow the fifth card in the input data deck as shown in Table 7.2.

The station order of the historical storage-yield characteristics must be the same as for their synthetic counterparts. This may be ensured by maintaining the same station order in the historical and synthetic data sets analysed by STATS.

It will usually be necessary to perform the historical data storage-yield analysis a number of times to obtain deficit, drought, draw and fill frequency distributions, which do not have either a concentration of values in only a few bins or values beyond the upper bound. For the first analysis, the upper bounds of these distributions may be estimated from previous experience or guessed. The results of the analysis may be used to modify the upper bounds for the next analysis and so on until satisfactory frequency distributions are obtained. Two analyses have generally proven to be sufficient.

7.4.3 Synthetic Data Analysis

Each of the NSAM synthetic data sets of NYEAR years of concurrent flows at each of the NSTN stations is analysed by program STATS using the historical storage-yield analysis parameters to determine the synthetic data Rippl storages and drought durations, the deficit, drought, draw and fill frequency distributions.

Program STATS writes the synthetic data Rippl storages and drought durations to output file NF30, the synthetic data deficit frequency distributions to output file NF33, the synthetic data drought frequency distributions to output file NF34, the synthetic data draw frequency distributions to output file NF35 and the synthetic data fill frequency distributions to output file NF36. The corresponding input file variables in this program are NFILE(5), NFILE(1), NFILE(2), NFILE(3) and NFILE(4) respectively.

7.4.4 Input Data Deck

The input data file NIN specifies program options, file reference numbers and problem data. These cards may be prepared from Table 7.2 and are followed by the historical storage-yield analysis parameters and characteristics.

7.5 OUTPUT INTERPRETATION

A brief description of the program output is given as the printout is clearly labelled. Sample printout from the example problem described in section 9 is given in Figure 9.17.

The first output page is entitled "SYNTHETIC STORAGE-YIELD ANALYSIS" and provides an echo check of the program options, file reference numbers, problem data, historical Rippl storages and drought durations, historical totals, means, upper bounds and number of bins, for each of the deficit, drought, draw and fill frequency distributions.

The Rippl storage and drought duration analysis results for each yield level, are printed out for each station in turn, under the heading "RIPPL STORAGE ANALYSIS STATION XXX", where "XXX" is a user assigned station code number.

The results of the comparison of the historical deficit, drought, draw and fill probability distributions with their synthetic counterparts make up the remaining output. The format of the printed results for each of these distributions differs only in the page headings, which identify a particular distribution. For this reason, only the printout of the results for the deficit distribution will be described.

The historical deficit relative and cumulative probability distributions for each station are printed out at each yield level under the heading "HISTORICAL DEFICIT PROBABILITIES --- YIELD Z.ZZ", where "Z.ZZ" is the appropriate yield level.

The synthetic deficit frequency analysis data is echo checked for each station, each synthetic sample and each yield level, under the heading "DEFICIT FREQUENCY ANALYSIS", if requested.

The results of the comparison of each station's historical deficit probability distribution with its synthetic counterparts are printed out for each station in turn, under the heading "DEFICIT FREQUENCY ANALYSIS", with the subheading "STATION XXX", where "XXX" is the station code number. The results for each station are subdivided by yield level by the subheading "YIELD Z.ZZ", where "Z.ZZ" is the appropriate yield level. Each yield level's results are further subdivided into individual sample results and population results by the subheadings "INDIVIDUAL SAMPLES" and "POPULATION RESULTS" respectively. The results printed under these two subheadings are the numerical values of the measures of difference discussed in section 7.2.

Sample-station rejection summary tables are printed for each yield level, showing which stations had synthetic deficit distributions statistically rejected and for which samples this occurred, for the two user specified significance levels, under the heading "DEFICIT FREQUENCY ANALYSIS", with subheadings "YIELD Z.ZZ" and "SUMMARY AT YY P/C", where "Z.ZZ" is the appropriate yield level and "YY" the appropriate significance level.

Sample-station rejection summary tables are also printed for all yield levels combined, for the two user specified significance levels, under the heading "DEFICIT FREQUENCY ANALYSIS" with subheadings "ALL YIELD LEVELS" and "SUMMARY AT YY P/C", where "YY" is the appropriate significance level.

8. PROGRAM *CORREL* DESCRIPTION

8.1 INTRODUCTION

Auto (or serial) and cross (or spatial) correlation functions have been defined in section 3.2.9. Program CORREL is used to compare an historical correlation function with a number of synthetic correlation functions calculated from synthetic data samples. The comparison is made by evaluating a number of measures of the difference between the historical and the synthetic correlation functions. The values of these measures may be used to assess whether the historical correlations have been preserved satisfactorily in the synthetic data.

The measures of difference calculated for each lag of the correlation function include the range of synthetic correlation coefficients, the mean synthetic correlation coefficient, the number of synthetic correlation coefficients outside the statistically acceptable sampling interval, and the number of coefficients statistically different from zero. The results of these calculations are printed out for each correlation function.

Auto and cross correlation functions may be analysed. The historical data auto/cross correlation functions may be obtained as punched output from program STATS. The synthetic data auto/cross correlation functions may be written to specified disk/tape files by program STATS.

Punched output comparing the historical and mean synthetic auto/cross correlation functions may be obtained for input to the user's graphical presentation system.

8.2 FUNDAMENTAL CALCULATIONS

The fundamental purpose of this program is to compare the auto/cross correlation functions of the historical data with those of the synthetic data. The auto and cross correlation functions may be calculated to any number of lags less than or equal to 60 by program STATS, but each correlation function (i.e., auto or cross) must be calculated to the same number of lags for both the historical and the synthetic data. For the historical data sample and each synthetic data sample there are NSTN auto correlation functions and NSTN(NSTN-1)/2 cross correlation functions (section 3.3.6). There are NSAM synthetic data samples.

The calculations are performed within two loops. The inner loop is over each lag of the correlation function evaluating the measures of difference (described in 8.1) between the synthetic correlation function and the historical correlation function. The outer loop is over the number of synthetic samples (NSAM). Distinct vectors are assigned for each measure of difference to record the results of the calculations at each lag of the correlation function. These vectors are defined in the listing of subprogram CORREL.

The minimum and maximum synthetic correlation coefficients at each lag are stored in vectors SYNMIN and SYNMAX respectively.

The synthetic population correlation function cannot be calculated from the known synthetic sample correlation functions as is done for other statistical properties in the other programs. Instead, the synthetic sample correlation coefficients are summed in vector SYNAV for each lag over the NSAM synthetic samples. The mean synthetic correlation coefficient at each lag is then calculated as an approximation to the synthetic population correlation coefficients.

To help assess the significance of the difference at each lag between the historical and the mean synthetic correlation function, the actual difference and the relative percentage error are calculated and printed out. The difference between two correlation coefficients at any lag is defined as the mean synthetic correlation coefficient minus the historical correlation coefficient. The definition of the relative percentage error was made dependent on the size of the correlations concerned. If the absolute value of the historical correlation is greater than 0.05, the relative percentage error is 100 times the difference in correlations divided by the historical correlation. If the absolute values of both the synthetic and historical correlations are less than 0.05, the relative percentage error is made zero. Otherwise, the relative percentage error is defined as 100 times the difference in correlations divided by the mean of the synthetic and historical correlations. The actual difference and relative percentage error are not calculated till output time and thus require only temporary vector space provided by WS1 and WS2 respectively.

The optional punched output consists of the historical and the mean synthetic correlation functions with the minimum and maximum synthetic correlations at each lag.

The sampling distribution of a correlation function coefficient at any lag k (r(k)) is not normal. However, if it can be assumed that the two series from which the coefficient is calculated have a bivariate normal distribution, then the variable Z, defined as:

$$Z = 0.5 \, \mathrm{LOG}_e\{[1.0 + r(k)]/[1.0 - r(k)]\} \tag{8.1}$$

is approximately normally distributed with mean:

$$\bar{Z} = 0.5 \, \mathrm{LOG}_e\{[1.0 + r_h(k)]/[1.0 - r_h(k)]\} \tag{8.2}$$

where $r_h(k)$ is the population correlation, that is, the historical correlation, and with a variance of:

$$1/(N-3) \tag{8.3}$$

where N is the number of pairs of values used in the calculation of r(k) [*Goodman* 1966, pp. 193-195].

It can be shown from the sampling properties that the interval within which the Z transforms of the synthetic correlations (Z_i) should lie, if they are to be considered statistically indistinguishable from the historical value, is:

$$\bar{Z} - \frac{N_{\alpha/2}}{\sqrt{N-3}} < Z_i < \bar{Z} + \frac{N_{\alpha/2}}{\sqrt{N-3}} \tag{8.4}$$

where $N_{\alpha/2}$ is the normal statistic at the specified significance level. The number of synthetic correlation coefficients whose Z transform lies below the lower interval bound or above the upper interval bound, for the specified significance level, are counted to give a measure of the statistical acceptability of the synthetic correlation coefficients. Whenever the Z transform of a synthetic correlation coefficient lies outside of the interval, the synthetic sample in which this has occurred is also noted. When all synthetic samples have been processed, the program prints a summary table showing which synthetic samples had a correlation coefficient rejected at the specified significance level.

Correlation coefficients are also tested for significant difference from zero. If r(k) is the correlation coefficient at lag k calculated from N pairs of observations from two series with a bivariate normal distribution, then it can be shown that the statistic:

$$r(k) \,.\, \sqrt{\frac{(N-2)}{(1-r(k)^2)}} \tag{8.5}$$

is distributed as t with N-2 degrees of freedom [*Goodman* 1966, pp. 192-193]. For the large sample sizes of this application, the t distribution will approximate the normal distribution. A correlation coefficient is statistically different from zero if the absolute value of the above statistic is greater than $N_{\alpha/2}$, the normal statistic for the specified significance level. Historical and synthetic correlations are tested and the number of correlations statistically different from zero are counted and the results printed.

In this description of the statistical tests, a single significance level was used to simplify the discussion. The program, however, performs the above tests at two significance levels simultaneously. The significance levels are selected from the commonly used values of 1%, 5% and 10% by means of the input variables IPCEN1 and IPCEN2. The results of the tests at each specified significance level are printed out.

8.3 PROGRAMMING FEATURES

8.3.1 Introduction

The calculation of the measures of difference used in comparing an historical correlation function with NSAM synthetic correlation functions has been described. How these calculations are organised for many historical correlation functions is now described.

8.3.2 Number of Correlation Functions

There are NSTN auto correlation functions, each calculated to NLAG1 lags. As the auto correlation coefficient at lag 0 is always 1.0 by definition (section 3.2.9), this lag is not included in the analysis and therefore the results vectors for the above measures of difference are all NLAG1 elements long, corresponding to lags + 1 to + NLAG1. Negative lags are not calculated because of the symmetry of auto correlation functions about lag 0 (section 3.2.9).

There are NSTN(NSTN-1)/2 distinct cross correlation functions between pairs of the NSTN stations in the analysis (section 3.3.6), each calculated to NLAG2 lags. The results vectors for the above measures of difference are each 2NLAG2 + 1 elements long corresponding to lags -NLAG2 to +NLAG2.

8.3.3 Core Storage Problem

There are 15 results vectors required for <u>each</u> correlation function to save the results of the calculated measures of difference. The amount of core required to allocate fixed storage for 15 vectors for each correlation function is unacceptably high. For instance, for the 300 cross correlation functions of 25 stations, calculated to say 60 lags, approximately 550 thousand bytes would be needed.

As the synthetic correlation functions are analysed sequentially, it is only necessary to have the particular set of 15 results vectors corresponding to the correlation function being analysed, available in core. Auxiliary storage may be used to hold result vector sets not required until they are needed. When the analysis of the present correlation function is completed, the updated results vectors may be returned to auxiliary storage and the appropriate next set read in.

In the main program, the vector SPACE is dimensioned 2100 elements long and the vector OUTPUT is equivalenced to it. The lengths of the results vectors are calculated (section 8.3.2), and then the indices of vector elements in SPACE, to correspond to the first element of each of the 15 results vectors, on the assumption that the results vectors sequentially and contiguously occupy the elements of SPACE, are determined. These 15 contiguous subdivisions of SPACE are passed to CORREL as 15 results vectors. Their total contiguous length is calculated. OUTPUT is passed to CORREL and dimensioned with this length. OUTPUT in subprogram CORREL, is thus equivalenced to the set of 15 results vectors and all input/output of the results vector set is done by reference to OUTPUT for maximum efficiency.

Two sequential files, NWF1 and NWF2 are used. One file contains the most recent copy of all the results vector sets. Each results vector set is read into core from this file as required. After each results vector set is updated, it is written out to the second file. After all results vector sets have been processed from the first file, the second file will then contain the most recent copy of all the results vector sets, and thus becomes the input file for the next synthetic sample set. And so on for all synthetic sample sets.

Three additional subdivisions of SPACE are similarly made and passed as three vectors to CORREL. Subprogram CORREL sequentially reads each historical correlation function into vector HIST. For each function, the coefficients are tested for a significant difference from zero at two significance levels, and the two results vectors and HIST are written to sequential file NWF3. This file saves the results for later output.

These scratch files are described in Table 8.1. All references to these files are illustrated in the program flowchart (Figure 8.1).

8.3.4 Problem Size

The size of problem that may be analysed with this program is limited by the dimensions of the vector SPACE and the work vectors WS1, WS2, WS3, NWS1, NWS2, and NWS3 of subprogram CORREL.

The vector SPACE must be long enough to contain the 18 results vectors. This condition is fulfilled when the sum of 16 times the length of a correlation function, plus 2 times the number of synthetic samples plus one, is less than or equal to the assigned dimension. The present assigned dimension is 2100.

The work vector dimensions are all equal and should be greater than or equal to the length of the correlation functions to be analysed. The present assigned dimension is 125.

These dimensions may be increased if required for particular problems.

8.3.5 Sampling Interval Calculations

The synthetic correlation coefficients could be tested by calculating their Z transforms (as described in section 8.2) and testing that they were within the acceptable limits. However, as there is a unique correspondence between a correlation coefficient and its Z transform, an alternative and faster test procedure may be used. An inverse transform can be used to determine the correlation coefficient values corresponding to the Z transform limits and the synthetic correlation coefficient compared against these correlation coefficient limits to determine its statistical acceptability.

The inverse transformation equation is:

$$r(k) = \{e^{2Z} - 1\}/\{1 + e^{2Z}\} \qquad (8.6)$$

which is obtained by solving for r(k) in the Fisher Z transformation equation (equation 8.1). This equation is used with the Z transform sampling limits, calculated from the historical correlation coefficient Z transforms, to determine the acceptable range of synthetic correlation coefficients. These limiting values of correlation coefficients are calculated for each correlation function, and stored in the vectors BDLS, BDLL, BDUL, and BDUS which belong to the results vector set of each function.

8.3.6 Data Initialisation

The normal statistics corresponding to 1%, 5% and 10% significance levels are assigned to vector RNOR by a data statement in the main program.

8.3.7 Library Subprograms

The Fortran IV library subprograms required by this program are ABS, ALOG, EXP, FLOAT, MOD and SQRT.

8.4 INPUT DATA PREPARATION

8.4.1 Introduction

To use this program to compare historical correlation functions with their synthetic counterparts, it is necessary to:

(a) analyse the historical data set and determine the historical auto/cross correlation functions,

(b) analyse the synthetic data sets and determine the synthetic auto/cross correlation functions for each data set,

(c) prepare the input data deck defining program options and problem data.

The results of these analyses and the input data deck comprise the program input which is described in Tables 8.1 and 8.2.

8.4.2 Historical Data Analysis

The historical auto correlation function for each station and the historical cross correlation function for each station combination are obtained by an analysis of the historical data by program STATS, which punches these correlation functions for input to this program.

The historical correlation function data follows the program option and problem data cards in input file NIN. If both auto and cross correlation analyses have been specified, the auto correlation data physically precede the cross correlation data. If only one analysis has been specified, only the correlation data for that analysis would be provided.

The station order of the historical auto correlation functions, and the station combination order of the historical cross correlation functions, must be the same as for their synthetic counterparts. This can be ensured by maintaining the same station order in the historical and synthetic data sets analysed by STATS.

8.4.3 Synthetic Data Analysis

Each of the NSAM synthetic data sets of NYEAR years of concurrent flows at each of the NSTN stations must be analysed by program STATS to determine the synthetic data auto and cross correlation functions. The calculated synthetic auto correlation functions are written to output file NF31, and the calculated synthetic cross correlation functions to output file NF32. These output files of STATS serve as the synthetic correlation function data input files for this program.

8.4.4 Input Data Deck

The input data file NIN specifies the program options and problem data. These cards may be prepared from Table 8.2 and are followed by the historical correlation function cards.

8.5 OUTPUT INTERPRETATION

A brief description of the program output is given as the printout is clearly labelled. Sample printout from the example problem described in section 9 is given in Figure 9.19.

The first output page is entitled "SYNTHETIC AUTO AND CROSS CORRELATION ANALYSIS" and provides an echo check of the program options and problem data as read from the input data deck by the program.

If a listing of the synthetic auto/cross correlation function data has been requested, it is provided under the heading "SYNTHETIC CORRELATION ANALYSIS DATA".

The results of the comparison of an historical correlation function with its synthetic counterparts are printed under the heading "AUTO CORRELATION ANALYSIS RESULTS STATION XXX" for auto correlation functions, and under the heading "CROSS CORRELATION ANALYSIS STATION XXX & STATION ZZZ" for cross correlation functions, where "XXX" and "ZZZ" are the appropriate station code numbers. The subheadings on these pages describe each column of results.

After the results for each correlation function have been printed, a summary table showing which synthetic correlation functions have had at least one correlation coefficient rejected, and in which synthetic sample this occurred, is printed for the two specified significance efficients are calculated for each correlation function, and stored in the vectors BDLS, BDLL, BDUL, and BDUS which belong to the results vector set of each function.

8.3.6 Data Initialisation

The normal statistics corresponding to 1%, 5% and 10% significance levels are assigned to vector RNOR by a data statement in the main program.

8.3.7 Library Subprograms

The Fortran IV library subprograms required by this program are ABS, ALOG, EXP, FLOAT, MOD and SQRT.

8.4 INPUT DATA PREPARATION

8.4.1 Introduction

To use this program to compare historical correlation functions with their synthetic counterparts, it is necessary to:

(a) analyse the historical data set and determine the historical auto/cross correlation functions,

(b) analyse the synthetic data sets and determine the synthetic auto/cross correlation functions for each data set,

(c) prepare the input data deck defining program options and problem data.

The results of these analyses and the input data deck comprise the program input which is described in Tables 8.1 and 8.2.

8.4.2 Historical Data Analysis

The historical auto correlation function for each station and the historical cross correlation function for each station combination are obtained by an analysis of the historical data by program STATS, which punches these correlation functions for input to this program.

The historical correlation function data follows the program option and problem data cards in input file NIN. If both auto and cross correlation analyses have been specified, the auto correlation data physically precede the cross correlation data. If only one analysis has been specified, only the correlation data for that analysis would be provided.

The station order of the historical auto correlation functions, and the station combination order of the historical cross correlation functions, must be the same as for their synthetic counterparts. This can be ensured by maintaining the same station order in the historical and synthetic data sets analysed by STATS.

8.4.3 Synthetic Data Analysis

Each of the NSAM synthetic data sets of NYEAR years of concurrent flows at each of the NSTN stations must be analysed by program STATS to determine the synthetic data auto and cross correlation functions. The calculated synthetic auto correlation functions are written to output file NF31, and the calculated synthetic cross correlation functions to output file NF32. These output files of STATS serve as the synthetic correlation function data input files for this program.

8.4.4 Input Data Deck

The input data file NIN specifies the program options and problem data. These cards may be prepared from Table 8.2 and are followed by the historical correlation function cards.

8.5 OUTPUT INTERPRETATION

A brief description of the program output is given as the printout is clearly labelled. Sample printout from the example problem described in section 9 is given in Figure 9.19.

The first output page is entitled "SYNTHETIC AUTO AND CROSS CORRELATION ANALYSIS" and provides an echo check of the program options and problem data as read from the input data deck by the program.

If a listing of the synthetic auto/cross correlation function data has been requested, it is provided under the heading "SYNTHETIC CORRELATION ANALYSIS DATA".

The results of the comparison of an historical correlation function with its synthetic counterparts are printed under the heading "AUTO CORRELATION ANALYSIS RESULTS STATION XXX" for auto correlation functions, and under the heading "CROSS CORRELATION ANALYSIS STATION XXX & STATION ZZZ" for cross correlation functions, where "XXX" and "ZZZ" are the appropriate station code numbers. The subheadings on these pages describe each column of results.

After the results for each correlation function have been printed, a summary table showing which synthetic correlation functions have had at least one correlation coefficient rejected, and in which synthetic sample this occurred, is printed for the two specified significance levels. This summary page is titled "CORRELATION ANALYSIS RESULTS AT YY P/C", where "YY" is the appropriate significance level.

9. AN EXAMPLE APPLICATION OF THE PROGRAM SUITE

9.1 INTRODUCTION

A typical river valley problem is introduced to illustrate one application of the program suite and to demonstrate how the suite may be implemented on a computer system. Synthetic streamflow and rainfall data are generated for the example river valley system and the program suite is used to compare the statistical properties of the synthetic and historical hydrologic data. The success of the generation model in preserving the statistical properties of the historical data, and thus the suitability of the synthetic data for use in a river valley simulation model, may then be assessed.

The detailed description of this application of the program suite demonstrates how the suite may be run on a computer system, and provides a documented example, which may be used to test the program suite on particular computers.

Although this example application was performed on an IBM 370/158 computer system, only minor program modifications would be necessary to run the suite on a computer system that supports American National Standards Institute Fortran IV.

9.2 TYPICAL RIVER VALLEY SYSTEM PROBLEM

Simulation is commonly adopted as a decision making aid in water resource system studies because of the multitude of factors which need to be included in these studies. Historical records of streamflow and rainfall have commonly been used as inputs to such simulation models.

A synthetic generation technique may alternatively be used to generate many sequences of streamflow and rainfall records as inputs for the simulation model. Such synthetic sequences need to be examined to ensure that they preserve the statistical properties of the historical data.

To demonstrate how this program suite may be used to validate such synthetic data, assume that the hydrologic inputs to a river valley simulation model may be characterised by two streamflow series and one rainfall series.

9.3 PREPARATION OF HISTORICAL DATA

The program suite requires the same number of concurrent years of record for each hydrologic series. Typically, however, the historical records will not have been recorded over the same period for each series. All series must then be truncated to the shortest common period of record, or extended by one of several techniques to the longest period of record. For this example, 50 years of concurrent records, covering the period 1922-1971 are available for each of the three hydrologic series.

Each series is identified to the program suite by a user assigned three digit integer code number. These code numbers are used by the program suite to associate printed results and punched output with the appropriate series. These code numbers may be chosen randomly or systematically by the user. In this example the identifying code numbers were chosen systematically as 150, 180 and 330. The hundreds digit indicates the type of hydrologic series, 1 corresponding to a streamflow series and 3 to a rainfall series. The tens digit indicates the relative upstream position of the recording station, a smaller digit indicating a more upstream station than a larger digit.

The historical data is read in one series at a time under a user specified format, which enables the use of a single card format for punching the historical data, which will be acceptable to both the user's synthetic data generation program and this program suite. The historical data records of the three stations used in this example are shown in Figure 9.7. This historical data is read by program STATS under the specified format: (8X,12F6.0).

9.4 GENERATION OF SYNTHETIC DATA

Any synthetic data generation program may be used to generate synthetic records for the hydrologic input series of the simulation model. It will be necessary to specify the number of synthetic data sets (NSAM) and the number of concurrent years of record (NYRS) to be generated for each series. The number of concurrent years of record (NYRS) will usually reflect the economic life of the project being investigated, and the number of synthetic samples (NSAM) will usually be a function of the variance of the decision criterion of the simulation model.

The synthetic data generation program HEC-4 [*U.S. Army Corps of*

Engineers 1971] was used to generate 5 synthetic data sets, each consisting of 50 years of concurrent record at each station. (These values of NSAM and NYRS were arbitrarily selected.) The five generated synthetic data sets are shown in Figure 9.1. Each synthetic data set consists of 50 years of concurrent synthetic record for each station, one station's 50 years of record being followed by the next station's 50 years of record, and so on. Each line of a station's record corresponds to 1 synthetic year. The left-most number is the station identification code number. This number is followed by the synthetic year number and the twelve monthly synthetic streamflow/ rainfall values. This synthetic data is read by program STATS within a double DO loop structure. The inner DO loop is over the number of series, and the outer DO loop is over the number of synthetic sample data sets.

9.5 ALLOCATION OF PROGRAM SUITE FILES

Before proceeding with the statistical analysis of the historical and synthetic data sets, and the comparison of their statistical properties, it is necessary to choose the computer storage media and the file reference numbers for the files used by this program suite. The file number variables in the various suite programs may then be assigned and the necessary job control language files to run the suite on the computer system prepared.

The system, scratch and permanent files used by this program suite are summarised in Table 9.1. The files have been grouped according to their function and a brief description of each file is given. The individual programs which reference each file, and their program variable names for the file are also tabulated.

The scratch files are automatically allocated by the system. The synthetic data file, the synthetic data statistics files, and the files for the comparison of historical and synthetic data statistics are permanent files and are allocated by the user. In this example, the synthetic data file is allocated to a tape and the other files to a private mountable disk pack. This allocation is the most convenient as a number of the above permanent files will be open simultaneously during processing.

A file reference number is chosen next for each of the program suite files. A systematic method of choosing unique file numbers for each file will avoid errors. The file numbers adopted for this example are tabulated in Table 9.1.

These file numbers must now be assigned to the appropriate variables in each of the suite programs. The standard input/output units are initialised by assignment statements at the beginning of each program. The remaining file variables are initialised by data cards in the program's input data deck, or in a data statement in the program's mainline. Further details are provided in the "Input Data Preparation" and "Programming Features" sections of each program's description.

Partitioned data sets are used for the permanent program suite files and these may now be allocated on the user's private mountable disk pack. This allocation is achieved by running a dummy system program (IEFBR14) with the appropriate data definition statements for each permanent file. The job used to allocate these permanent files on a 2316 disk pack on the IBM 370/158 system is shown in Figure 9.2.

9.6 PREPARATION OF STANDARD SETS OF JOB CONTROL LANGUAGE

The implementation of this program suite on any computer system may be simplified by the preparation of a few standard sets of job control language which will be suitable for most problems and which may be used to execute more than one program. Each job control language set actually consists of two subsets of job control language when executing programs on the IBM 370/158 system. The first job control language subset is inserted before the Fortran program to be executed, and specifies the system resources required by the program, and the Fortran catalogued procedure to be invoked. The second job control language subset follows the program, and contains the data definition statements for the permanent and scratch files which may be referenced by the program. The program's input data deck follows next, completing the requirements for the execution of the program by the system.

Three standard job control language sets are needed for this example application, two of which are required for program STATS. The first set is used for the statistical analysis of the historical data and the second set for the statistical analysis of the synthetic data. The second job control language set is shown in Figures 9.3 and 9.4. The first job control language set may be obtained from this set as described below. The third job control language set is used by programs MOMENT, FREQ, RUNS, YIELD and CORREL in the comparison of historical and synthetic data statistics, and is shown in Figures 9.5 and 9.6.

Two different job control language sets are needed for program STATS because of the different system resources required in the analysis of historical and synthetic data. One disk and one tape drive are required for the analysis of synthetic data as shown in Figure 9.3, whereas, no disk or tape drives are required for the analysis of the historical data. Synthetic data analysis also requires data definition statements FT08F001 and FT20F001 through FT36F001 inclusive, as shown in Figure 9.4. These data definition statements are not needed for the analysis of historical data. The job control language set shown in Figures 9.3 and 9.4, when modified as described above, may be used for the analysis of historical data.

The Fortran compile, link and go catalogued procedure is used in each job control language set.

9.7 HISTORICAL DATA ANALYSIS

After the file reference numbers have been selected and assigned in each program and the standard job control language sets prepared, the statistical analysis of the historical and synthetic data sets and the comparison of their statistical properties may be considered.

The user must decide which statistical properties are to be used in the comparison of the synthetic and historical data. The comparison of the values of these statistical properties for the historical and synthetic data sets determines the success of the generation model in preserving these statistical characteristics in the synthetic data, and thus the suitability of the synthetic data for use in the system simulation model.

For this example application, choose (1) the monthly, annual and time series first three moments, (2) the time series frequency distribution, (3) the run length frequency distributions, (4) the Rippl storage-yield behaviour and (5) the monthly auto and cross correlations, as the statistical properties to be compared for the historical and synthetic data.

The calculated values of these five statistical properties for the historical and the synthetic data sets may be compared by programs MOMENT, FREQ, RUNS, YIELD and CORREL respectively. Program STATS can also calculate some additional statistical properties of hydrologic data. These properties will be included in this historical data analysis to demonstrate these options.

The calculated historical data values of these five statistical properties will be punched onto cards for subsequent input to the appropriate comparison programs as part of their input data decks.

Appropriate frequency analysis upper bounds for the time series frequency analysis and the storage-yield behaviour frequency analyses are specified from previous experience with these three particular hydrologic series.

The historical analysis input data deck was prepared from Table 3.3 taking these considerations into account and is shown in Figure 9.7.

The historical data analysis may now be performed using program STATS, the historical data analysis input data deck and job control language set 1. Some example printed output from this analysis is shown in Figure 9.8.

The punched output from this analysis is split into the selected statistical property groups for subsequent use in the input data decks of the statistic comparison programs (Figures 9.10, 9.12, 9.14, 9.16 and 9.18).

9.8 SYNTHETIC DATA ANALYSIS

Five synthetic data sets, each consisting of fifty years of concurrent record for each of the three hydrologic series, have been generated. Each of these synthetic data sets will now be statistically analysed. The same statistical properties as were specified for the analysis of the historical data set are specified for the analysis of each synthetic data set. The calculated values of each of these statistical properties for each hydrologic series and for each synthetic data set are written to permanent disk files. These files are later read by the statistic comparison programs, which compare the synthetic data statistic values with their historical data counterparts, which have been read in the programs' input data decks.

The same statistical property analysis options are used in the synthetic data analysis as were used in the historical data analysis. For example, if the maximum run length specified in the historical data run length frequency distribution analysis was thirty, then this limit is also specified for the synthetic data analysis. This convention must be adhered to as the comparison programs expect the same analysis options to have been used in each analysis.

The same statistical property parameters must be used in the synthetic data analysis as were used in the historical data analysis if the calculated statistic values are to be validly compared. This requirement applies to the time series frequency distribution analysis, the run length frequency distribution analysis and the Rippl storage-yield analysis. The historical data time series and Rippl storage-yield frequency distribution analyses were performed with specified upper bounds for their distributions. These same upper bounds must be specified for the corresponding synthetic data analyses.

The historical data run length frequency distributions for each series were determined about each series median as no other value was specified. The synthetic data run length frequency distributions must therefore be determined about each series historical data median, which may be obtained from the historical data analysis printout.

The historical data Rippl storage deficits of each series were scaled by the series time series mean as no other scaling value was specified. The synthetic data Rippl storage deficits must likewise be scaled by the appropriate historical data time series mean, which may be read from the historical data analysis printout.

The synthetic data analysis input data deck was prepared from Table 3.3 to meet these requirements, and is shown in Figure 9.9. The synthetic data analysis may now be performed using program STATS, this input data deck and job control language set 2 (Figures 9.3 and 9.4).

9.9 COMPARISON OF HISTORICAL AND SYNTHETIC DATA MOMENTS

The monthly, annual and time series first three moments and the monthly extremes of the historical and the synthetic data sets have now been calculated. The historical data moments and extremes have been punched onto cards. The synthetic data moments and extremes have been written to disk files. These statistics of the historical and synthetic data may now be compared by program MOMENT.

The input data deck for program MOMENT is prepared from Table 4.2. The historical data moment and extreme cards are separated from the historical analysis punched output and are placed behind the other input data cards to complete the input data deck, which is shown in

Figure 9.10.

The comparison of the historical and synthetic data moments and extremes may now be performed using program MOMENT, the prepared input data deck and job control language set 3 (Figures 9.5 and 9.6). Some sample pages of output from this comparison are shown in Figure 9.11.

9.10 COMPARISON OF HISTORICAL AND SYNTHETIC DATA TIME SERIES FREQUENCY DISTRIBUTIONS

The time series frequency distributions of the historical and the synthetic data have been calculated. The historical data frequency distributions have been punched onto cards and the synthetic data frequency distributions have been written to disk files. As the historical and synthetic data frequency analyses were performed with the same distribution upper bounds, the frequency distributions may be validly compared by program FREQ.

The input data deck for program FREQ is prepared from Table 5.2. The historical data frequency distributions are separated from the historical analysis punched output and are placed behind the other input data cards to complete the input data deck, which is shown in Figure 9.12.

The comparison of the historical and the synthetic data frequency distributions is performed using program FREQ, the prepared input data deck and job control language set 3 (Figures 9.5 and 9.6). Some sample pages of output from this analysis are shown in Figure 9.13.

9.11 COMPARISON OF HISTORICAL AND SYNTHETIC DATA RUN LENGTH FREQUENCY DISTRIBUTIONS

The run length frequency distributions of the historical and the synthetic data have been calculated. The historical data run length frequency distributions have been punched onto cards and the synthetic data distributions have been written to disk files. As the historical and synthetic data run length frequency distribution analyses were both performed about the historical data medians, these distributions may be validly compared using program RUNS.

The input data deck for program RUNS is prepared from Table 6.2. The historical data run length frequency distributions are separated from the historical analysis punched output and are placed behind the other input data cards to complete the input data deck, which is shown in Figure 9.14.

The comparison of the historical and the synthetic data run length frequency distributions is performed using program RUNS, the prepared input data deck and job control language set 3 (Figures 9.5 and 9.6). Some sample pages of output from this analysis are shown in Figure 9.15.

9.12 COMPARISON OF HISTORICAL AND SYNTHETIC DATA STORAGE-YIELD BEHAVIOURS

The Rippl storage-yield statistics of the historical and synthetic data sets have been calculated. The historical data Rippl storage-yield statistics have been punched onto cards. The synthetic data Rippl storage-yield statistics have been written to disk files. As the Rippl storage-yield analysis of both the historical and the synthetic data sets was based on the historical data time series means, the historical and synthetic data Rippl storage-yield statistics may be validly compared by program YIELD.

The input data deck for program YIELD is prepared from Table 7.2. The historical data Rippl storage-yield statistics are separated from the historical analysis punched output and placed behind the other input data cards to complete the input data deck, which is shown in Figure 9.16.

The comparison of the historical and the synthetic data Rippl storage-yield statistics is performed using program YIELD, the prepared input data deck and job control language set 3 (Figures 9.5 and 9.6). Some sample pages of output from this analysis are shown in Figure 9.17.

9.13 COMPARISON OF HISTORICAL AND SYNTHETIC DATA CORRELATIONS

The auto and cross correlation functions of the historical and synthetic data sets have been calculated. The historical data auto and cross correlation functions have been punched onto cards and the synthetic data auto and cross correlation functions have been written to disk files. The correlations of the historical and synthetic data sets may now be compared by program CORREL.

The input data deck for program CORREL is prepared from Table 8.2. The historical data auto and cross correlation function cards are separated from the historical analysis punched output and are placed behind the other input data cards to complete the input data deck, which is shown in Figure 9.18.

The comparison of the historical and synthetic data auto and cross correlation functions is now performed using program CORREL, the prepared input data deck and job control language set 3 (Figures 9.5 and 9.6). Some sample pages of output from this analysis are shown in Figure 9.19.

9.14 VALIDATION OF THE SYNTHETIC DATA

The synthetic data values of the selected statistical properties have been calculated and compared with their historical data counterparts. Measures of the difference between the historical and synthetic data values of each statistical property have been calculated and printed. The user must now assess these differences and determine if they are significant.

Satisfactory agreement between the historical and synthetic data values of the selected statistical properties confirms that the generation model was successful in preserving these statistical properties in the synthetic data. The synthetic data may thus be confidently used in the water resource system simulation model.

Alternatively, poor agreement between the statistical property values means that these characteristics of the historical data were not being preserved in the synthetic data. The generation was unsuccessful and the synthetic data should not be used in the system simulation model.

10. REFERENCES

Association for Computer Aided Design, 1974. Recommended Standards for Documentation and Checking of Computer Aided Engineering Computations. Policy Document 74/1. August 1974.

Association for Computer Aided Design, 1977. Recommended Minimum Program Documentation and Details of Abstracts, Status Reports. Technical Note DO/1. May 1977.

Biswas, A.K. (Ed.), 1976. Systems Approach to Water Management. New York, McGraw Hill.

Blainey, J.M., 1970. Description of Computer Program for River Murray System Behaviour Studies. Report to the River Murray Commission Technical Committee. Canberra.

Ciriani, T.A., Maione, U. and Wallis, J.R. (Eds.), 1977. Mathematical Models for Surface Water Hydrology. New York, John Wiley.

Dixon, W.J. (Ed.), 1973. BMD: Biomedical Computer Programs. (3rd Edition), Berkeley, California University Press.

Fiering, M.B. and Jackson, B.B., 1971. Synthetic Streamflows. Water Resources Monograph 1. Washington, American Geophysical Union.

Finzini, G., Todini, E. and Wallis, J.R., 1977. SPUMA: Simulation Package Using Matalas Algorithm. In (Ciriani et al. (Eds.), 1977), pp. 249-276.

Goodman, R., 1966. Statistics. London, English Universities Press Ltd.

Hufschmidt, M.M. and Fiering, M.B., 1966. Simulation Techniques for Design of Water Resource Systems. Cambridge, Harvard University Press.

International Business Machines, 1970. System/360 Scientific Subroutine Package, Version III, Programmer's Manual. Document GH20-0205-4. August 1970.

Jackson, B.B., 1975. The Use of Streamflow Models in Planning. Water Resources Research, Vol. 11, No. 1, February 1975, pp. 54-63.

Lindner, M.A., 1978. Water Resources Optimisation in a Complex River Valley. Unpublished doctoral dissertation, The University of New South Wales. Manuscript in process.

Linsley, R.K., Kohler, M.A. and Paulhus, J.L., 1975. Hydrology for Engineers. New York, McGraw Hill.

Matalas, N.C., 1975. Developments in Stochastic Hydrology. Journal of Geophysics and Space Physics, Vol. 13, No. 3, July 1975, pp. 67-73.

Matalas, N.C. and Wallis, J.R., 1976. Generation of Synthetic Flow Sequences. In (Biswas, A.K. (Ed.), 1976), pp. 54-79.

Miller, A.J., 1974. Accuracy of Regression Packages. The Statistical Society of Australia, New South Wales Branch, Symposium on Statistical Computing, Sydney, August 29-20, 1974.

Perrens, S.J. and Howell, D.T., 1972. Effects of Serially Correlated Inflows on Reservoir Behaviour. Water Resources Research, Vol. 8, No. 3, June 1972, pp. 642-647.

Reti, G.A., 1973. Engineering Computer Program Documentation Standards. Proc. ASCE, Journal of the Soil Mechanics and Foundation Division, Vol. 99, No. SM3, March 1973, pp. 249-266.

Siegel, S., 1956. Nonparametric Statistics for the Behavioural Sciences. New York, McGraw Hill.

Texas Water Development Board, 1974. Analytic Techniques for Planning Complex Water Resource Systems. Report 183, April 1974.

Thompson, C.M., 1941. Tables of Percentage Points of the Incomplete Beta Function. Biometrika, Vol. 32, 1941, p. 189.

U.S. Army Corps of Engineers, Hydrologic Engineering Center, 1971. HEC-4, Monthly Streamflow Simulation. Computer program 723-X6-L2340, February 1971.

Yevjevich, V., 1972a. Probability and Statistics in Hydrology. Fort Collins, Water Resources Publications.

Yevjevich, V., 1972b. Stochastic Processes in Hydrology. Fort Collins, Water Resources Publications.

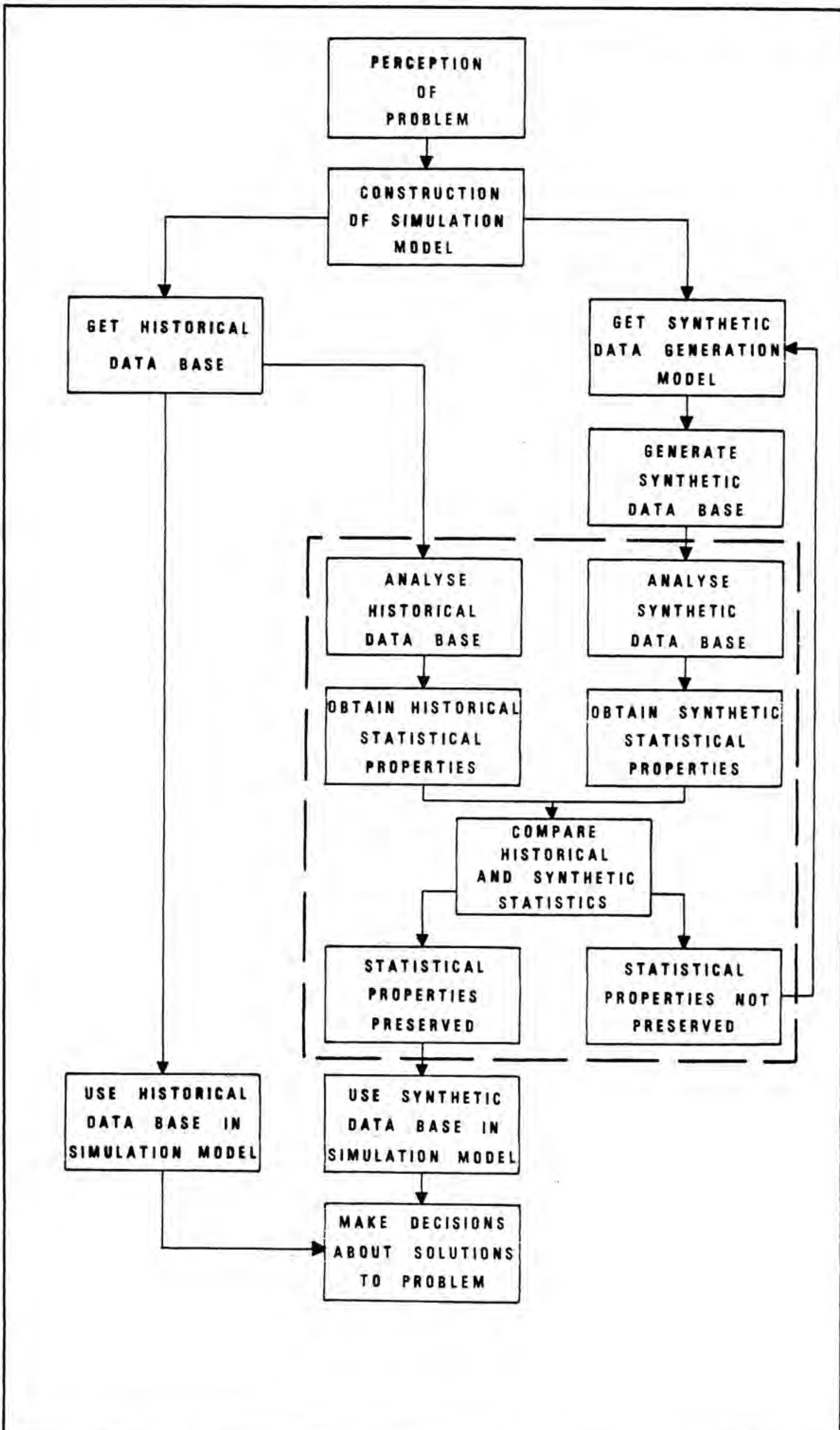


FIGURE 1.1: THE ANALYSIS AND COMPARISON OF HYDROLOGIC RECORDS IN WATER RESOURCES SYSTEM STUDIES

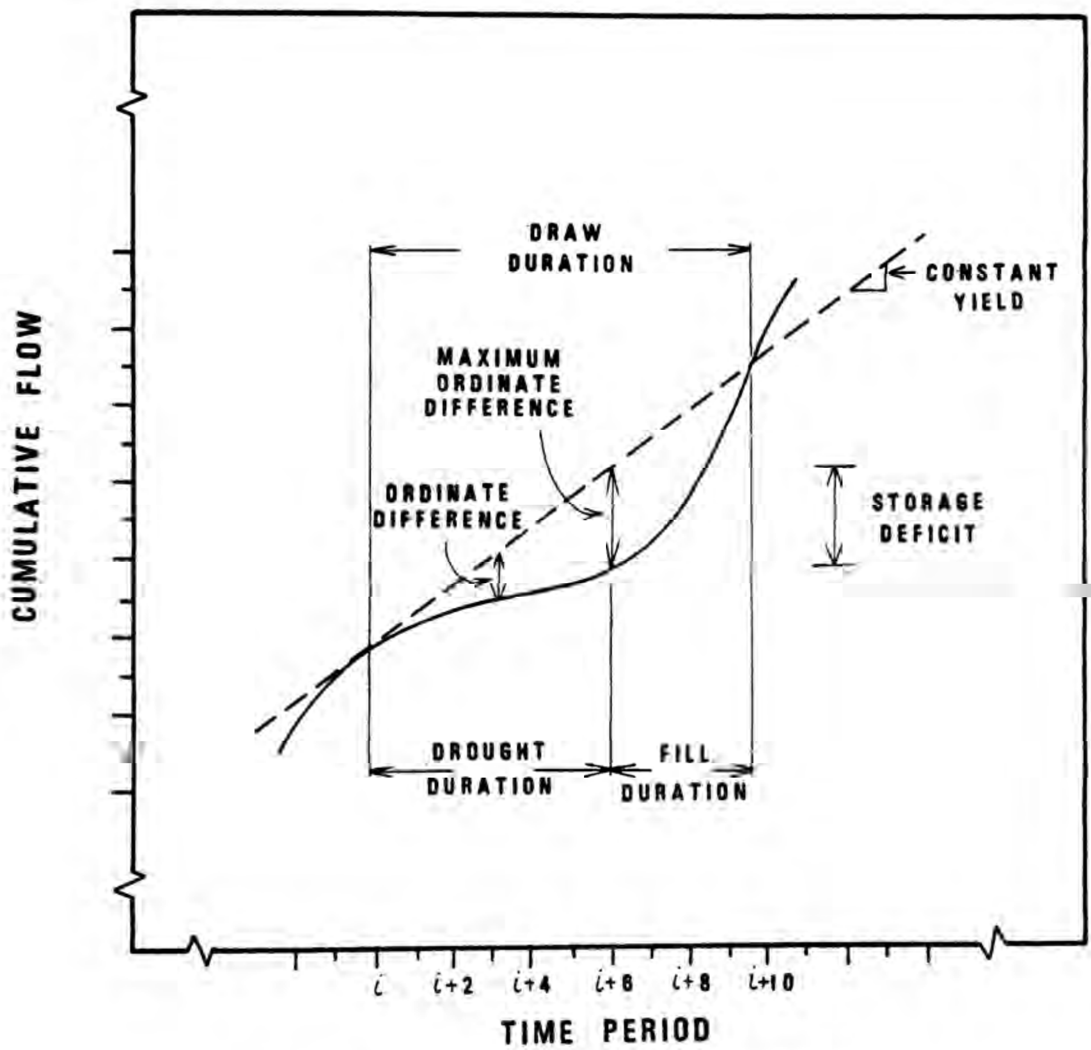
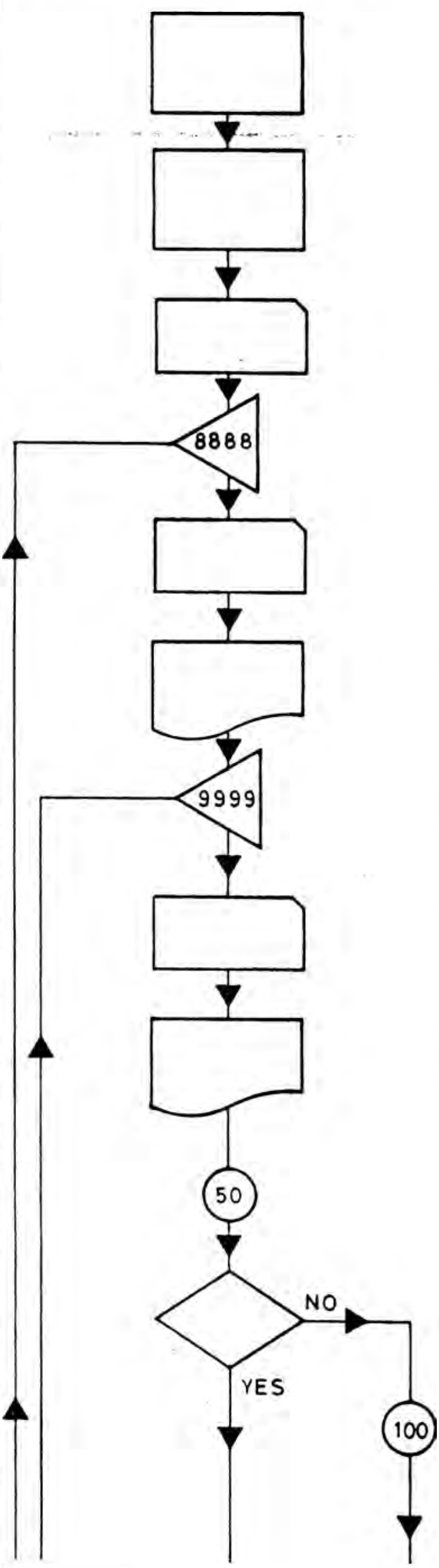


FIGURE 3.1: HYPOTHETICAL MASS CURVE

PROGRAM: STATS	SUBPROGRAM: MAIN



Initialise:
(a) monthly water year vector,
(b) scratch and synthetic data statistic file unit numbers,
(c) upper bounds for drought, draw and fill duration distributions (BDCOM) and storage deficit distributions (BDDEF), and appropriate numbers of bins.

Assign installation unit numbers for printer, card reader/punch, and default unit number for hydrologic data file.

Read the number of sets of hydrologic data to be analysed and whether problem data and problem options are the same for each set.

Loop over the number of sets of hydrologic data to be analysed.

Read problem data and program options. (This step bypassed in subsequent iterations if problem data and program options are the same for each set).

Echo check problem data, program options, scratch and synthetic data statistic file unit numbers.

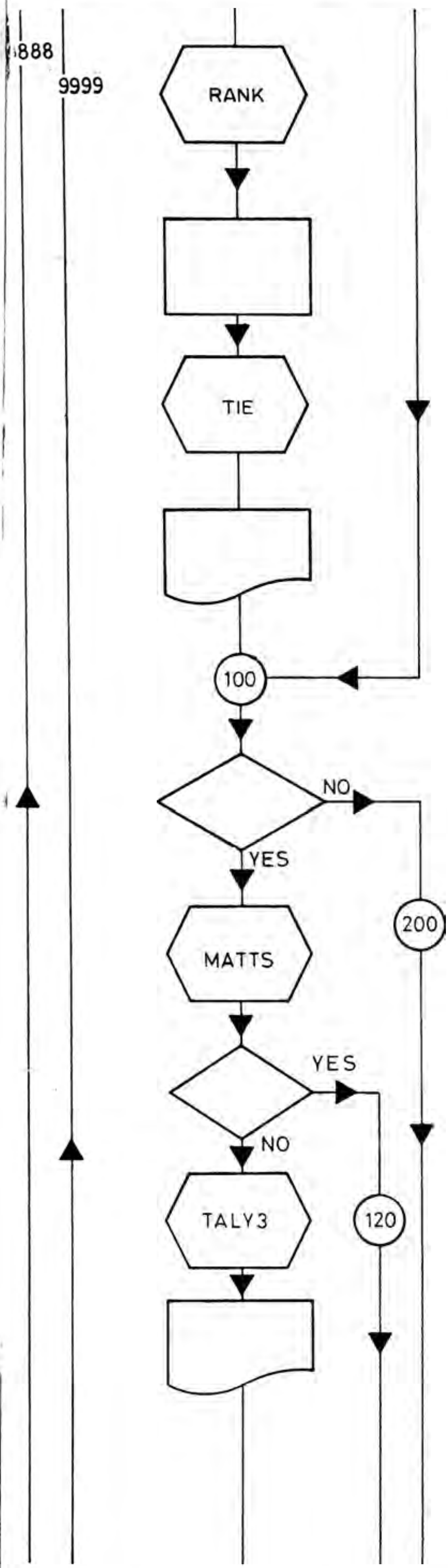
Loop over the number of stations in this set (NSTN).

Read this station's data from the user specified file, according to the user specified format.

Echo check this station's data if required.

Is the Kendall tau test for trend required?

PROGRAM: STATS	SUBPROGRAM: MAIN



Calculate the rank of each time series value.

Calculate the number of pairs of ranks which are in natural order.

Calculate the correction factor for tied ranks.

Calculate, standardise and print Kendall's tau.

Are monthly moments and extrema to be calculated?

Take the monthly time series values, currently stored by year in a vector, and store by month in another vector for analysis by TALY3 or TAB4.

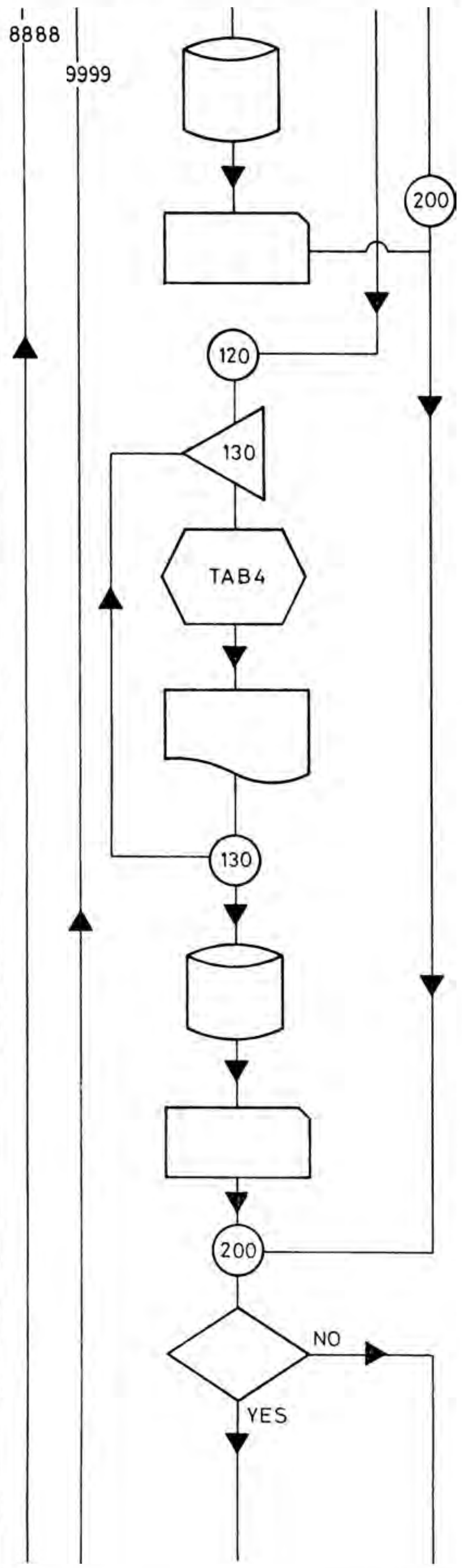
Are monthly frequency distributions to be calculated and plotted as line printer histograms?

Calculate monthly moments and extrema.

Print monthly moments and extrema.

FIGURE 3.2 (CONT'D.): PROGRAM STATS FLOWCHART

PROGRAM: STATS	SUBPROGRAM: MAIN



If specified, write monthly means, standard deviations and skew coefficients to files NF20, NF21 and NF22 respectively; and monthly minima and maxima to file NF23.

If the monthly moments are to be punched, temporarily write them to file NPU1. If the monthly extrema are to be punched, temporarily write them to file NPU1.

Loop over each month.

Calculate the moments, extrema and frequency distribution for this month.

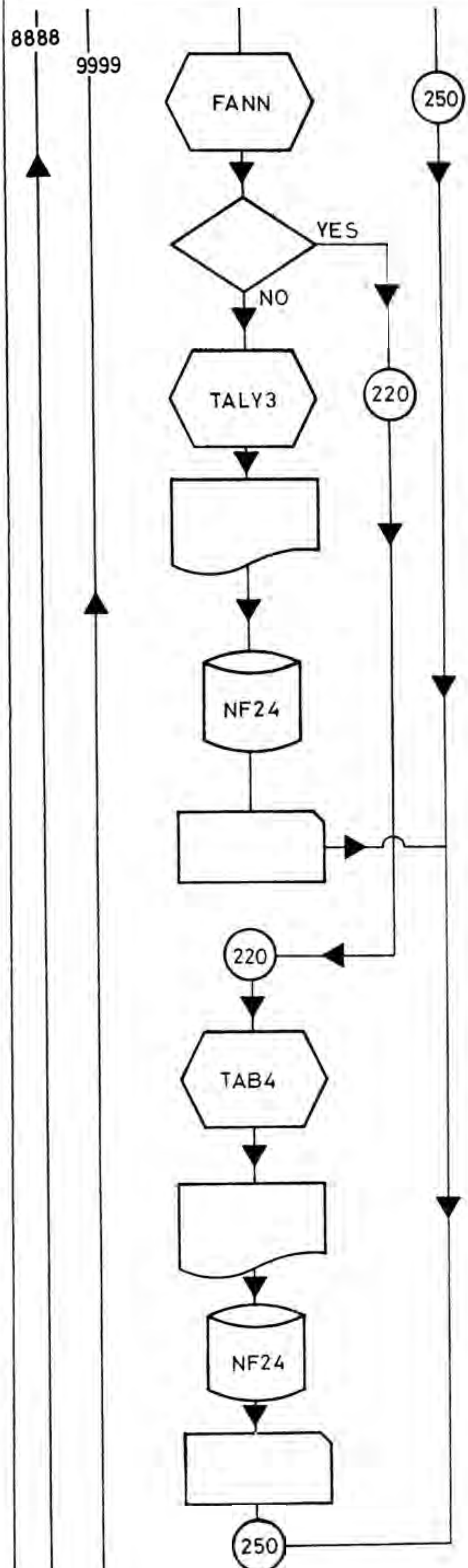
Print the moments, extrema and frequency distribution for this month, and plot the frequency distribution as a line printer histogram.

If specified, write monthly means, standard deviations and skew coefficients to files NF20, NF21 and NF22 respectively; and monthly minima and maxima to file NF23.

If the monthly moments are to be punched, temporarily write them to file NPU1. If the monthly extrema are to be punched, temporarily write them to file NPU1.

Are annual moments and extrema to be calculated?

PROGRAM: STATS	SUBPROGRAM: MAIN



Take the monthly time series values, currently stored by year in a vector, and sum by year to form an annual series in another vector, for analysis by TALY3 or TAB4.

Is the annual frequency distribution to be calculated and plotted as a line printer histogram?

Calculate the annual moments and extrema.

Print the annual moments and extrema.

If specified, write the annual mean, standard deviation and skew coefficient to file NF24.

If the annual moments are to be punched, temporarily write them to file NPU2.

Calculate the annual moments, extrema and frequency distribution.

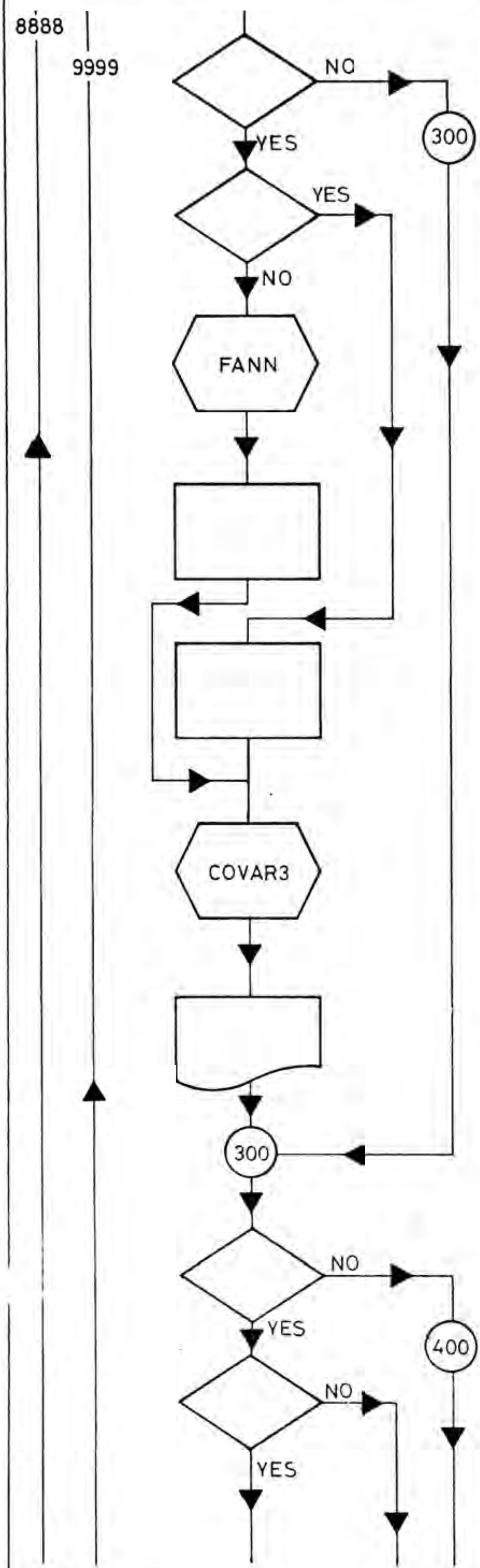
Print the annual moments, extrema and frequency distribution, and plot the frequency distribution as a line printer histogram.

If specified, write the annual mean, standard deviation and skew coefficient to file NF24.

If the annual moments are to be punched, temporarily write them to file NPU2.

FIGURE 3.2 (CONT'D): PROGRAM STATS FLOWCHART

PROGRAM: STATS	SUBPROGRAM: MAIN



Is the annual auto correlation function to be calculated?

Have the annual moments and extrema been calculated?

Form the annual time series from the monthly time series.

Calculate the annual time series mean and variance, and subtract the mean from each value of the annual time series.

Subtract the previously calculated annual mean from each value of the annual time series.

Calculate the annual auto correlation function to the specified number of lags.

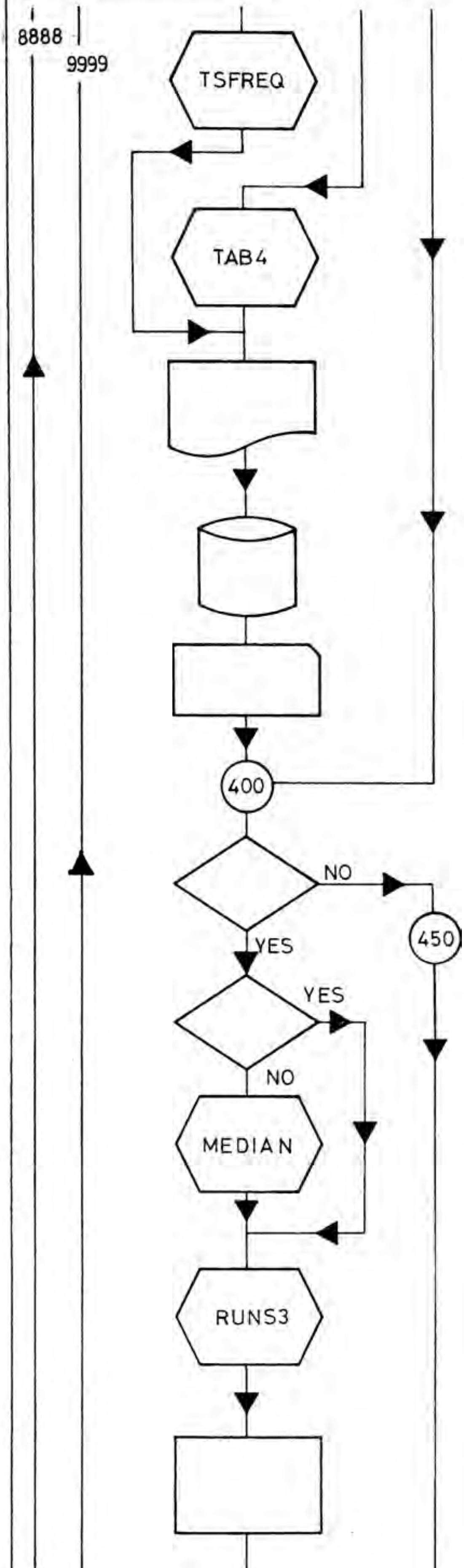
Print the annual auto correlation and covariance functions, and if required, a line printer plot of the annual auto correlation function.

Are the time series moments and frequency distribution to be calculated?

Are user specified bounds to be used in calculating the frequency distribution?

FIGURE 3.2 (CONT'D): PROGRAM STATS FLOWCHART

PROGRAM: STATS	SUBPROGRAM: MAIN



Calculate the time series moments and frequency distribution using the user specified bounds.

Calculate the time series moments and frequency distribution using program selected bounds.

Print the calculated time series moments and frequency distribution, and if required, a line printer histogram plot of the frequency distribution.

If specified, write the time series moments to file NF25. If specified, write the time series frequency distribution to file NF26.

If the time series moments are to be punched, temporarily write them to file NPU3. If the time series frequency distribution is to be punched, temporarily write it to file NPU4.

Are run length frequency distributions to be calculated?

Are the run length frequency distributions to be calculated above and below a user specified value?

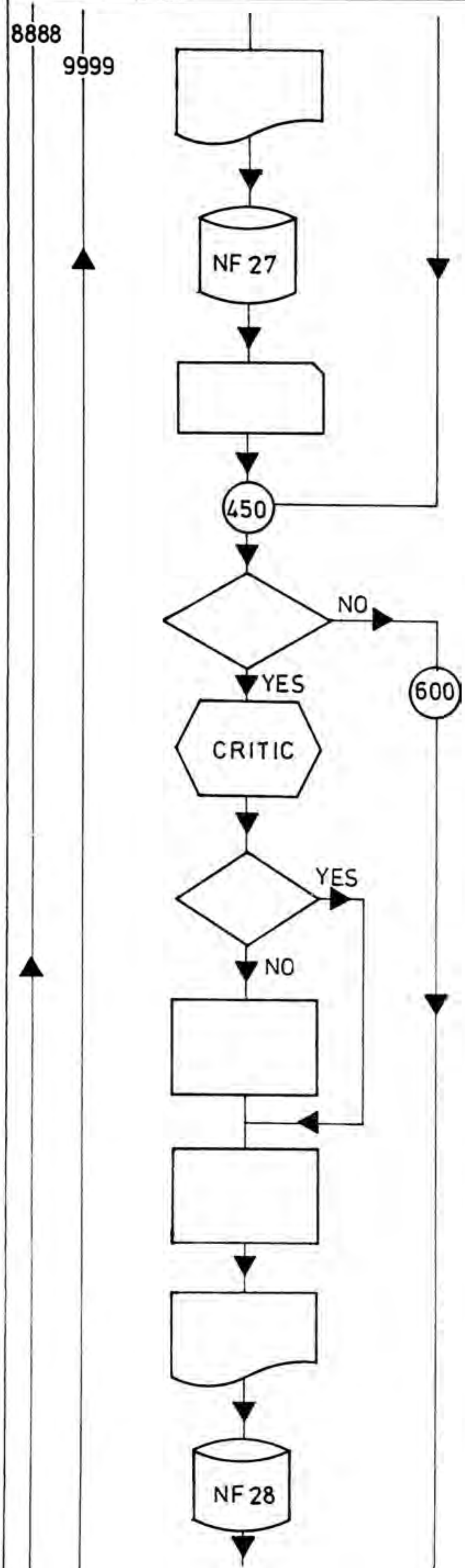
Calculate the monthly time series median.

Calculate the run length frequency distributions above and below the user specified value or the time series median as appropriate.

Calculate the run length frequency distribution about the user specified value or time series median; and the total number and average run length above, below and about the specified value or time series median.

FIGURE 3.2 (CONT'D): PROGRAM STATS FLOWCHART

PROGRAM: STATS	SUBPROGRAM: MAIN



Print the run length frequency distributions, and the total number and average run lengths above, below and about the specified value or median; and if required, line printer histogram plots of run length frequency distributions above and below the user specified value or time series median.

If specified, write the run length frequency distributions to file NF27.

If run length frequency distributions are to be punched, temporarily write them to file NPU9.

Are aggregate minima and maxima to be calculated?

Calculate aggregate minima and maxima from one time period to the specified number of time periods.

Are the aggregate minima and maxima to be scaled by a user specified value?

Calculate the time series mean if it has not been previously calculated.

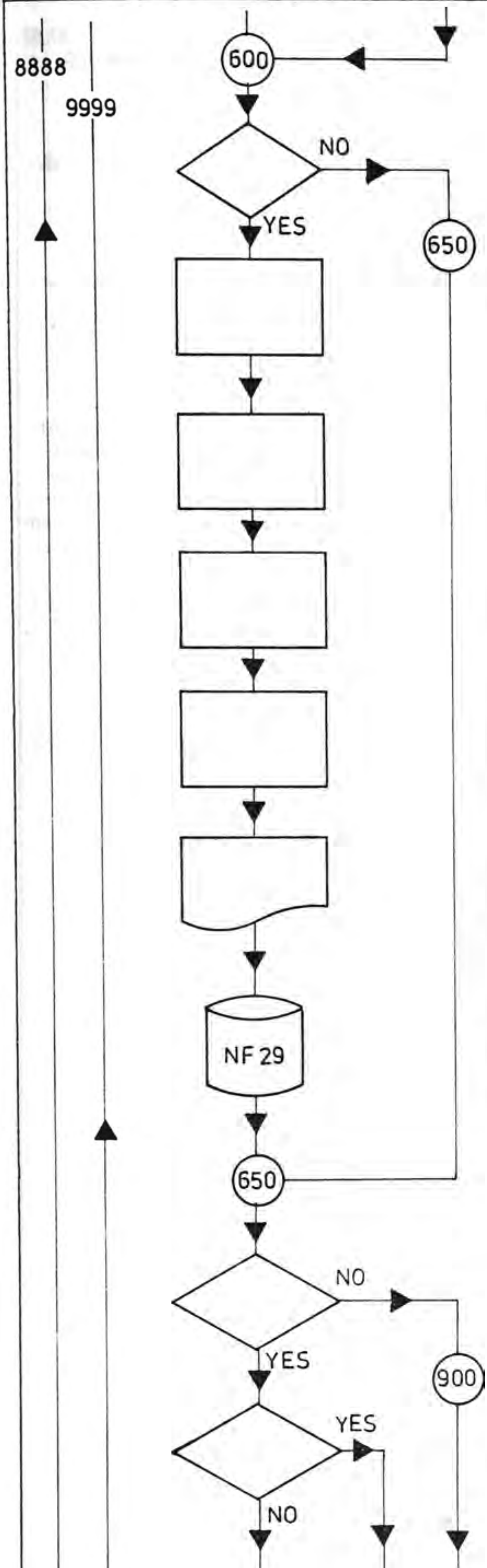
Scale the aggregate minima and maxima values by the user specified value or time series mean as appropriate.

Print the unscaled and scaled aggregate minima and maxima values, and if required, line printer plots of the aggregate minima and maxima values.

If specified, write the aggregate minima and maxima values to file NF28.

FIGURE 3.2 (CONT'D): PROGRAM STATS FLOWCHART

PROGRAM: STATS	SUBPROGRAM: MAIN



Is range analysis required?

Calculate the time series mean if it has not been previously calculated.

Calculate the maximum positive cumulative deviation from the mean (surplus), and the minimum negative cumulative deviation from the mean (deficit), and thus the range.

Calculate the theoretical range.

Scale the surplus, deficit, range and theoretical range by the user specified value or time series mean as appropriate.

Print the unscaled and scaled surplus, deficit, range and theoretical range.

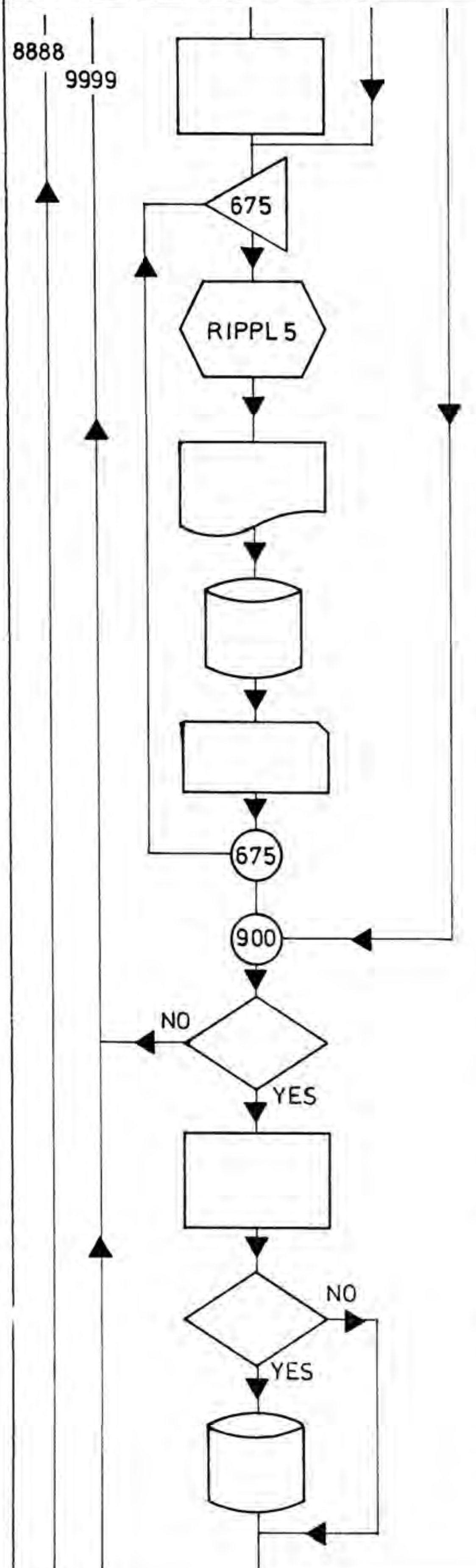
If specified, write the surplus and deficit values to file NF29.

Is storage-yield analysis required?

Are storage deficits to be scaled by a user specified value?

FIGURE 3.2 (CONT'D): PROGRAM STATS FLOWCHART

PROGRAM: STATS	SUBPROGRAM: MAIN



Calculate the time series mean if it has not been previously calculated.

Loop over the number of yield levels (NYIELD).

Calculate for this yield level the Rippl storage and drought duration, the drought, draw and fill duration frequency distributions and statistics, and the storage deficit distribution and statistics.

Print for this yield level, the Rippl storage and drought duration, the drought, draw and fill duration frequency distributions and statistics, the storage deficit distribution and statistics, and if required, line printer histogram plots of the drought, draw and fill duration frequency distributions, and the storage deficit frequency distribution.

If specified, write the Rippl storage and drought duration to file NF30, the storage deficit distribution to file NF33, and the drought draw and fill duration distributions to files NF34, NF35 and NF36 respectively.

If storage-yield analysis results are to be punched, temporarily write the Rippl storage and drought duration, the frequency distribution bounds, number of bins and statistics to file NPU10, the storage deficit distribution to file NPU11, and the drought, draw and fill duration distributions to files NPU12, NPU13 and NPU14 respectively.

Has auto correlation or cross correlation analysis been specified?

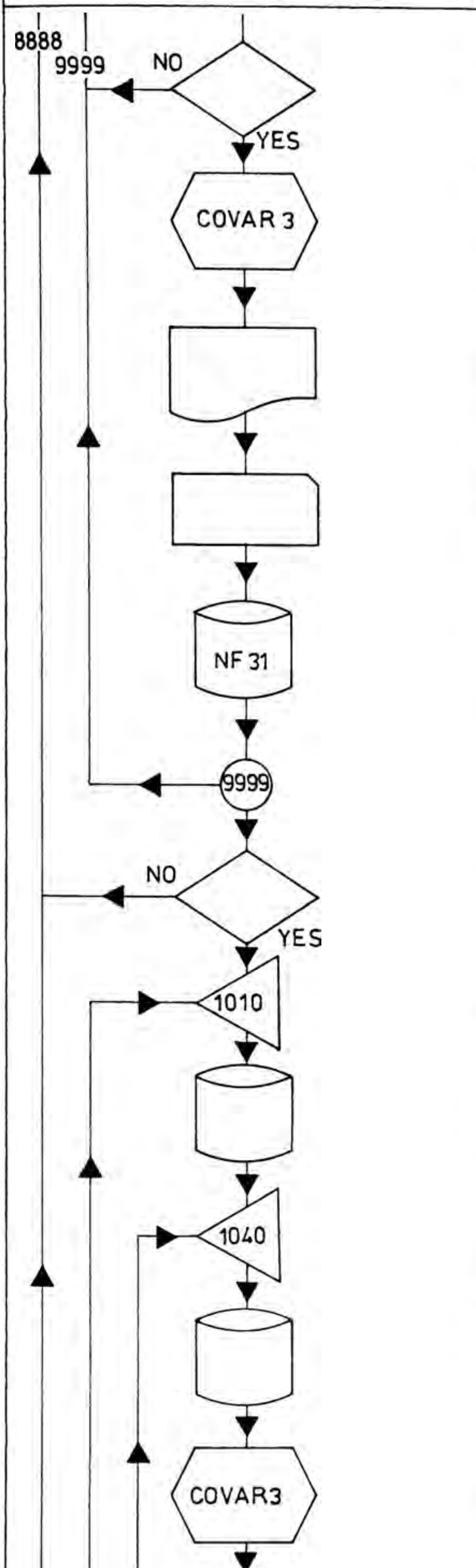
Calculate the time series mean and/or variance if they have not been previously calculated, and subtract the mean from each time series value.

Has cross correlation analysis been specified?

Write the modified time series for this station to the appropriate disk scratch file for later use in the cross correlation analysis.

FIGURE 3.2 (CONT'D): PROGRAM STATS FLOWCHART

PROGRAM: STATS	SUBPROGRAM: MAIN



Has auto correlation analysis been specified?

Calculate the auto correlation function to the specified number of lags.

Print the auto correlation and covariance functions, and if required, a line printer plot of the auto correlation function.

If the auto correlation function is to be punched, temporarily write it to file NPU15.

If specified, write the auto correlation function to file NF31.

Has cross correlation analysis been specified?

Loop from the first station to the penultimate station. Identify this station as station I.

Read in station I's modified time series from the appropriate disk scratch file.

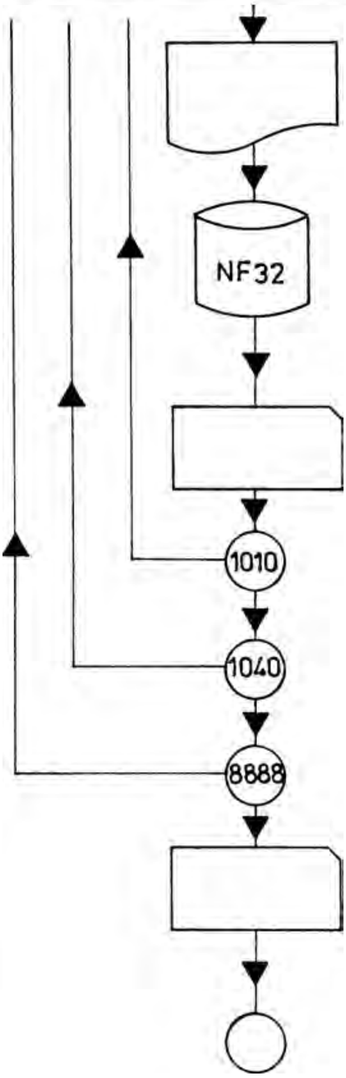
Loop from station I+1 to the last station. Identify this station as station J.

Read in station J's modified time series from the appropriate disk scratch file.

Calculate the cross correlation function of stations I and J to the specified number of lags.

FIGURE 3.2 (CONT'D): PROGRAM STATS FLOWCHART

PROGRAM: STATS	SUBPROGRAM: MAIN



Print the cross correlation and covariance functions, and if required, a line printer plot of the cross correlation function.

If specified, write the cross correlation function to file NF32.

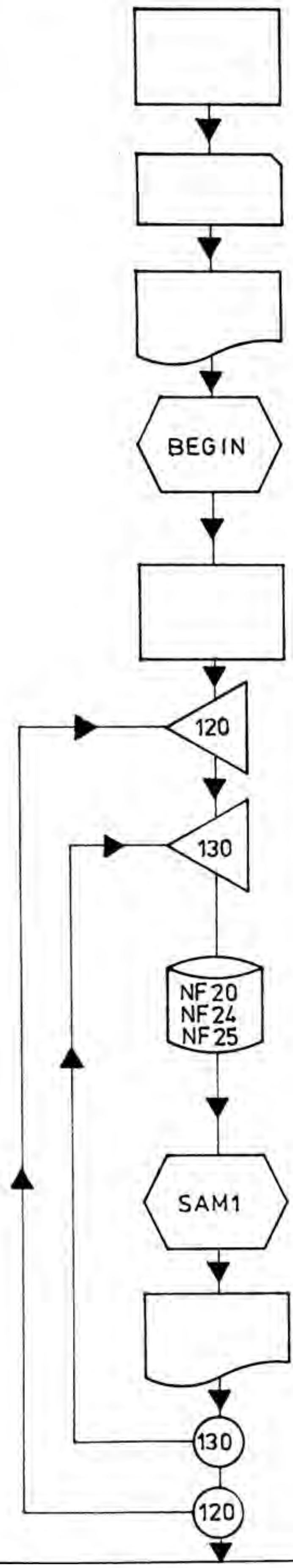
If the cross correlation function is to be punched, temporarily write it to file NPU15.

All user specified punch output was temporarily written to disk scratch files in card image format. Each of these files is now read in turn and the contents punched.

Stop.

FIGURE 3.2 (CONT'D): PROGRAM STATS FLOWCHART

PROGRAM: MOMENT	SUBPROGRAM: MAIN



Assign installation unit numbers for printer, card reader and punch.

Read in problem data, program options, data file reference numbers and historical first three moments and extremes data.

Echo check problem data, program options, data file numbers, historical moments and extremes data.

Initialise results vectors and sample rejection matrices for first moment analysis.

Determine for each station, the statistically acceptable sampling intervals for the synthetic sample first moments and the synthetic population first moments.

Loop over the number of synthetic samples (NSAM).

Loop over the number of stations (NSTN).

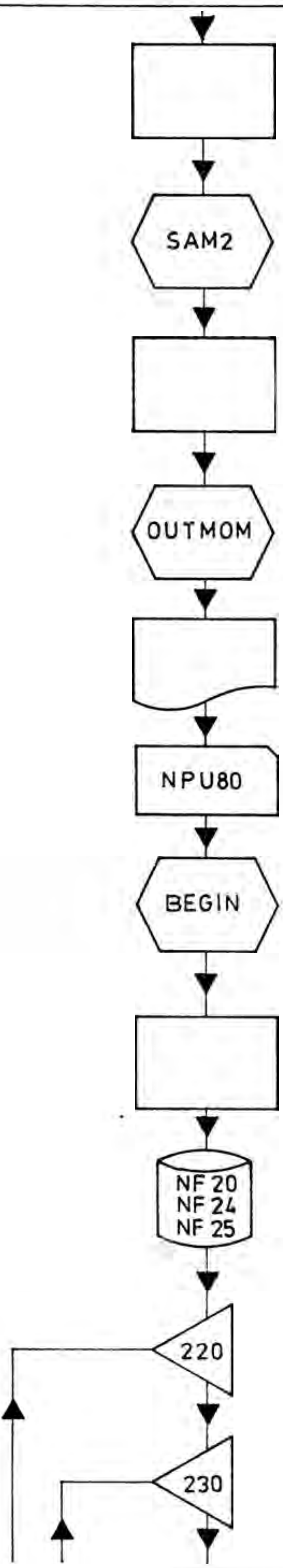
Read this station's synthetic monthly first moments from file NF20 and, if required, synthetic annual and time series first moments from files NF24 and NF25 respectively.

Update the range of synthetic moments, sums for calculation of synthetic monthly, annual and time series first moment sampling distributions, and check synthetic moments against sample confidence intervals noting violations and sample rejections.

Print this station's synthetic first moment data for this synthetic sample.

FIGURE 4.1: PROGRAM MOMENT FLOWCHART

PROGRAM: MOMENT	SUBPROGRAM: MAIN



Calculate the synthetic population first moments, and save cumulative sums of synthetic sample first moments squared and cubed (for each station), for use in the calculation of synthetic population 2nd and 3rd moments.

Check synthetic population first moments against population confidence intervals. Calculate: (1) the sampling distribution moments of the synthetic data, (2) the percentage errors/month (between historical and synthetic population moments), (3) the mean absolute percentage error/month and /wettest and/driest 6 months.

For each station, sum the number of synthetic samples rejected. For each sample, sum the number of stations rejecting the sample.

For all stations, output the results of the confidence interval tests, the sampling distribution and error indice calculations, and the determined range of synthetic values.

Print station-sample rejection matrices for the two specified significance levels.

For all stations punch the historical and synthetic first moments with confidence intervals, if required.

Initialise results vectors and sample rejection matrices for second moment analysis.

Determine for each station, the statistically acceptable sampling intervals for the synthetic sample standard deviations and the synthetic population standard deviations.

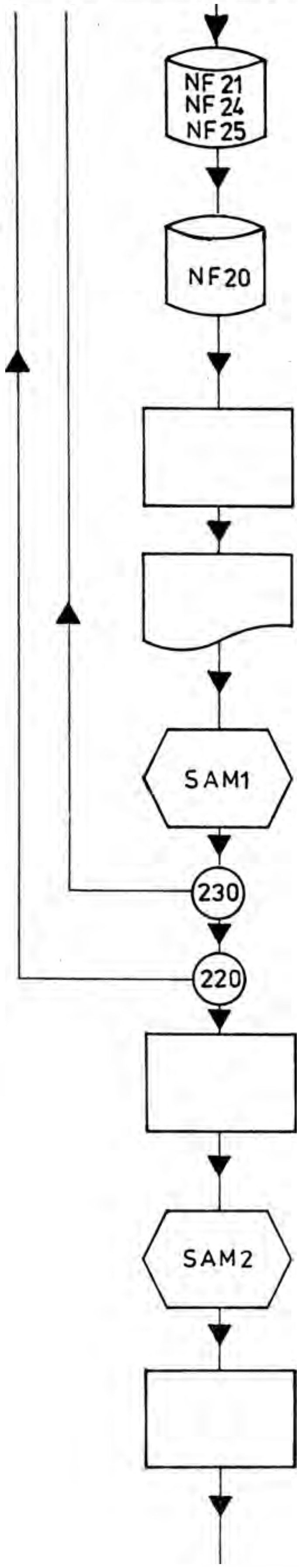
Rewind the synthetic monthly first moments file (NF20), and the synthetic annual (NF24) and time series (NF25) moments files.

Loop over the number of synthetic samples (NSAM).

Loop over the number of stations (NSTN).

FIGURE A.1 (CONT'D): PROGRAM MOMENT FLOWCHART

PROGRAM: MOMENT	SUBPROGRAM: MAIN



Read this station's synthetic monthly standard deviations from file NF21 and, if required, synthetic annual and time series standard deviations from files NF24 and NF25 respectively.

Reread synthetic first moments for use in calculation of factors used in determining population third moments.

Update the values of the mean synthetic second moments and calculate third moment population correction factors from products of first and second synthetic sample moments.

Print this station's synthetic standard deviation data for this synthetic sample.

Update the range of synthetic standard deviations, sums for calculation of synthetic monthly, annual and time series standard deviation sampling distributions, and check synthetic standard deviations against sample confidence intervals noting violations and sample rejections.

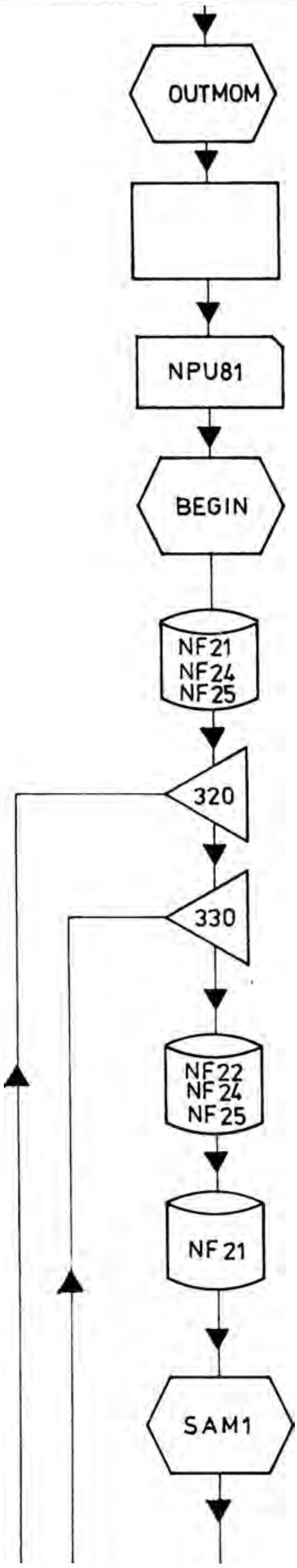
Calculate the synthetic population standard deviations from the mean synthetic sample second moments and results from first moment analysis.

Check synthetic population standard deviations against population confidence intervals. Calculate: (1) the sampling distribution moments of the synthetic data, (2), the percentage errors/month (between historical and synthetic population moments), (3) the mean absolute percentage error/month and/wettest and/driest 6 months.

Form sums for second moment station-sample rejection matrices, and form combined 1st and 2nd moment rejection matrices by combining with 1st moment rejection results.

FIGURE 4.1 (CONT'D): PROGRAM MOMENT FLOWCHART

PROGRAM: MOMENT	SUBPROGRAM: MAIN



For all stations output the results of the confidence interval tests, the sampling distribution and error indice calculations, and the determined range of synthetic values.

Print station-sample rejection matrices for the second moment, and for the first and second moment combined results at the two specified significance levels.

For all stations punch the historical and synthetic standard deviations with confidence intervals, if required.

Initialise results vectors for third moment analysis.

Rewind the synthetic monthly standard deviations file (NF21), and the synthetic annual (NF24) and time series (NF25) moments files.

Loop over the number of synthetic samples (NSAM).

Loop over the number of stations (NSTN).

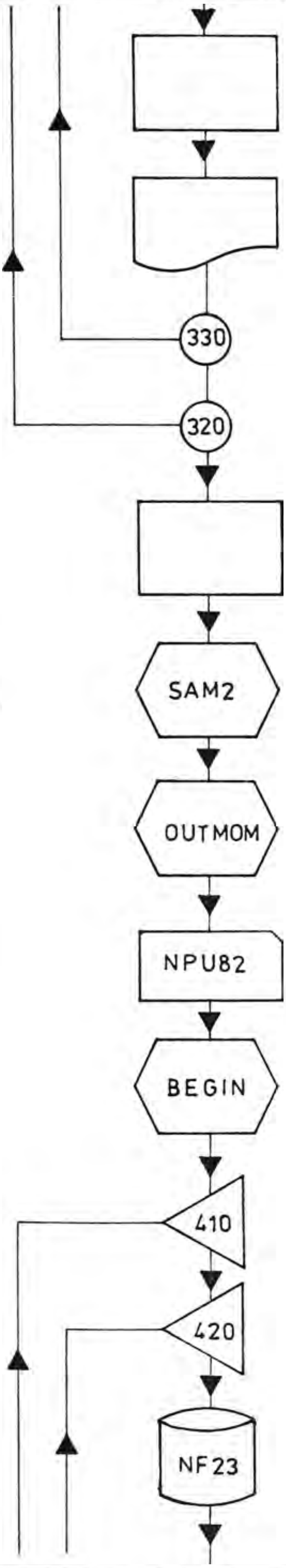
Read this station's synthetic monthly skew coefficients from file NF22 and, if required, synthetic annual and time series skew coefficients from files NF24 and NF25 respectively.

Reread standard deviations for use in calculation of third moments from skew coefficients.

Update the range of synthetic skew coefficients, sums for calculation of synthetic monthly, annual and time series skew coefficient sampling distributions.

FIGURE 4.1 (CONT'D): PROGRAM MOMENT FLOWCHART

PROGRAM: MOMENT | SUBPROGRAM: MAIN



Calculate third moments from skew coefficients and standard deviations and update the values of mean synthetic third moments.

Print this station's synthetic skew coefficient data for this synthetic sample.

Calculate the synthetic population third moments from the mean synthetic sample third moments and previous results, and then the synthetic population skew coefficients.

Calculate: (1) the sampling distribution moments of the synthetic data, (2) the percentage errors/month (between historical and synthetic population moments), (3) the mean absolute percentage error/month and/ wettest and/driest 6 months.

Output the results of the sampling distribution and error indice calculations and the determined range of synthetic values for all stations.

For all stations punch the historical and synthetic monthly skew coefficients if required.

Initialise vectors required for comparison of historical and synthetic monthly extremes.

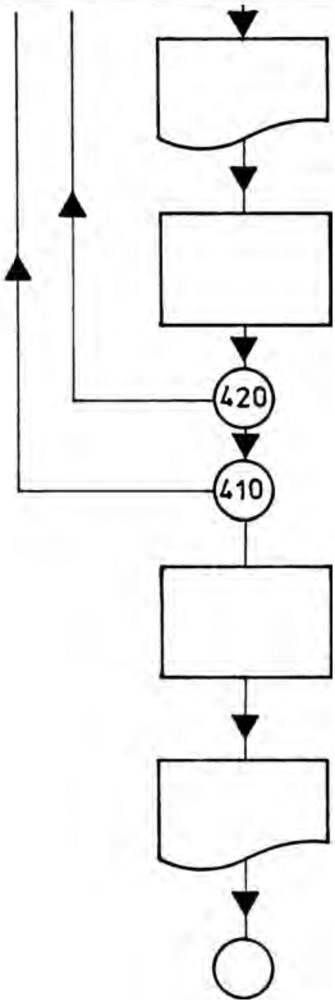
Loop over each synthetic sample (NSAM).

Loop over each station (NSTN).

Read monthly minimum and maximum extremes data.

FIGURE 4.1 (CONT'D): PROGRAM MOMENT FLOWCHART

PROGRAM: MOMENT	SUBPROGRAM: MAIN



Print this station's monthly minimum and maximum extremes data for this synthetic sample.

For both minimum and maximum monthly extremes, accumulate statistics for estimating sampling distribution moments, and determine the range of synthetic extremes and count the number of synthetic values more extreme than the historical extremes.

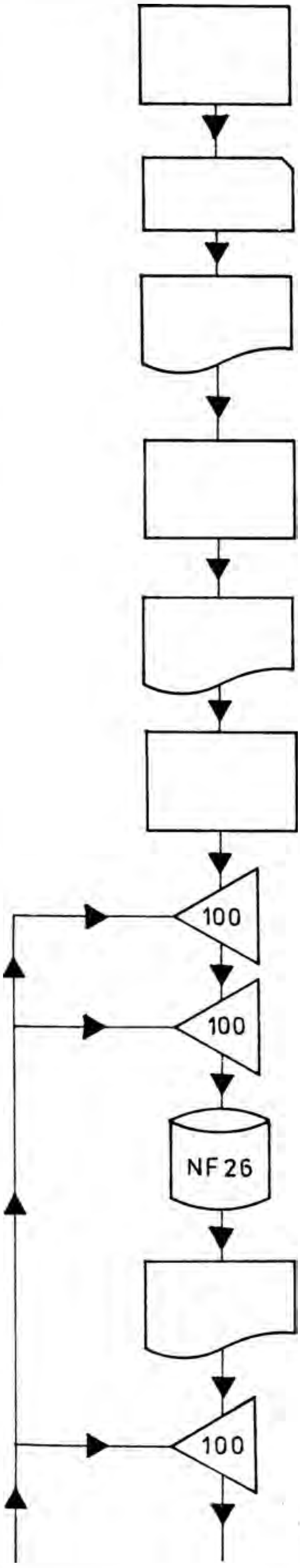
For each station and for each extreme, calculate the sampling distribution moments, and the percentage of synthetic extremes more extreme than historical extremes for each month.

For each station and for each extreme, output the range of synthetic extremes, the percentage of extremes more extreme than the historical extremes, and the sampling distribution moments for each month.

Stop

FIGURE 4.1 (CONT'D): PROGRAM MOMENT FLOWCHART

PROGRAM: FREQ	SUBPROGRAM: MAIN



Assign card reader/punch and printer installation unit numbers, and initialise Kolmogorov-Smirnov (K-S) statistic vector (by data statement).

Read in problem data, program options, data file reference number, and historical data time series frequency distributions.

Echo check the problem data, program options, data file reference number and K-S statistics.

Initialise the matrices used in the analysis of the synthetic sample time series frequency data, and calculate the time series cumulative probability distribution and the expected number of values in each bin or interval of the synthetic sample time series relative and cumulative frequency distributions for each station.

Print out for each station, the historical relative and cumulative probability distributions, and the expected number of values in each bin or interval of the synthetic sample relative and cumulative frequency distributions.

Initialise the sample-station rejection matrices.

Loop over the number of synthetic data samples (NSAM).

Loop over the number of stations (NSTN).

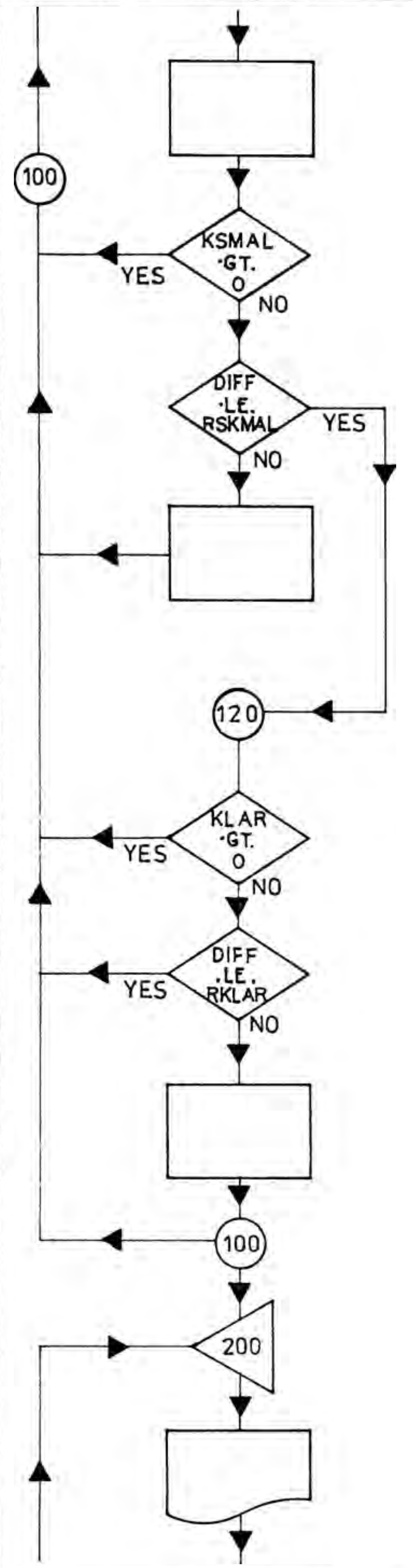
Read in the synthetic frequency distribution of this station for this sample.

Print out the synthetic frequency distribution of this station for this sample.

Loop over the common number of bins or class intervals of this synthetic frequency distribution (INT).

FIGURE 5.1: PROGRAM FREQ FLOWCHART

PROGRAM: FREQ	SUBPROGRAM: MAIN



Increment the population relative frequency distribution and update the range of synthetic sample relative frequencies for this bin of this station. Calculate the cumulative frequency distribution to this bin, and update the range of synthetic sample cumulative frequencies for this bin of this station.

Test whether this synthetic sample distribution has already been statistically rejected at a previous bin by the K-S test for the smaller significance level.

Calculate the K-S difference between this sample and the corresponding historical distribution for this bin, and test if it is less than the K-S statistic for the smaller significance level.

Increment the total number of samples rejected at this bin by the K-S test at both the smaller and larger significance levels. Record that this sample was rejected at both the small and large significance levels and increment station and sample rejection totals.

Test whether this synthetic sample distribution has already been statistically rejected at a previous bin by the K-S test for the larger significance level.

Test whether the K-S difference for this bin is less than the K-S statistic for the larger significance level.

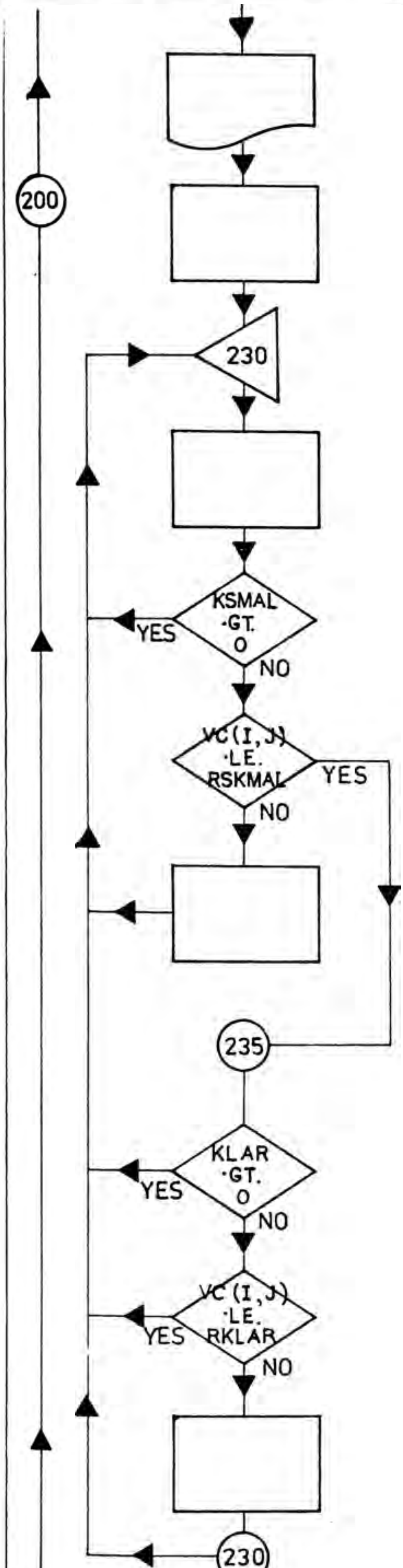
Increment the total number of samples rejected at this bin by the K-S test at the large significance level. Record that this sample was rejected at the large significance level and increment station and sample rejection totals.

Loop over the number of stations (NSTN).

Print out the historical relative frequencies and probabilities, and the range of synthetic sample relative frequencies and probabilities for this station.

FIGURE 5.1 (CONT'D): PROGRAM FREQ FLOWCHART

PROGRAM: FREQ	SUBPROGRAM: MAIN



Print out the historical cumulative frequencies and probabilities, the range of synthetic sample cumulative frequencies and probabilities, and the number of synthetic samples rejected at the small and large significance levels.

Calculate the K-S statistics for the synthetic population at the two significance levels.

Loop over the common number of bins or class intervals of the synthetic frequency distributions (INT).

Calculate the synthetic population relative and cumulative probability for this bin, and the percentage difference between the synthetic population cumulative probability and the historical cumulative probability for this bin.

Test whether this synthetic population distribution has already been statistically rejected at a previous bin by the K-S test for the smaller significance level.

Test whether the K-S difference for this bin is less than the K-S statistic for the smaller significance level.

Note that the synthetic population distribution is rejected at both significance levels for this bin.

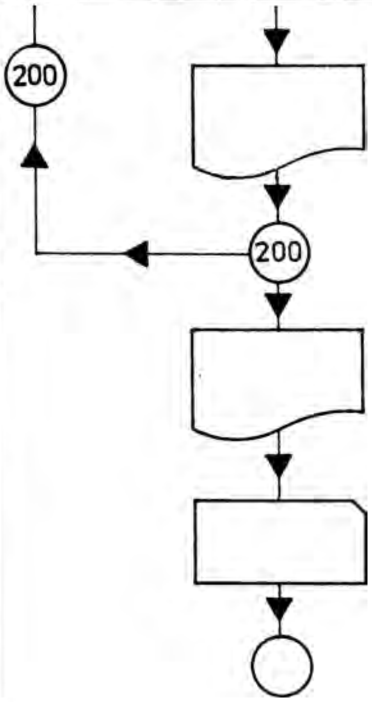
Test whether the synthetic population distribution has already been statistically rejected at a previous bin by the K-S test for the large significance level.

Test whether the K-S difference for this bin is less than the K-S statistic for the large significance level.

Note that the synthetic population distribution is rejected at the large significance level.

FIGURE 5.1 (CONT'D): PROGRAM FREQ FLOWCHART

PROGRAM: FREQ	SUBPROGRAM: MAIN



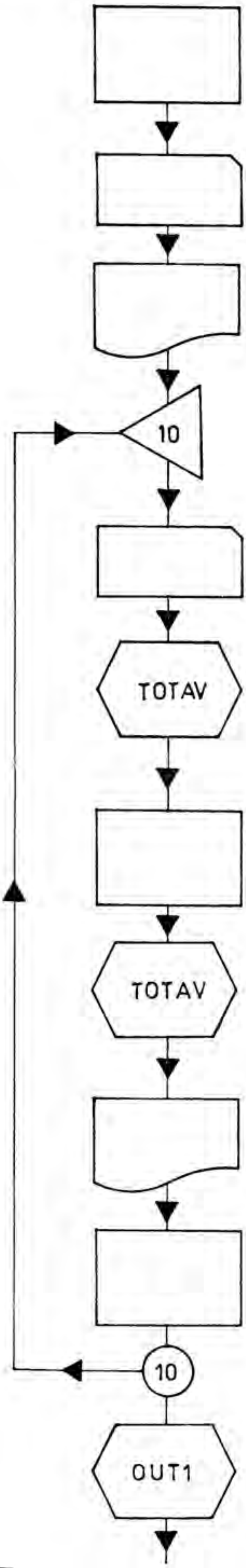
Print out the historical and synthetic population relative frequency and probability distributions; the historical and synthetic population cumulative frequency and probability distributions; the K-S differences, and whether they were statistically significant; the percentage differences for each bin, and the mean absolute percentage difference per bin.

Print out sample-station rejection matrices at the two specified significance levels.

Write the relative and cumulative historical and synthetic population probability distributions to file NPU83 if required.

Stop

PROGRAM: RUNS	SUBPROGRAM: MAIN



Assign card reader/punch and printer installation unit numbers (by data statement) and initialise Kolmogorov-Smirnov (K-S) statistic vector (by assignment statements).

Read in problem data, data file reference numbers and program options.

Echo check the problem data, data file reference numbers and program options.

Loop over the number of stations (NSTN).

Read in the historical data run length frequencies above and below the median and add to form run length frequencies about the median.

Determine the overall total number of runs and the average run length for each frequency distribution type.

Determine the longest run length for which the sum of the number of runs about the median, for this and all longer run lengths, first exceeds 'ILRUN' percent (P/C) of the total number of runs about the median (i.e., the minimum long run length).

Determine the total number of long runs and the average long run length for each frequency distribution type.

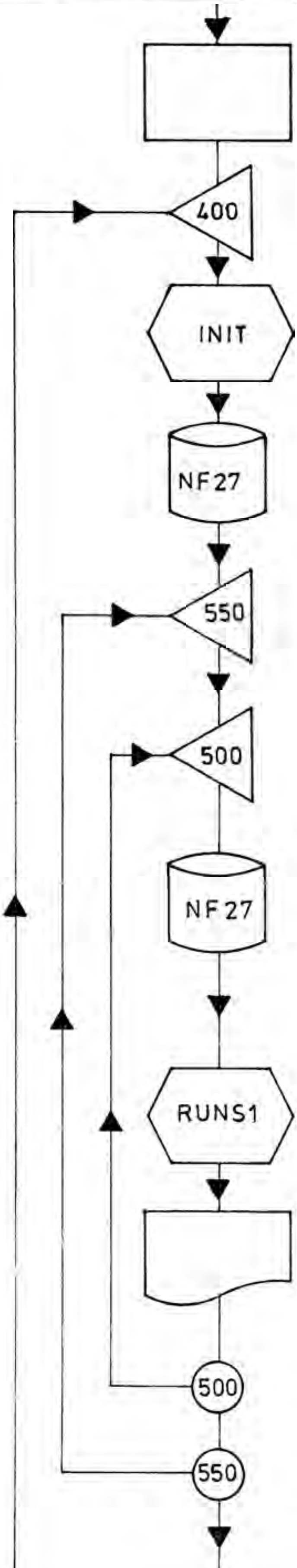
Output the overall total number of runs and average run length, and the total number of long runs and average long run length, for each frequency distribution type.

Calculate the relative and cumulative run length probability distributions for each distribution type.

Output each station's historical relative and cumulative run length probability distributions for each distribution type.

PROGRAM: RUNS

SUBPROGRAM: MAIN



Initialise the combined sample-station rejection matrices which combine the sample-station rejections of runs above, below and about the median.

Loop over the run length distribution types (i.e. runs above, below and about the median distributions).

Initialise the matrices used in the analysis of the synthetic sample run length frequency data.

Rewind the synthetic sample run length frequency data file.

Loop over the number of synthetic samples (NSAM).

Loop over the number of stations (NSTN).

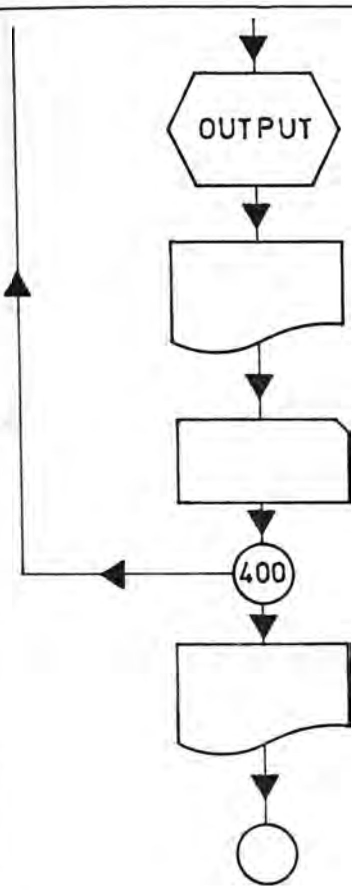
Read the synthetic sample run length frequency data above and below the median.

Test the appropriate distribution type against the corresponding historical, update the range of synthetic probabilities and the totals for the calculation of the sampling distribution moments of the synthetic distribution's overall total and average, and long run total and average.

Echo check the synthetic sample run length frequency distribution data for this distribution type.

FIGURE 6.1 (CONT'D): PROGRAM RUNS FLOWCHART

PROGRAM: RUNS | SUBPROGRAM: MAIN



Calculate the synthetic population relative and cumulative run length probability distributions for this distribution type and apply the K-S test. Calculate the sampling distribution moments of the distribution overall and long run totals and averages, and output the results for each station.

Output the sample-station rejection matrices for this distribution type at the two specified significance levels.

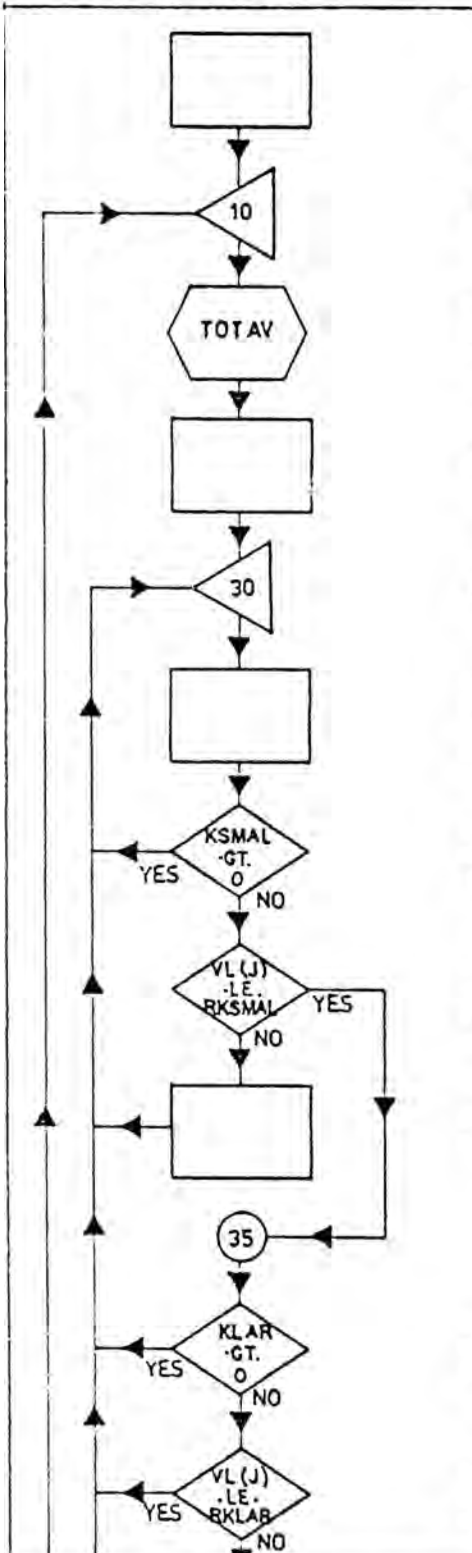
Punch the relative and cumulative historical and synthetic population run length probability distributions of this distribution type if required.

Output the combined sample-station rejection matrices at the two specified significance levels.

Stop.

FIGURE 6.1 (CONT'D): PROGRAM RUNS FLOWCHART

PROGRAM: RUNS	SUBPROGRAM: OUTPUT



Assign the specified significance percentage levels to variables used in printing results.

Loop over the number of stations in the analysis (NSTN)

Determine the synthetic population frequency distribution overall and long run total number of runs and average run lengths.

Calculate the K-S statistics for the synthetic population at the two significance levels.

Loop over each run length of the synthetic population distribution (NRUNS)

Calculate the relative probability of this run length and the cumulative probability to it, and the percentage difference between the synthetic population cumulative probability and the historical cumulative probability.

Test if this synthetic population distribution has already been statistically rejected at a previous run length by the K-S test for the smaller significance level.

Test whether the K-S difference for this run length is less than the K-S statistic for the smaller significance level.

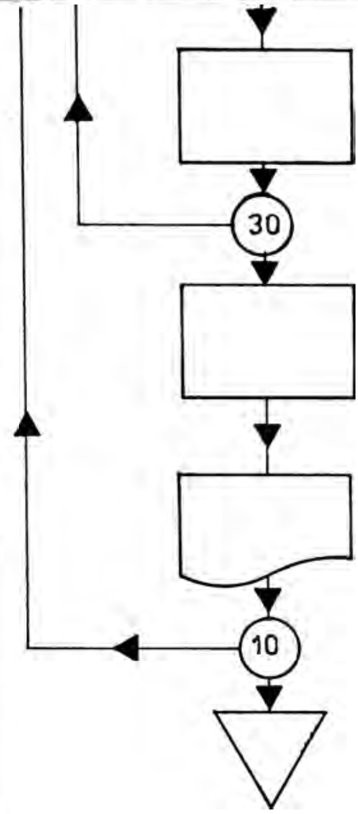
Note that the synthetic population distribution is rejected at both significance levels for this run length.

Test if this synthetic population distribution has already been statistically rejected at a previous run length by the K-S test for the large significance level.

Test whether the K-S difference for this run length is less than the K-S statistic for the large significance level.

FIGURE 6.2: SUBPROGRAM OUTPUT FLOWCHART

PROGRAM: RUNS	SUBPROGRAM: OUTPUT



Note that the synthetic population distribution is rejected at the large significance level.

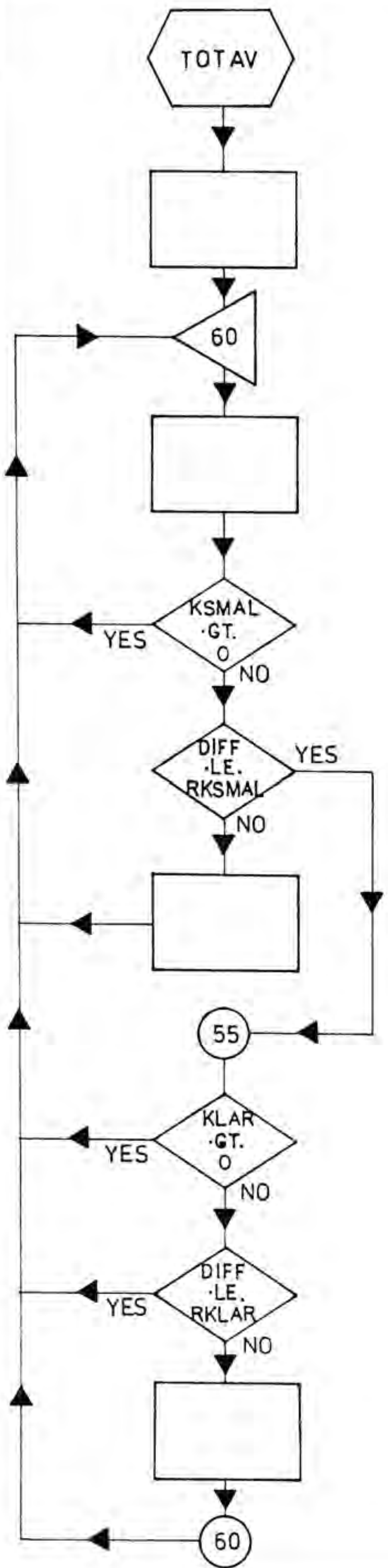
Calculate the sampling distribution moments of the distribution overall and long run totals and average run lengths.

Output the results of all calculations for the individual synthetic sample distributions, the synthetic population distribution, and the sampling distribution moments of the synthetic sample overall and long run totals and averages.

Return to the calling program.

FIGURE 6.2 (CONT'D): SUBPROGRAM OUTPUT FLOWCHART

PROGRAM: RUNS	SUBPROGRAM: RUNS1



Determine the total number of runs and the average run length of this synthetic sample frequency distribution.

Calculate the Kolmogorov-Smirnov (K-S) statistics for this synthetic sample run length distribution at the two specified significance levels.

Loop over the number of run lengths (NRUNS).

Calculate the relative probability of this run length and the cumulative probability to it. Update the ranges of synthetic sample relative and cumulative probabilities for this run length.

Test if this sample distribution has already been statistically rejected at a previous run length by the K-S test for the smaller significance level.

Calculate the K-S difference between this sample and the corresponding historical distribution for this run length, and test if it is less than the K-S statistic for the smaller significance level.

Increment the total number of samples rejected at this run length by the K-S test at both the small and large significance levels. Record that this sample was rejected at both the small and large significance levels and increment station and sample rejection totals.

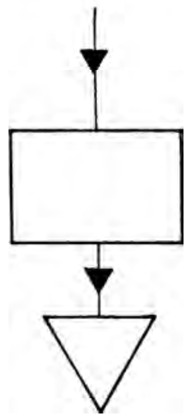
Test if this sample distribution has already been statistically rejected at a previous run length by the K-S test for the larger significance level.

Test whether the K-S difference for this run length is less than the K-S statistic for the larger significance level.

Increment the total number of samples rejected at this run length by the K-S test at the large significance level. Record that this sample was rejected at the large significance level and increment station and sample rejection totals.

FIGURE 6.3: SUBPROGRAM RUNS1 FLOWCHART

PROGRAM: RUNS	SUBPROGRAM: RUNS1

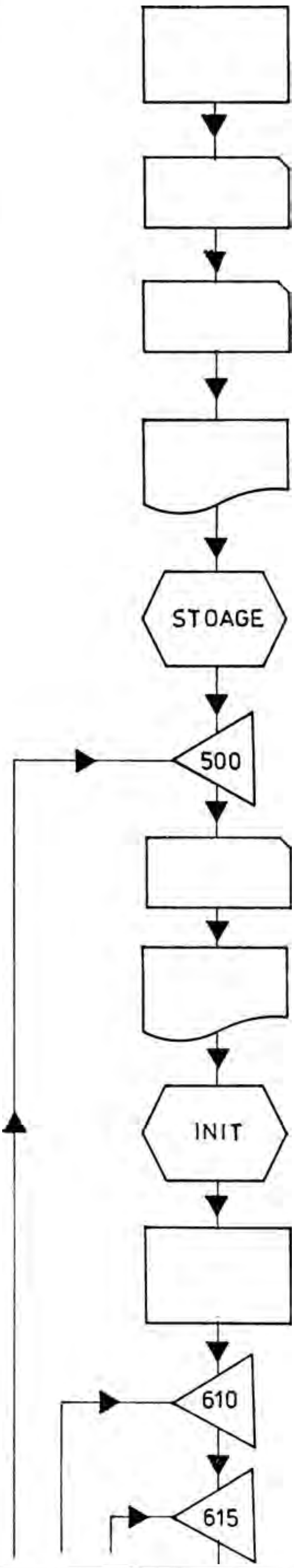


Collect statistics from which the moments of the sampling distribution of the distribution overall and long run totals and averages may be calculated. Update the range of distribution overall and long run totals and averages.

Return to the calling program.

FIGURE 6.3 (CONT'D): SUBPROGRAM RUNS 1 FLOWCHART

PROGRAM: YIELD	SUBPROGRAM: MAIN



Assign installation unit numbers for printer, card reader/punch.

Read in problem data, data file reference numbers and program options.

Read in the historical Rippl storages and drought durations, the distribution bounds and the number of bins used in the frequency analyses, and the calculated distribution statistics (totals and averages).

Echo check the problem data, data file reference numbers, program options, historical Rippl analysis results, frequency analyses parameters, and distribution statistics.

Analyse the synthetic data Rippl analysis results calculating: (1) the moments of the sampling distributions of the synthetic Rippl storages and drought durations, (2) the number of standard deviations the historical Rippl storage is from the synthetic Rippl mean storage, (3) the probability of storage being less than or equal to each synthetic Rippl estimate. Print results.

Loop over the four frequency distributions (deficits, drought, draw and fill) to be analysed.

Read in each station's historical frequency distributions at each yield level and hence calculate its relative and cumulative probability distributions.

Echo check the relative and cumulative probability distributions at each yield level for each station.

Initialise the matrices used in the analysis of the synthetic sample distributions.

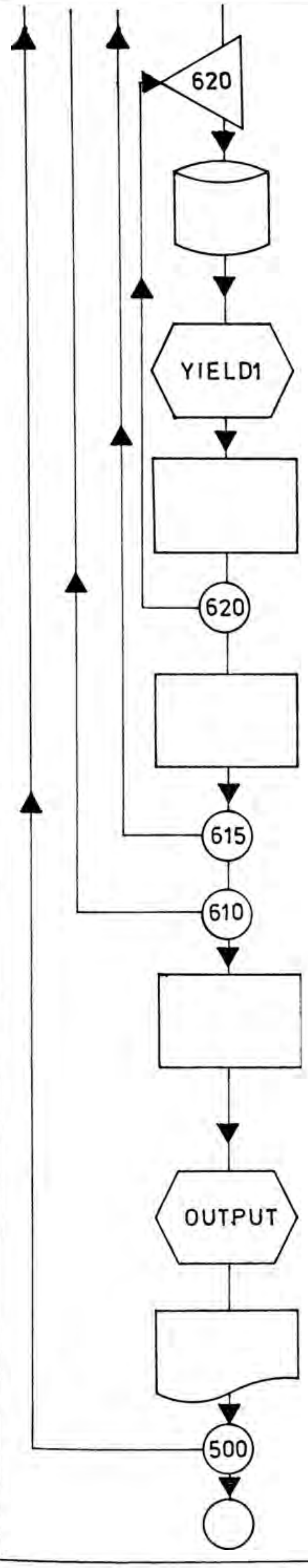
If the synthetic sample distributions are to be echo checked, then initialise program output control parameters.

Loop over the number of synthetic samples (NSAM).

Loop over the number of stations (NSTN).

FIGURE 7.1: PROGRAM YIELD FLOWCHART

PROGRAM: YIELD	SUBPROGRAM: MAIN



Loop over the number of yield levels (NYIELD).

Read in a synthetic frequency distribution and its total number of values together with the calculated mean value.

Test this synthetic distribution against the corresponding historical, update the range of synthetic probabilities and the totals for the calculation of the sampling distribution moments of the synthetic distributions' totals and averages.

If the synthetic sample distributions are to be echo checked, then fill in this sample's output area in the page image matrix IDATA.

If the synthetic sample distributions are to be echo checked, then test if the page image is complete, and if so, then print a physical page and reset the program output control parameters.

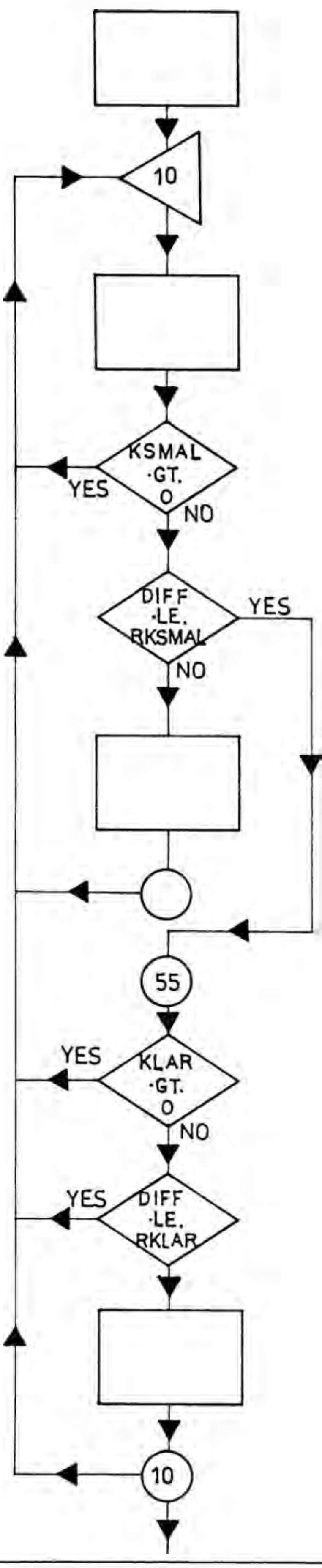
If the synthetic sample distributions are to be echo checked, then check if an incomplete page image remains to be printed, and if so, then print it.

Calculate the sampling distribution moments of the distribution total and average, the synthetic population relative and cumulative probability distributions and apply the Kolmogorov-Smirnov test. Output all station results and all the sample rejection matrices.

Stop

FIGURE 7.1 (CONT'D): PROGRAM YIELD FLOWCHART

PROGRAM: YIELD	SUBPROGRAM: YIELD1



Calculate the Kolmogorov-Smirnov (K-S) statistics for this synthetic sample frequency distribution at the two specified significance levels.

Loop over this sample's frequency distribution bins.

Calculate the relative probability of this bin and the cumulative probability to it. Update the ranges of synthetic sample relative and cumulative probabilities for this bin.

Test if this sample distribution has already been statistically rejected at a previous bin by the K-S test for the smaller significance level.

Calculate the K-S difference between this sample and the historical distribution for this bin and test if it is less than the K-S statistic for the smaller significance level.

Increment the total number of samples rejected at this bin by the K-S test at both the small and large significance levels. Record that this sample was rejected at both the small and large significance levels and increment station and sample rejection totals.

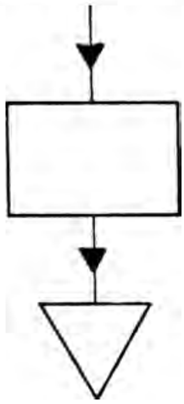
Test if this sample distribution has already been statistically rejected at a previous bin by the K-S test for the larger significance level.

Test whether the K-S difference for this bin is less than the K-S statistic for the larger significance level.

Increment the total number of samples rejected at this bin by the K-S test at the large significance level. Record that this sample was rejected at the large significance level and increment station and sample rejection totals.

FIGURE 7.2: SUBPROGRAM YIELD1 FLOWCHART

PROGRAM: YIELD	SUBPROGRAM: YIELD1

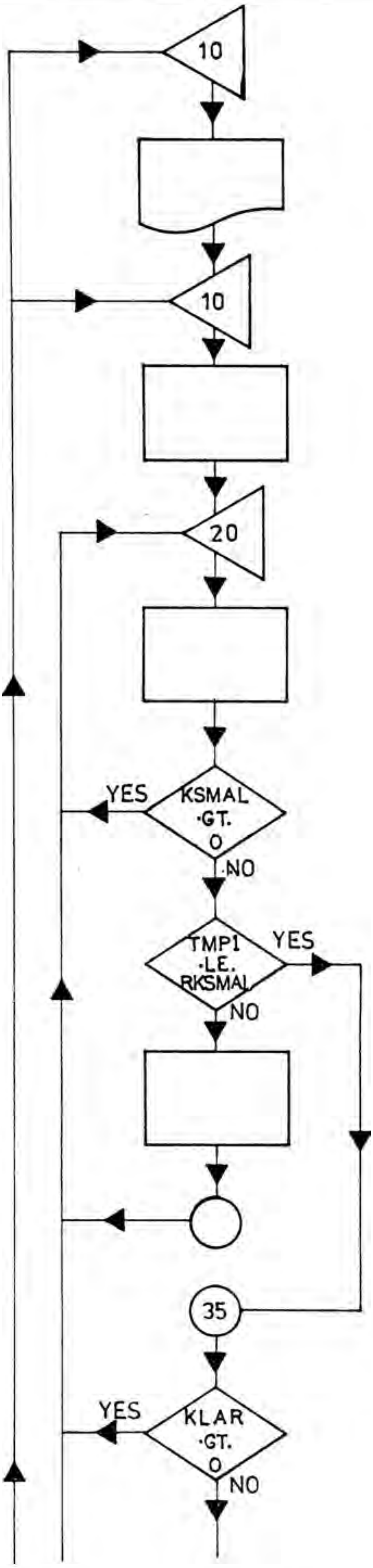


Collect statistics from which the moments of the sampling distribution of the distribution total and average may be calculated. Update the range of distribution totals and averages.

Return to the calling program.

FIGURE 7.2 (CONT'D): SUBPROGRAM YIELD1 FLOWCHART

PROGRAM: YIELD	SUBPROGRAM: OUTPUT



Loop over all the stations in this analysis (NSTN).

Output the page headings for the analysis results for this station.

Loop over all the yield levels in this analysis (NYIELD).

Sum the number of observations in the synthetic population and hence calculate the K-S statistics for the synthetic population at the two significance levels.

Loop over each bin of the synthetic population frequency distribution.

Calculate the relative probability of this bin and the cumulative probability to it, and the percentage difference between the synthetic population cumulative probability and the historical cumulative probability.

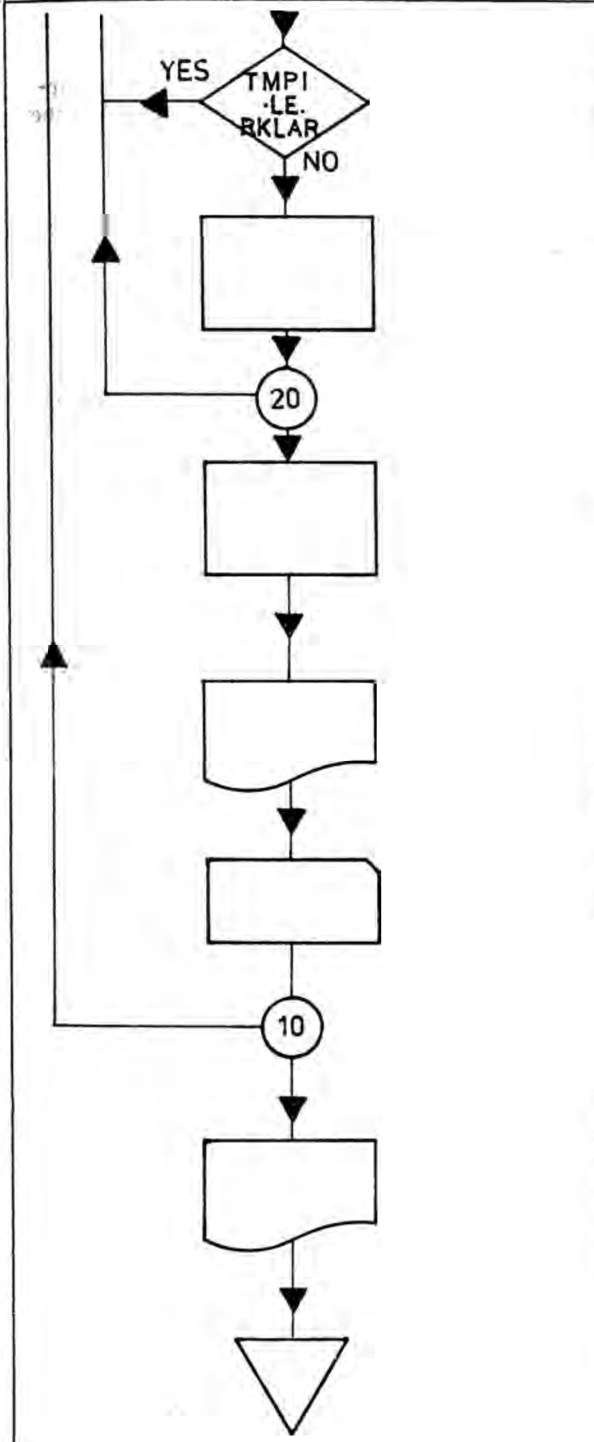
Test if this population distribution has already been statistically rejected at a previous bin by the K-S test for the smaller significance level.

Test whether the K-S difference for this bin is less than the K-S statistic for the smaller significance level.

Note that the synthetic population distribution is rejected at both significance levels at this bin.

Test if this population distribution has already been statistically rejected at a previous bin by the K-S test for the larger significance level.

PROGRAM: YIELD	SUBPROGRAM: OUTPUT



Test whether the K-S difference for this bin is less than the K-S statistic for the larger significance level.

Note that the synthetic population distribution is rejected at the larger significance level.

Calculate the sampling distribution moments of the total and average of the synthetic sample distributions.

Output the results of all calculations for the individual synthetic sample distributions, the synthetic population distribution, and the moments of the sampling distributions of the synthetic sample total and average statistics.

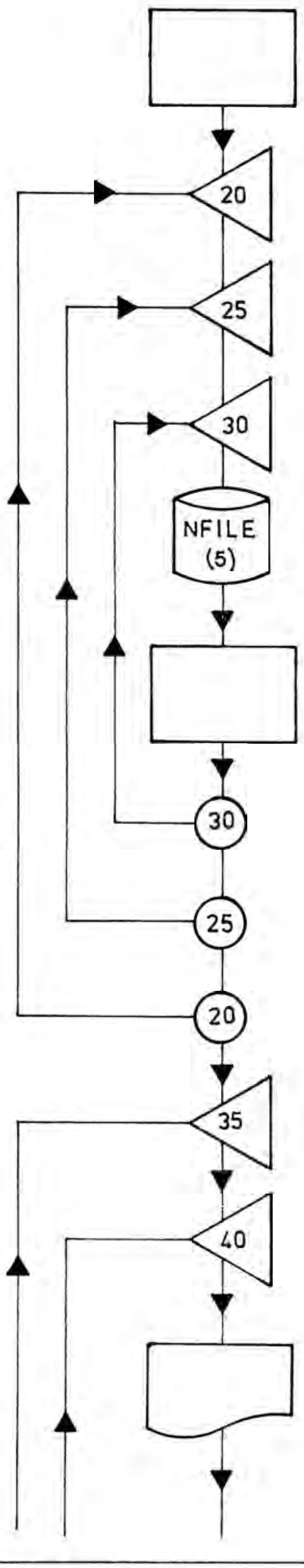
Punch the historical and synthetic population relative and cumulative probability distributions to the specified output file if required.

Output for each yield level and each significance level, the station-sample rejection matrices, showing for which samples a station's synthetic frequency distribution was statistically unacceptable. Then output for each significance level, the station-sample rejection matrices that result by combining the rejections at all yield levels.

Return to the calling program.

FIGURE 7.3 (CONT'D): SUBPROGRAM OUTPUT FLOWCHART

PROGRAM: YIELD	SUBPROGRAM: STOAGE



Initialise the matrix used for the moments of the sampling distributions of the synthetic Rippl storages and the drought durations.

Loop over the number of synthetic samples (NSAM).

Loop over the number of stations (NSTN).

Loop over the number of yield levels (NYIELD).

Read this synthetic sample Rippl storage and drought duration and save their values together with the sample number.

Save the statistics for the calculation of the sampling distribution moments of the Rippl storage and drought duration

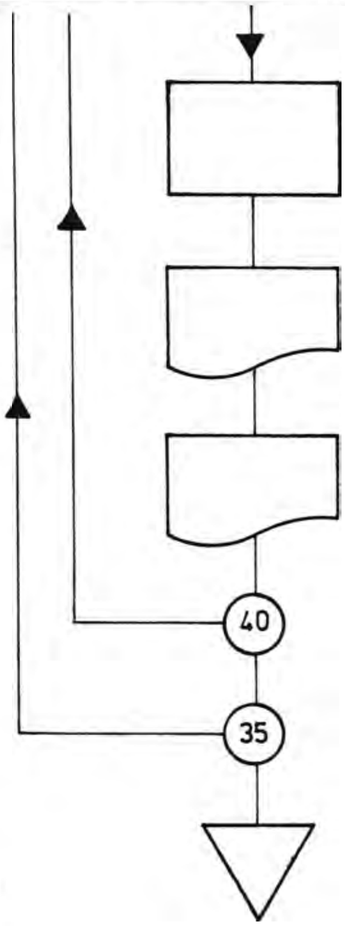
Loop over the number of stations (NSTN).

Loop over the number of yield levels (NYIELD).

Calculate the moments of the sampling distributions of the Rippl storage and drought duration and output them.

FIGURE 7.4: SUBPROGRAM STOAGE FLOWCHART

PROGRAM: YIELD	SUBPROGRAM: STOAGE



Order the Rippl storages from the smallest to the largest and sort the drought duration and sample numbers similarly, so as to preserve the correspondence between a storage, its drought duration and its sample number.

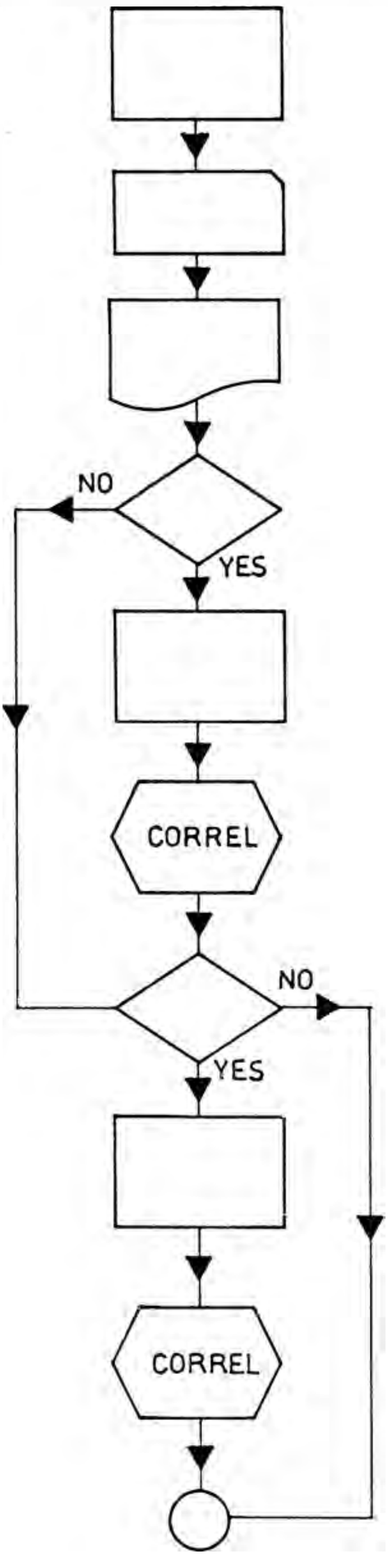
Output the ordered Rippl storages and their drought durations and sample numbers.

Calculate the number of standard deviations the historical Rippl storage is from the synthetic Rippl storage mean and print.

Return to the calling program.

FIGURE 7.4 (CONT'D): SUBPROGRAM STOAGE FLOWCHART

PROGRAM: CORREL	SUBPROGRAM: MAIN



Assign installation unit numbers for printer, card reader and punch.

Read in problem data, data file reference numbers and program options.

Echo check the problem data, data file reference numbers and program options.

Is auto correlation analysis required?

Set up auto correlation results vectors contiguously in core by calculating their starting elements as vector element numbers in the vector 'SPACE'.

Call subprogram 'CORREL' to perform the analysis and pass across the contiguous auto correlation results vectors.

Is cross correlation analysis required?

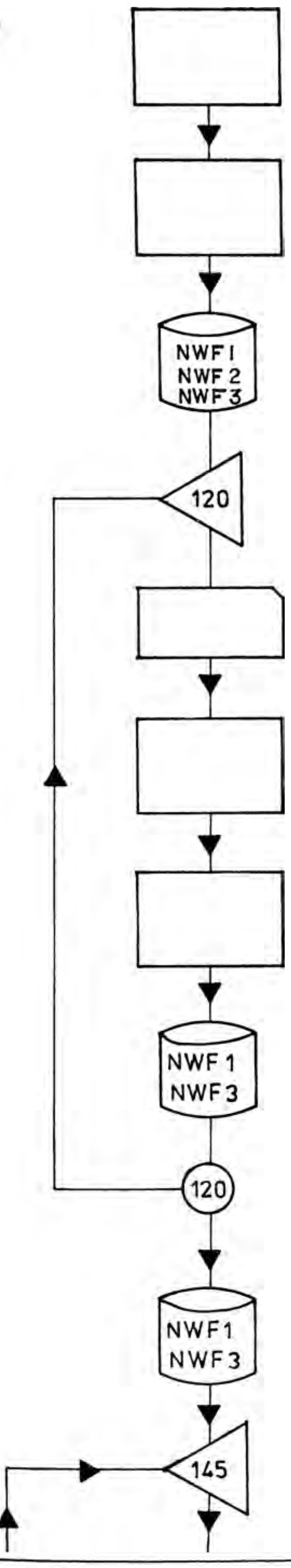
Set up cross correlation results vectors contiguously in core by calculating their starting elements as vector element numbers in the vector 'SPACE'.

Call subprogram 'CORREL' to perform the analysis and pass across the contiguous cross correlation results vectors.

Stop

FIGURE 8.1: PROGRAM CORREL FLOWCHART

PROGRAM: CORREL	SUBPROGRAM: CORREL



Initialise the results vectors common to each auto/cross correlation function.

Calculate the variance of the significance test statistics for the auto/cross correlation function.

Rewind the work files.

Loop over each auto/cross correlation function.

Read in the historical auto/cross correlation function.

Calculate the statistically acceptable range of the synthetic auto/cross correlation coefficient at each lag and assign to appropriate results vectors.

Test each historical auto/cross correlation coefficient of this function for significant difference from zero and note result in appropriate results vectors.

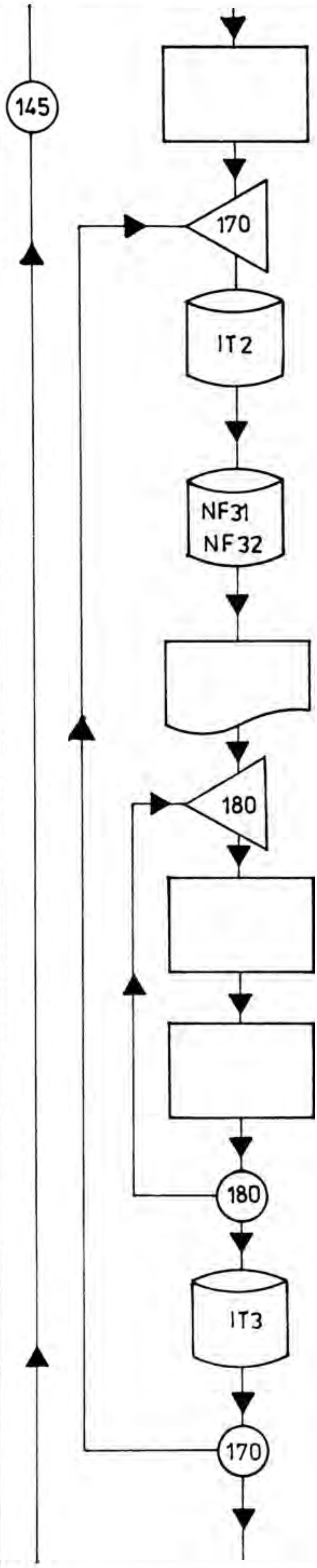
Write this auto/cross correlation function's synthetic results vectors to work file NWF1 and the historical results vectors to work file NWF3.

All historical auto/cross correlation functions have been processed: end file and rewind work files.

Loop over the number of synthetic samples (NSAM).

FIGURE 8.2: SUBPROGRAM CORREL FLOWCHART

PROGRAM: CORREL	SUBPROGRAM: CORREL



Determine which work file, NWF1 or NWF2, contains the latest updated set of results vectors. (The results vectors are read back from this file (IT2), updated, and then written to the other file (IT3)).

Loop over the number of auto/cross correlation functions (NSTN auto correlation functions; NSTN (NSTN-1)/2 cross correlation functions).

Read the latest updated results vectors for this auto/cross correlation function.

Read the synthetic auto/cross correlation function.

Print the synthetic auto/cross correlation function, if required.

Loop over each lag of this auto/cross correlation function.

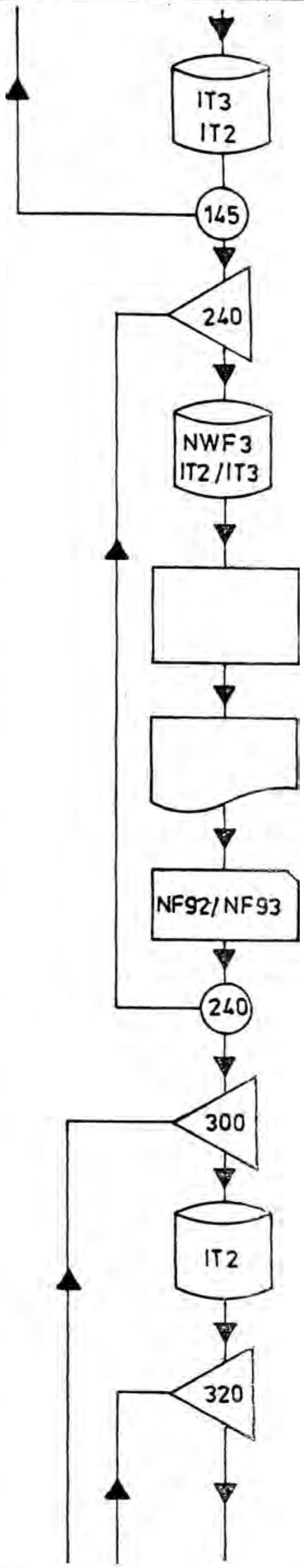
Update the synthetic average, minimum and maximum correlation coefficients in the appropriate results vectors.

Test whether the correlation coefficient is within acceptable limits and whether it is significantly different from zero and then update the appropriate results vectors accordingly.

Write out the updated results vectors. (Note that this output work file becomes the input results file (IT2) for the next synthetic sample.

FIGURE 8.2 (CONT'D): SUBPROGRAM CORREL FLOWCHART

PROGRAM: CORREL	SUBPROGRAM: CORREL



All updated results vectors have been written out; endfile and rewind IT3. Output file for next synthetic sample set becomes IT2.

Loop over the number of auto/cross correlation functions (NSTN or NSTN (NSTN-1)/2 respectively).

Read in the historical results vectors from file NWF3 and the synthetic results vectors from IT2 or IT3 as appropriate.

Calculate the mean synthetic auto/cross correlation coefficients and the relative percentage errors with respect to historical coefficients.

Print the results vectors for each auto/cross correlation function.

Punch historical and synthetic, mean, min and max auto/cross correlation coefficients to the appropriate file, if required.

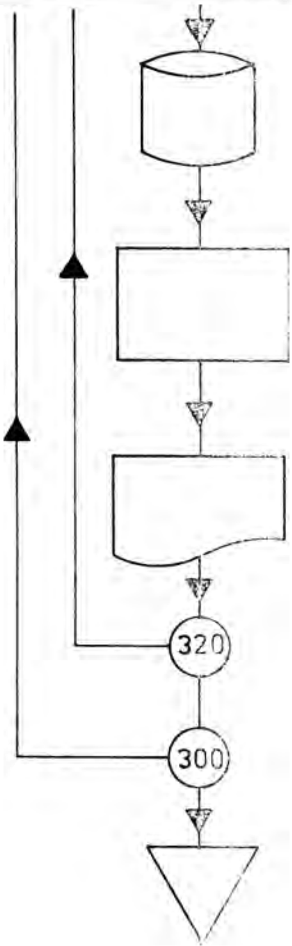
Loop over the two statistical confidence levels specified.

Rewind the results file for reprocessing.

Loop over the necessary number of pages of output.

FIGURE 8.2 (CONT'D): SUBPROGRAM CORREL FLOWCHART

PROGRAM: CORREL	SUBPROGRAM: CORREL



Re-read station-sample rejection matrix data.

Set up a full page of station-sample rejection matrix results.

Print page of station-sample rejection matrix results.

Return to the calling program.

FIGURE 8.2 (CONT'D): SUBPROGRAM CORREL FLOWCHART

FILE: DEMO SYNDATA A CMS REL3 PLC12 CANBERRA 370/158

150	1	50	65	266	359	691	428	192	92	49	39	26	41
150	2	63	128	347	468	561	466	122	165	82	56	59	88
150	3	41	47	111	414	526	995	1024	494	398	176	235	258
150	4	148	339	464	620	763	611	270	174	78	47	138	182
150	5	328	138	130	311	363	893	478	133	72	21	27	73
150	6	267	300	877	860	671	333	232	182	219	101	65	96
150	7	207	301	334	526	730	671	583	357	251	119	71	107
150	8	138	262	649	809	958	788	243	137	81	87	45	68
150	9	97	228	117	110	101	75	123	37	31	15	20	39
150	10	162	140	241	761	592	595	352	275	133	150	105	141
150	11	247	455	386	862	852	746	523	238	159	62	51	44
150	12	138	115	141	534	536	1143	714	512	320	202	184	268
150	13	726	877	368	866	893	1141	579	286	297	166	108	171
150	14	272	338	271	563	729	426	232	61	73	100	76	67
150	15	125	153	186	355	227	228	154	88	63	44	15	24
150	16	133	219	180	87	427	630	271	337	163	114	212	377
150	17	136	238	151	180	214	394	244	256	109	99	79	126
150	18	202	791	484	698	683	570	354	180	135	120	133	96
150	19	116	383	1222	737	782	927	564	256	210	73	45	82
150	20	313	463	379	818	891	1501	1013	342	260	103	206	274
150	21	294	390	896	995	1005	1159	637	278	170	111	92	54
150	22	77	90	52	184	466	672	461	97	132	65	45	70
150	23	1216	1369	460	844	643	610	238	92	80	79	104	149
150	24	89	153	260	237	430	790	312	105	70	81	124	92
150	25	190	135	314	440	412	195	42	88	68	31	25	34
150	26	81	88	325	195	455	694	500	363	129	61	67	72
150	27	193	100	152	296	321	388	190	102	46	29	22	51
150	28	84	339	1107	1234	876	761	631	488	250	151	112	136
150	29	83	189	1259	777	815	1052	1237	773	352	174	102	113
150	30	107	161	176	136	298	512	795	219	75	33	49	39
150	31	94	150	222	329	472	371	185	206	109	86	122	109
150	32	109	156	269	216	231	222	279	238	73	95	104	71
150	33	80	105	564	981	790	1344	1113	468	153	59	82	75
150	34	128	179	183	286	276	633	422	146	67	59	37	75
150	35	124	174	134	700	720	946	401	226	79	51	37	36
150	36	58	88	111	104	100	92	95	32	16	34	62	57
150	37	118	172	831	752	661	880	352	168	99	102	231	302
150	38	237	257	330	371	557	705	260	230	90	66	92	73
150	39	79	104	305	367	483	650	751	237	155	160	96	128
150	40	522	1605	535	830	762	885	414	291	136	82	89	122
150	41	172	355	983	1113	756	608	474	233	121	73	59	81
150	42	133	323	459	485	701	855	367	218	168	105	133	116
150	43	146	394	445	506	642	559	610	274	98	90	62	77
150	44	429	442	1275	1590	1389	1414	715	482	362	282	367	811
150	45	332	531	1063	784	1174	907	434	180	75	46	75	118
150	46	178	281	643	1092	1056	1151	596	253	127	69	96	223
150	47	238	595	1202	651	675	792	785	385	235	126	138	166
150	48	428	484	429	702	803	812	564	225	148	58	100	70
150	49	196	160	161	597	776	709	568	477	178	163	168	231
150	50	247	257	514	674	677	532	158	71	51	24	13	31
180	1	7	6	56	128	224	107	97	9	8	4	2	5
180	2	6	22	77	166	227	222	27	28	6	3	12	15
180	3	4	9	28	123	109	295	236	141	54	52	70	167
180	4	35	196	247	247	304	159	53	43	7	8	28	33
180	5	282	87	47	127	109	368	153	60	12	1	1	4
180	6	29	96	428	460	264	123	60	34	41	19	15	14
180	7	144	117	130	470	285	174	162	135	55	25	22	27
180	8	24	152	450	310	411	230	45	26	25	5	3	5
180	9	22	65	56	57	22	4	16	8	10	1	0	0
180	10	12	42	55	400	401	122	121	43	5	8	14	30
180	11	47	245	191	541	431	357	135	78	60	8	7	3
180	12	28	36	67	349	208	478	257	145	66	48	49	155
180	13	134	586	289	433	262	497	281	37	47	28	23	25
180	14	92	197	181	316	209	68	35	7	8	13	10	11
180	15	23	47	68	118	65	64	38	11	6	7	1	0
180	16	15	25	51	32	131	220	53	95	40	18	40	141
180	17	44	152	71	96	107	109	82	82	18	24	19	36
180	18	47	412	445	229	268	94	44	37	44	15	31	20
180	19	29	136	545	529	331	321	156	43	32	12	6	15
180	20	157	136	363	510	514	423	319	98	41	16	29	100
180	21	148	44	461	655	453	410	98	63	57	25	12	8
180	22	14	28	22	99	132	113	66	13	18	16	2	6
180	23	353	1086	402	725	330	141	38	20	11	9	15	35
180	24	18	86	163	118	94	249	83	29	2	10	20	28
180	25	104	55	150	252	134	68	12	9	3	2	2	3
180	26	10	16	66	74	139	203	104	99	41	6	5	10
180	27	72	57	87	193	138	86	36	20	1	1	0	2

FIGURE 9.1: SYNTHETIC DATA SETS FOR EXAMPLE PROBLEM

180	28	12	103	600	694	403	212	222	192	62	24	12	49
180	29	31	37	341	509	390	505	487	230	63	32	16	36
180	30	25	74	94	69	117	97	229	72	11	1	3	3
180	31	9	28	60	101	134	99	33	18	18	16	35	29
180	32	12	69	133	147	79	44	76	108	22	15	15	14
180	33	18	48	226	621	350	513	388	159	11	7	7	15
180	34	39	116	123	158	61	179	76	17	16	11	4	8
180	35	38	163	111	380	247	221	83	80	9	4	3	5
180	36	16	10	29	60	21	7	7	4	0	2	7	8
180	37	24	44	474	427	272	201	83	42	16	21	54	89
180	38	49	58	150	266	173	145	78	36	17	8	21	14
180	39	6	18	63	77	231	171	184	90	40	12	17	18
180	40	106	316	451	258	383	252	135	79	46	5	7	13
180	41	93	97	427	378	277	156	109	56	32	24	4	6
180	42	15	67	241	274	296	223	158	64	26	10	12	31
180	43	25	100	146	261	234	300	185	90	22	18	10	10
180	44	60	271	772	773	680	374	244	131	66	66	108	291
180	45	322	272	657	352	495	319	123	21	19	8	9	29
180	46	43	77	359	769	516	365	167	52	29	4	9	30
180	47	63	128	622	348	335	280	214	111	53	9	30	49
180	48	229	399	316	448	246	386	139	21	15	10	21	32
180	49	49	214	65	326	352	134	148	110	28	29	46	46
180	50	132	155	226	374	321	162	29	4	9	4	1	1
330	1	22	21	123	122	55	33	59	3	45	33	7	7
330	2	34	21	84	14	81	24	8	45	5	17	49	16
330	3	1	24	45	86	66	44	76	68	30	101	127	74
330	4	18	184	57	88	18	31	48	50	19	125	30	105
330	5	52	62	81	42	70	89	14	86	9	0	8	46
330	6	27	98	120	60	114	22	91	80	74	16	23	74
330	7	76	53	29	91	37	27	36	120	58	110	101	57
330	8	60	58	74	102	74	54	9	56	51	2	12	18
330	9	56	16	21	119	18	9	3	42	18	0	1	6
330	10	121	32	33	105	55	39	53	7	17	20	55	105
330	11	72	84	91	116	96	65	40	88	29	21	11	11
330	12	111	55	37	116	47	130	74	57	9	98	175	76
330	13	67	156	61	34	43	157	72	57	145	6	30	58
330	14	114	40	105	54	61	24	15	20	1	21	22	58
330	15	12	34	52	49	19	100	16	36	9	0	19	2
330	16	81	54	39	14	69	82	109	45	45	132	43	70
330	17	18	88	54	43	50	113	16	14	121	45	55	57
330	18	120	70	116	85	26	46	59	69	120	38	40	87
330	19	7	126	105	89	67	76	57	45	34	5	48	40
330	20	101	42	76	121	65	126	75	67	11	31	132	87
330	21	116	30	138	65	59	118	23	132	30	73	32	7
330	22	51	42	23	71	15	12	9	25	42	17	2	68
330	23	128	146	30	115	62	20	90	62	4	23	20	46
330	24	43	77	53	113	47	102	71	6	16	11	111	112
330	25	19	62	33	114	40	34	0	10	10	1	18	0
330	26	57	41	53	24	58	77	63	93	37	13	48	10
330	27	85	14	53	20	106	32	69	14	3	1	1	35
330	28	65	70	160	58	81	142	46	120	153	82	14	97
330	29	8	69	113	98	85	154	58	87	20	71	26	78
330	30	45	76	32	20	18	79	90	8	33	0	10	22
330	31	52	25	52	38	61	45	3	3	36	108	104	46
330	32	41	110	47	49	33	87	74	50	53	9	28	16
330	33	34	109	87	73	82	124	55	56	2	19	29	21
330	34	47	124	51	31	40	53	103	15	23	27	1	39
330	35	51	150	43	55	20	112	99	42	26	11	12	19
330	36	23	11	35	34	31	7	7	0	3	8	22	84
330	37	17	56	82	62	38	87	31	25	3	89	108	136
330	38	19	27	94	27	38	73	18	12	13	14	157	54
330	39	8	44	73	77	49	83	71	150	75	35	15	85
330	40	44	64	100	93	131	50	37	123	30	45	37	17
330	41	130	44	88	101	31	68	52	66	31	23	4	29
330	42	93	52	77	66	40	69	30	57	103	11	32	81
330	43	71	49	66	28	122	69	74	5	15	7	3	57
330	44	57	71	226	87	83	142	87	24	275	235	152	149
330	45	20	120	71	112	59	77	63	35	116	56	62	66
330	46	70	113	97	105	107	68	43	17	18	73	55	42
330	47	82	143	102	105	43	42	95	189	22	26	60	109
330	48	106	56	89	127	11	22	7	27	7	34	111	25
330	49	86	67	47	131	35	109	61	8	51	80	107	135
330	50	41	82	54	117	116	44	14	26	3	6	3	5
150	51	281	126	267	527	622	1265	590	355	182	105	92	75
150	52	109	204	805	274	750	923	528	238	161	171	198	88
150	53	126	204	380	641	1220	1149	669	376	209	175	100	113
150	54	133	204	261	151	282	1042	823	339	237	109	72	47
150	55	130	175	282	628	465	174	108	89	44	109	36	48
150	56	256	206	546	1097	1286	690	463	178	76	75	88	115
150	57	79	95	117	324	484	617	516	238	129	171	130	68
150	58	306	168	309	382	421	530	249	106	133	25	37	47
150	59	213	363	306	524	510	348	274	185	75	115	223	727
150	60	631	491	225	377	928	1134	639	368	141	37	30	43

FIGURE 9.1 (CONT'D).

150	61	136	181	397	395	635	1034	1220	543	319	203	229	344
150	62	203	246	155	435	680	956	723	436	148	78	83	132
150	63	108	97	101	273	227	381	220	72	26	26	54	77
150	64	129	223	199	295	790	559	403	235	54	48	25	47
150	65	73	119	89	434	442	241	74	130	100	100	40	47
150	66	106	517	456	307	328	263	79	51	57	69	36	195
150	67	470	1265	881	1615	774	878	540	192	198	66	56	89
150	68	197	207	294	1276	1123	1332	765	399	379	128	73	91
150	69	210	305	616	791	571	375	275	132	252	159	153	179
150	70	280	312	992	819	604	734	727	271	98	66	88	52
150	71	205	339	665	1230	1248	485	134	192	115	108	77	66
150	72	49	78	213	231	379	817	413	265	141	59	24	51
150	73	83	124	169	329	517	674	454	236	136	85	53	58
150	74	94	126	565	768	710	971	358	317	178	144	61	78
150	75	181	135	396	672	781	724	556	240	156	48	66	119
150	76	185	270	694	558	230	499	483	433	126	60	162	123
150	77	105	139	94	461	406	350	312	194	159	129	57	133
150	78	185	507	1114	809	660	908	636	441	188	121	86	75
150	79	95	315	911	323	462	494	749	363	155	135	213	535
150	80	1615	1552	625	1269	908	583	189	96	44	59	67	182
150	81	126	132	241	175	453	554	453	331	207	109	128	114
150	82	228	275	353	859	709	789	694	608	298	107	159	359
150	83	383	685	762	1451	1361	1501	670	380	283	137	134	206
150	84	106	155	341	217	316	699	189	136	90	86	144	80
150	85	58	116	165	301	327	404	182	101	80	28	67	114
150	86	92	139	138	224	365	419	318	189	131	39	269	520
150	87	228	226	598	722	438	373	93	47	31	11	7	28
150	88	34	51	136	110	141	245	186	95	84	68	58	75
150	89	450	974	220	281	330	659	394	152	67	60	58	76
150	90	91	73	109	274	555	516	264	85	86	33	144	310
150	91	130	93	118	167	418	237	189	134	44	23	35	66
150	92	186	118	228	198	289	176	134	58	37	50	68	85
150	93	248	306	479	365	495	1197	969	499	256	115	216	154
150	94	463	1007	2252	1330	1304	504	415	230	146	63	77	112
150	95	320	1022	589	977	780	471	279	54	46	41	45	77
150	96	186	272	526	663	638	607	542	266	74	19	32	46
150	97	124	492	859	820	853	1322	398	207	134	98	66	53
150	98	220	575	1353	1450	1142	1077	432	171	51	34	44	25
150	99	123	193	385	730	646	1444	1049	402	303	89	28	56
150	100	156	192	230	569	650	1149	860	467	217	250	222	160
180	51	21	25	55	239	284	467	208	86	29	23	13	11
180	52	16	65	248	162	155	231	118	53	48	37	58	36
180	53	26	57	167	289	479	481	214	106	40	14	13	35
180	54	44	35	98	81	121	377	182	79	51	21	40	27
180	55	47	73	184	290	143	66	9	17	20	4	7	6
180	56	104	122	233	521	618	177	186	44	14	1	18	19
180	57	23	85	50	142	209	237	88	41	26	29	20	4
180	58	136	150	196	212	172	127	50	11	20	5	2	2
180	59	50	172	200	466	261	109	95	40	10	8	48	174
180	60	225	167	129	145	324	590	163	86	20	4	1	1
180	61	29	51	103	170	269	291	405	137	63	37	93	67
180	62	123	152	108	264	275	282	310	226	63	21	6	7
180	63	17	9	32	110	42	136	45	6	15	1	4	27
180	64	42	99	64	137	301	85	114	53	10	12	1	1
180	65	11	37	43	220	148	66	13	30	26	10	5	5
180	66	32	101	186	115	123	47	32	5	9	13	3	12
180	67	95	326	521	569	290	229	107	29	40	16	22	37
180	68	23	114	131	932	477	344	194	102	60	29	8	13
180	69	99	132	463	408	200	111	58	22	50	16	19	42
180	70	75	111	555	389	255	136	222	74	11	5	13	8
180	71	34	210	274	569	555	165	46	55	11	15	4	10
180	72	7	22	127	126	166	273	96	65	19	19	4	3
180	73	6	31	50	123	139	160	82	34	25	22	5	6
180	74	17	27	272	441	378	245	78	70	38	28	4	11
180	75	19	15	128	320	281	290	115	44	15	4	7	20
180	76	145	195	390	224	105	86	137	102	29	20	57	92
180	77	30	80	51	224	184	70	48	24	20	7	6	8
180	78	22	139	532	434	206	262	190	135	76	27	26	36
180	79	29	99	454	170	147	53	220	102	43	15	28	88
180	80	800	1063	587	524	478	112	52	25	1	10	2	17
180	81	54	52	96	94	123	152	197	96	33	13	20	22
180	82	41	188	194	676	307	179	206	195	62	24	60	231
180	83	182	574	462	904	561	688	223	155	31	22	38	69
180	84	25	68	132	138	201	197	35	17	25	40	33	19
180	85	12	25	55	135	142	88	45	38	22	10	7	37
180	86	12	21	50	73	76	108	72	47	28	17	86	299
180	87	36	87	219	356	177	155	12	5	3	2	2	1
180	88	1	5	19	53	17	58	32	12	17	15	9	11
180	89	187	733	158	113	110	249	61	36	4	2	5	15
180	90	23	22	27	108	186	77	88	9	6	0	16	62

FIGURE 9.1 (CONT'D): SYNTHETIC DATA SETS FOR EXAMPLE PROBLEM

180	91	52	58	62	75	135	39	73	47	1	1	2	4
180	92	22	26	116	95	80	35	16	7	4	1	7	23
180	93	67	137	296	132	129	356	305	174	45	23	68	86
180	94	193	355	1361	889	608	92	69	15	20	7	3	15
180	95	71	244	339	509	426	196	59	7	2	2	1	4
180	96	29	240	344	382	246	142	92	64	16	6	2	5
180	97	21	72	399	583	458	568	128	75	35	7	8	16
180	98	27	158	764	1110	585	423	88	49	2	3	2	1
180	99	19	34	169	401	300	663	471	101	40	9	4	4
180	100	18	89	158	405	208	434	257	161	45	56	28	41
330	51	100	37	77	117	55	101	77	14	16	16	52	16
330	52	91	46	112	74	22	67	56	40	91	185	21	79
330	53	12	83	62	30	85	86	67	134	57	58	34	50
330	54	30	84	46	7	78	40	37	18	10	37	53	85
330	55	61	81	198	47	118	11	4	35	14	4	27	15
330	56	55	90	95	52	103	115	39	63	53	13	65	17
330	57	42	41	51	78	84	4	54	26	39	10	25	75
330	58	82	57	54	73	55	51	2	15	76	22	21	5
330	59	44	119	41	85	108	115	10	31	5	83	46	92
330	60	128	47	37	104	45	154	107	23	22	12	0	22
330	61	52	115	44	109	52	76	99	42	130	277	46	19
330	62	114	26	51	92	56	148	78	21	64	38	82	44
330	63	19	8	95	26	30	9	5	6	1	5	20	41
330	64	51	85	58	95	25	71	9	27	51	0	8	2
330	65	45	14	31	97	14	37	65	17	15	4	26	30
330	66	105	66	76	53	72	113	56	1	8	8	1	22
330	67	87	84	88	80	36	90	31	72	50	43	49	56
330	68	45	50	181	118	19	122	80	77	129	32	25	128
330	69	56	100	41	120	61	141	21	79	7	5	70	107
330	70	75	53	177	68	59	87	52	17	18	93	48	20
330	71	22	102	101	49	92	89	11	9	19	6	65	61
330	72	8	20	72	74	93	60	96	22	47	1	13	8
330	73	77	37	37	46	23	73	73	38	14	46	69	5
330	74	22	58	27	113	82	62	5	53	30	6	14	110
330	75	26	40	105	94	46	47	25	67	7	3	105	84
330	76	28	126	97	60	13	43	58	15	44	218	75	11
330	77	47	57	56	68	51	43	48	11	10	4	12	41
330	78	43	73	56	61	42	143	65	45	39	66	41	93
330	79	21	190	75	47	38	38	110	124	95	22	61	152
330	80	129	66	70	107	74	24	51	35	0	73	39	21
330	81	72	54	57	74	40	79	36	56	90	20	18	142
330	82	46	134	65	82	69	33	68	49	63	85	141	134
330	83	128	142	70	44	49	101	77	258	13	29	115	21
330	84	98	44	129	17	126	34	27	54	55	227	125	20
330	85	6	39	62	35	31	12	37	39	9	21	10	48
330	86	42	43	25	70	20	87	37	148	100	108	63	58
330	87	64	33	64	115	99	30	32	16	19	3	3	5
330	88	3	16	72	12	43	31	56	43	108	27	6	54
330	89	125	113	48	8	54	33	63	9	10	9	13	16
330	90	88	17	21	127	35	66	7	1	35	38	67	109
330	91	11	50	26	41	40	73	44	10	1	5	3	20
330	92	49	36	25	98	6	28	21	15	11	4	67	80
330	93	87	47	85	42	16	70	84	149	26	34	217	98
330	94	55	198	95	131	73	26	64	27	44	7	1	52
330	95	118	51	50	117	100	93	2	10	0	28	5	61
330	96	9	80	109	67	62	72	42	64	48	17	4	25
330	97	63	58	72	123	82	100	56	255	92	7	11	93
330	98	46	122	109	119	67	91	22	74	20	3	64	52
330	99	11	51	78	125	82	150	105	52	44	5	6	47
330	100	22	75	107	99	77	39	62	108	25	88	185	47
150	101	189	451	577	405	554	667	596	157	71	38	24	45
150	102	81	95	83	336	658	477	359	142	106	78	102	71
150	103	122	172	421	1051	568	1096	824	313	145	116	44	90
150	104	739	1301	1390	1871	959	1359	929	545	239	70	59	81
150	105	394	1623	1241	840	903	1029	782	231	122	120	91	76
150	106	106	186	497	873	1124	1453	752	296	147	74	19	38
150	107	33	52	53	239	593	727	324	349	257	163	277	138
150	108	136	263	152	213	457	579	248	84	57	61	60	91
150	109	302	432	278	984	1277	1422	394	149	129	110	41	30
150	110	117	137	143	295	240	400	432	195	78	40	102	75
150	111	119	284	591	888	1236	929	322	197	150	130	123	222
150	112	218	297	716	423	480	315	63	26	18	5	21	33
150	113	125	494	681	569	420	650	548	126	64	59	51	80
150	114	50	162	229	462	707	662	164	80	34	6	15	26
150	115	84	92	43	113	266	528	434	260	190	61	61	130
150	116	365	358	663	1162	798	1348	880	436	192	61	65	184
150	117	94	92	98	295	545	943	422	507	250	180	169	166
150	118	185	202	221	382	630	491	332	76	38	66	57	58
150	119	110	172	576	1035	592	410	417	187	113	94	91	134
150	120	393	455	623	587	913	990	835	393	347	158	92	118

FIGURE 9.1 (CONT'D): SYNTHETIC DATA SETS FOR EXAMPLE PROBLEM

150	121	151	165	240	453	466	159	40	51	27	46	85	60
150	122	254	324	384	192	416	663	309	158	106	82	125	179
150	123	107	126	319	948	741	852	419	159	112	125	76	99
150	124	165	212	707	396	364	300	336	295	304	109	132	208
150	125	527	747	608	799	860	858	795	646	146	103	92	109
150	126	102	101	245	339	739	676	438	217	127	90	27	41
150	127	260	177	886	809	1082	1319	322	196	122	147	110	103
150	128	115	202	604	883	998	662	406	223	143	158	187	175
150	129	107	175	290	180	190	668	504	181	108	127	154	254
150	130	262	296	221	341	554	852	644	540	224	76	155	122
150	131	209	325	681	709	676	942	1030	554	363	238	216	537
150	132	705	1178	609	971	677	379	304	125	38	56	95	48
150	133	216	331	359	255	469	618	502	201	172	86	36	39
150	134	109	129	452	1557	759	809	528	262	116	70	58	64
150	135	82	82	151	351	640	543	568	574	264	95	121	116
150	136	147	111	134	256	522	840	541	293	67	14	33	47
150	137	164	395	585	315	163	277	195	139	66	46	33	97
150	138	95	159	135	282	307	448	281	133	99	64	93	180
150	139	570	1028	450	1730	1238	909	821	322	170	55	19	52
150	140	107	171	413	488	391	574	408	196	113	59	28	61
150	141	58	94	379	357	190	202	47	140	127	55	73	72
150	142	144	85	103	118	278	411	365	129	42	77	88	141
150	143	218	142	306	204	303	184	91	84	81	99	287	898
150	144	400	641	714	1262	1211	770	305	242	152	108	74	67
150	145	139	230	214	409	960	1221	889	562	387	124	148	90
150	146	71	121	393	459	445	989	614	222	166	64	69	55
150	147	94	384	658	475	443	119	154	150	139	161	205	282
150	148	201	338	702	388	425	908	494	218	110	147	213	391
150	149	108	341	556	843	832	494	325	217	95	91	142	104
150	150	275	177	405	397	439	356	126	91	131	99	70	80
180	101	41	140	286	190	273	132	80	37	2	4	0	1
180	102	26	24	28	175	213	159	84	32	25	2	7	10
180	103	15	38	169	481	198	417	294	85	51	4	3	7
180	104	181	623	1061	1237	519	387	311	254	29	9	5	17
180	105	92	795	682	606	498	299	199	52	39	19	23	26
180	106	17	75	170	453	505	734	316	64	41	19	5	6
180	107	2	3	12	68	182	159	73	84	53	45	109	175
180	108	76	143	93	116	125	105	44	13	6	11	9	8
180	109	91	165	187	362	566	453	98	14	7	15	3	1
180	110	11	31	71	102	52	94	192	116	9	2	3	7
180	111	31	67	337	432	454	360	82	44	46	9	14	97
180	112	114	132	322	185	156	104	9	2	0	0	1	0
180	113	8	107	360	422	154	252	118	31	4	7	4	14
180	114	5	21	73	173	345	304	35	9	6	1	1	1
180	115	14	28	11	70	94	108	123	32	43	4	3	8
180	116	190	122	275	526	290	502	352	147	24	28	6	28
180	117	50	23	43	126	279	123	104	157	36	30	19	33
180	118	34	85	121	129	314	194	108	19	9	8	7	6
180	119	13	51	203	593	302	55	103	49	45	29	20	27
180	120	257	682	541	361	377	301	207	66	68	37	13	17
180	121	45	65	108	444	188	56	14	4	7	5	21	18
180	122	67	104	214	97	170	164	57	14	13	14	10	28
180	123	27	46	209	518	258	195	94	35	17	14	10	13
180	124	43	137	220	221	54	98	44	90	59	26	42	93
180	125	162	265	346	362	321	307	384	210	34	14	7	11
180	126	19	30	56	87	185	82	113	53	11	13	2	2
180	127	57	49	509	509	485	743	72	21	50	12	15	11
180	128	26	64	363	723	420	152	158	64	28	17	42	76
180	129	39	97	169	60	60	204	111	59	41	33	43	80
180	130	165	198	150	164	208	160	184	125	31	13	96	126
180	131	25	129	209	298	235	284	475	241	65	38	58	128
180	132	184	472	212	542	345	156	68	34	3	5	21	13
180	133	92	149	231	211	183	157	106	44	39	9	7	10
180	134	40	54	299	741	313	260	134	30	23	3	1	4
180	135	10	19	66	96	279	57	147	134	30	13	18	16
180	136	23	21	31	226	280	336	185	100	7	0	3	3
180	137	23	64	198	252	56	67	23	30	4	4	2	11
180	138	19	109	57	154	114	120	70	31	46	17	14	31
180	139	125	630	259	960	478	408	191	106	46	22	1	3
180	140	13	39	195	289	154	174	98	29	22	17	6	12
180	141	5	22	134	165	53	47	17	11	8	1	4	9
180	142	53	31	20	60	71	171	44	33	2	12	8	26
180	143	99	37	155	82	69	34	28	18	10	3	65	237
180	144	109	403	385	648	529	168	42	29	52	15	10	12
180	145	37	200	85	164	515	502	262	143	47	33	46	34
180	146	23	37	205	297	223	299	182	49	58	14	11	5
180	147	10	96	383	288	102	22	23	44	24	39	44	98
180	148	82	113	286	168	202	356	101	76	13	19	26	113
180	149	32	245	554	385	386	82	50	49	11	5	11	25
180	150	101	51	214	161	105	79	36	14	24	22	26	21

FIGURE 9.1 (CONT'D): SYNTHETIC DATA SETS FOR EXAMPLE PROBLEM

330	101	75	58	48	117	72	23	40	91	36	0	13	31
330	102	63	16	37	19	68	41	21	127	49	8	11	54
330	103	40	69	117	24	39	101	88	11	41	0	27	94
330	104	64	96	85	127	59	114	72	129	14	3	35	71
330	105	59	67	78	111	87	58	31	15	19	71	94	18
330	106	22	68	72	44	130	107	94	18	54	10	19	4
330	107	2	14	33	78	38	56	9	112	95	96	144	106
330	108	31	68	45	104	37	28	5	14	16	73	11	88
330	109	94	52	87	35	136	31	1	8	7	67	11	37
330	110	2	143	15	60	15	150	95	78	35	31	28	15
330	111	78	35	98	110	83	98	15	136	36	49	35	52
330	112	79	46	58	72	59	21	2	16	5	3	10	7
330	113	97	32	68	93	73	36	34	13	4	33	29	19
330	114	61	15	61	13	111	32	34	40	37	9	0	1
330	115	87	44	31	55	22	92	6	27	3	0	6	39
330	116	129	31	85	128	60	143	103	100	25	32	27	25
330	117	60	28	35	80	45	59	107	44	72	43	62	126
330	118	22	115	67	129	34	100	28	5	35	12	21	40
330	119	26	97	92	88	36	50	36	16	11	46	41	105
330	120	69	121	56	64	22	51	85	88	252	26	21	65
330	121	128	35	35	119	61	34	32	8	3	53	58	45
330	122	72	64	39	81	42	99	72	4	14	83	11	31
330	123	75	33	92	102	39	107	10	60	18	8	50	22
330	124	120	46	130	21	52	16	74	85	29	166	66	27
330	125	126	120	86	94	34	140	104	40	43	8	20	16
330	126	3	21	50	101	9	67	37	32	5	12	5	26
330	127	46	28	98	126	108	99	18	6	12	14	67	33
330	128	101	131	69	101	112	141	65	27	5	68	130	37
330	129	108	77	50	37	48	17	40	101	13	43	134	150
330	130	29	82	36	95	25	71	72	36	45	65	171	133
330	131	67	42	237	31	25	114	90	84	101	218	49	143
330	132	40	92	98	118	59	70	25	18	0	18	36	51
330	133	65	129	81	5	58	66	76	37	53	2	3	39
330	134	58	53	101	82	41	35	81	71	15	4	20	41
330	135	18	53	49	39	59	117	30	66	46	14	63	92
330	136	13	41	29	79	98	113	61	32	3	4	3	8
330	137	60	48	121	23	61	60	22	25	2	87	13	42
330	138	80	80	55	65	39	84	56	8	6	13	13	93
330	139	58	91	114	71	68	20	30	22	131	6	0	12
330	140	90	100	39	120	33	139	105	5	27	17	34	6
330	141	15	14	104	85	34	90	31	21	65	8	21	54
330	142	41	32	41	84	40	30	12	28	37	5	100	91
330	143	32	49	49	112	15	65	23	54	87	76	150	130
330	144	64	56	185	49	75	7	17	4	137	2	12	50
330	145	13	164	45	96	82	121	89	148	29	142	70	55
330	146	44	57	49	90	77	91	75	62	130	35	8	7
330	147	5	189	58	101	36	7	59	146	72	124	111	117
330	148	64	154	94	31	36	14	19	258	31	19	48	37
330	149	12	74	80	40	96	47	66	6	53	58	118	30
330	150	114	96	93	48	62	63	39	37	9	47	136	67
150	151	140	225	441	690	849	489	238	184	59	57	36	53
150	152	77	106	104	621	1079	944	819	389	175	86	61	62
150	153	197	266	662	475	591	563	709	325	112	72	80	135
150	154	201	237	875	939	931	1284	1026	365	203	91	154	156
150	155	141	240	636	552	705	379	212	149	95	85	121	165
150	156	93	118	101	250	759	1214	1146	267	217	139	282	595
150	157	712	1020	689	920	864	750	461	201	78	61	170	261
150	158	112	102	497	1142	818	1398	721	483	188	121	79	60
150	159	34	65	125	356	481	687	521	230	100	64	32	34
150	160	179	154	275	429	901	817	324	150	104	78	37	48
150	161	226	298	334	409	659	642	600	245	72	129	149	146
150	162	156	277	330	520	513	702	423	338	180	42	42	59
150	163	92	125	335	586	714	1368	267	193	194	55	52	49
150	164	304	403	344	694	780	1332	732	449	176	46	96	116
150	165	228	187	116	213	501	331	68	119	111	42	64	57
150	166	125	272	619	477	791	1186	401	352	240	51	102	169
150	167	161	293	340	503	596	487	588	216	156	150	120	124
150	168	241	204	229	276	109	142	165	340	197	172	113	103
150	169	45	52	55	127	341	599	205	238	177	117	219	250
150	170	624	1736	1335	1177	792	714	272	183	60	62	152	114
150	171	161	195	327	320	710	1544	966	382	125	66	52	44
150	172	147	235	317	411	335	619	617	291	188	207	81	71
150	173	73	106	181	362	238	839	373	71	51	34	57	76
150	174	298	264	392	661	486	594	722	428	118	49	38	43
150	175	172	420	719	448	595	802	520	279	123	109	101	195
150	176	190	165	187	399	398	283	123	158	83	79	82	151
150	177	105	354	630	592	849	471	187	53	52	85	66	82
150	178	93	164	244	259	217	350	120	63	39	52	150	597
150	179	507	1616	863	806	467	135	169	69	50	63	63	76

FIGURE 9.1 (CONT'D): SYNTHETIC DATA SETS FOR EXAMPLE PROBLEM

150	180	141	161	262	382	250	214	62	52	76	116	76	76
150	181	223	333	892	1310	760	837	407	99	33	19	36	50
150	182	172	151	410	400	733	540	394	249	289	221	139	98
150	183	102	107	174	246	255	248	87	31	16	7	15	45
150	184	155	1049	971	985	1106	991	615	432	264	169	128	126
150	185	275	461	1391	2313	1040	879	445	309	138	57	58	83
150	186	234	646	1017	1387	878	962	477	410	255	135	92	116
150	187	294	258	649	815	596	480	480	219	187	117	57	60
150	188	206	210	213	265	490	516	349	299	255	79	67	181
150	189	373	286	678	607	790	706	525	116	74	49	20	40
150	190	130	196	174	415	742	562	227	142	155	70	16	28
150	191	72	86	201	298	594	1142	787	559	215	133	156	132
150	192	66	122	108	225	333	703	533	192	160	164	256	321
150	193	357	393	649	2617	1168	1127	540	256	169	90	38	63
150	194	364	838	480	397	547	383	214	45	18	8	9	25
150	195	39	71	84	85	122	194	279	102	132	93	112	118
150	196	84	116	202	401	444	766	695	123	53	129	104	71
150	197	157	301	169	152	223	714	1024	521	294	169	268	554
150	198	212	340	826	779	1021	624	335	202	77	46	58	77
150	199	179	192	334	473	595	443	293	278	213	113	139	92
150	200	119	201	481	381	877	778	413	402	218	116	96	82
180	151	73	142	245	332	247	160	53	63	12	14	3	6
180	152	13	21	43	383	332	223	166	65	53	24	9	3
180	153	28	116	235	195	179	182	193	201	63	15	25	21
180	154	30	57	418	518	558	237	276	79	40	2	13	96
180	155	48	53	251	218	259	117	67	25	21	10	16	36
180	156	16	47	42	130	237	426	369	41	14	3	49	249
180	157	354	337	660	381	446	170	69	23	12	4	18	64
180	158	44	42	191	943	436	567	176	122	23	18	11	8
180	159	4	8	42	176	137	126	134	82	20	1	2	4
180	160	49	51	140	177	398	281	185	28	46	15	10	12
180	161	63	243	168	255	246	288	169	77	12	44	64	45
180	162	52	130	184	365	299	211	131	150	58	9	7	12
180	163	18	60	155	277	282	582	101	34	46	14	4	11
180	164	137	230	166	390	328	518	344	190	34	20	14	26
180	165	163	102	68	72	150	39	20	11	43	7	16	9
180	166	15	50	279	291	373	367	131	70	37	9	12	24
180	167	46	143	210	181	139	214	163	57	18	13	21	13
180	168	75	48	118	125	35	18	18	70	62	22	25	56
180	169	13	7	9	71	109	122	87	65	55	21	18	33
180	170	191	841	825	580	297	137	82	45	3	13	32	38
180	171	62	44	131	168	197	453	309	212	33	16	5	11
180	172	21	106	145	174	98	48	110	60	13	56	37	44
180	173	12	17	85	280	49	299	56	13	3	4	2	4
180	174	50	89	186	396	212	97	164	136	26	4	3	4
180	175	42	264	331	187	239	311	135	89	35	19	15	68
180	176	25	65	91	251	153	130	19	25	9	6	8	48
180	177	23	127	255	259	296	90	44	7	22	2	4	16
180	178	35	41	109	176	102	90	34	18	16	8	21	113
180	179	429	521	448	482	200	34	20	5	7	12	8	7
180	180	33	32	113	148	151	87	6	16	13	9	10	12
180	181	36	213	603	682	393	321	77	10	9	3	6	3
180	182	55	40	228	224	290	93	82	46	36	29	24	33
180	183	16	31	59	93	48	69	17	4	2	4	0	1
180	184	16	436	529	416	427	486	304	116	65	32	27	21
180	185	28	128	730	1165	533	293	93	81	50	6	2	5
180	186	64	138	452	940	380	387	187	97	49	16	8	14
180	187	50	107	351	445	263	142	158	77	26	28	7	14
180	188	29	140	84	143	132	190	73	55	14	8	16	59
180	189	97	112	278	220	375	236	220	22	5	2	4	4
180	190	12	101	63	195	319	124	83	18	23	5	0	0
180	191	5	25	69	140	133	554	261	113	61	26	24	50
180	192	16	36	42	92	145	222	97	17	18	40	149	224
180	193	214	444	573	1638	752	405	102	39	12	5	1	9
180	194	111	460	321	185	347	38	35	10	0	0	0	0
180	195	1	5	17	36	19	32	21	7	13	9	11	40
180	196	27	62	82	207	132	196	160	27	10	4	10	5
180	197	30	71	80	44	97	172	295	140	24	47	40	59
180	198	142	140	506	403	401	237	59	72	4	1	2	4
180	199	60	98	121	279	276	134	91	56	51	25	32	22
180	200	19	62	243	281	255	200	116	116	38	29	32	18
330	151	34	61	45	51	92	50	46	36	9	78	20	65
330	152	8	30	44	101	76	41	96	82	20	3	1	43
330	153	77	63	58	12	66	94	93	137	58	19	44	99
330	154	26	113	88	129	73	68	97	18	2	19	37	12
330	155	89	28	72	78	111	106	16	4	57	158	109	84
330	156	37	22	51	110	56	134	65	42	84	32	150	126
330	157	80	94	92	83	52	22	39	8	1	10	40	88
330	158	24	27	85	123	49	49	70	45	4	55	37	3
330	159	10	44	36	93	7	122	77	14	8	32	7	8

FIGURE 9.1 (CONT'D): SYNTHETIC DATA SETS FOR EXAMPLE PROBLEM

330	160	81	48	89	41	64	118	65	10	84	43	64	31
330	161	97	73	28	89	122	131	95	27	62	174	107	74
330	162	14	105	53	80	86	81	10	76	109	3	28	25
330	163	25	30	95	77	85	103	33	86	44	1	45	57
330	164	109	135	60	74	81	145	9	47	41	10	72	96
330	165	91	72	33	88	45	112	29	64	89	15	3	14
330	166	14	59	54	115	59	64	30	20	30	3	12	27
330	167	97	129	76	43	35	86	20	190	37	24	36	71
330	168	84	50	18	111	14	12	16	69	144	18	143	136
330	169	2	13	20	39	35	105	29	101	11	113	33	104
330	170	91	161	132	84	49	49	70	6	12	92	25	51
330	171	87	30	45	68	48	62	109	93	11	57	5	21
330	172	38	47	57	128	32	20	101	115	59	51	23	29
330	173	8	69	65	108	88	84	22	32	121	14	18	94
330	174	53	69	116	14	23	85	75	156	35	0	2	17
330	175	121	157	63	47	96	79	105	35	2	2	47	64
330	176	52	79	23	111	27	31	15	9	7	30	98	38
330	177	10	92	80	65	48	43	6	27	1	26	56	15
330	178	66	59	74	57	23	36	9	20	11	12	151	103
330	179	113	135	77	85	80	20	62	13	29	30	16	99
330	180	36	53	65	36	65	10	61	16	15	2	95	57
330	181	75	78	97	55	52	60	4	5	4	62	16	21
330	182	109	33	63	112	27	25	55	18	6	35	174	28
330	183	17	30	79	60	41	21	3	14	22	24	0	4
330	184	115	116	50	107	134	92	64	127	175	42	38	65
330	185	62	54	286	113	54	27	63	25	34	164	12	32
330	186	46	85	122	104	72	70	13	31	35	20	17	94
330	187	26	122	105	111	29	88	71	154	103	16	61	57
330	188	31	103	50	92	34	25	14	100	35	79	30	14
330	189	125	57	80	86	90	146	48	16	12	4	1	58
330	190	40	39	40	125	75	111	39	92	19	1	55	0
330	191	14	14	62	59	61	136	71	14	169	51	26	36
330	192	66	43	86	22	36	51	32	18	42	159	73	121
330	193	118	141	92	84	78	16	67	23	60	1	12	34
330	194	119	107	134	100	57	41	92	0	4	1	0	1
330	195	2	17	57	0	4	15	28	22	24	22	70	110
330	196	35	40	35	25	50	105	7	29	28	15	71	28
330	197	60	17	67	98	51	125	11	81	10	136	61	123
330	198	92	58	135	22	67	15	37	22	21	3	8	25
330	199	81	38	54	92	74	70	29	43	37	10	56	85
330	200	54	127	90	53	18	136	96	172	22	97	51	105
150	201	195	454	559	434	259	266	128	124	26	21	52	63
150	202	235	251	287	254	228	792	625	209	173	162	130	101
150	203	298	246	177	220	523	643	283	114	95	71	28	42
150	204	138	632	625	629	798	1238	609	381	141	74	46	67
150	205	150	121	52	205	393	483	285	186	146	110	91	150
150	206	63	70	57	194	512	947	708	422	149	49	40	47
150	207	72	141	637	785	565	910	808	217	67	19	42	56
150	208	83	168	840	508	567	399	59	29	22	68	158	159
150	209	315	185	157	118	150	361	451	190	153	123	183	102
150	210	148	678	622	958	867	1316	1170	483	134	93	42	109
150	211	74	151	235	342	285	276	149	115	81	21	32	41
150	212	54	93	174	142	97	128	270	114	40	20	24	38
150	213	80	92	329	416	659	697	215	138	69	48	31	31
150	214	348	793	526	728	891	947	506	426	185	174	108	112
150	215	137	304	229	504	642	958	401	84	115	56	70	171
150	216	193	355	503	843	672	744	756	259	148	77	42	63
150	217	64	84	218	278	468	383	384	286	138	175	140	87
150	218	109	193	224	220	525	486	586	142	196	134	116	174
150	219	548	946	485	849	1082	959	593	472	268	123	83	55
150	220	91	111	121	138	183	101	94	48	60	50	36	91
150	221	103	75	187	438	835	1011	530	373	253	143	114	58
150	222	135	206	724	552	571	717	576	252	99	110	156	149
150	223	168	291	770	787	632	653	812	469	146	60	18	51
150	224	415	457	806	914	899	788	905	580	255	143	75	241
150	225	374	426	252	586	659	731	386	265	146	85	213	561
150	226	533	461	468	1177	1149	1493	1064	424	181	106	143	164
150	227	332	424	662	468	701	776	403	377	306	91	118	136
150	228	341	663	536	948	625	792	356	334	324	193	377	1580
150	229	142	164	361	373	632	823	377	99	33	34	93	77
150	230	191	327	191	287	374	653	274	112	85	80	114	166
150	231	93	178	620	610	639	991	484	267	34	18	23	38
150	232	115	179	489	370	821	895	912	280	148	99	112	162
150	233	235	429	352	255	328	201	73	88	70	28	29	76
150	234	58	92	135	328	547	361	309	178	62	52	51	65
150	235	212	139	175	1326	1375	1023	665	329	270	180	124	112
150	236	133	259	727	1616	944	1009	662	399	284	164	128	139
150	237	885	783	492	630	888	642	357	185	89	37	62	48
150	238	69	61	73	229	323	196	57	20	25	48	74	137
150	239	185	296	810	577	780	393	270	168	170	114	121	64

FIGURE 9.1 (CONT'D): SYNTHETIC DATA SETS FOR EXAMPLE PROBLEM

150	240	98	139	276	283	522	927	488	252	118	83	120	76
150	241	111	279	338	704	740	905	482	266	156	137	90	52
150	242	150	119	567	778	676	399	110	115	74	49	16	44
150	243	97	201	380	330	430	412	223	100	110	105	104	92
150	244	142	801	2206	1492	1247	1336	559	392	163	112	196	134
150	245	166	182	167	517	675	1045	502	193	96	126	150	328
150	246	379	856	1186	793	615	961	438	331	209	143	117	125
150	247	156	136	443	1289	746	377	240	140	100	100	130	563
150	248	427	399	299	369	600	692	334	288	215	89	52	48
150	249	52	81	163	315	289	423	538	312	271	155	164	128
150	250	247	248	813	2065	998	840	399	203	134	31	21	41
180	201	30	138	296	228	103	69	42	42	1	2	3	6
180	202	36	97	126	124	71	253	209	32	22	25	28	28
180	203	130	82	133	125	147	175	46	26	24	9	3	2
180	204	28	68	220	250	383	383	179	132	29	9	4	5
180	205	29	40	10	70	114	79	67	30	23	13	15	41
180	206	10	9	10	98	140	227	244	112	10	2	2	1
180	207	10	37	223	362	180	353	231	73	8	2	4	4
180	208	20	52	374	354	325	90	19	3	1	13	21	33
180	209	78	202	113	79	57	112	76	29	51	19	51	22
180	210	38	182	347	532	415	442	487	141	26	10	7	19
180	211	14	55	138	142	91	51	73	14	9	3	3	4
180	212	6	11	33	47	17	58	52	30	3	2	2	3
180	213	17	26	93	213	307	230	60	17	28	13	5	3
180	214	103	140	262	257	329	170	97	96	35	40	20	32
180	215	31	126	77	310	389	262	136	14	10	8	3	14
180	216	74	122	315	487	302	245	133	47	45	21	6	12
180	217	3	22	104	131	144	73	89	85	15	34	40	22
180	218	37	189	92	91	159	192	99	10	34	28	21	23
180	219	175	319	320	459	454	299	98	127	62	23	31	14
180	220	15	27	54	73	64	4	6	5	33	3	1	6
180	221	18	27	72	288	284	400	123	79	50	26	16	18
180	222	32	205	293	212	197	103	179	108	28	20	27	99
180	223	52	83	307	513	216	147	178	192	48	5	4	4
180	224	34	317	577	557	466	239	394	275	67	10	2	17
180	225	223	409	176	389	368	228	135	44	40	16	43	201
180	226	383	147	274	741	496	451	277	87	50	28	19	19
180	227	96	268	528	175	187	305	126	79	36	30	44	106
180	228	48	204	420	445	225	312	100	73	43	32	80	282
180	229	45	39	132	186	207	207	180	17	1	2	15	16
180	230	51	112	115	96	115	199	51	47	12	17	32	90
180	231	22	73	348	201	253	484	119	86	1	1	1	1
180	232	7	24	128	212	403	304	204	56	24	12	18	38
180	233	70	597	147	159	109	81	13	12	13	0	0	7
180	234	6	9	60	180	293	42	68	61	4	1	4	12
180	235	49	57	68	1347	723	248	162	90	66	38	25	29
180	236	67	56	322	765	454	247	158	87	49	32	34	79
180	237	340	709	339	430	380	166	79	46	28	10	15	17
180	238	13	13	31	138	148	22	11	1	2	6	4	7
180	239	52	56	314	258	263	125	20	23	34	27	27	8
180	240	12	47	150	145	142	395	124	65	20	13	25	23
180	241	46	243	224	374	260	355	159	51	20	3	8	7
180	242	35	44	339	297	289	81	17	33	13	1	1	2
180	243	11	109	125	99	126	99	79	20	19	13	11	7
180	244	28	326	1537	783	522	576	312	121	32	22	33	85
180	245	55	91	65	224	347	389	114	24	7	24	27	118
180	246	212	198	617	337	182	256	160	95	64	25	22	32
180	247	48	63	192	966	401	108	29	28	21	19	21	139
180	248	330	279	155	229	280	194	106	91	50	16	9	5
180	249	9	15	71	139	50	125	163	48	38	19	25	35
180	250	52	68	297	1049	341	427	73	65	19	1	0	1
330	201	62	53	27	72	54	40	62	23	1	13	11	76
330	202	41	52	40	61	57	117	73	34	31	145	74	50
330	203	99	39	55	38	38	33	10	36	18	0	9	30
330	204	29	51	60	124	30	67	5	38	49	5	17	13
330	205	115	11	66	73	28	53	11	45	8	33	87	20
330	206	46	37	67	14	79	144	18	20	10	1	3	16
330	207	59	28	110	79	66	97	86	44	4	10	13	80
330	208	7	95	104	116	61	3	36	14	6	79	70	73
330	209	69	168	25	91	13	47	106	6	40	19	39	66
330	210	99	64	102	68	87	121	52	43	7	20	19	70
330	211	15	39	44	28	81	48	10	18	9	3	21	100
330	212	5	41	44	61	6	63	41	5	3	3	33	12
330	213	37	39	124	99	59	123	39	47	109	55	62	45
330	214	55	69	28	18	70	41	46	58	22	29	105	98
330	215	40	130	52	93	15	105	65	82	63	32	4	128

FIGURE 9.1 (CONT'D): SYNTHETIC DATA SETS FOR EXAMPLE PROBLEM

330	216	41	136	86	57	67	60	70	88	49	53	5	0
330	217	52	35	27	78	53	50	57	132	55	36	129	59
330	218	26	61	53	41	33	33	5	68	32	40	29	49
330	219	87	94	115	89	57	90	21	46	20	72	21	128
330	220	2	86	26	14	9	4	1	26	16	1	25	26
330	221	43	92	113	113	71	46	21	191	39	146	17	41
330	222	62	96	69	123	22	85	63	35	10	59	69	112
330	223	19	59	90	109	41	24	89	8	43	11	1	13
330	224	102	124	159	74	77	153	105	142	103	7	46	81
330	225	121	74	40	130	32	102	20	45	25	98	41	127
330	226	120	54	67	118	107	103	103	43	149	20	16	89
330	227	52	113	57	85	45	85	55	106	60	58	35	94
330	228	58	129	61	88	107	34	11	95	142	138	233	124
330	229	6	71	103	20	34	121	24	2	7	19	7	89
330	230	96	11	38	60	20	83	16	106	36	26	28	32
330	231	13	91	89	98	79	98	53	3	12	11	1	12
330	232	11	26	30	7	129	42	94	9	31	96	86	74
330	233	52	105	93	56	44	82	6	4	53	0	5	72
330	234	4	60	26	99	45	24	64	33	45	6	5	33
330	235	88	23	96	126	98	60	33	197	138	55	88	63
330	236	102	98	119	80	72	70	82	46	3	29	51	69
330	237	128	107	47	114	30	38	60	26	84	28	48	21
330	238	64	6	39	30	23	8	21	4	1	40	30	42
330	239	121	93	91	62	98	51	7	56	11	115	29	6
330	240	56	61	59	69	87	79	85	17	13	5	123	13
330	241	101	92	160	95	57	103	63	26	18	8	31	27
330	242	29	104	74	97	72	32	50	1	13	0	3	28
330	243	83	31	80	16	38	105	83	60	22	6	41	16
330	244	93	115	120	122	99	159	96	78	19	46	166	29
330	245	67	34	52	109	90	66	5	14	2	84	61	92
330	246	94	57	73	64	22	87	72	36	175	74	25	16
330	247	46	39	87	123	73	10	42	59	8	18	130	125
330	248	97	43	49	110	81	73	51	127	79	47	6	3
330	249	32	48	62	35	53	56	93	60	114	100	130	101
330	250	8	48	166	54	94	110	87	59	71	0	6	1

FIGURE 9.1 (CONT'D): SYNTHETIC DATA SETS FOR EXAMPLE PROBLEM

```
FILE: ALLOCATE JCL1     A1  CMS REL3 PLC12 CANBERRA 370/158

//CSALNDNR JOB (1009,10,001,06,2,52467,C0000),LINDNER,
// PROFILE='TYPE=RJE,TAPE=0,D2314=1',
// MSGLEVEL=(1,1),MSGCLASS=Z
//*
//MESSAGE EXEC PGM=MESSAGE
//UTIN DD *
.....PLS MOUNT 2316 PACK 'MIKES1'.....
/*
//*
//STEP EXEC PGM=IEFBR14
//FT20F001 DD DSN=BEARD.MEANS.MONTHLY,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(2,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT21F001 DD DSN=BEARD.STDDEV.MONTHLY,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(2,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT22F001 DD DSN=BEARD.SKEW.MONTHLY,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(2,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT23F001 DD DSN=BEARD.EXT.MONTHLY,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(4,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT24F001 DD DSN=BEARD.MOMENTS.ANNUAL,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(1,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT25F001 DD DSN=BEARD.SERIES.MOMENTS,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(1,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT26F001 DD DSN=BEARD.SERIES.FREQ,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(4,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT27F001 DD DSN=BEARD.RUNS.MEDIAN,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(9,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT28F001 DD DSN=BEARD.SERIES.VOLUMES,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(8,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT29F001 DD DSN=BEARD.SERIES.RANGE,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(1,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT30F001 DD DSN=BEARD.YIELD.STORAGE,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(1,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT31F001 DD DSN=BEARD.AUTO.CORREL,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(2,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT32F001 DD DSN=BEARD.CROSS.CORREL,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(5,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT33F001 DD DSN=BEARD.YIELD.DEFICIT,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(4,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT34F001 DD DSN=BEARD.YIELD.DROUGHT,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(4,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT35F001 DD DSN=BEARD.YIELD.DRAW,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
//       SPACE=(CYL,(4,1,10)),
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT36F001 DD DSN=BEARD.YIELD.FILL,UNIT=DISK14,
//       DCB=(RECFM=VBS,BLKSIZE=7294,DSORG=PO),
```

FIGURE 9.2: JOB CONTROL LANGUAGE FILE TO ALLOCATE PERMANENT SYNTHETIC DATA STATISTIC FILES

```
//         SPACE=(CYL,(4,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT80F001 DD DSN=BRDPOP.MOM.MEAN,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT81F001 DD DSN=BRDPOP.MOM.STDDEV,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT82F001 DD DSN=BRDPOP.MOM.SKEW,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT83F001 DD DSN=BRDPOP.SERIES.FREQ,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT84F001 DD DSN=BRDPOP.RUNS.UP,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT85F001 DD DSN=BRDPOP.RUNS.DOWN,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT86F001 DD DSN=BRDPOP.RUNS.TOTAL,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT87F001 DD DSN=BRDPOP.YIELD.DEFICIT,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT88F001 DD DSN=BRDPOP.YIELD.DROUGHT,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT89F001 DD DSN=BRDPOP.YIELD.DRAW,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT90F001 DD DSN=BRDPOP.YIELD.FILL,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT91F001 DD DSN=BRDPOP.RIPPL.STORAGE,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT92F001 DD DSN=BRDPOP.AUTO.CORREL,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//FT93F001 DD DSN=BRDPOP.CROSS.CORREL,UNIT=DISK14,
//         DCB=(RECFM=FB,LRECL=80,BLKSIZE=7200,DSORG=PO),
//         SPACE=(CYL,(1,1,10)),
//         VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(NEW,KEEP)
//
```

FIGURE 9.2 (CONT'D): JOB CONTROL LANGUAGE FILE TO ALLOCATE PERMANENT SYNTHETIC DATA STATISTIC FILES

```
FILE: ANALYSES JCL1      A    CMS REL3 PLC12 CANBERRA 370/158

//CSALNDNR JOB (1009,10,001,06,2,52467,00000),LINDNER,
// PROFILE='TYPE=RJE,TAPE=1,D2314=1',
// MSGLEVEL=(1,1),MSGCLASS=Z
//*
//MESSAGE EXEC PGM=MESSAGE
//UTIN DD *
.....PLS MOUNT 1. 2316 PACK MIKES1   AND   ....
.....          2. TAPE UCC912 WITH RING OUT.....
/*
//*
// EXEC FTG1CLG,PARM.FORT='MAP,ID',TIME=5
//FORT.SYSIN DD *
```

FIGURE 9.3: JOB CONTROL LANGUAGE SUBSET 1 FOR PROGRAM STATS

FILE: ANALYSES JCL2   A   CMS REL3 PLC12 CANBERRA 370/158

```
/*
//*  USE NEW.FORTRAN LIBRARY
//LKED.SYSLIB DD DSN=PPFORT.FORTLIB,DISP=SHR
//*  DD TO SUPERESS ALL PRINTER OUTPUT EXCEPT JCL
//*GO.FT06F001 DD DUMMY,SYSOUT=
//*             MONTHLY MOMENTS (1,2,3) AND EXTREMA
//GO.FT01F001 DD UNIT=SYSDA,SPACE=(TRK,(6,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*             ANNUAL MOMENTS (1,2,3)
//GO.FT02F001 DD UNIT=SYSDA,SPACE=(TRK,(1,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*             TIME SERIES MOMENTS (1,2,3)
//GO.FT03F001 DD UNIT=SYSDA,SPACE=(TRK,(1,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*             TIME SERIES FREQUENCY DISTRIBUTION
//GO.FT04F001 DD UNIT=SYSDA,SPACE=(TRK,(1,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*             RUNS UP AND DOWN
//GO.FT09F001 DD UNIT=SYSDA,SPACE=(TRK,(3,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*             RIPPL STORAGE  DURATION  BOUNDS  TOTAL OBS  AND  MEANS
//GO.FT10F001 DD UNIT=SYSDA,SPACE=(TRK,(9,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*             DEFICIT FREQ DISTRIBUTION
//GO.FT11F001 DD UNIT=SYSDA,SPACE=(TRK,(3,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*             DURATION FREQ DISTRIBUTION
//GO.FT12F001 DD UNIT=SYSDA,SPACE=(TRK,(3,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*             DRAW FREQ DISTRIBUTION
//GO.FT13F001 DD UNIT=SYSDA,SPACE=(TRK,(3,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*             FILL FREQ DISTRIBUTION
//GO.FT14F001 DD UNIT=SYSDA,SPACE=(TRK,(3,1)),
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*               AUTO AND CROSS CORRELATIONS
//GO.FT15F001 DD UNIT=SYSDA,SPACE=(TRK,(20,2)),
//       DCB=(RECFM=FB,LRECL=80,BLKSIZE=6400)
//*
//GO.FT50F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(TRK,(2,2))
//GO.FT51F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(TRK,(2,2))
//GO.FT52F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(TRK,(2,2))
//GO.FT53F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT54F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT55F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT56F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT57F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT58F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT59F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT60F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(TRK,(2,2))
//GO.FT61F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(TRK,(2,2))
//GO.FT62F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(TRK,(2,2))
//GO.FT63F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT64F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT65F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT66F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT67F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT68F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
```

FIGURE 9.4: JOB CONTROL LANGUAGE SUBSET 2 FOR PROGRAM STATS

```
//GO.FT69F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT70F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(TRK,(2,2))
//GO.FT71F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(TRK,(2,2))
//GO.FT72F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(TRK,(2,2))
//GO.FT73F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//GO.FT74F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//     SPACE=(TRK,(2,2))
//*
//*
//GO.FT08F001 DD DSN=DEMO1,LABEL=(16,SL,,IN),UNIT=TAPE,VOL=SER=UCC912,
//         DCB=(DEN=3,RECFM=FB,LRECL=80,BLKSIZE=7200,BUFNO=3),
//         DISP=(OLD,KEEP)
//*
//*
//GO.FT20F001 DD DSN=BEARD.MEANS.MONTHLY(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT21F001 DD DSN=BEARD.STDDEV.MONTHLY(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT22F001 DD DSN=BEARD.SKEW.MONTHLY(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT23F001 DD DSN=BEARD.EXT.MONTHLY(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT24F001 DD DSN=BEARD.MOMENTS.ANNUAL(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT25F001 DD DSN=BEARD.SERIES.MOMENTS(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT26F001 DD DSN=BEARD.SERIES.FREQ(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT27F001 DD DSN=BEARD.RUNS.MEDIAN(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT28F001 DD DSN=BEARD.SERIES.VOLUMES(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT29F001 DD DSN=BEARD.SERIES.RANGE(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT30F001 DD DSN=BEARD.YIELD.STORAGE(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT31F001 DD DSN=BEARD.AUTO.CORREL(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT32F001 DD DSN=BEARD.CROSS.CORREL(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT33F001 DD DSN=BEARD.YIELD.DEFICIT(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT34F001 DD DSN=BEARD.YIELD.DROUGHT(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT35F001 DD DSN=BEARD.YIELD.DRAW(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT36F001 DD DSN=BEARD.YIELD.FILL(DEMO1),UNIT=DISK14,
//       VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.SYSIN DD *
```

FIGURE 9.4 (CONT'D): JOB CONTROL LANGUAGE SUBSET 2 FOR PROGRAM STATS

```
FILE: COMPARE  JCL1      A1  CMS REL3 PLC12 CANBERRA 370/158

//CSALNDNR JOB (1009,10,001,06,2,52467,C0000),LINDNER,
// PROFILE='TYPE=RJE,TAPE=0,D2314=1',
// MSGLEVEL=(1,1),MSGCLASS=Z
//*
//MESSAGE EXEC PGM=MESSAGE
//OTIN DD *
.....PLS MOUNT 2316 PACK MIKES1.....
/*
//*
// EXEC FTG1CLG,PARM.FORT='MAP,ID',TIME=5
//FORT.SYSIN DD *
```

FIGURE 9.5: JOB CONTROL LANGUAGE SUBSET 1 FOR STATISTIC COMPARISON PROGRAMS

FILE: COMPARE JCL2 A CMS REL3 PLC12 CANBERRA 370/158

```
/*
//*  USE NEW FORTRAN COMPILER LIBRARY
//LKED.SYSLIB DD DSN=PPFORT.FORTLIB,DISP=SHR
//*  WORK FILES FOR CORRELATION
//GO.FT01F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(CYL,(10,2))
//GO.FT02F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(CYL,(10,2))
//GO.FT03F001 DD UNIT=SYSDA,DCB=(RECFM=VBS,BLKSIZE=6447),
//    SPACE=(CYL,(1,1))
//GO.FT20F001 DD DSN=BEARD.MEANS.MONTHLY(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT21F001 DD DSN=BEARD.STDDEV.MONTHLY(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT22F001 DD DSN=BEARD.SKEW.MONTHLY(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT23F001 DD DSN=BEARD.EXT.MONTHLY(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT24F001 DD DSN=BEARD.MOMENTS.ANNUAL(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT25F001 DD DSN=BEARD.SERIES.MOMENTS(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT26F001 DD DSN=BEARD.SERIES.FREQ(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT27F001 DD DSN=BEARD.RUNS.MEDIAN(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT28F001 DD DSN=BEARD.SERIES.VOLUMES(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT29F001 DD DSN=BEARD.SERIES.RANGE(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT30F001 DD DSN=BEARD.YIELD.STORAGE(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT31F001 DD DSN=BEARD.AUTO.CORREL(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT32F001 DD DSN=BEARD.CROSS.CORREL(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT33F001 DD DSN=BEARD.YIELD.DEFICIT(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT34F001 DD DSN=BEARD.YIELD.DROUGHT(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT35F001 DD DSN=BEARD.YIELD.DRAW(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT36F001 DD DSN=BEARD.YIELD.FILL(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,IN),DISP=(OLD,KEEP)
//GO.FT80F001 DD DSN=BRDPOP.MOM.MEAN(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT81F001 DD DSN=BRDPOP.MOM.STDDEV(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT82F001 DD DSN=BRDPOP.MOM.SKEW(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT83F001 DD DSN=BRDPOP.SERIES.FREQ(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT84F001 DD DSN=BRDPOP.RUNS.UP(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT85F001 DD DSN=BRDPOP.RUNS.DOWN(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT86F001 DD DSN=BRDPOP.RUNS.TOTAL(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT87F001 DD DSN=BRDPOP.YIELD.DEFICIT(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT88F001 DD DSN=BRDPOP.YIELD.DROUGHT(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT89F001 DD DSN=BRDPOP.YIELD.DRAW(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT90F001 DD DSN=BRDPOP.YIELD.FILL(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT91F001 DD DSN=BRDPOP.RIPPL.STORAGE(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT92F001 DD DSN=BRDPOP.AUTO.CORREL(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.FT93F001 DD DSN=BRDPOP.CROSS.CORREL(DEMO1),UNIT=DISK14,
//        VOL=(PRIVATE,SER=MIKES1),LABEL=(,,,OUT),DISP=(OLD,KEEP)
//GO.SYSIN DD *
```

FIGURE 9.6: JOB CONTROL LANGUAGE SUBSET 2 FOR STATISTIC COMPARISON PROGRAMS

```
FILE: HISTDATA ANALYSIS A1  CMS EFL3 PLC12 CANBERRA 370/158

  1  1
401  1  5
 ( 8X,12F6.0 )
  3 50
  1
  1  0  0  1  1
  1  0  1
  1 12  0
  1  1  0  0  0  1  1
 0.0        2160.0     0.0        1080.0     0.0        270.0
  1 30  0  0  0  1
  1 60  0  0  0
  1  0  0
  1  2  0  0  0  1
 0.6        0.8
  2  6  4  7
  3  7  6 11
  1  2  3  4
  1 12  0  0  1
  1 12  1  0  1
150180330
H1501922  124.   92.  485.  387.  459.  455.  203.   93.   73.   36.   27.   26.
H1501923  153.  689.  640.  531.  615. 1089.  615.  394.  307.  258.  170.  184.
H1501924  172.  288.  225.  814.  856.  824.  959.  396.  332.  228.  153.  111.
H1501925  184.  203.  266.  625.  553.  601.  295.  119.   81.   38.   70.  145.
H1501926  449.  498.  683.  819.  679.  732.  305.  177.  141.   59.   57.   50.
H1501927  109.  113.  273.  486.  385.  763.  301.  130.   70.  129.  209.  234.
H1501928  252.  375.  369.  271.  363.  658.  335.  128.   59.   53.   47.  101.
H1501929  119.  202.  169.  307.  418.  467.  226.  165.   95.   39.   31.   38.
H1501930  101.  119.  153.  363.  335.  861.  375.  360.  242.   98.  210.  197.
H1501931  523. 1692. 1168.  945.  830.  603.  387.  186.   74.   64.  128.  271.
H1501932  145.  375.  482.  656. 1045.  520.  311.  151.   80.   44.   39.   48.
H1501933   96.  155.  397.  269.  910.  523.  198.  325.  267.  113.   92.  155.
H1501934  123.  101.  417.  646.  590. 1414.  929.  738.  245.  165.  114.  246.
H1501935  295.  258.  430.  824.  726.  689.  367.  215.  155.   90.  107.  118.
H1501936  100.  221.  534. 1396.  667.  503.  283.  292.  213.  102.   93.   66.
H1501937   97.   97.  102.  154.  601.  394.  180.  103.   68.   52.   47.   55.
H1501938   75.  109.  132.  182.  274.  177.   74.   32.   21.   66.  384.  310.
H1501939  228.  485.  624. 1144.  972. 1074.  836.  283.  125.   59.   43.  109.
H1501940  139.  143.  128.  194.  247.  157.   87.   87.  247.   77.  106.   65.
H1501941   61.  107.  277.  171.  304.  390.  166.   93.   39.   38.   33.   33.
H1501942  351.  424.  996.  667.  978.  788.  461.  197.  135.   68.   54.  221.
H1501943  159.  150.  299.  379.  590.  824.  379.  166.   69.   37.   41.   59.
H1501944  230.  154.  186.  132.  128.  156.   96.   59.   41.   44.   23.   50.
H1501945   48.  156.  128.  470.  427.  411.  305.  101.   53.  123.  192.  130.
H1501946  121.  230. 1140.  937.  513.  615.  421.  197.   89.   66.  122.  109.
H1501947  100.  226.  691.  689.  925.  916.  640.  387.  251.  146.   91.  103.
H1501948  285.  255.  208.  250.  391.  584.  868.  215.  124.   64.   81.   74.
H1501949  101.  124.  228.  292.  419.  787.  713.  246.  106.  148.  192.  301.
H1501950  132.  173.  289.  386.  529.  732.  560.  240.  130.   84.   57.  125.
H1501951  335.  504.  815.  935.  726.  763.  410.  199.   85.   43.   69.  150.
H1501952  397. 1695.  951.  683. 1310.  935. 1033.  787.  260.  160.   84.   82.
H1501953  160.  184.  482.  795.  863. 1193.  841.  292.  171.  169.   81.   81.
H1501954  134.  164.  208.  389.  358.  283.  616.  328.  130.  129.  108.   73.
H1501955  149.  364.  502. 1727. 1140. 1469.  729.  389.  400.  194.  341. 1207.
H1501956 1124. 1529. 1585. 1264. 1151. 1309.  825.  430.  180.  111.  111.   85.
H1501957  122.  184.  378.  262.  298.  435.  232.  134.  133.   75.   53.   65.
H1501958  304.  353.  642. 1501.  648. 1403.  571.  257.  105.   90.  113.  118.
H1501959   73.  118.  127.  258.  553.  529.  283.  123.   75.   46.   34.   69.
H1501960  514.  509.  804.  980.  953.  839.  497.  315.  138.   60.   82.  149.
H1501961  117.  177.  276.  328.  405.  341.  192.  154.   98.   70.   53.   47.
H1501962  107.  374.  248.  375.  381.  513.  285.  181.  129.   76.   47.   44.
H1501963  151.  141.  198.  401.  458.  462.  320.  164.   64.   44.   44.   68.
H1501964   77.  212. 1245.  793. 1015. 1444.  605.  315.  124.   57.   47.   58.
H1501965   68.   70.   84.  271.  469.  258.  235.  169.   55.   46.   69.   52.
H1501966  143.  226.  256.  453.  701.  777.  497.  573.  175.   73.   50.   48.
H1501967   36.   42.   65.  123.  182.  271.   64.   32.   27.    6.    6.   26.
H1501968  256.  456.  193.  759.  520. 1080.  632.  258.  129.   77.  114.  164.
H1501969  204.  335.  795.  493.  685.  418.  446.  244.  225.   86.   84.  199.
H1501970  335.  415.  642. 1337. 1421.  983.  696.  310.  167.  224.  102.  113.
H1501971  305.  248.  234.  383.  611.  892.  824.  315.  247.  154.  125.  108.
H1801922   21.   21.  170.  133.  130.  154.   53.   17.   10.    1.    0.    0.
H1801923   21.  355.  365.  378.  192.  378.  113.   63.   52.   37.   32.   61.
H1801924   52.  146.   81.  331.  354.  271.  316.  116.   90.   60.   39.   31.
H1801925   33.   50.   95.  228.  257.  138.   48.   22.    5.    0.   17.   28.
H1801926  315.  235.  295.  530.  295.  300.   90.   34.   20.    6.    5.    4.
H1801927   12.   27.   98.  264.  154.  140.   76.   27.   11.   47.   44.   61.
```

FIGURE 9.7: PROGRAM STATS INPUT DATA DECK FOR HISTORICAL DATA ANALYSIS

H1801928	107.	346.	231.	121.	112.	428.	119.	33.	11.	7.	5.	23.
H1801929	31.	130.	106.	175.	130.	135.	44.	32.	20.	4.	2.	2.
H1801930	25.	15.	48.	198.	112.	322.	151.	123.	53.	16.	37.	39.
H1801931	266.	624.	615.	557.	433.	169.	107.	44.	14.	10.	33.	213.
H1801932	71.	307.	298.	241.	367.	125.	66.	32.	12.	5.	4.	4.
H1801933	11.	41.	149.	127.	386.	219.	48.	102.	68.	21.	18.	32.
H1801934	25.	18.	171.	221.	124.	492.	390.	220.	48.	21.	17.	46.
H1801935	77.	122.	330.	501.	385.	172.	96.	41.	22.	11.	10.	12.
H1801936	14.	39.	231.	681.	244.	87.	64.	80.	50.	22.	15.	11.
H1801937	21.	23.	32.	71.	177.	85.	37.	18.	7.	9.	4.	5.
H1801938	10.	50.	75.	63.	81.	27.	9.	2.	0.	7.	119.	314.
H1801939	140.	402.	363.	724.	533.	317.	223.	89.	38.	7.	4.	10.
H1801940	21.	28.	47.	63.	50.	28.	17.	9.	41.	11.	36.	23.
H1801941	15.	33.	121.	82.	63.	148.	46.	18.	6.	5.	4.	2.
H1801942	122.	272.	664.	437.	486.	192.	96.	42.	20.	9.	5.	23.
H1801943	25.	43.	130.	240.	257.	148.	53.	25.	7.	2.	1.	5.
H1801944	43.	46.	97.	64.	28.	20.	18.	9.	2.	4.	1.	1.
H1801945	1.	28.	34.	235.	180.	73.	130.	34.	10.	20.	69.	43.
H1801946	28.	66.	514.	474.	171.	89.	63.	36.	16.	6.	11.	12.
H1801947	15.	81.	386.	386.	304.	337.	228.	92.	39.	18.	11.	10.
H1801948	43.	87.	107.	155.	161.	107.	264.	57.	25.	10.	26.	18.
H1801949	22.	34.	84.	103.	161.	290.	326.	91.	22.	16.	22.	26.
H1801950	18.	32.	86.	108.	161.	188.	155.	42.	16.	6.	4.	15.
H1801951	89.	205.	562.	415.	204.	192.	118.	41.	15.	4.	4.	14.
H1801952	60.	595.	584.	278.	522.	264.	370.	221.	59.	26.	12.	11.
H1801953	16.	70.	306.	553.	384.	448.	191.	60.	25.	25.	11.	12.
H1801954	33.	82.	73.	180.	188.	58.	189.	277.	44.	74.	74.	41.
H1801955	48.	237.	472.	1107.	588.	506.	283.	106.	109.	38.	38.	558.
H1801956	518.	641.	1005.	696.	476.	427.	203.	91.	34.	18.	17.	15.
H1801957	25.	47.	129.	86.	68.	135.	61.	26.	18.	11.	9.	7.
H1801958	66.	76.	252.	935.	236.	514.	182.	65.	20.	20.	28.	52.
H1801959	22.	26.	25.	134.	154.	111.	47.	21.	11.	4.	2.	9.
H1801960	236.	218.	315.	485.	373.	247.	108.	54.	15.	4.	7.	23.
H1801961	20.	41.	121.	155.	128.	52.	28.	18.	22.	7.	2.	9.
H1801962	64.	279.	157.	263.	139.	196.	91.	32.	39.	28.	11.	10.
H1801963	37.	87.	138.	394.	293.	187.	97.	33.	9.	4.	4.	7.
H1801964	14.	91.	676.	402.	413.	616.	151.	59.	16.	2.	2.	5.
H1801965	16.	22.	36.	230.	295.	82.	48.	57.	7.	9.	7.	22.
H1801966	31.	39.	92.	323.	371.	301.	137.	328.	66.	15.	14.	16.
H1801967	9.	10.	16.	54.	60.	53.	12.	2.	1.	1.	1.	1.
H1801968	198.	424.	140.	498.	258.	394.	151.	60.	18.	11.	26.	38.
H1801969	36.	66.	245.	181.	346.	121.	48.	38.	28.	5.	7.	76.
H1801970	116.	117.	337.	488.	520.	170.	91.	41.	21.	18.	7.	16.
H1801971	113.	170.	157.	236.	314.	403.	309.	79.	54.	31.	23.	18.
H3301922	44.	70.	80.	60.	57.	44.	0.	39.	12.	0.	4.	0.
H3301923	70.	168.	102.	52.	37.	66.	40.	71.	58.	78.	71.	54.
H3301924	55.	66.	37.	118.	47.	62.	169.	82.	217.	117.	27.	10.
H3301925	27.	85.	70.	33.	60.	18.	16.	1.	14.	0.	75.	118.
H3301926	96.	107.	66.	84.	44.	54.	14.	51.	76.	7.	8.	2.
H3301927	66.	21.	83.	107.	13.	90.	17.	37.	59.	227.	132.	32.
H3301928	60.	55.	69.	21.	28.	76.	4.	4.	14.	33.	38.	58.
H3301929	34.	49.	30.	61.	87.	19.	35.	52.	3.	3.	6.	30.
H3301930	66.	34.	59.	97.	30.	102.	42.	202.	35.	26.	161.	74.
H3301931	158.	212.	60.	91.	43.	45.	96.	11.	1.	50.	59.	121.
H3301932	15.	112.	53.	90.	34.	51.	24.	22.	33.	0.	20.	30.
H3301933	72.	65.	52.	50.	110.	27.	66.	173.	81.	83.	22.	77.
H3301934	0.	47.	100.	131.	63.	222.	125.	58.	22.	64.	43.	124.
H3301935	21.	53.	90.	68.	110.	113.	16.	32.	55.	36.	29.	42.
H3301936	18.	107.	163.	96.	21.	29.	16.	158.	66.	2.	28.	16.
H3301937	41.	42.	25.	59.	86.	82.	25.	35.	12.	42.	3.	54.
H3301938	11.	84.	41.	62.	22.	9.	4.	3.	3.	205.	143.	198.
H3301939	27.	128.	69.	189.	44.	85.	105.	7.	9.	5.	0.	99.
H3301940	34.	17.	42.	33.	58.	2.	34.	54.	145.	58.	59.	15.
H3301941	5.	40.	70.	14.	40.	39.	29.	19.	10.	47.	19.	23.
H3301942	161.	105.	41.	90.	55.	56.	67.	13.	64.	21.	11.	96.
H3301943	18.	54.	60.	118.	64.	35.	47.	4.	7.	4.	71.	44.
H3301944	77.	20.	26.	3.	9.	66.	16.	44.	61.	1.	9.	13.
H3301945	22.	102.	52.	145.	24.	85.	48.	10.	66.	114.	75.	45.
H3301946	33.	52.	132.	61.	5.	18.	125.	51.	8.	42.	65.	26.
H3301947	20.	63.	125.	69.	39.	76.	50.	81.	29.	57.	2.	37.
H3301948	62.	38.	25.	29.	53.	104.	81.	48.	8.	35.	60.	29.
H3301949	36.	15.	53.	22.	62.	184.	102.	11.	16.	76.	117.	19.
H3301950	65.	25.	62.	81.	65.	119.	54.	53.	39.	29.	4.	97.
H3301951	108.	106.	89.	74.	53.	87.	28.	16.	7.	5.	58.	108.
H3301952	125.	137.	81.	63.	70.	78.	73.	56.	54.	18.	0.	27.
H3301953	58.	87.	70.	88.	96.	59.	54.	10.	113.	8.	10.	59.
H3301954	34.	46.	45.	80.	44.	51.	129.	133.	35.	76.	79.	7.
H3301955	78.	102.	120.	123.	83.	159.	53.	65.	101.	10.	165.	111.
H3301956	127.	114.	134.	45.	78.	112.	28.	14.	1.	12.	45.	24.
H3301957	17.	118.	68.	29.	41.	55.	20.	76.	10.	15.	53.	39.

FIGURE 9.7 (CONT'D): PROGRAM STATS INPUT DATA DECK FOR HISTORICAL DATA ANALYSIS

H3301958	112.	19.	117.	82.	57.	111.	47.	13.	11.	15.	84.	91.
H3301959	20.	41.	40.	70.	62.	127.	36.	15.	69.	20.	14.	69.
H3301960	138.	22.	129.	80.	113.	36.	39.	78.	25.	19.	45.	63.
H3301961	11.	76.	95.	99.	28.	40.	62.	90.	57.	5.	13.	22.
H3301962	112.	74.	42.	98.	54.	64.	26.	43.	66.	12.	34.	21.
H3301963	113.	101.	96.	77.	106.	71.	39.	41.	17.	29.	43.	115.
H3301964	45.	83.	165.	44.	120.	94.	37.	52.	1.	6.	8.	26.
H3301965	54.	18.	54.	128.	79.	48.	39.	39.	15.	43.	60.	15.
H3301966	56.	41.	58.	61.	133.	136.	54.	215.	25.	10.	24.	4.
H3301967	36.	40.	38.	92.	34.	46.	3.	7.	23.	5.	19.	71.
H3301968	154.	69.	69.	85.	28.	81.	57.	63.	13.	78.	88.	83.
H3301969	66.	31.	133.	30.	71.	52.	30.	44.	96.	6.	71.	119.
H3301970	31.	56.	18.	130.	86.	15.	53.	38.	7.	39.	22.	70.
H3301971	74.	36.	57.	79.	45.	41.	98.	25.	48.	75.	5.	27.

FIGURE 9.7 (CONT'D): PROGRAM STATS INPUT DATA DECK FOR HISTORICAL DATA ANALYSIS

```
*****   JOB DESCRIPTION   *****                         RUN NO 401          NO OF SUB-JOBS    1

                                        CONSTANT ANALYSIS OPTIONS   1            SUB-JOB NO    1

*****   DATA DESCRIPTION  *****                         NO OF SERIES   3        NO OF YEARS   50

                                                      SERIES NOS 150  180  330

                                                 DATA INPUT UNIT    5           DATA PRINT    1

                                                      DATA FORMAT ( 8X,12F6.0 )

*****   ANALYSIS OPTIONS  *****

        KENDALL TREND                            1

        MONTHLY: MOMENTS,EXTREMUM,FREQUENCY 1                   PLOT  0  WRITE  0  TO FILES 20 21 22 23
                                                                         PUNCH 1  TO FILES 1

        ANNUAL  : MOMENTS,EXTREMUM,FREQUENCY 1                  PLOT  0  WRITE  0  TO FILES 24
                                                                         PUNCH  1  TO FILES 2

                : CORRELATION               1  LAGS     12      PLOT  0

        SERIES  : MOMENTS,EXTREMUM,FREQUENCY 1  BOUNDS   1       PLOT 0   WRITE  0 TO FILES 25 26
                                                                          PUNCH  1 TO FILES 3   4

                : RUNS                      1 LENGTH 30  SCALE 0  PLOT 0   WRITE  0 TO FILES 27
                                                                          PUNCH  1 TO FILES  9

                : VOLUMES                   1 LENGTH 60  SCALE 0  PLOT 0   WRITE  0 TO FILES 28

                :  RANGE                    1            SCALE 0           WRITE  0 TO FILES 29

                :  MASS CURVE               1 YIELDS   2 SCALE 0  PLOT 0   WRITE  0 TO FILES 30 33 34 35 36
                                                                          PUNCH  1 TO FILES 10 11 12 13 14

                : AUTO-CORRELATION          1 LAGS     12         PLOT 0   WRITE 0 TO FILES 31
                                                                          PUNCH 1 TO FILES 15

                : CROSS-CORRELATION         1 LAGS     12         PLOT 1   WRITE 0 TO FILES  32
                                                                          PUNCH 1 TO FILES 15
```

FIGURE 6.6 SAMPLE OUTPUT FROM HISTORICAL DATA ANALYSIS

TREND ANALYSIS ----------- KENDALL TAU TEST ------------ STATION NO 150

ACTUAL SCORE :	- 371.0000
MAXIMUM SCORE :	179700.000
CORRECTION FOR TIES :	381.0000
TAU :	- 0.0021
STD DEV OF TAU:	0.0273
STANDARDISED STATISTIC	- 0.0757

FIGURE 9.8 (CONT'D): SAMPLE OUTPUT FROM HISTORICAL DATA ANALYSIS

MOMENT ANALYSIS ---------- MONTHLY ------------- STATION NO. 150

MONTH	5	6	7	8	9	10	11	12	1	2	3	4
MEAN	203.660	330.280	456.980	597.920	631.540	706.080	454.160	244.880	140.980	91.560	96.600	134.800
STD DEV	177.356	362.893	344.533	386.069	296.106	346.161	257.627	155.602	85.615	55.883	73.587	170.646
SKEW	3.205	2.959	1.315	1.063	0.684	0.564	0.520	1.593	0.980	1.224	2.116	5.286
KURTOSIS	17.092	11.783	4.553	3.761	3.113	2.877	2.435	6.545	3.672	4.165	8.620	35.341
COEF VAR	0.871	1.099	0.754	0.646	0.469	0.490	0.567	0.635	0.607	0.610	0.762	1.266

EXTREMUM ANALYSIS ------------ MONTHLY ------------- STATION NO 150

MONTH	5	6	7	8	9	10	11	12	1	2	3	4
MINIMUM	36.000	42.000	65.000	123.000	128.000	156.000	64.000	32.000	21.000	6.000	6.000	26.000
MAXIMUM	1124.000	1695.000	1585.000	1727.000	1421.000	1469.000	1033.000	787.000	400.000	258.000	384.000	1207.000

FIGURE 9.8 (CONT'D.):SAMPLE OUTPUT FROM HISTORICAL DATA ANALYSIS

FIGURE 9.8 (CONT'D). SAMPLE OUTPUT FROM HISTORICAL DATA ANALYSIS

FREQUENCY ANALYSIS ------------- OF THE TIME SERIES --------------- STATION NO 150

INTERVAL SPECIFICATIONS WERE : LOWER BOUND = 0.0

NO OF CLASSES = 20.000

UPPER BOUND = 2160.000

CALCULATED STATISTICS WERE : MEAN = 340.787

STD DEVIATION = 329.557

SKEWNESS COEFF = 1.655

KURTOSIS COEFF = 5.645

COEFF OF VAR = 0.967

CALCULATED FREQUENCIES WERE :

CLASS	U BOUND	FREQ	% FREQ	CUM FREQ	CLASS	U BOUND	FREQ	% FREQ	CUM FREQ	CLASS	U BOUND	FREQ	% FREQ	CUM FREQ
1	0.0	0.	0.0	0.0	2	120.00	174.	29.00	174.00	3	240.00	140.	23.33	314.00
4	360.00	81.	13.50	395.00	5	480.00	59.	9.83	454.00	6	600.00	32.	5.33	486.00
7	720.00	33.	5.50	519.00	8	840.00	28.	4.67	547.00	9	960.00	17.	2.83	564.00
10	1080.00	9.	1.50	573.00	11	1200.00	9.	1.50	582.00	12	1320.00	5.	0.83	587.00
13	1440.00	5.	0.83	592.00	14	1560.00	4.	0.67	596.00	15	1680.00	1.	0.17	597.00
16	1800.00	3.	0.50	600.00	17	1920.00	0.	0.0	600.00	18	2040.00	0.	0.0	600.00
19	2160.00	0.	0.0	600.00	20	99999.00	0.	0.0	600.00					

EXTREMUM ANALYSIS ------------- OF THE TIME SERIES --------------- STATION NO 150

MINIMUM VALUE = = 6.000

MAXIMUM VALUE = 1727.000

RUNS ANALYSIS --------------- ABOUT THE MEDIAN --------------- STATION NO 150

(THE MEDIAN OF THIS TIME SERIES IS 224.500)

RUN LENGTH	POSITIVE RUNS	NEGATIVE RUNS	TOTAL RUNS	RUN LENGTH	POSITIVE RUNS	NEGATIVE RUNS	TOTAL RUNS
1	8	6	14	2	5	3	8
3	0	6	6	4	9	5	14
5	6	11	17	6	9	7	16
7	6	11	17	8	8	5	13
9	4	1	5	10	2	1	3
11	0	0	0	12	0	0	0
13	0	0	0	14	0	1	1
15	0	0	0	16	0	0	0
17	0	0	0	18	0	0	0
19	0	0	0	20	0	0	0
21	0	0	0	22	0	0	0
23	0	0	0	24	0	0	0
25	0	0	0	26	0	0	0
27	0	0	0	28	0	0	0
29	0	0	0	30	0	0	0

TOTALS: POSITIVE = 57 ; NEGATIVE = 57 ; TOTAL = 114.

MEANS : = 5 ; = 5 ; = 5.

FIGURE 9.8 (CONT'D): SAMPLE OUTPUT FROM HISTORICAL DATA ANALYSIS

VOLUME PERIOD ANALYSIS ----------- MAX AND MIN ------------- STATION NO 180

(THE SAMPLE MEAN USED IN SCALING IS 128.348)

PERIOD LENGTH	MINIMUM VOLUMES ACTUAL	MINIMUM VOLUMES SCALED	MAXIMUM VOLUMES ACTUAL	MAXIMUM VOLUMES SCALED	PERIOD LENGTH	MINIMUM VOLUMES ACTUAL	MINIMUM VOLUMES SCALED	MAXIMUM VOLUMES ACTUAL	MAXIMUM VOLUMES SCALED
1	0.0	0.0	1107.0	8.62	2	0.0	0.0	1701.0	13.25
3	1.0	0.008	2342.0	18.25	4	4.0	0.031	2860.0	22.28
5	6.0	0.047	3418.0	26.63	6	18.0	0.140	3894.0	30.34
7	36.0	0.280	4321.0	33.67	8	56.0	0.436	4524.0	35.25
9	84.0	0.654	4615.0	35.96	10	112.0	0.873	4653.0	36.25
11	146.0	1.138	4709.0	36.69	12	210.0	1.636	5497.0	42.83
13	236.0	1.839	6193.0	48.25	14	250.0	1.948	6669.0	51.96
15	265.0	2.065	7141.0	55.64	16	331.0	2.579	7568.0	58.96
17	349.0	2.719	7805.0	60.81	18	374.0	2.914	8008.0	62.39
19	402.0	3.132	8099.0	63.10	20	436.0	3.397	8147.0	63.48
21	489.0	3.810	8188.0	63.80	22	637.0	4.963	8262.0	64.37
23	824.0	6.420	8363.0	65.16	24	839.0	6.537	8566.0	66.74
25	861.0	6.708	8755.0	68.21	26	911.0	7.098	8846.0	68.92
27	938.0	7.308	9001.0	70.13	28	976.0	7.604	9181.0	71.53
29	1065.0	8.298	9272.0	72.24	30	1187.0	9.248	9345.0	72.81
31	1311.0	10.214	9427.0	73.45	32	1377.0	10.729	9461.0	73.71
33	1525.0	11.882	9494.0	73.97	34	1719.0	13.393	9512.0	74.11
35	1919.0	14.951	9529.0	74.24	36	2015.0	15.699	9557.0	74.46
37	2051.0	15.980	9972.0	77.69	38	2071.0	16.136	10448.0	81.40
39	2113.0	16.463	10875.0	84.73	40	2182.0	17.001	11181.0	87.11
41	2211.0	17.227	11384.0	88.70	42	2253.0	17.554	11475.0	89.41
43	2319.0	18.068	11545.0	89.95	44	2415.0	18.816	11579.0	90.22
45	2607.0	20.312	11597.0	90.36	46	3093.0	24.098	11614.0	90.49
47	3510.0	27.347	11630.0	90.61	48	3602.0	28.064	11966.0	93.23
49	3626.0	28.251	12325.0	96.03	50	3637.0	28.337	12752.0	99.35
51	3649.0	28.430	13306.0	103.67	52	3664.0	28.547	13782.0	107.38
53	3684.0	28.703	14209.0	110.71	54	3726.0	29.030	14412.0	112.29
55	3807.0	29.661	14503.0	113.00	56	3903.0	30.409	14563.0	113.46
57	4095.0	31.905	14597.0	113.73	58	4408.0	34.344	14615.0	113.87
59	4458.0	34.734	14632.0	114.00	60	4521.0	35.224	14647.0	114.12

FIGURE 9.8 (CONT'D): SAMPLE OUTPUT FROM HISTORICAL DATA ANALYSIS

MASS CURVE ANALYSIS -------- RIPPL METHOD --------- STATION NO 180

(THE SCALING VALUE IS THE SAMPLE MEAN WHICH IS 128.348)

*** YIELD PROPORTION 0.600 ***

FREQUENCY ANALYSIS -- UPPER BOUNDS ----

1. DROUGHT/DRAW DURATION ANALYSIS 28.000 MONTHS
2. FILL DURATION ANALYSIS 14.000 MONTHS
3. DEFICIT INTENSITY ANALYSIS 10.000 X MEAN

RIPPL STORAGE 9.196

DROUGHT DURATION 28

FREQ BIN	DEFICIT FREQ	FILL FREQ	DRAW FREQ	DROUGHT FREQ	MEAN DEFICIT	STD DEV DEFICIT	MIN DEFICIT	MAX DEFICIT
1	0	0	0	0	0.0	0.0	0.0	0.0
2	1	3	0	2	0.740	0.187	0.608	0.873
3	6	8	2	2	1.516	0.138	1.418	1.613
4	4	9	1	14	2.046	0.627	0.811	3.047
5	9	5	7	8	2.909	0.501	2.213	3.437
6	6	2	8	0	0.0	0.0	0.0	0.0
7	2	3	5	0	0.0	0.0	0.0	0.0
8	3	0	0	0	0.0	0.0	0.0	0.0
9	0	0	0	1	7.021	0.0	7.021	7.021
10	0	0	0	3	4.447	0.316	4.193	4.801
11	1	0	4	2	4.053	0.463	3.726	4.380
12	0	1	2	1	8.790	0.0	8.790	8.790
13	0	1	1	0	0.0	0.0	0.0	0.0
14	3	1	0	0	0.0	0.0	0.0	0.0
15	0	0	0	2	8.888	0.435	8.580	9.196
16	0	1	4	0	0.0	0.0	0.0	0.0
TOTALS :	35.000	34.000	34.000	35.000				
MEANS :	3.184	4.559	14.265	9.543				
STD DEVS :	2.178	4.627	9.752	7.043				
MAXIMUMS ABOVE UPPER BDS	0	25	46	0				

*** YIELD PROPORTION 0.800 ***

FREQUENCY ANALYSIS -- UPPER BOUNDS ----

1. DROUGHT/DRAW DURATION ANALYSIS 64.000 MONTHS
2. FILL DURATION ANALYSIS 24.000 MONTHS
3. DEFICIT INTENSITY ANALYSIS 30.000 X MEAN

RIPPL STORAGE 19.155

DROUGHT DURATION 79

FREQ BIN	DEFICIT FREQ	FILL FREQ	DRAW FREQ	DROUGHT FREQ	MEAN DEFICIT	STD DEV DEFICIT	MIN DEFICIT	MAX DEFICIT
1	0	0	0	0	0.0	0.0	0.0	0.0
2	10	8	5	13	2.967	1.125	1.008	4.805
3	6	8	6	1	3.910	0.0	3.910	3.910
4	5	2	5	6	7.212	1.413	5.186	8.401
5	1	2	2	1	14.302	0.0	14.302	14.302
6	0	0	1	0	0.0	0.0	0.0	0.0
7	1	0	2	1	10.454	0.0	10.454	10.454
8	0	0	0	0	0.0	0.0	0.0	0.0
9	0	0	0	0	0.0	0.0	0.0	0.0
10	0	2	1	1	19.155	0.0	19.155	19.155
TOTALS :	23.000	22.000	22.000	23.000				
MEANS :	5.637	9.000	24.636	15.174				
STD DEVS :	4.352	15.799	31.541	16.876				
MAXIMUMS ABOVE UPPER BDS	0	75	154	79				

FIGURE 9.8 (CONT'D): OUTPUT FROM HISTORICAL DATA ANALYSIS

MONTHLY AUTOCORRELATION ANALYSIS ----------- LARGE SAMPLE FORMULA -------------

AUTOCORRELATION FUNCTION OF STATION NO 180

LAG	AUTOCOVARIANCE	AUTOCORRELATION
0	26760.457031	1.000019
1	18192.093750	0.679825
2	11658.742188	0.435679
3	4432.800781	0.165651
4	- 2470.182373	- 0.092309
5	- 6667.503906	- 0.249160
6	- 8561.007813	- 0.319919
7	- 7750.238281	- 0.289621
8	- 4071.080566	- 0.152133
9	- 310.213379	- 0.011592
10	4238.460938	0.158388
11	8120.390625	0.303453
12	9012.406250	0.336787

FIGURE 9.8 (CONT'D): SAMPLE OUTPUT FROM HISTORICAL DATA ANALYSIS

CROSS CORRELATION ANALYSIS ----------- LARGE SAMPLE FORMULA -------------

CROSS CORRELATION FUNCTION OF STN NO 150 WITH STN NO 180 FOR 12 LAGS

LAG	COVARIANCE -K	CORRELATION -K	COVARIANCE +K	CORRELATION +K
0	49167.53516	0.91278	49167.53516	0.91278
1	32351.49219	0.60060	39019.24609	0.72438
2	17157.48438	0.31852	31697.77344	0.58846
3	3701.15381	0.06871	17135.42188	0.31812
4	-9434.92188	-0.17516	1485.59375	0.02758
5	-16125.64063	-0.29937	-9305.15234	-0.17275
6	-17795.39844	-0.33037	-15693.56641	-0.29135
7	-14429.89844	-0.26789	-16263.81641	-0.30193
8	-5205.91016	-0.09665	-11049.38672	-0.20513
9	4947.07422	0.09184	-4619.74609	-0.08576
10	13955.91016	0.25909	4820.89844	0.08950
11	18392.45313	0.34145	13541.44922	0.25139
12	17478.87109	0.32449	18018.38281	0.33451

(+K MEANS STN NO 150 IS LEADING STN NO 180)

```
FILE: SYNDATA  ANALYSIS A1  CMS REL3 PLC12 CANBERRA 370/158

  5   1
402   0   0
 ( 8X,12F6.0 )
  3  50
  0
  1   1   1   0   0
  1   1   0
  0
  1   1   0   1   1   0   0
 0.0          2160.0     0.0         1080.0     0.0         270.0
  1  30   1   0   1   0
 224.50        57.000    52.00
  0
  0
  1   2   1   0   1   0
 0.6          0.8
 340.787      128.348    57.342
  2   6   4   7
  3   7   6  11
  1   2   3   4
  1  12   0   1   0
  1  12   0   1   0
150180330
```

FIGURE 9.9: PROGRAM STATS INPUT DATA DECK FOR SYNTHETIC DATA ANALYSIS

```
FILE: MOPS     COMPARE  A   CMS REL3 PLC12 CANBERRA 370/158

403  3 50  2  5 14
  1 80 81 82
 20 21 22 23 24 25
  CEB01
  2  1
150180330
  8  8  7
  2  2  1
   203.65999   330.27979   456.97998   597.91992   631.53979   706.07983150 10 1
   454.15991   244.87999   140.98000    91.56000    96.59999   134.79999150 10 2
   177.35564   362.89331   344.53296   386.06885   296.10596   346.16064150 11 1
   257.62671   155.60191    85.61528    55.88261    73.58655   170.64595150 11 2
     3.20489     2.95946     1.31504     1.06314     0.68351     0.56435150 12 1
     0.51990     1.59272     0.98018     1.22410     2.11583     5.28626150 12 2
    36.00000    42.00000    65.00000   123.00000   128.00000   156.00000150 13 1
    64.00000    32.00000    21.00000     6.00000     6.00000    26.00000150 13 2
  1124.00000  1695.00000  1585.00000  1727.00000  1421.00000  1469.00000150 14 1
  1033.00000   787.00000   400.00000   258.00000   384.00000  1207.00000150 14 2
    67.43999   144.87999   237.21999   319.67993   256.35986   221.12000180 10 1
   127.21959    65.17999    27.31999    15.06000    18.01999    40.67999180 10 2
    93.99416   163.91971   211.06470   232.30824   145.51064   147.77672180 11 1
    97.93484    66.76633    23.12560    15.04471    22.04494    91.06186180 11 2
     3.04497     1.71459     1.54806     1.30364     0.47605     0.79392180 12 1
     1.11000     2.40512     1.51809     2.06132     2.62323     4.58915180 12 2
     1.00000    10.00000    16.00000    54.00000    28.00000    20.00000180 13 1
     9.00000     2.00000     0.0         0.0         0.0         0.0     180 13 2
   518.00000   641.00000  1005.00000  1107.00000   588.00000   616.00000180 14 1
   390.00000   328.00000   109.00000    74.00000   119.00000   558.00000180 14 2
    59.65999    69.06000    73.09999    75.81999    57.81999    70.81999330 10 1
    49.43999    51.17999    40.34000    39.75999    46.01999    55.07999330 10 2
    42.68663    41.73259    36.16135    36.78716    30.05757    44.24271330 11 1
    36.70355    49.24603    41.82990    47.27106    42.15109    42.30591330 11 2
     0.86301     1.08743     0.81739     0.45922     0.57604     1.19337330 12 1
     1.24987     1.83704     2.01176     2.30365     1.22415     1.02004330 12 2
     0.0        15.00000    18.00000     3.00000     5.00000     2.00000330 13 1
     0.0         1.00000     1.00000     0.0         0.0         0.0     330 13 2
   161.00000   212.00000   165.00000   189.00000   133.00000   222.00000330 14 1
   169.00000   215.00000   217.00000   227.00000   165.00000   198.00000330 14 2
  4089.43954  1947.20972     0.76878                                     150 15 1
  1540.17993   891.42432     0.98338                                     180 15 1
   688.09985   179.15903     0.19975                                     330 15 1
   340.78662   329.55688     1.65520                                     150 16 1
   128.34833   163.72121     2.06868                                     180 16 1
    57.34166    42.60582     1.12156                                     330 16 1
```

FIGURE 9.10: PROGRAM MOMENT INPUT DATA DECK FOR EXAMPLE PROBLEM

***** FIRST MOMENT DATA *****

SAMPLE : 1

150	207.420	316.120	459.260	588.120	632.240	709.220	457.060	244.940	141.900	90.780	96.100	129.080	4072.240	339.353
180	66.320	140.000	237.640	316.540	258.300	221.000	128.080	64.000	26.960	14.480	17.780	34.480	1525.580	127.132
330	57.580	69.240	73.440	75.320	57.440	69.840	49.460	50.840	42.060	40.580	47.440	54.880	688.120	57.343

SAMPLE : 2

150	211.900	326.380	464.120	601.400	633.460	709.580	457.620	245.400	141.420	91.260	96.200	135.000	4113.738	342.812
180	67.980	142.240	240.820	324.940	258.000	224.080	128.120	64.040	26.860	14.660	18.760	35.800	1546.300	128.858
330	57.200	69.160	73.400	75.800	58.040	70.160	48.660	52.280	39.480	43.100	46.740	53.860	687.880	57.323

SAMPLE : 3

150	202.480	326.140	443.580	603.780	633.360	709.540	457.160	245.380	141.280	91.220	96.160	133.140	4083.220	340.268
180	60.460	146.620	231.340	323.580	258.740	223.640	128.700	64.540	27.360	14.700	18.500	35.160	1533.340	127.778
330	58.420	69.320	73.500	75.940	57.000	70.700	48.720	52.380	41.340	40.620	47.300	53.640	688.880	57.407

SAMPLE : 4

150	195.760	328.340	453.340	610.940	632.660	709.440	457.520	244.900	140.280	91.280	95.820	131.780	4092.060	341.005
180	63.840	137.660	235.480	328.180	258.020	223.100	127.240	64.040	27.180	14.660	17.540	33.560	1530.500	127.542
330	59.220	69.320	73.560	77.200	57.820	70.740	48.280	52.080	41.180	41.360	47.120	55.840	693.720	57.810

SAMPLE : 5

150	202.780	308.380	454.500	603.420	632.520	709.980	457.300	244.600	140.640	91.660	95.980	148.280	4090.040	340.836
180	66.600	136.640	235.260	327.320	258.360	221.640	127.120	63.380	27.360	14.960	17.240	35.960	1531.840	127.653
330	59.080	68.640	73.280	76.000	58.060	70.560	49.360	51.220	41.560	39.980	46.680	55.680	690.100	57.508

FIGURE 9.11: SAMPLE OUTPUT FROM MOMENTS COMPARISON

***** MOMENT 1 ***** STATION 150 *****

INDIVIDUAL SAMPLES :

TIME PERIOD :	5	6	7	8	9	10	11	12	1	2	3	4	ANN	TS
MIN SYNTHETIC VALUE :	195.76	308.38	443.58	588.12	632.24	709.22	457.06	244.60	140.28	90.78	95.82	129.08	4072.24	339.35
1 PER CENT LOWER LIMIT :	139.07	198.13	331.51	457.33	523.71	580.02	360.34	188.22	109.80	71.21	69.80	72.66	3380.34	306.14
NO BELOW THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5 PER CENT LOWER LIMIT :	154.50	229.69	361.48	490.91	549.46	610.13	382.75	201.75	117.25	76.07	76.20	87.50	3549.70	314.42
NO BELOW THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HISTORICAL VALUES :	203.66	330.28	456.98	597.92	631.54	706.08	454.16	244.88	140.98	91.56	96.60	134.80	4089.44	340.79
5 PER CENT UPPER LIMIT :	252.82	430.87	552.48	704.93	713.62	802.03	525.57	288.01	164.71	107.05	117.00	182.10	4629.18	367.16
NO ABOVE THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1 PER CENT UPPER LIMIT :	268.25	462.43	582.44	738.51	739.37	832.14	547.98	301.54	172.16	111.91	123.40	196.94	4798.54	375.43
NO ABOVE THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MAX SYNTHETIC VALUE :	211.90	328.34	464.12	610.94	633.46	709.98	457.62	245.40	141.90	91.66	96.20	148.28	4113.74	342.81

POPULATION RESULTS :

TIME PERIOD :	5	6	7	8	9	10	11	12	1	2	3	4	ANN	TS
1 PER CENT LOWER LIMIT :	174.78	271.18	400.87	535.05	583.32	649.71	412.20	219.54	127.04	82.46	84.62	107.01	3772.32	325.29
NO BELOW THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5 PER CENT LOWER LIMIT :	181.67	285.29	414.27	550.06	594.83	663.17	422.22	225.59	130.37	84.63	87.48	113.65	3848.06	328.99
NO BELOW THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HISTORICAL VALUES :	203.66	330.28	456.98	597.92	631.54	706.08	454.16	244.88	140.98	91.56	96.60	134.80	4089.44	340.79
POP SYNTHETIC VALUES:	204.07	321.07	454.96	601.53	632.85	709.55	457.33	245.04	141.10	91.24	96.05	135.46	4090.26	340.85
PER CENT DIFFERENCE :	0.200	-2.788	-0.442	0.604	0.207	0.492	0.698	0.067	0.088	-0.350	-0.567	0.487	0.020	0.020
SAMPLE VALUES MEAN :	204.07	321.07	454.96	601.53	632.85	709.55	457.33	245.04	141.10	91.24	96.05	135.46	4090.26	340.85
SAMPLE VALUES STD DEV :	6.04	8.55	7.67	8.32	0.0	0.50	0.35	0.41	0.67	0.35	0.19	7.49	15.36	1.31
SAMPLE VALUES SKEW :	-0.15	-1.19	-1.42	0.19	1000.00	2560.00	*******	-374.24	-21.86	-69.22	-408.00	1.78	22.60	-47.34
5 PER CENT UPPER LIMIT :	225.65	375.26	499.69	645.78	668.25	748.99	486.10	264.17	151.59	98.49	105.72	155.95	4330.82	352.58
NO ABOVE THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1 PER CENT UPPER LIMIT :	232.54	389.38	513.09	660.79	679.76	762.45	496.12	270.22	154.92	100.66	108.58	162.59	4406.55	356.28
NO ABOVE THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MEAN PER CENT ERROR/MONTH IS : 0.582
MEAN WET 6 MONTHS ERROR/MONTH IS: 0.872
MEAN DRY 6 MONTHS ERROR/MONTH IS: 0.293

FIGURE 9.11 (CONT'D): SAMPLE OUTPUT FROM MOMENTS COMPARISON

FIGURE 9.11 (CONT'D): SAMPLE OUTPUT FROM MOMENTS COMPARISON

***** MOMENT 2 ***** STATION 150 *****

INDIVIDUAL SAMPLES

TIME PERIOD :	5	6	7	8	9	10	11	12	1	2	3	4	ANN	TS
MIN SYNTHETIC VALUE :	136.72	235.40	283.48	325.91	266.23	324.41	248.08	134.56	74.62	46.33	63.85	125.85	1530.48	315.19
1 PER CENT LOWER LIMIT:	134.04	274.27	260.40	291.79	223.80	261.63	194.71	117.60	64.71	42.24	55.62	128.97	1471.69	305.16
NO BELOW THIS LIMIT:	0	1	0	0	0	0	0	0	0	0	0	1	0	0
5 PER CENT LOWER LIMIT:	144.13	294.91	279.99	313.74	240.63	281.31	209.36	126.45	69.58	45.41	59.80	138.68	1582.41	310.89
NO BELOW THIS LIMIT:	1	1	0	0	0	0	0	0	0	0	0	2	1	0
HISTORICAL VALUES:	177.36	362.89	344.53	386.07	296.11	346.16	257.63	155.60	85.62	55.88	73.59	170.65	1947.21	329.56
5 PER CENT UPPER LIMIT:	214.12	438.12	415.95	466.10	357.49	417.92	311.03	187.86	103.36	67.47	88.84	206.02	2350.85	348.21
NO ABOVE THIS LIMIT:	1	0	0	1	0	0	0	0	0	0	0	1	0	0
1 PER CENT UPPER LIMIT:	225.89	462.21	438.82	491.73	377.14	440.90	328.13	198.19	109.05	71.18	93.73	217.35	2480.11	354.20
NO ABOVE THIS LIMIT:	1	0	0	0	0	0	0	0	0	0	0	1	0	0
MAX SYNTHETIC VALUE:	235.97	349.92	392.08	486.18	308.85	357.09	273.35	154.27	90.74	53.20	67.78	233.16	1748.80	335.40

POPULATION RESULTS:

TIME PERIOD:	5	6	7	8	9	10	11	12	1	2	3	4	ANN	TS
1 PER CENT LOWER LIMIT:	157.05	321.35	305.09	341.88	262.21	306.54	228.14	137.79	75.81	49.49	65.16	151.11	1724.31	318.62
NO BELOW THIS LIMIT:	0	1	0	0	0	0	0	0	0	0	0	0	1	0
5 PER CENT LOWER LIMIT:	161.78	331.02	314.27	352.16	270.10	315.76	235.00	141.94	78.10	50.97	67.12	155.66	1776.19	321.22
NO BELOW THIS LIMIT:	0	1	0	0	0	0	0	0	0	1	1	0	1	0
HISTORICAL VALUES :	177.36	362.89	344.53	386.07	296.11	346.16	257.63	155.60	85.62	55.88	73.59	170.65	1947.21	329.56
POP SYNTHETIC VALUES:	180.71	308.89	342.28	410.27	283.80	340.69	262.29	143.28	83.05	49.67	65.39	160.01	1672.43	325.88
PER CENT DIFFERENCE:	1.894	-14.882	-0.653	6.268	-4.157	-1.580	1.810	-7.917	-2.995	-11.115	-11.143	-6.233	-14.112	-1.115
SAMPLE VALUES MEAN:	177.24	306.33	340.27	406.99	283.30	340.47	262.14	143.08	82.85	49.62	65.37	154.85	1670.67	325.79
SAMPLE VALUES STD DEV :	38.98	43.49	40.74	57.30	18.85	13.62	9.85	8.54	6.35	2.53	1.47	44.42	84.45	8.43
SAMPLE VALUES SKEW :	0.92	-1.34	-0.28	-0.08	0.63	-0.17	-0.74	0.29	-0.19	0.28	0.78	2.08	-1.51	0.0
5 PER CENT UPPER LIMIT :	192.91	394.73	374.76	419.94	322.08	376.53	280.23	169.25	93.13	60.78	80.04	185.62	2118.03	337.89
NO ABOVE THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1 PER CENT UPPER LIMIT :	197.97	405.07	384.58	430.94	330.52	386.40	287.57	173.69	95.57	62.38	82.14	190.48	2173.55	340.54
NO ABOVE THIS LIMIT :	0	0	0	0	0	0	0	0	0	0	0	0	0	0

MEAN PER CENT ERROR/MONTH IS : 5.887
MEAN WET 6 MONTHS ERROR/MONTH IS: 4.892
MEAN DRY 6 MONTHS ERROR/MONTH IS: 6.883

***** MOMENT 3 ***** STATION 150 *****

INDIVIDUAL SAMPLES :

TIME PERIOD :	5	6	7	8	9	10	11	12	1	2	3	4	ANN	TS
MIN SYNTHETIC VALUE :	1.89	1.22	1.06	0.58	-0.11	0.10	0.44	0.36	0.20	0.24	1.01	2.64	0.38	1.39
HISTORICAL VALUES:	3.20	2.96	1.32	1.06	0.68	0.56	0.52	1.59	0.98	1.22	2.12	5.29	0.79	1.66
MAX SYNTHETIC VALUE :	4.60	2.78	2.58	2.48	0.83	0.51	0.91	1.17	1.18	1.22	1.70	5.13	1.00	2.04

POPULATION RESULTS:

TIME PERIOD :	5	6	7	8	9	10	11	12	1	2	3	4	ANN	TS
HISTORICAL VALUES :	3.20	2.96	1.32	1.06	0.68	0.56	0.52	1.59	0.98	1.22	2.12	5.29	0.79	1.66
POP SYNTHETIC VALUES :	3.66	2.51	1.83	1.64	0.42	0.37	0.69	0.78	0.85	0.75	1.37	4.78	0.59	1.71
PER CENT DIFFERENCE :	14.078	-15.212	39.458	54.282	-38.608	-34.212	32.270	-51.307	-12.826	-38.912	-35.402	-9.535	-25.160	3.080
SAMPLE VALUES MEAN :	2.79	2.29	1.63	1.36	0.36	0.36	0.67	0.72	0.78	0.72	1.35	3.54	0.59	1.69
SAMPLE VALUES STD DEV :	1.18	0.63	0.75	0.71	0.36	0.17	0.17	0.39	0.41	0.35	0.32	1.02	0.24	0.24
SAMPLE VALUES SKEW :	1.15	-1.60	0.67	1.00	0.01	-1.04	0.09	0.47	-0.50	0.13	0.33	1.02	1.81	0.37

MEAN PER CENT ERROR/MONTH IS : 31.342
MEAN WET 6 MONTHS ERROR/MONTH IS : 35.674
MEAN DRY 6 MONTHS ERROR/MONTH IS : 27.010

FIGURE 9.11 (CONT'D): SAMPLE OUTPUT FROM MOMENTS COMPARISON

***** EXTREMES ***** STATION NO 150 *****

MINIMUM EXTREMES:

TIME PERIOD :	5	6	7	8	9	10	11	12	1	2	3	4
SYNTHETIC MIN MIN VALUES :	33.00	47.00	43.00	85.00	97.00	75.00	40.00	20.00	16.00	5.00	7.00	24.00
PER CENT VALUES.LE.MIN HIST :	60.00	0.0	80.00	100.00	60.00	80.00	80.00	80.00	60.00	20.00	0.0	80.00
HISTORICAL MIN VALUES :	36.00	42.00	65.00	123.00	128.00	156.00	64.00	32.00	21.00	6.00	6.00	26.00
PER CENT VALUES.GT.MIN HIST :	40.00	100.00	20.00	0.0	40.00	20.00	20.00	20.00	40.00	80.00	100.00	20.00
SYNTHETIC MAX MIN VALUES :	52.00	61.00	89.00	118.00	163.00	174.00	74.00	47.00	26.00	18.00	16.00	31.00
DISTRIBUTION :												
MEANS :	38.80	52.60	58.20	102.60	122.00	120.80	55.00	31.20	19.60	11.20	12.00	26.20
STD DEVS :	8.04	5.13	17.80	15.44	28.81	37.16	14.21	10.03	4.34	5.40	3.87	2.77
SKEWS :	1.52	1.27	1.86	-0.46	0.82	0.40	0.24	0.99	0.91	0.15	-0.43	1.88

MAXIMUM EXTREMES :

TIME PERIOD :	5	6	7	8	9	10	11	12	1	2	3	4
SYNTHETIC MIN MAX VALUES :	712.00	946.00	1275.00	1590.00	1168.00	1453.00	1030.00	559.00	294.00	193.00	269.00	597.00
PER CENT VALUES.LT.MAX HIST :	60.00	80.00	60.00	40.00	100.00	20.00	20.00	100.00	100.00	80.00	100.00	80.00
HISTORICAL MAX VALUES :	1124.00	1695.00	1585.00	1727.00	1421.00	1469.00	1033.00	787.00	400.00	258.00	384.00	1207.00
PER CENT VALUES.GE.MAX HIST :	40.00	20.00	40.00	60.00	0.0	80.00	80.00	0.0	0.0	20.00	0.0	20.00
SYNTHETIC MAX MAX VALUES :	1615.00	1736.00	2252.00	2617.00	1389.00	1544.00	1237.00	773.00	398.00	282.00	377.00	1580.00
DISTRIBUTION :												
MEANS :	1033.40	1492.40	1702.80	1951.60	1314.00	1498.40	1160.60	633.20	356.40	236.80	316.40	922.60
STD DEVS :	381.97	312.72	482.94	420.17	92.52	32.35	81.74	84.68	45.06	33.12	51.30	383.90
SKEWS :	1.05	- 1.97	0.57	1.18	- 1.27	- 0.10	- 1.21	1.48	-0.75	0.08	0.55	1.77

FIGURE 9.11 (CONT'D): SAMPLE OUTPUT FROM MOMENTS COMPARISON

```
FILE: FREQ      COMPARE  A   CMS REL3 PLC12 CANBERRA 370/158

405  3 50 50  5
150180330
  2  1 20 26   1 83
  DEMO1
 2160.000    1080.000    270.0
    0.0          0.0          0.0
  0174140 81 59 32 33 28 17  9  9  5  5  4  1  3  0  0  0  0  0  0  0  0150  1 1
  0304 92 55 36 26 24 21  9 14  7  4  4  1  0  0  1  1  0  1  0  0  0  0180  1 1
  0 88 93 90 85 74 54 32 33 23  6  7  5  2  3  4  1  0  0  0  0  0  0  0330  1 1
```

FIGURE 9.12: PROGRAM FREQ INPUT DATA DECK FOR EXAMPLE PROBLEM

***** FREQUENCY ANALYSIS DATA *****

SAMPLE: 1																				
150	0.	183.	140.	72.	50.	37.	33.	33.	19.	10.	9.	7.	4.	1.	2.	0.	0.	0.	0.	0.
180	0.	295.	97.	62.	32.	30.	24.	19.	16.	10.	3.	5.	2.	4.	0.	0.	0.	0.	0.	1.
330	0.	90.	90.	84.	91.	66.	56.	29.	42.	28.	9.	9.	1.	2.	0.	0.	2.	0.	0.	1.
SAMPLE: 2																				
150	0.	176	148.	73.	54.	39.	33.	23.	12.	11.	8.	10.	5.	5.	2.	0.	0.	0.	0.	1.
180	0.	298.	96.	67.	43.	26.	14.	10.	15.	6.	11.	2.	3.	2.	1.	1.	2.	0.	1.	2.
330	0.	97.	84.	82.	91.	70.	55.	40.	35.	22.	8.	3.	1.	4.	2.	2.	1.	0.	2.	1.
SAMPLE: 3																				
150	0.	177.	142.	72.	60.	41.	32.	18.	24.	11.	5.	9.	4.	2.	1.	1.	1.	0.	0.	0.
180	0.	302.	93.	58.	44.	28.	17.	16.	10.	13.	5.	4.	2.	4.	1.	0.	0.	1.	1.	1.
330	0.	96.	79.	105.	68.	69.	45.	56.	28.	27.	14.	5.	2.	2.	0.	1.	1.	1.	1.	0.
SAMPLE: 4																				
150	0.	173.	147.	79.	48.	43.	35.	26.	14.	13.	8.	3.	6.	1.	1.	1.	0.	0.	0.	2.
180	0.	300.	88.	64.	43.	35.	15.	18.	12.	8	6.	1.	2.	2.	1.	1.	2.	0.	0.	2.
330	0.	94.	95.	80.	70.	65.	58.	49.	37.	23.	13.	9.	5.	1.	0.	0.	0.	0.	0.	1.
SAMPLE: 5																				
150	0.	173.	142.	72.	54.	41.	37.	30.	25.	8.	5.	4.	3.	2.	2.	0.	0.	0.	1.	1.
180	0.	295.	95.	58.	42.	31.	29.	18.	10.	8.	4.	1.	1.	3.	1.	0.	0.	1.	1.	2.
330	0.	101.	77.	82.	79.	75.	48.	57.	31.	30.	9.	4.	4.	1.	1.	0.	1.	0.	0.	0.

FIGURE 9.13: SAMPLE OUTPUT FROM FREQUENCY DISTRIBUTION COMPARISON

***** FREQUENCY ANALYSIS STATION 150 *****

INDIVIDUAL SAMPLES:

RELATIVE FREQUENCY HISTOGRAMS

INTERVAL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
MIN OBS	0	173	140	72	48	37	32	18	12	8	5	3	3	1	1	0	0	0	0	0
MIN PROB (X 100)	0.0	28.83	23.33	12.00	8.00	6.17	5.33	3.00	2.00	1.33	0.83	0.50	0.50	0.17	0.17	0.0	0.0	0.0	0.0	0.0
HIST OBS	0	174	140	81	59	32	33	28	17	9	9	5	5	4	1	3	0	0	0	0
HIST PROB (X 100)	0.0	29.00	23.33	13.50	9.83	5.33	5.50	4.67	2.83	1.50	1.50	0.83	0.83	0.67	0.17	0.50	0.0	0.0	0.0	0.0
MAX OBS	0	183	148	79	60	43	37	33	25	13	9	10	6	5	2	1	1	0	1	2
MAX PROB (X 100)	0.0	30.50	24.67	13.17	10.00	7.17	6.17	5.50	4.17	2.17	1.50	1.67	1.00	0.83	0.33	0.17	0.17	0.0	0.17	0.33

CUMULATIVE FREQUENCY HISTOGRAMS

INTERVAL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
MIN OBS	0	173	315	387	441	482	515	542	558	569	577	587	592	596	597	598	598	598	598	600
MIN PROB (X 100)	0.0	28.83	52.50	64.50	73.50	80.33	85.83	90.33	93.00	94.83	96.17	97.83	98.67	99.33	99.50	99.67	99.67	99.67	99.67	100.00
HIST OBS	0	174	314	395	454	486	519	547	564	573	582	587	592	596	597	600	600	600	600	600
HIST PROB (X 100)	0.0	29.00	52.33	65.83	75.67	81.00	86.50	91.17	94.00	95.50	97.00	97.83	98.67	99.33	99.50	100.00	100.00	100.00	100.00	100.00
MAX OBS	0	183	324	399	451	492	525	551	574	582	587	593	597	598	600	600	600	600	600	600
MAX PROB (X 100)	0.0	30.50	54.00	66.50	75.17	82.00	87.50	91.83	95.67	97.00	97.83	98.83	99.50	99.67	100.00	100.00	100.00	100.00	100.00	100.00
REJECT 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P/C REJ.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
REJECT 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P/C REJ.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

TOTAL P/C REJECTION AT 5 P/C IS: 0.0

TOTAL P/C REJECTION AT 1 P/C IS: 0.0

FIGURE 9.13 (CONT'D): SAMPLE OUTPUT FROM FREQUENCY DISTRIBUTION COMPARISON

POPULATION RESULTS:

RELATIVE FREQUENCY HISTOGRAMS -

INTERVAL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
SYN OBS	0	176	144	74	53	40	34	26	19	11	7	7	4	2	2	0	0	0	0	1
SYN PROB (X 100)	0.0	29.40	23.97	12.27	8.87	6.70	5.67	4.33	3.13	1.77	1.17	1.10	0.73	0.37	0.27	0.07	0.03	0.0	0.03	0.13
HIST PROB	0.0	29.00	23.33	13.50	9.83	5.33	5.50	4.67	2.83	1.50	1.50	0.83	0.83	0.67	0.17	0.50	0.0	0.0	0.0	0.0
HIST OBS	0	174	140	81	59	32	33	28	17	9	9	5	5	4	1	3	0	0	0	0
PROB DIFF	0.0	0.40	0.63	-1.23	-0.97	1.37	0.17	-0.33	0.30	0.27	-0.33	0.27	-0.10	-0.30	0.10	-0.43	0.03	0.0	0.03	0.13

CUMULATIVE FREQUENCY HISTOGRAMS-

INTERVAL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
HIST OBS	0	174	314	395	454	486	519	547	564	573	582	587	592	596	597	600	600	600	600	600
HIST PROB	0.0	29.00	52.33	65.83	75.67	81.00	86.50	91.17	94.00	95.50	97.00	97.83	98.67	99.33	99.50	100.00	100.00	100.00	100.00	100.00
SYN PROB (X 100)	0.0	29.40	53.37	65.63	74.50	81.20	86.87	91.20	94.33	96.10	97.27	98.37	99.10	99.47	99.73	99.80	99.83	99.83	99.87	100.00
SYN OBS	0	176	320	394	447	487	521	547	566	577	584	590	595	597	598	599	599	599	599	600
K-S DIFF (X 100)	0.0	0.400	1.033	0.200	1.167	0.200	0.367	0.033	0.333	0.600	0.267	0.533	0.433	0.133	0.233	0.200	0.167	0.167	0.133	0.000
REJECT 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
REJECT 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P/C DIFF	0.0	1.38	1.97	-0.30	-1.54	0.25	0.42	0.04	0.35	0.63	0.27	0.55	0.44	0.13	0.23	-0.20	-0.17	-0.17	-0.13	0.00

MEAN ABSOLUTE P/C DIFFERENCE/CLASS INTERVAL 0.483

FIGURE 9.13 (CONT'D): SAMPLE OUTPUT FROM FREQUENCY DISTRIBUTION COMPARISON

FILE: RUNS COMPARE A CMS FEL3 FLC12 CANBERRA 370/158

```
406  3 50 50  5
150180330
  2  1  5 30 27
  1 84 85 86
  DEMO1
  8  5  0  9  6  9  6  8  4  2  0  0  0  0  0  0  0  0  0  0  0  0  0 0150  2 1
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 0150  2 2
  6  3  6  5 11  7 11  5  1  1  0  0  0  1  0  0  0  0  0  0  0  0  0 0150  3 1
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 0150  3 2
  5  3  5  6  6  5 14  7  3  1  0  0  0  0  0  0  0  0  0  0  0  0  0 0180  2 1
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 0180  2 2
  4  4  2  6 10 11 13  2  1  1  1  0  0  0  0  0  0  0  0  0  0  0  0 0180  3 1
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 0180  3 2
 64 23 21  8  4  6  3  0  1  0  1  0  0  0  0  0  0  0  0  0  0  0  0 0330  2 1
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 0330  2 2
 61 29 17 10  7  1  4  1  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0 0330  3 1
  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 0330  3 2
```

FIGURE 9.14: PROGRAM RUNS INPUT DATA DECK FOR EXAMPLE PROBLEM

***** RUNS DATA BELOW THE MEDIAN *****

SAMPLE	1																													
STN RUNS																														
150	6	4	6	6	7	7	8	6	2	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
180	9	2	3	8	13	9	8	2	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
330	64	25	8	9	8	4	2	2	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SAMPLE	2																													
STN RUNS																														
150	13	1	4	7	7	12	8	5	3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
180	11	3	6	3	11	9	10	4	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
330	55	31	13	9	9	1	1	0	4	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SAMPLE	3																													
STN RUNS																														
150	12	3	2	4	11	10	15	2	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
180	14	3	3	5	9	9	16	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
330	67	22	14	13	5	4	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SAMPLE	4																													
STN RUNS																														
150	6	2	8	5	11	5	10	8	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
180	8	4	5	11	11	7	9	2	2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
330	61	23	15	11	8	6	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SAMPLE	5																													
STN RUNS																														
150	3	3	4	7	6	12	8	6	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
180	9	2	6	4	13	11	6	5	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
330	53	31	7	12	5	4	1	1	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

FIGURE 9.15: SAMPLE OUTPUT FROM RUN LENGTH FREQUENCY DISTRIBUTION COMPARISON

***** RUNS BELOW THE MEDIAN : STATION 150

RUN LENGTH	INDIVIDUAL SAMPLES: RELATIVE FREQ MIN	RELATIVE FREQ MAX	CUMULATIVE FREQ MIN	CUMULATIVE FREQ MAX	REJECTION SMALL	REJECTION LARGE	POPULATION RESULTS: RELATIVE FREQ HIST	RELATIVE FREQ SYN	CUMULATIVE FREQ HIST	CUMULATIVE FREQ SYN	K-S	REJECTION SMALL	REJECTION LARGE	P/C DIFF
1	5.66	20.97	5.66	20.97	0	0	10.53	13.79	10.53	13.79	3.27	0	0	31.03
2	1.61	7.27	11.32	24.19	0	0	5.26	4.48	15.79	18.28	2.49	0	0	15.75
3	3.23	13.79	18.87	29.09	0	0	10.53	8.28	26.32	26.55	0.24	0	0	0.90
4	6.45	13.21	32.08	40.32	0	0	8.77	10.00	35.09	36.55	1.46	0	0	4.17
5	11.29	18.97	43.40	55.17	0	0	19.30	14.48	54.39	51.03	3.35	0	0	- 6.16
6	8.62	22.64	63.79	70.97	0	0	12.28	15.86	66.67	66.90	0.23	0	0	0.34
7	12.90	24.19	80.00	91.94	0	0	19.30	16.90	85.96	83.79	2.17	0	0	- 2.53
8	3.23	13.79	90.91	95.16	0	0	8.77	9.31	94.74	93.10	1.63	0	0	- 1.72
9	1.61	4.84	94.34	98.28	0	0	1.75	3.10	96.49	96.21	0.28	0	0	- 0.29
10	0.0	1.89	96.23	98.39	0	0	1.75	1.38	98.25	97.59	0.66	0	0	- 0.67
11	0.0	1.61	96.23	100.00	0	0	0.0	0.34	98.25	97.93	0.31	0	0	- 0.32
12	0.0	1.89	98.11	100.00	0	0	0.0	1.38	98.25	99.31	1.06	0	0	1.08
13	0.0	0.0	98.11	100.00	0	0	0.0	0.0	98.25	99.31	1.06	0	0	1.08
14	0.0	0.0	98.11	100.00	0	0	1.75	0.0	100.00	99.31	0.69	0	0	- 0.69
15	0.0	0.0	98.11	100.00	0	0	0.0	0.0	100.00	99.31	0.69	0	0	- 0.69
16	0.0	0.0	98.11	100.00	0	0	0.0	0.0	100.00	99.31	0.69	0	0	- 0.69
17	0.0	0.0	98.11	100.00	0	0	0.0	0.0	100.00	99.31	0.69	0	0	- 0.69
18	0.0	1.89	100.00	100.00	0	0	0.0	0.69	100.00	100.00	0.00	0	0	- 0.00
19	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
20	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
21	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
22	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
23	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
24	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
25	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
26	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
27	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
28	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
29	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
30	0.0	0.0	100.00	100.00	0	0	0.0	0.0	100.00	100.00	0.00	0	0	- 0.00
TOTALS:								100.00						68.82

STATISTIC	MEAN	STD DEV	SKEW	MIN	MAX	POP
TOTAL RUNS :	58.000	4.062	-0.149	53.000	62.000	290.000
AVERAGE RUN LENGTH :	5.242	0.331	0.777	4.935	5.736	5.224
TOTAL LONG RUNS :	4.000	1.000	0.0	3.000	5.000	20.000
AVERAGE LONG RUN LENGTH :	10.757	1.121	0.554	9.600	12.250	10.800
K-S STAT 5 :						7.986
K-S STAT 1 :						9.572

FIGURE 9.15 (CONT'D): SAMPLE OUTPUT FROM RUN LENGTH FREQUENCY DISTRIBUTION COMPARISO

```
FILE: YIELD    COMPARE  A   CMS REL3 FLC12 CANBERRA 370/158

407  3  5  2 50 50  1  2  1
150180330
  1 87 88 89 90 91
 33 34 35 36 30
  DFM01
     0.600     6.122     20      20                                                     150   8 1
    18.000    18.000     6.000                                                          150   8 2
    43.000     1.744    43.000     7.047    42.000    10.524    42.000      3.381150   8 3
     0.800    14.244     78      10                                                     150   8 1
    56.000    24.000    14.000                                                          150   8 2
    22.000     4.218    22.000    16.091    21.000    24.905    21.000      8.238150   8 3
     0.600     9.196     29      16                                                     180   8 1
    28.000    14.000    10.000                                                          180   8 2
    35.000     3.184    35.000     9.543    34.000    14.265    34.000      4.559180   8 3
     0.800    19.155     79      10                                                     180   8 1
    64.000    24.000    20.000                                                          180   8 2
    23.000     5.637    23.000    15.174    22.000    24.636    22.000      9.000180   8 3
     0.600     2.641      6      12                                                     330   8 1
    10.000    10.000     4.000                                                          330   8 2
    93.000     0.623    93.000     2.172    92.000     4.043    92.000      1.870330   8 3
     0.800     5.903     20      20                                                     330   8 1
    18.000    18.000     6.000                                                          330   8 2
    63.000     1.329    63.000     4.476    62.000     7.968    62.000      3.452330   8 3
  0  4  6  4  9  4  4  4  2  0  1  1  0  0  0  0  1  0  1  2  0  0  0  0150   4 1
  0  7  6  4  0  1  2  1  0  1  0  0  0  0  0  0  0  0  0  0  0  0  9  0150   4 1
  0  1  6  4  9  6  2  3  0  0  1  0  0  3  0  0  0  0  0  0  0  0  0  0180   4 1
  0 10  6  5  1  0  1  0  0  0  0  0  0  0  0  0  0  0  9  0  0  0  0  0180   4 1
  0 48 23  6  7  3  5  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0330   4 1
  0 16 13  3 11  5  3  2  0  2  2  0  1  0  0  2  1  1  1  0  0  0  0  0330   4 1
  0  2  2  6  4  9  7  6  1  0  0  0  0  0  0  0  1  1  0  4  0  0  0  0150   5 1
  0 14  0  3  1  0  2  1  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0150   5 1
  0  2  2 14  8  0  0  0  1  3  2  1  0  0  2  0  0  0  0  0  0  0  0  0180   5 1
  0 13  1  6  1  0  1  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0180   5 1
  0 53 16  9  6  3  2  2  0  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0330   5 1
  0 25  9  8  1  5  2  3  4  0  0  0  1  0  0  1  0  0  0  4  0  0  0  0330   5 1
  0  0  2  1  2  2  2  5 10  2  6  3  1  0  0  0  0  0  0  6  0  0  0  0150   6 1
  0  4  8  0  1  1  0  1  1  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0150   6 1
  0  0  2  1  7  8  5  0  0  0  4  2  1  0  0  4  0  0  0  0  0  0  0  0180   6 1
  0  5  6  5  2  1  2  0  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0  0180   6 1
  0  0 42 14 13  3  3  4  4  3  2  4  0  0  0  0  0  0  0  0  0  0  0  0330   6 1
  0  0 18 10  5  5  1  6  1  1  2  1  0  1  1  1  2  0  0  7  0  0  0  0330   6 1
  0  4 13  9  7  6  1  1  0  0  0  0  0  0  1  0  0  0  0  0  0  0  0  0150   7 1
  0  9  7  1  1  0  1  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0150   7 1
  0  3  8  9  5  2  3  0  0  0  0  1  1  1  0  1  0  0  0  0  0  0  0  0180   7 1
  0  8  8  2  2  0  0  0  0  2  0  0  0  0  0  0  0  0  0  0  0  0  0  0180   7 1
  0 55 20  8  3  2  1  1  1  0  1  0  0  0  0  0  0  0  0  0  0  0  0  0330   7 1
  0 25 18  5  3  0  0  0  5  0  0  1  1  0  1  1  0  2  0  0  0  0  0  0330   7 1
```

FIGURE 9.16: PROGRAM YIELD INPUT DATA DECK FOR EXAMPLE PROBLEM

***** RIPPL STORAGE ANALYSIS STATION 150 *****

***** YIELD 0.6000 *****

DISTRIBUTION/		PARAMETERS		
	MEAN	STD DEV	SKEW	COEFF VAR
SYN STORAGE :	6.8086	0.6795	-0.1614	0.0998
HISTORICAL :	6.1220			
SYN DROUGHT :	34.6000	19.4628	1.3146	0.5625
HISTORICAL :	20.0000			

ORDERED RIPPL STORAGES --------

RIPPL STORAGE	PROBABILITY .LE. RIPPL	DROUGHT DURATION	SAMPLE NUMBER	RIPPL STORAGE	PROBABILITY .LE. RIPPL	DROUGHT DURATION	SAMPLE NUMBER	RIPPL STORAGE	PROBABILITY .LE. RIPPL	DROUGHT DURATION	SAMPLE NUMBER
5.9578	0.1667	39	3	6.3132	0.3333	20	4	6.8621	0.5000	18	1
7.3154	0.6667	66	2	7.5946	0.8333	30	5				

NO OF STANDARD DEVIATIONS HISTORICAL RIPPL
STORAGE IS FROM SYNTHETIC DISTRIBUTION MEAN
IS - 1.0104

***** YIELD 0.8000 *****

DISTRIBUTION/		PARAMETERS		
	MEAN	STD DEV	SKEW	COEFF VAR
SYN STORAGE :	15.5333	5.8770	1.8880	0.3783
HISTORICAL :	14.2440			
SYN DROUGHT :	58.4000	30.9887	1.8806	0.5306
HISTORICAL :	78.0000			

ORDERED RIPPL STORAGES---------

RIPPL STORAGE	PROBABILITY .LE. RIPPL	DROUGHT DURATION	SAMPLE NUMBER	RIPPL STORAGE	PROBABILITY .LE. RIPPL	DROUGHT DURATION	SAMPLE NUMBER	RIPPL STORAGE	PROBABILITY .LE. RIPPL	DROUGHT DURATION	SAMPLE NUMBER
10.5697	0.1667	35	4	13.2975	0.3333	55	1	13.8684	0.5000	40	5
14.2008	0.6667	50	3	25.7300	0.8333	112	2				

NO OF STANDARD DEVIATIONS HISTORICAL RIPPL
STORAGE IS FROM SYNTHETIC DISTRIBUTION MEAN
IS - 0.2194

FIGURE 9.17: SAMPLE OUTPUT FROM STORAGE-YIELD BEHAVIOUR COMPARISON

***** D E F I C I T F R E Q U E N C Y A N A L Y S I S *****

STN	150	150	180	180	330	330	150	150	180	180	330	330	150	150	180	180	330	330	150	150	180	180	330	330	150	150	180	180	330	330
SAM	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5
YLD	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	5	6	2	9	39	20	5	6	3	12	37	17	3	6	3	9	33	20	4	5	1	11	43	18	3	10	3	14	38	17
3	2	11	4	2	17	14	4	11	4	6	14	17	5	12	4	6	16	7	1	16	1	6	21	14	5	5	6	3	20	14
4	5	2	6	2	6	8	7	6	8	0	12	5	5	6	9	6	10	8	5	4	11	3	10	5	9	5	7	1	7	8
5	7	0	11	1	5	4	3	1	12	1	1	1	5	3	9	1	9	4	7	2	11	2	8	5	4	2	3	2	5	2
6	5	1	3	3	3	1	8	0	6	0	4	5	7	2	3	0	4	5	7	4	6	1	7	3	5	0	9	1	5	3
7	2	2	0	0	3	4	4	0	0	1	5	2	8	0	4	1	4	3	5	0	4	0	0	5	2	3	0	0	1	3
8	4	2	1	0	2	1	4	0	0	0	0	2	1	0	4	0	0	5	4	1	1	0	1	4	6	0	0	0	3	2
9	3	1	1	0	2	1	4	1	1	0	2	2	5	0	0	0	1	1	4	0	0	0	0	2	3	1	2	0	0	2
10	1	0	0	0	0	2	0	1	0	1	1	1	2	1	1	0	0	2	2	0	1	0	0	2	2	0	0	0	2	2
11	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	2	0	0	2	0	0	2	0	0	0
12	2	0	2	0	1	0	2	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	1	0	1	0	1	0	0	1
13	1	0	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
14	0	0	2	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
15	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0
16	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
19	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	2	0	0	0	0	5	1	0	0	0	0	2	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	2

FIGURE 9.17 (CONT'D): SAMPLE OUTPUT FROM STORAGE-YIELD BEHAVIOUR COMPARISON

***** D E F I C I T   F R E Q U E N C Y   A N A L Y S I S *****

S T A T I O N   150

***** Y I E L D: 0.600 *****   DEFICIT UPPER BOUND 6.00

BIN NUMBER	INDIVIDUAL SAMPLES RELATIVE FREQ MIN	MAX	CUMULATIVE FREQ MIN	MAX	REJECTION SMALL	LARGE	POPULATION RESULTS RELATIVE FREQ HIST	SYN	CUMULATIVE FREQ HIST	SYN	K-S	REJECTION SMALL	LARGE	P/C DIFF
1	0.0	0.0	0.0	0.0	0	0	0.0	0.0	0.0	0.0	0.0	0	0	0.0
2	6.67	12.20	6.67	12.20	0	0	9.30	9.30	9.30	9.30	0.0	0	0	0.0
3	2.33	11.63	11.63	20.93	0	0	13.95	7.91	23.26	17.21	6.05	0	0	-26.00
4	11.11	20.93	23.26	39.53	0	0	9.30	14.42	32.56	31.63	0.93	0	0	- 2.86
5	6.98	17.07	39.53	48.84	0	0	20.93	12.09	53.49	43.72	9.77	0	1	- 18.26
6	11.63	18.60	55.56	62.79	0	0	9.30	14.88	62.79	58.60	4.19	0	0	- 6.67
7	4.65	17.78	63.41	73.33	0	0	9.30	9.77	72.09	68.37	3.72	0	0	- 5.16
8	2.22	13.95	73.17	81.40	0	0	9.30	8.84	81.40	77.21	4.19	0	0	- 5.14
9	6.98	11.11	80.49	90.70	0	0	4.65	8.84	86.05	86.05	0.00	0	0	- 0.00
10	0.0	4.65	82.93	91.11	0	0	0.0	3.26	86.05	89.30	3.26	0	0	3.78
11	0.0	2.44	85.37	93.33	0	0	2.33	0.93	88.37	90.23	1.86	0	0	2.11
12	2.22	4.88	90.24	95.56	0	0	2.33	3.26	90.70	93.49	2.79	0	0	3.08
13	0.0	2.44	92.68	97.67	0	0	0.0	1.40	90.70	94.88	4.19	0	0	4.62
14	0.0	2.33	92.68	97.78	0	0	0.0	0.93	90.70	95.81	5.12	0	0	5.64
15	0.0	0.0	92.68	97.78	0	0	0.0	0.0	90.70	95.81	5.12	0	0	5.64
16	0.0	2.44	93.02	97.78	0	0	0.0	0.47	90.70	96.28	5.58	0	0	6.15
17	0.0	2.33	95.12	97.78	0	0	2.33	0.47	93.02	96.74	3.72	0	0	4.00
18	0.0	2.33	95.12	97.78	0	0	0.0	0.47	93.02	97.21	4.19	0	0	4.50
19	0.0	2.22	95.12	100.00	0	0	2.33	0.47	95.35	97.67	2.33	0	0	2.44
20	0.0	4.88	100.00	100.00	0	0	4.65	2.33	100.00	100.00	0.00	0	0	- 0.00
TOTAL :														106.04

STATISTIC	MEAN	STD DEV	SKEW	MIN	MAX	POP
TOTAL :	43.000	1.414	0.0	41.000	45.000	215.000
AVERAGE :	1.714	0.105	- 0.065	1.571	1.855	
K-S STAT 5:						9.275
K-S STAT 1:						11.117

***** Y I E L D :   0.800 *****   DEFICIT UPPER BOUND 14.00

BIN NUMBER	INDIVIDUAL SAMPLES RELATIVE FREQ MIN	MAX	CUMULATIVE FREQ MIN	MAX	REJECTION SMALL	LARGE	POPULATION RESULTS RELATIVE FREQ HIST	SYN	CUMULATIVE FREQ HIST	SYN	K-S	REJECTION SMALL	LARGE	P/C DIFF
1	0.0	0.0	0.0	0.0	0	0	0.0	0.0	0.0	0.0	0.0	0	0	0.0
2	15.63	38.46	15.63	38.46	0	0	31.82	23.74	31.82	23.74	8.08	0	0	- 25.39
3	19.23	50.00	57.69	68.00	0	0	27.27	39.57	59.09	63.31	4.22	0	0	7.14
4	8.00	23.08	76.00	88.46	0	0	18.18	16.55	77.27	79.86	2.58	0	0	3.34
5	0.0	10.00	76.00	92.31	0	0	0.0	5.76	77.27	85.61	8.34	0	0	10.79
6	0.0	12.50	80.00	96.88	0	0	4.55	5.04	81.82	90.65	8.83	0	0	10.79
7	0.0	11.54	88.00	96.88	0	0	9.09	3.60	90.91	94.24	3.34	0	0	3.67
8	0.0	8.00	92.31	100.00	0	0	4.55	2.16	95.45	96.40	0.95	0	0	0.99
9	0.0	4.00	96.15	100.00	0	0	0.0	2.16	95.45	98.56	3.11	0	0	3.25
10	0.0	3.85	100.00	100.00	0	0	4.55	1.44	100.00	100.00	0.00	0	0	0.00
TOTAL :														65.37

STATISTIC	MEAN	STD DEV	SKEW	MIN	MAX	POP
TOTAL :	27.800	3.033	0.760	25.000	32.000	139.000
AVERAGE :	3.862	0.238	0.549	3.600	4.190	
K-S STAT 5 :						11.535
K-S STAT 1 :						13.825

FIGURE 9.17 (CONT'D). SAMPLE OUTPUT FROM STORAGE-YIELD BEHAVIOUR COMPARISON

```
FILE: CORREL   COMPARE  A   CMS FEL3 PLC12 CANBERRA 370/158

404  1 12  1 12  2  1
  1  1  1  1 50 50  5  3
150180330
 31 32 92 93  1  2  3
  DEMO1
     0.72523     0.48541     0.20937    -0.06333    -0.22520    -0.30122150 17 1
    -0.28583    -0.15657     0.00163     0.18090     0.32022     0.37877150 17 2
     0.67983     0.43568     0.16565    -0.09231    -0.24916    -0.31992180 17 1
    -0.28962    -0.15213    -0.01159     0.15839     0.30345     0.33679180 17 2
     0.22233     0.12806     0.04197    -0.04322    -0.02672    -0.07999330 17 1
    -0.06744    -0.01820    -0.03527     0.02145     0.03095    -0.00059330 17 2
     0.32449     0.34145     0.25909     0.09184    -0.09665    -0.26789  1 18 1
    -0.33037    -0.29937    -0.17516     0.06871     0.31852     0.60060  1 18 2
     0.91278     0.72438     0.58846     0.31812     0.02758    -0.17275  1 18 3
    -0.29135    -0.30193    -0.20513    -0.08576     0.08950     0.25139  1 18 4
     0.33451     0.0         0.0         0.0         0.0         0.0      1 18 5
     0.09409     0.12388     0.10483     0.05392    -0.02331    -0.10167  2 18 1
    -0.18640    -0.17077    -0.14898    -0.08745     0.03445     0.13249  2 18 2
     0.46840     0.45200     0.36454     0.32046     0.16821     0.07094  2 18 3
    -0.01113    -0.09683    -0.09014    -0.06776     0.01379     0.06479  2 18 4
     0.10994     0.0         0.0         0.0         0.0         0.0      2 18 5
     0.10607     0.14949     0.09615     0.01047    -0.05763    -0.12315  3 18 1
    -0.20965    -0.18171    -0.11127    -0.03060     0.10420     0.15112  3 18 2
     0.47102     0.46967     0.34343     0.28247     0.12960     0.00470  3 18 3
    -0.08747    -0.12917    -0.08277    -0.04887     0.04594     0.09264  3 18 4
     0.12687     0.0         0.0         0.0         0.0         0.0      3 18 5
```

FIGURE 9.18: PROGRAM CORREL INPUT DATA DECK FOR EXAMPLE PROBLEM

FIGURE 9.19: SAMPLE OUTPUT FROM CORRELATIONS COMPARISON

***** SYNTHETIC CORRELATION ANALYSIS DATA *****

SAMPLE 1

1	0.73	0.43	0.15	-0.11	-0.29	-0.35	-0.30	-0.17	0.03	0.24	0.43	0.49
2	0.69	0.42	0.11	-0.14	-0.30	-0.35	-0.30	-0.16	0.02	0.24	0.39	0.48
3	0.23	0.17	0.14	0.08	0.08	0.03	0.03	0.04	0.03	0.03	0.06	0.09

SAMPLE 2

1	0.72	0.41	0.10	-0.14	-0.30	-0.36	-0.32	-0.20	0.01	0.21	0.40	0.48
2	0.68	0.36	0.06	-0.13	-0.27	-0.30	-0.28	-0.18	-0.01	0.17	0.34	0.45
3	0.21	0.11	0.12	0.02	-0.03	-0.01	-0.04	0.06	0.04	0.03	0.09	0.05

SAMPLE 3

1	0.72	0.43	0.14	-0.12	-0.29	-0.37	-0.35	-0.22	-0.01	0.23	0.39	0.46
2	0.69	0.40	0.11	-0.13	-0.27	-0.33	-0.30	-0.20	-0.00	0.22	0.36	0.40
3	0.15	0.15	0.04	0.03	-0.06	-0.07	-0.05	-0.08	0.07	0.01	0.10	0.10

SAMPLE 4

1	0.70	0.40	0.08	-0.17	-0.32	-0.39	-0.35	-0.22	-0.04	0.18	0.32	0.42
2	0.68	0.38	0.06	-0.15	-0.28	-0.33	-0.30	-0.20	-0.03	0.16	0.29	0.36
3	0.21	0.12	0.02	-0.01	-0.06	-0.03	-0.03	-0.10	-0.08	0.03	0.11	0.08

SAMPLE 5

1	0.70	0.39	0.12	-0.12	-0.29	-0.34	-0.31	-0.17	0.01	0.20	0.33	0.39
2	0.61	0.30	0.09	-0.12	-0.26	-0.31	-0.27	-0.15	0.04	0.17	0.31	0.35
3	0.24	0.22	0.10	0.04	-0.02	-0.01	-0.02	-0.02	-0.01	0.02	0.01	0.00

**** A U T O   C O R R E L A T I O N   A N A L Y S I S   R E S U L T S   S T A T I O N   150 *****

COMPARISON OF SYNTHETIC SAMPLE AND HISTORICAL POPULATION CORRELATIONS

LAG	MIN SYN	BELOW 1P/C BOUND	1P/C BOUND	BELOW 5P/C BOUND	5P/C BOUND	HIST	5P/C BOUND	ABOVE 5P/C BOUND	1P/C BOUND	ABOVE 1P/C BOUND	MAX SYN	STATISTICALLY DIFFERENT FROM 0.0 SYNTHETIC 1P/C	SYNTHETIC 5P/C	HISTORICAL 1P/C	HISTORICAL 5P/C	MEAN SYN	SYN -HIST	RELATIVE ERROR
1	0.698	0	0.671	0	0.685	0.725	0.761	0	0.772	0	0.728	5	5	1	1	0.714	-0.011	-1.541
2	0.393	2	0.401	3	0.422	0.485	0.544	0	0.562	0	0.433	5	5	1	1	0.413	-0.072	-14.884
3	0.083	2	0.106	3	0.131	0.209	0.285	0	0.308	0	0.150	3	5	1	1	0.120	-0.089	-42.746
4	-0.168	1	-0.168	1	-0.143	-0.063	0.017	0	0.042	0	-0.108	5	5	0	0	-0.130	-0.067	105.433
5	-0.322	0	-0.323	1	-0.300	-0.225	-0.147	0	-0.123	0	-0.285	5	5	1	1	-0.296	-0.070	31.254
6	-0.386	0	-0.394	2	-0.373	-0.301	-0.226	0	-0.202	0	-0.342	5	5	1	1	-0.363	-0.062	20.530
7	-0.348	0	-0.380	0	-0.358	-0.286	-0.210	0	-0.186	0	-0.301	5	5	1	1	-0.326	-0.040	13.991
8	-0.220	0	-0.258	0	-0.234	-0.157	-0.077	0	-0.052	0	-0.166	5	5	1	1	-0.195	-0.039	24.772
9	-0.036	0	-0.104	0	-0.079	0.002	0.082	0	0.107	0	0.027	0	0	0	0	0.000	-0.001	0.0
10	0.179	0	0.076	0	0.102	0.181	0.258	0	0.281	0	0.243	5	5	1	1	0.212	0.031	17.299
11	0.323	0	0.222	0	0.246	0.320	0.391	3	0.412	1	0.431	5	5	1	1	0.376	0.055	17.291
12	0.391	0	0.284	0	0.307	0.379	0.446	3	0.466	2	0.491	5	5	1	1	0.447	0.068	18.059
TOTALS:																		307.800

FIGURE 9.19 (CONT'D): SAMPLE OUTPUT FROM CORRELATIONS COMPARISON

**** C R O S S   C O R R E L A T I O N   A N A L Y S I S -- S T A T I O N 150 & S T A T I O N 180 ******

COMPARISON OF SYNTHETIC SAMPLE AND HISTORICAL POPULATION CORRELATIONS

LAG	MIN SYN	BELOW 1 P/C BOUND	1 P/C BOUND	BELOW 5 P/C BOUND	5 P/C BOUND	HIST	5 P/C BOUND	ABOVE 5 P/C BOUND	1 P/C BOUND	ABOVE 1 P/C BOUND	MAX SYN	STATISTICALLY DIFFERENT FROM 0.0 SYNTHETIC 1P/C	SYNTHETIC 5P/C	HISTORICAL 1P/C	HISTORICAL 5P/C	MEAN SYN	MEAN SYN -HIST	P/C RELATIVE ERROR
-12	0.339	0	0.226	0	0.250	0.324	0.395	3	0.416	2	0.444	5	5	1	1	0.398	0.074	22.761
-11	0.332	0	0.244	0	0.268	0.341	0.411	2	0.432	1	0.453	5	5	1	1	0.398	0.057	16.555
-10	0.256	0	0.158	0	0.182	0.259	0.333	2	0.355	1	0.357	5	5	1	1	0.312	0.053	20.536
-9	0.064	0	-0.014	0	0.011	0.092	0.171	0	0.196	0	0.162	4	4	0	1	0.121	0.029	31.566
-8	-0.131	0	-0.200	0	-0.176	-0.097	-0.016	0	0.009	0	-0.063	2	3	0	1	-0.097	-0.001	0.846
-7	-0.288	0	-0.363	0	-0.341	-0.268	-0.192	0	-0.167	0	-0.235	5	5	1	1	-0.261	0.007	-2.694
-6	-0.354	0	-0.421	0	-0.400	-0.330	-0.257	0	-0.233	0	-0.312	5	5	1	1	-0.336	-0.005	1.580
-5	-0.322	0	-0.392	0	-0.371	-0.299	-0.224	0	-0.200	0	-0.302	5	5	1	1	-0.313	-0.014	4.612
-4	-0.209	0	-0.275	0	-0.252	-0.175	-0.096	0	-0.071	0	-0.180	5	5	1	1	-0.193	-0.018	10.375
-3	-0.012	0	-0.037	0	-0.012	0.069	0.148	0	0.173	0	0.041	0	0	0	0	0.012	-0.056	-82.143
-2	0.233	0	0.221	1	0.245	0.319	0.389	0	0.410	0	0.320	5	5	1	1	0.285	-0.034	-10.564
-1	0.559	0	0.529	0	0.547	0.601	0.649	0	0.664	0	0.640	5	5	1	1	0.612	0.011	1.864
0	0.881	2	0.893	2	0.898	0.913	0.925	0	0.929	0	0.912	5	5	1	1	0.899	-0.014	-1.499
1	0.689	0	0.670	0	0.684	0.724	0.760	0	0.771	0	0.730	5	5	1	1	0.711	-0.013	-1.830
2	0.459	4	0.515	4	0.533	0.588	0.639	0	0.653	0	0.552	5	5	1	1	0.497	-0.091	-15.537
3	0.174	2	0.220	3	0.244	0.318	0.388	0	0.410	0	0.270	5	5	1	1	0.221	-0.097	-30.424
4	-0.085	1	-0.078	2	-0.053	0.028	0.108	0	0.133	0	-0.034	0	1	0	0	-0.052	-0.079	654.585
5	-0.270	0	-0.273	2	-0.250	-0.173	-0.094	0	-0.069	0	-0.229	5	5	1	1	-0.246	-0.073	42.357
6	-0.362	0	-0.385	0	-0.363	-0.291	-0.216	0	-0.192	0	-0.326	5	5	1	1	-0.344	-0.053	18.082
7	-0.354	0	-0.395	0	-0.373	-0.302	-0.227	0	-0.203	0	-0.327	5	5	1	1	-0.340	-0.038	12.652
8	-0.271	0	-0.304	0	-0.281	-0.205	-0.127	0	-0.102	0	-0.219	5	5	1	1	-0.252	-0.047	23.033
9	-0.111	0	-0.190	0	-0.165	-0.086	-0.005	0	0.020	0	-0.068	1	4	0	1	-0.092	-0.006	7.170
10	0.067	0	-0.017	0	0.009	0.090	0.169	0	0.194	0	0.123	2	4	0	1	0.099	0.009	10.522
11	0.225	0	0.149	0	0.174	0.251	0.326	0	0.348	0	0.320	5	5	1	1	0.274	0.023	9.187
12	0.347	0	0.237	0	0.261	0.335	0.404	2	0.425	1	0.459	5	5	1	1	0.396	0.061	18.321
																		1051.294

FIGURE 9.19 (CONT'D): SAMPLE OUTPUT FROM CORRELATIONS COMPARISON

STATISTICAL PROPERTY CALCULATED	HISTORICAL DATA SET ANALYSIS	SYNTHETIC DATA SET ANALYSIS	STATISTICAL PROPERTY COMPARISON	TOTAL TIME
MOMENTS	0.37	7.55	1.45	9.37
FREQUENCY	0.34	7.65	0.47	8.46
RUNS	0.79	6.90	1.50	9.19
STORAGE-YIELD	0.49	12.87	4.11	17.47
AUTO & CROSS CORREL-ATIONS	1.93 (0.25)	81.55 (9.40)	4.55 (0.50)	88.03 (10.15)
TOTAL TIME	3.92 (2.24)	116.52 (44.37)	12.08 (8.03)	132.52 (54.64)

Note: Auto and cross correlation times are unrepresentative as explained in text. Representative times are given in brackets.

TABLE 1.1: SAMPLE PROGRAM EXECUTION TIMES (MINUTES)

PROGRAM	STORAGE REQUIREMENTS (K Bytes)
STATS	156
MOMENT	186
FREQ	144
RUNS	176
YIELD	244
CORREL	144

Note: Storage requirements given are the maximum of the link-edit and go step requirements.

TABLE 1.2: PROGRAM STORAGE REQUIREMENTS

STATISTICS	FILE PROGRAM NAME	ASSIGNED FORTRAN UNIT NO.	STATISTIC CODE NUMBER	COMPARISON PROGRAM
Monthly first three moments and extrema	NPU1	1	10 - 14	
Annual first three moments	NPU2	2	15	MOMENT
Time series first three moments	NPU3	3	16	
Time series frequency distributions	NPU4	4	1	FREQ
Run length frequency distributions	NPU9	9	2 - 3	RUNS
Yield level and corresponding Rippl storage and drought duration, frequency distribution bounds, statistics, and number of bins	NPU10	10	8	
Storage deficit frequency distribution	NPU11	11	4	YIELD
Drought duration frequency distribution	NPU12	12	5	
Draw duration frequency distribution	NPU13	13	6	
Fill duration frequency distribution	NPU14	14	7	
Auto correlation and cross correlation functions	NPU15	15	17 - 18	CORREL

TABLE 3.1: PROGRAM STATS PUNCHED CARD OUTPUT

FILE TYPE	FILE PROGRAM NAME	ASSIGNED FORTRAN UNIT NO.	ASSIGNED DEVICE TYPE	DESCRIPTION
INPUT	NIN	5	Card reader	A sequential card input file defining program options, problem data and, optionally, the hydrologic data to be analysed.
	INFLOW	8	Card, tape or disk	An optional, sequential, card image input file containing the hydrologic data to be analysed.
OUTPUT	NOUT	6	Line printer	A line printer file consisting of (a) an echo check of the problem specification, (b) an optional echo check of the analysed hydrologic data, (c) tabulated and, optionally, plotted results of the user specified analyses of the hydrologic data.
	NPUN	7	Card punch	An optional, sequential, card output file of user selected hydrologic data statistics, such as, moments, frequencies, correlations, runs and storage-yields suitable for input to programs MOMENT, FREQ, CORREL, RUNS and YIELD.
	NF20	20	Disk	An output file of synthetic hydrologic data monthly means suitable for input to program MOMENT.
	NF21	21	Disk	An output file of synthetic hydrologic data monthly standard deviations suitable for input to program MOMENT.
	NF22	22	Disk	An output file of synthetic hydrologic data monthly skew coefficients suitable for input to program MOMENT.
	NF23	23	Disk	An output file of synthetic hydrologic data monthly extrema suitable for input to program MOMENT.
	NF24	24	Disk	An output file of synthetic hydrologic data annual means, standard deviations and skew coefficients suitable for input to program MOMENT.
	NF25	25	Disk	An output file of synthetic hydrologic data time series means, standard deviations and skew coefficients suitable for input to program MOMENT.
	NF26	26	Disk	An output file of synthetic hydrologic data time series frequency distributions suitable for input to program FREQ.
	NF27	27	Disk	An output file of synthetic hydrologic data run length frequency distributions suitable for input to program RUNS.
	NF28	28	Disk	An output file of synthetic hydrologic data aggregate minima and maxima values.
	NF29	29	Disk	An output file of synthetic hydrologic data surplus and deficit statistics.
	NF30	30	Disk	An output file of synthetic hydrologic data Rippl storages and drought durations suitable for input to program YIELD.
	NF31	31	Disk	An output file of synthetic hydrologic data auto correlation functions suitable for input to program CORREL.
	NF32	32	Disk	An output file of synthetic hydrologic data cross correlation functions suitable for input to program CORREL.
	NF33	33	Disk	An output file of synthetic hydrologic data deficit frequency distributions suitable for input to program YIELD.
	NF34	34	Disk	An output file of synthetic hydrologic data drought duration frequency distributions suitable for input to program YIELD.
	NF35	35	Disk	An output file of synthetic hydrologic data draw duration frequency distributions suitable for input to program YIELD.

Note: Files NF20, NF21,....., NF36 are all sequential, unformatted, variable record length files.

FILE TYPE	FILE PROGRAM NAME	ASSIGNED FORTRAN UNIT NO.	ASSIGNED DEVICE TYPE	DESCRIPTION
OUTPUT	NF36	36	Disk	An output file of synthetic hydrologic data fill duration frequency distributions suitable for input to program YIELD.
SCRATCH	NFILE(1) to NFILE(25)	50 to 74	System scratch disk	Sequential, unformatted, variable record length scratch files, used to store modified time series of station 1 to station NSTN respectively, used in the calculation of cross correlation functions.
	NPU1 to NPU4; NPU9 to NPU15	1 to 4; 9 to 15	System scratch disk	Sequential card image files used to temporarily store hydrologic data statistics, such as, moments, frequencies, correlations, runs and storage-yield distributions for subsequent punching in appropriate order for statistic comparison programs.

Note: Files NF20, NF21, ..., NF36 are all sequential, unformatted, variable record length files.

TABLE 3.2 (CONT'D): PROGRAM STATS FILE DESCRIPTION

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
1	1 - 3	Specify the number of sets of hydrologic data to be analysed.	NJOB	1-999	I3
	6	Specify whether the program options and problem data description for each set of hydrologic data are the same by coding '1', else '0'.	NOPT	0, 1	I3
2	1 - 3	A user integer code number to identify this run.	NRUN	0-999	I3
	6	Specify an echo check of the hydrologic data analysed by coding '1', else '0'.	IDATA	0, 1	I3
	8 - 9	Specify the unit number of the file from which the hydrologic data are to be read if different from the default unit number, else code '0'.	ITP	0-99	I 3
3	1 -80	Specify format of station data in the hydrologic data file.	FMT	-	20A4
4	2 - 3	Specify the number of stations in the analysis.	NSTN	1 -25	I 3
	4 - 6	Specify the common number of years of data for each station in the analysis.	NYRS	≤100	I 3
5	3	Specify the Kendall Tau test for trend by coding '1', else '0'.	IKEN	0, 1	I3
6	3	Specify calculation of monthly moments and extrema by coding '1'. Specify calculation of monthly moments and extrema, and monthly frequency distributions by coding '2'. Else code '0'.	MOMON	0, 1, 2	I 3
	6	Specify that monthly means, standard deviations and skew coefficients are to be written to disk files NF20, NF21 and NF22 respectively by coding '1', else code '0'.	KKMOM	0, 1	I 3
	9	Specify that monthly minima and maxima are to be written to disk file NF23 by coding '1', else code ' 0'.	KKMEXT	0, 1	I 3
	12	Specify that monthly means, standard deviations and skew coefficients are to be punched by coding '1' , else code '0'.	KPMOM	0, 1	I 3
	15	Specify that monthly minima and maxima are to be punched by coding '1', else code '0'.	KPMEXT	0, 1	I 3
7	3	Specify calculation of annual moments and extrema by coding '1'. Specify calculation of annual moments and extrema, and frequency distribution by coding '2'. Else code '0'.	MOAN	0, 1, 2	I 3
	6	Specify that annual mean, standard deviation and skew coefficient are to be written to disk file NF24 by coding '1', else code '0'.	KKANN	0, 1	I 3
	9	Specify that annual mean, standard deviation and skew coefficient are to be punched by coding '1', else code '0'.	KPANN	0, 1	I 3
8	3	Specify calculation of annual auto correlation function by coding '1', else code '0'.	IANAUT	0, 1	I 3
	4 - 6	Specify the number of lags of the annual auto correlation function.	NLAG3	1-120	I 3
	9	Specify a line printer plot of the annual auto correlation function by coding '1', else code '0'.	IANPL	0, 1	I 3

TABLE 3.3: PROGRAM STATS INPUT DATA DECK

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
	3	Specify calculation of time series, moments and frequency distribution by coding '1', else code '0'.	IDISTS	0, 1	I3
	6	Specify that user lower and upper bounds are to be used in frequency distribution analysis by coding '1', else code '0' and the program will select appropriate lower and upper bounds.	IDISBD	0, 1	I 3
	9	Specify a line printer histogram plot of the time series frequency distribution by coding '1', else code '0'.	IFRQPL	0, 1	I 3
9	12	Specify that the time series mean, standard deviation and skew coefficient are to be written to disk file NF25 by coding '1', else code '0'.	KKTSM	0, 1	I 3
	15	Specify that the time series frequency distribution is to be written to disk file NF26 by coding '1', else code '0'.	KKTSF	0, 1	I 3
	18	Specify that the time series mean standard deviation and skew coefficient are to be punched by coding '1', else code '0'.	KPTSM	0, 1	I 3
	21	Specify that the time series frequency distribution is to be punched by coding '1', else code '0'.	KPTSF	0, 1	I 3
-	-	If user lower and upper bounds are to be used in the time series frequency distribution analysis, provide for each station in turn, the lower and upper bound to be used.	BDLOW BDUP	-	8F10.3
	3	Specify calculation of run length frequency distributions by coding '1', else code '0'.	IRUN	0, 1	I 3
	5 - 6	Specify the upper bound or maximum run length for the frequency distribution analysis.	NRUNS	$\leqslant 60$	I3
	9	Specify that run length frequency distributions are to be calculated above and below a user value by coding '1', else code '0' and the program will calculate run length frequency distributions above and below the median value.	IRUN50	0, 1	I 3
10	12	Specify a line printer histogram plot of the run length frequency distributions above and below the specified value or median by coding '1', else code '0'.	IRUNPL	0, 1	I 3
	15	Specify that run length frequency distributions above and below the specified value or median are to be written to disk file NF27 by coding '1', else code '0'.	KKRUN	0, 1	I 3
	18	Specify that run length frequency distributions above and below the specified value or median are to be punched by coding '1', else code '0'.	KPRUN	0, 1	I 3
-	-	If run length frequency distributions are to be calculated above and below a user specified value, provide for each station in turn, the specified value to be used.	VRUN50	-	8F10.3
11	3	Specify the calculation of aggregate minima and maxima by coding '1', else code '0'.	IVOL	0, 1	I 3
	4 - 6	Specify the maximum time period or number of months for which aggregate minima and maxima are to be calculated.	LVOL	$\leqslant 120$	I 3

TABLE 3.3 (CONT'D): PROGRAM STATS INPUT DATA DECK

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
	9	Specify that the aggregate minima and maxima are to be scaled by a user specified value by coding '1', else code '0' and the program will scale aggregate minima and maxima by the time series mean.	IVOLX	0, 1	I 3
	12	Specify a line printer histogram plot of the aggregate minima and maxima up to the maximum time period by coding '1', else code '0'.	IVOLPL	0, 1	I 3
	15	Specify that aggregate minima and maxima are to be written to disk file NF28 by coding '1', else code '0'.	KKVOL	0, 1	I 3
-	-	If aggregate minima and maxima are to be scaled by a user specified value, provide for each station in turn, the specified value to be used.	VOL	-	8F10.3
12	3	Specify range analysis by coding '1', else code '0'.	IRANGE	0, 1	I 3
	6	Specify that the range is to be scaled by a user specified value by coding '1', else code '0' and the program will scale the range by the time series mean.	IRANX	0, 1	I 3
	9	Specify that the surplus and deficit are to be written to disk file NF29 by coding '1', else code '0'.	KKRAN	0, 1	I 3
-	-	If the range is to be scaled by a user specified value, provide for each station in turn, the specified value to be used.	VRANGE	-	8F10.3
13	3	Specify storage-yield analysis by coding '1', else code '0'.	IMASS	0, 1	I 3
	6	Specify the number of yield levels in the analysis.	NYIELD	1-5	I 3
	9	Specify that storage deficits are to be scaled by a user specified value by coding '1', else code '0' and the program will scale storage deficits by the time series mean. (Note that yield levels are calculated as specified proportions of the scale value used).	ISYLD	0, 1	I 3
	12	Specify a line printer histogram plot of the storage deficit distribution, drought, draw and fill duration distributions, and a line printer plot of the storage-yield curve by coding '1', else code '0'	IMASPL	0, 1	I 3
	15	Specify that the Rippl storage and associated drought duration, storage deficit distribution, drought, draw and fill duration distributions, are to be written to disk files NF30, NF33, NF34, NF35 and NF36 respectively by coding '1', else code '0'.	KKMASS	0, 1	I 3
	18	Specify that the Rippl storage and associated drought duration, storage deficit distribution, drought, draw and fill duration distributions, and all distribution bounds, bins and statistics are to be punched by coding '1', else code '0'.	KPMASS	0, 1	I 3

TABLE 3.3 (CONT'D.): PROGRAM STATS INPUT DATA DECK

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
-	-	If storage-yield analysis was specified, specify NYIELD yield levels as decimal fractions. (For example, yield levels equal to 20%, 50% and 80% of the scale value are specified as 0.2, 0.5, 0.8).	YIELD	0.0-1.0	8F10.3
-	-	If storage-yield analysis was specified, and if storage deficits are to be scaled by a user specified value, provide for each station in turn, the specified value to be used.	VSYLD	-	8F10.3
-	-	If storage-yield analysis was specified, provide for each station in turn, for each yield level in turn, the row subscript of matrix BDCOM which identifies the upper bounds required for drought, draw and fill duration distribution analyses, and the subscript of matrix BDDEF which identifies the upper bound required for storage deficit distribution analysis.	IBDCOM IBDDEF	-	10I3
14	3	Specify calculation of the station auto correlation function by coding '1', else code '0'.	IAUTO	0, 1	I 3
	4-6	Specify the number of lags to which the auto correlation function is to be calculated.	NLAG1	1-120	I 3
	9	Specify a line printer plot of the station auto correlation function by coding '1', else code '0'.	IAUTPL	0, 1	I 3
	12	Specify that the station auto correlation function be written to disk file NF31 by coding '1', else code '0'.	KKAUTO	0, 1	I 3
	15	Specify that the station auto correlation function be punched by coding '1', else code '0'.	KPAUTO	0, 1	I 3
15	3	Specify calculation of the cross correlation function between all station pairs by coding '1', else code '0'.	ICROSS	0, 1	I 3
	5-6	Specify the number of lags to which the cross correlation functions are to be calculated.	NLAG2	1-60	I 3
	9	Specify a line printer plot of each cross correlation function by coding '1', else code '0'.	ICRSPL	0, 1	I 3
	12	Specify that each cross correlation function be written to disk file NF32 by coding '1', else code '0'.	KKCROS	0, 1	I 3
	15	Specify that each cross correlation function be punched by coding '1', else code '0'.	KPCROS	0, 1	I 3
16	1-3 4-6 ...	Specify NSTN three digit, user integer code numbers to identify each station.	NSTNNO	1-999	24 I 3
—	—	If the user specified on CARD 2 that the hydrologic data were to be read from the input data card deck, provide for each station in turn, according to the format specified on CARD 3, the hydrologic data for this analysis.	—	—	—

TABLE 3.3 (CONT'D): PROGRAM STATS INPUT DATA DECK

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
—	—	If the user specified on CARD 1 that a number of sets of hydrologic data were to be analysed, and that: (a) the program options and problem data description were the same for each set of hydrologic data, then provide for each hydrologic data set in turn, for each station in turn, according to the format specified on CARD 3, the hydrologic data for the analyses; (b) the program options and problem data description were different for each set of hydrologic data, then provide for each hydrologic data set in turn, (1) the applicable program options and problem data description (CARDS 2-16), and (2) for each station in turn, according to the format specified on CARD 3, the hydrologic data for the analysis.	—	—	—

TABLE 3.3 (CONT'D): PROGRAM STATS INPUT DATA DECK

FILE TYPE	FILE PROGRAM NAME	ASSIGNED FORTRAN UNIT NO.	ASSIGNED DEVICE TYPE	DESCRIPTION
INPUT	NIN	5	Card reader	A sequential card input file defining program options, problem data, file logical unit numbers and the historical moment and extrema data (which may be obtained as punched output from STATS).
	NF20	20	Disk	An input file of synthetic monthly mean data written by program STATS.
	NF21	21	Disk	An input file of synthetic monthly standard deviation data written by program STATS.
	NF22	22	Disk	An input file of synthetic monthly skew coefficient data written by program STATS.
	NF23	23	Disk	An input file of synthetic monthly extrema data written by program STATS.
	NF24	24	Disk	An input file of synthetic annual mean, standard deviation and skew coefficient data written by program STATS.
	NF25	25	Disk	An input file of synthetic time series mean, standard deviation and skew coefficient data written by program STATS.
OUTPUT	NOUT	6	Line printer	A line printer file consisting of (a) an echo check of the problem specification, (b) an optional echo check of the synthetic data, (c) results of the comparison of historical and synthetic moments and extrema, and (d) a display of synthetic sample number against station number showing statistically rejected samples.
	NPUN	7	Card punch	An optional, sequential, hard card copy output file for historical moments and confidence intervals and corresponding synthetic population moments for subsequent input to a data presentation system for graphical display.
	NPU80	80	Disk	An optional, sequential, card image, output file of historical means and confidence intervals and corresponding synthetic population means for subsequent input to a data presentation system. This file is produced when the program option 'KPPUN' is specified as '1'. (If NPU80 is equated to the installation punch (NPUN), hard copy is produced).
	NPU81	81	Disk	An optional, sequential, card image, output file of historical standard deviations and confidence intervals and corresponding synthetic population standard deviations for subsequent input to a data presentation system. This file is produced when the program option 'KPPUN' is specified as '1'. (If NPU81 is equated to the installation punch (NPUN), hard card copy is produced).
	NPU82	82	Disk	An optional, sequential, card image, output file of historical skew coefficients and corresponding synthetic population skew coefficients for subsequent input to a data presentation system. This file is produced when the program option 'KPPUN' is specified as '1'. (If NPU82 is equated to the installation punch (NPUN), hard card copy is produced).

Note: Files NF20, NF21, ... , NF25 are all sequential, unformatted, variable record length files.

TABLE 4.1: PROGRAM MOMENT FILE DESCRIPTION

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
1	1 - 3	A user integer code number to identify this run.	NRUN	0 - 999	I3
	5 - 6	Specify the number of stations in the analysis.	NSTN	1 - 25	I 3
	7 - 9	Specify the common number of years of synthetic data at each station in each synthetic sample data set.	NYEAR	≤100	I 3
	12	Specify as '1' for NYEAR=75 and as '2' for NYEAR=50 (CH1**2 small sample statistics assigned in BLOCK DATA subprogram for these two small sample possibilities).	IYR	1,2	I3
	14-15	Specify the number of synthetic sample data sets.	NSAM	<50	I3
	17-18	Specify as "12' for monthly or as '14' for monthly, annual and time series moments and extremes analysis.	MONTHS	12,14	I3
2	3	To obtain punched comparison of historical and synthetic moments code '1', else '0'.	KPPUN	0,1	I3
	5 - 6	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic first moments.	NPU80	1 - 99	I3
	8 - 9	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic second moments.	NPU81	1 - 99	I3
	11-12	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic third moments.	NPU82	1 - 99	I3
3	2 - 3	Specify the Fortran file number associated with the synthetic monthly first moment data file.	NF20	1 - 99	I3
	5 - 6	Specify the Fortran file number associated with the synthetic monthly standard deviation data file.	NF21	1 - 99	I3
	8 - 9	Specify the Fortran file number associated with the synthetic monthly skew coefficient data file.	NF22	1 - 99	I 3
	11-12	Specify the Fortran file number associated with the synthetic monthly extreme data file.	NF23	1 - 99	I3
	14-15	Specify the Fortran file number associated with the synthetic annual first moment, standard deviation and skew coefficient data file.	NF24	1 - 99	I3
	17-18	Specify the Fortran file number associated with the synthetic time series first moment, standard deviation and skew coefficient data file.	NF25	1 - 99	I3
4	1-12	Specify the common, partitioned data set member name, used to identify the synthetic moments and extremes data files to be used in the analysis.	-	-	3A4
5	3	Specify the larger significance level of 1%, 5%, 10% to be used by coding integers 1,2,3. (e.g., 2 corresponds to 5%).	IPCEN1	1,2,3	I3
	6	Specify the smaller significance level of 1%, 5%, 10% to be used by coding integers 1,2,3. (e.g., 1 corresponds to 1%).	IPCEN2	1,2,3	I3

TABLE 4.2: PROGRAM MOMENT INPUT DATA DECK

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
6	1 - 3 4 - 6 7 - 9 72-75	Specify NSTN user three digit integer code numbers used to identify each station. Identification numbers must be in station order as analysed by STATS to be correctly associated with synthetic data.	NSTNNO (25)	1 - 999	25I3
7	2 - 3 5 - 6 8 - 9 73-75	Specify NSTN vector subscripts each corresponding to the vector element of the starting month of the 6 driest contiguous months on average throughout the year for each station. Station order must be as for analysis by STATS.	ISDRY (25)	1 - 12	25 I3
8	2 - 3 5 - 6 8 - 9 73-75	Specify NSTN vector subscripts each corresponding to the vector element of the starting month of the 6 wettest contiguous months on average throughout the year for each station. Station order must be as for analysis by STATS.	ISWET (25)	1 - 12	25I3
		Provide for each station in turn in the same order as for analysis by STATS: NSTN sets { 12 monthly historical first moments. 12 monthly historical standard deviations. 12 monthly historical skew coefficients. 12 monthly historical minima. 12 monthly historical maxima. (Note: These cards may be obtained as punched output from STATS).	HISTAV HISTSD HISTSK HISTS HISTL	- - - - -	6F12.5 6F12.5 6F12.5 6F12.5 6F12.5
		If months = 14 was specified, provide for each station in turn in the same order as for analysis by STATS: NSTN cards { Annual mean, standard deviation and skew coefficient. (Note: These cards may be obtained as punched output from STATS).	HISTAV, HISTSD, HISTSK	-	3F12.5
		If months = 14 was specified, provide for each station in turn in the same order as for analysis by STATS: NSTN cards { Time series mean, standard deviation and skew coefficient	HISTAV, HISTSD, HISTSK	-	3F12.5

TABLE 4.2 (CONT'D): PROGRAM MOMENT INPUT DATA DECK

FILE TYPE	FILE PROGRAM NAME	ASSIGNED FORTRAN UNIT NO.	ASSIGNED DEVICE TYPE	DESCRIPTION
INPUT	NIN	5	Card reader	A sequential card input file defining program options, problem data, file logical unit numbers and historical time series frequency distributions (which may be obtained as punched output from STATS).
	NF26	26	Disk	A sequential, unformatted, variable record length, input file of synthetic time series frequency distributions written by program STATS.
OUTPUT	NOUT	6	Line printer	A line printer file consisting of (a) an echo check of the problem specification, (b) an optional echo check of the synthetic time series frequency distributions, (c) results of the comparison of the historical and synthetic time series frequency distributions, and (d) a display of synthetic sample number against station number showing statistically rejected samples.
	NPUN	7	Card punch	An optional, sequential, hard card copy output file for relative and cumulative historical and synthetic data time series probability distributions for subsequent input to a data presentation system for graphical display.
	NPU83	83	Disk	An optional, sequential, card image output file of historical and synthetic population relative and cumulative time series probability distributions for subsequent input to a data presentation system. This file is produced when the program option 'KPPUN' is specified as '1'. (If NPU83 is equated to the installation punch (NPUN), hard card copy is produced).

TABLE 5.1: PROGRAM FREQ FILE DESCRIPTION

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
1	1 - 3	A user integer code number to identify this run.	NRUN	0-999	I 3
	5 - 6	Specify the number of stations in the analysis.	NSTN	1-25	I 3
	7 - 9	Specify the common number of years of synthetic data at each station for each synthetic sample data set.	NYEAR	$\leqslant$100	I 3
	10-12	Specify the common number of years of historical data at each station.	NYEARH	$\leqslant$100	I 3
	14-15	Specify the number of synthetic sample data sets.	NSAM	$\leqslant$50	I 3
2	1 - 3 4 - 6 7 - 9 72 - 75	Specify NSTN user three digit integer code numbers used to identify each station. (Identification numbers must be in station order as analysed by STATS to be correctly associated with synthetic data).	NSTNNO (25)	1-999	25 I3
3	3	Specify the larger significance level of 1%, 5%, 10% to be used by coding integers 1,2,3 (e.g. 2 corresponds to 5%).	IPCEN1	1,2,3	I 3
	6	Specify the smaller significance level of 1%, 5%, 10% to be used by coding integers 1,2,3 (e.g. 1 corresponds to 1%).	IPCEN2	1,2,3	I 3
	8 - 9	Specify the number of class intervals or bins used.	INT	$\leqslant$20	I 3
	11-12	Specify the Fortran file number associated with the synthetic time series frequency distribution data.	NF26	1-99	I 3
	15	To obtain punched comparison of historical and synthetic population relative and cumulative time series probability distributions code '1', else code '0'	KPPUN	0, 1	I 3
	17-18	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic time series probability distributions.	NPU83	1-99	I 3
4	1 - 12	Specify the common partitioned data set member name used to identify the synthetic time series frequency distribution data to be used in the analysis.	—	—	3A4
—	—	Specify for each station (1 to NSTN) in turn, in the same order as analysed by program STATS, the upper bound of each station used in the time series frequency distribution analysis.	BDUP(25)	—	8F10.5
—	—	Specify for each station (1 to NSTN) in turn, in the same order as analysed by program STATS, the lower bound of each station used in the time series frequency distribution analysis.	BDL(25)	—	8F10.5
—	—	Provide for each station (1 to NSTN) in turn, in the same order as analysed by program STATS, the historical data time series frequency distributions. (Note: These cards may be obtained as punched output from program STATS).	HISTP	—	24F3.0

TABLE 5.2: PROGRAM FREQ INPUT DATA DECK

FILE TYPE	FILE PROGRAM NAME	ASSIGNED FORTRAN UNIT NO.	ASSIGNED DEVICE TYPE	DESCRIPTION
INPUT	NIN	5	Card reader	A sequential card input file defining program options, problem data, file logical unit numbers and historical run length frequencies above and below the median (which may be obtained as punched output from STATS).
	NF27	27	Disk	A sequential, unformatted, variable record length, input file of synthetic data run length frequencies above and below the historical data median written by program STATS.
OUTPUT	NOUT	6	Line printer	A line printer file consisting of (a) an echo check of the problem specification, (b) an optional echo check of the synthetic data run length frequencies, (c) results of the comparison of the historical and synthetic data run length distributions, and (d) a display of synthetic sample number against station number showing statistically rejected samples.
	NPUN	7	Card punch	An optional, sequential, hard card copy output file for relative and cumulative historical and synthetic data run length probability distributions for subsequent input to a data presentation system for graphical display.
	NPU84	84	Disk	An optional, sequential, card image output file of historical and synthetic population relative and cumulative run length probability distributions above the median for subsequent input to a data presentation system. This file is produced when the program option 'KPPUN' is specified as '1'. (If NPU84 is equated to the installation punch (NPUN), hard card copy is produced).
	NPU85	85	Disk	An optional, sequential, card image output file of historical and synthetic population relative and cumulative run length probability distributions below the median for subsequent input to a data presentation system. This file is produced when the program option 'KPPUN' is specified as '1'. (If NPU85 is equated to the installation punch (NPUN), hard card copy is produced).
	NPU86	86	Disk	An optional, sequential, card image output file of historical and synthetic population relative and cumulative run length probability distributions about the median for subsequent input to a data presentation system. This file is produced when the program option 'KPPUN' is specified as '1'. (If NPU86 is equated to the installation punch (NPUN), hard card copy is produced).

TABLE 6.1: PROGRAM RUNS FILE DESCRIPTION

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
1	1 - 3	A user integer code number to identify this run.	NRUN	0-999	I 3
	5 - 6	Specify the number of stations in the analysis.	NSTN	1 - 25	I 3
	7 - 9	Specify the common number of years of historical data at each station.	NYEARH	≤100	I 3
	10-12	Specify the common number of years of synthetic data at each station for each synthetic sample data set.	NYEAR	≤100	I 3
	14-15	Specify the number of synthetic sample data sets.	NSAM	≤50	I 3
2	1 - 3 4 - 6 7 - 9 72-75	Specify NSTN user three digit integer code numbers used to identify each station. Identification numbers must be in station order as analysed by STATS to be correctly associated with synthetic data.	NSTNNO (25)	1-999	25I3
3	3	Specify the larger significance level of 1%, 5%, 10% to be used by coding integers 1, 2, 3 (e.g.,2 corresponds to 5%).	IPCEN1	1,2,3	I3
	6	Specify the smaller significance level of 1%, 5%, 10% to be used by coding integers 1,2,3 (e.g., 1 corresponds to 1%).	IPCEN2	1,2,3	I3
	8 - 9	Specify the percentage of runs about the median used to define the minimum long run length.	ILRUN	1-100	I 3
	11-12	Specify the length of the runs frequency vectors.	NRUNS	≤60	I 3
	14-15	Specify the Fortran file number associated with the synthetic runs frequency distribution data.	NF27	1-99	I 3
4	3	To obtain punched comparison of historical and synthetic population relative and cumulative run length probability distributions above, below and about the median code '1', else '0'.	KPPUN	0, 1	I 3
	5 - 6	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic run length distributions above the median.	NPU84	1 - 99	I 3
	8 - 9	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic run length distributions below the median.	NPU85	1 - 99	I 3
	11-12	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic run length distributions about the median.	NPU86	1 - 99	I 3
5	1 - 12	Specify the common partitioned data set member name used to identify the synthetic run length frequency distribution data to be used in the analysis.	——	——	3A4

TABLE 6.2: PROGRAM RUNS INPUT DATA DECK

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
—	—	Provide for each station (1 to NSTN) in turn in the same order as for analysis by STATS:			
		NSTN sets: Historical run length frequency distribution above the median	HISTU	—	24 I3
		NSTN sets: Historical run length frequency distribution below the median	HISTD	—	24 I 3
		(Note: These cards may be obtained as punched output from program STATS)			

TABLE 6.2 (CONT'D): PROGRAM RUNS INPUT DATA DECK

FILE TYPE	FILE PROGRAM NAME	ASSIGNED FORTRAN UNIT NO.	ASSIGNED DEVICE TYPE	DESCRIPTION
INPUT	NIN	5	Card reader	A sequential card input file defining program options, problem data, file logical unit numbers, Rippl analysis parameters (bounds, bins) and results (which may be obtained as punched output from STATS).
	NFILE(1)	33	Disk	An input file of synthetic deficit frequency distribution data written by program STATS.
	NFILE(2)	34	Disk	An input file of synthetic drought duration frequency distribution data written by program STATS.
	NFILE(3)	35	Disk	An input file of synthetic draw duration frequency distribution data written by program STATS.
	NFILE(4)	36	Disk	An input file of synthetic fill duration frequency distribution data written by program STATS.
	NFILE(5)	30	Disk	An input file of synthetic Rippl storages and drought durations written by program STATS.
OUTPUT	NOUT	6	Line printer	A line printer file consisting of (a) an echo check of the problem specification, (b) an optional echo check of the synthetic frequency distribution data (c) results of the comparison of the historical and synthetic Rippl storages and frequency distributions, and(d) a display of synthetic sample number against station number showing statistically rejected samples.
	NPUN	7	Card punch	An optional, sequential, hard card copy output file for relative and cumulative historical and synthetic population probability distributions for subsequent input to a data presentation system for graphical display.
	NPU87	87	Disk	An output file of historical and synthetic population relative and cumulative deficit probability distributions for subsequent input to a data presentation system. (If NPU87 is equated to the installation punch (NPUN), hard card copy is produced).
	NPU88	88	Disk	An output file of historical and synthetic population relative and cumulative drought duration probability distributions for subsequent input to a data presentation system. (If NPU88 is equated to the installation punch (NPUN), hard card copy is produced).
	NPU89	89	Disk	An output file of historical and synthetic population relative and cumulative draw duration probability distributions for subsequent input to a data presentation system. (If NPU90 is equated to the installation punch (NPUN), hard card copy is produced).
	NPU90	90	Disk	An output file of historical and synthetic population relative and cumulative fill duration probability distributions for subsequent input to a data presentation system. (If NPU90 is equated to the installation punch (NPUN), hard card copy is produced).
	NPU91	91	Disk	An output file for synthetic Rippl storage distribution results. This file is not being used at present.

Note: Files NPU87, NPU88, ..., NPU90 are all optional, sequential, card image files produced when the program option 'KPPUN' is specified as '1'.

TABLE 7.1: PROGRAM YIELD FILE DESCRIPTION

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
1	1 - 3	A user integer code number to identify this run.	NRUN	0-999	I 3
	5 - 6	Specify the number of stations in the analysis.	NSTN	1 - 25	I 3
	8 - 9	Specify the number of synthetic sample data sets.	NSAM	⩽50	I 3
	12	Specify the number of yield levels in the analysis.	NYIELD	1 - 5	I 3
	13-15	Specify the common number of years of historical data at each station.	NYEARH	⩽100	I 3
	16-18	Specify the common number of years of synthetic data at each station for each synthetic sample data set.	NYEAR	⩽100	I 3
	21	To list synthetic frequency data code '1', else '0'.	IECHO	0, 1	I 3
	24	Specify the larger significance level of 1%, 5%, 10% to be used by coding integers 1,2,3. (e.g., 2 corresponds to 5%).	IPCEN1	1,2,3	I 3
	27	Specify the smaller significance level of 1%, 5%, 10% to be used by coding integers 1,2,3 (e.g., 1 corresponds to 1%).	IPCEN2	1,2,3	I 3
2	1 - 3 4 - 6 7 - 9 72-75	Specify NSTN user three digit integer code numbers used to identify each station. Identification numbers must be in station order as analysed by STATS to be correctly associated with synthetic data.	NSTNNO (25)	1 -999	25 I3
3	3	To obtain punched comparison of historical and synthetic relative and cumulative probability distributions code '1', else '0'.	KPPUN	0, 1	I 3
	5 - 6	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic deficit distributions.	NPU87	1 - 99	I 3
	8 - 9	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic drought duration distributions.	NPU88	1 - 99	I 3
	11-12	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic draw duration distributions.	NPU89	1 - 99	I 3
	14-15	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic fill duration distributions.	NPU90	1 - 99	I 3
	17-18	Specify the Fortran file number associated with the card image disk output file for synthetic Rippl storage analysis results. (Not used at present).	NPU91	1 - 99	I 3
4	2 - 3	Specify the Fortran file number associated with the synthetic deficit distribution data.	NFILE(1)	1 - 99	I 3
	5 - 6	Specify the Fortran file number associated with the synthetic drought duration distribution data.	NFILE(2)	1 - 99	I 3

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
4	8 - 9	Specify the Fortran file number associated with the synthetic draw duration distribution data.	NFILE(3)	1 - 99	I 3
	11 - 12	Specify the Fortran file number associated with the synthetic fill duration distribution data.	NFILE(4)	1 - 99	I 3
	14 - 15	Specify the Fortran file number associated with the synthetic Rippl storage analysis results.	NFILE(5)	1 - 99	I 3
5	1 - 12	Specify the common, partitioned data set member name, used to identify the synthetic frequency distribution data to be used in the analysis.	——	——	3A4
——	——	Provide for each station (1 to NSTN) in turn, and for each yield level (1 to NYIELD) in turn: NSTN sets { NYIELD sets { yield level, Rippl storage and drought duration, no. of frequency bins.	YIELD, RIPPL, CRIT, IBINS	——	2F9.3, F6.0, I6
		upper bounds for drought/draw, fill and deficit frequency analyses.	BOUNDS	——	3F9.3
		total and average for deficit, drought, draw and fill distributions in turn. (Note: These cards may be obtained as punched output from STATS.	HSTAT	——	8F9.3
——	——	Provide for each station (1 to NSTN) in turn, and for each yield level (1 to NYIELD) in turn: NSTN sets { NYIELD cards { Historical deficit frequency distribution data (Note: These cards may be obtained as punched output from STATS).	HISTP	——	24F3.0
——	——	Provide for each station (1 to NSTN) in turn, and for each yield level (1 to NYIELD) in turn: NSTN sets { NYIELD cards { Historical drought duration frequency distribution data. (Note: These cards may be obtained as punched output from STATS).	HISTP	——	24F3.0
——	——	Provide for each station (1 to NSTN) in turn, and for each yield level (1 to NYIELD) in turn: NSTN sets { NYIELD cards { Historical draw duration frequency distribution data. (Note: These cards may be obtained as punched output from STATS).	HISTP	——	24F3.0

TABLE 7.2 (CONT'D.): PROGRAM YIELD INPUT DATA DECK

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
——	——	Provide for each station (1 to NSTN) in turn, and for each yield level (1 to NYIELD) in turn: NSTN sets { NYIELD cards { Historical fill duration frequency distribution data. (Note: These cards may be obtained as punched output from STATS).	HISTP	——	24F3.0

TABLE 7.2 (CONT'D): PROGRAM YIELD INPUT DATA DECK

FILE TYPE	FILE PROGRAM NAME	ASSIGNED FORTRAN UNIT NO.	ASSIGNED DEVICE TYPE	DESCRIPTION
INPUT	NIN	5	Card reader	A sequential card input file defining program options, problem data, file logical unit numbers and the historical auto and or cross correlation functions. (The historical correlation data may be obtained as punched output from program STATS).
	NF31	31	Disk	A sequential, unformatted, variable record length, input file of synthetic data auto correlation functions written by program 'STATS').
	NF32	32	Disk	A sequential, unformatted, variable record length, input file of synthetic data cross correlation functions written by program 'STATS'.
OUTPUT	NOUT	6	Line printer	A line printer file consisting of (a) an echo check of the problem specification, (b) an optional echo check of synthetic correlation data, (c) results of the tests on the correlation coefficients, (d) a display of synthetic sample number against correlation function number showing rejected correlation functions.
	NPUN	7	Card punch	An optional, sequential, card output file of historical and mean, min and max synthetic correlations for subsequent input to a data presentation system for graphical display.
	NF92	92	Disk	An optional, sequential, card image, output file of historical mean, min and max synthetic auto correlations for subsequent input to a data presentation system for graphical display. This file is produced when the program option 'KPAUTO' is specified as '1'. (If NF92 is equated to the installation punch (NPUN) hard card copy is produced.
	NF93	93	Disk	An optional, sequential, card image, output file of historical mean, min and max synthetic cross correlations for subsequent input to a data presentation system for graphical display. This file is produced when the program option 'KPCROS' is specified as '1'. (If NF93 is equated to the installation punch (NPUN) hard card copy is produced.
SCRATCH	NWF1, NWF2, NWF3	1, 2, 3	System scratch disk	Sequential, unformatted, variable record length scratch files used to store a set of results vectors for each auto/cross correlation function.

TABLE 8.1: PROGRAM CORREL FILE DESCRIPTION

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
1	1 - 3	A user integer code number to identify this run.	NRUN	0-999	I 3
	6	Comparison of auto correlation functions requested by coding '1', else '0'.	IAUTO	0, 1	I 3
	7 - 9	Specify no. of lags of the auto correlation functions.	NLAG1	1-120	I 3
	12	Comparison of cross correlation functions requested by coding '1', else '0'.	ICROSS	0, 1	I 3
	14-15	Specify no. of lags of the cross correlation functions.	NLAG2	0-60	I 3
	18	Specify the larger significance level of 1%, 5%, 10% to be used by coding integers 1, 2, 3 (e.g., 2 corresponds to 5%).	IPCEN1	1, 2, 3	I 3
	21	Specify the smaller significance level of 1%, 5%, 10% to be used by coding integers 1, 2, 3 (e.g., 1 corresponds to 1%).	IPCEN2	1, 2, 3	I 3
2	3	To obtain punched comparison of mean synthetic and historical auto correlation functions code '1', else '0'.	KPAUTO	0, 1	I 3
	6	To obtain punched comparison of mean synthetic and historical cross correlation functions code '1', else '0'.	KPCRDS	0, 1	I 3
	9	To list synthetic auto correlation data code '1', else '0'.	IPAUTO	0, 1	I 3
	12	To list synthetic cross correlation data code '1', else '0'.	IPCROS	0, 1	I 3
	13-15	Specify common no. of years of historical data at each station.	NYEARH	≤100	I 3
	16-18	Specify common no. of years of synthetic data at each station for each synthetic sample data set.	NYEAR	≤100	I 3
	20-21	Specify number of synthetic sample data sets.	NSAM	≤50	I 3
	23-24	Specify number of stations in the analysis.	NSTN	1-25	I 3
3	1 - 3 4 - 6 72-75	Specify NSTN user three digit integer code numbers used to identify each station. Identification numbers must be in station order as analysed by STATS to be correctly associated with correlation functions.	NSTNNO (25)	1-999	25 I 3
4	2 - 3	Specify the Fortran file number associated with the synthetic auto correlation data file.	NF31	1-99	I 3
	5 - 6	Specify the Fortran file number associated with the synthetic cross correlation data file.	NF32	1-99	I 3
	8 - 9	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and mean synthetic auto correlation functions.	NF92	1-99	I 3
	11-12	Specify the Fortran file number associated with the card image disk output file for the comparison of historical and synthetic cross correlation functions.	NF93	1-99	I 3
	14-15	Specify the Fortran file number associated with work file 1.	NWF1	1-99	I 3

CARD NO.	COLS	DESCRIPTION	PROGRAM VARIABLE NAME	RANGE OF VALUES	FORMAT
4	17-18	Specify the Fortran file number associated with work file 2.	NWF2	1 - 99	I 3
	20-21	Specify the Fortran file number associated with work file 3.	NWF3	1 - 99	I 3
5	1 - 12	Specify the common, partitioned data set member name, used to identify the synthetic auto and cross correlation data files to be used in the analysis.	—	—	3A4
—	—	If auto correlation analysis was specified, provide for each station in turn in the same order as for analysis by STATS: The station historical auto correlation function for lags 1 to NLAG1 (Note: These cards may be obtained as punched output from STATS).	HIST	—	6F12.5
—	—	If cross correlation analysis was specified provide for each station combination in turn in the same order as analysed by STATS: The station combination historical cross correlation function for lags -NLAG2 to +NLAG2 (Note: These cards may be obtained as punched output from STATS).	HIST	—	6F12.5

TABLE 8.2 (CONT'D): PROGRAM CORREL INPUT DATA DECK

FILE TYPE	FILE DESCRIPTION	FILE REFERENCED BY PROGRAM(S)	FILE PROGRAM VARIABLE	ASSIGN-ED FILE NUMBER	ASSIGNED DEVICE TYPE
STANDARD INPUT/ OUTPUT UNITS	Card reader	ALL	NIN	5	——
	Line printer	ALL	NOUT	6	——
	Card punch	ALL	NPUN	7	——
SYNTHETIC DATA	Generated synthetic data sets	STATS	INFLOW	8	TAPE[a]
SYNTHETIC DATA STATISTICS FILES	Monthly means	STATS, MOMENT	NF20	20	DISK1[b]
	Monthly standard deviations	STATS, MOMENT	NF21	21	DISK1
	Monthly skew coefficients	STATS, MOMENT	NF22	22	DISK1
	Monthly extremes	STATS, MOMENT	NF23	23	DISK1
	Annual moments	STATS, MOMENT	NF24	24	DISK1
	Time series moments	STATS, MOMENT	NF25	25	DISK1
	Time series frequencies	STATS, FREQ	NF26	26	DISK1
	Run-length frequencies	STATS, RUNS	NF27	27	DISK1
	Cumulative volume run sums	STATS	NF28	28	DISK1
	Cumulative residuals range	STATS	NF29	29	DISK1
	Rippl storages and drought durations	STATS, YIELD	NF30	30	DISK1
	Auto correlation functions	STATS, CORREL	NF31	31	DISK1
	Cross correlation functions	STATS, CORREL	NF32	32	DISK1
	Rippl deficit frequencies	STATS, YIELD	NF33	33	DISK1
	Rippl drought frequencies	STATS, YIELD	NF34	34	DISK1
	Rippl draw frequencies	STATS, YIELD	NF35	35	DISK1
	RIPPL fill frequencies	STATS, YIELD	NF36	36	DISK1
COMPARISON OF HISTORICAL AND SYNTHETIC DATA STATISTICS	Monthly means	MOMENT	NPU80	80	DISK1
	Monthly standard deviations	MOMENT	NPU81	81	DISK1
	Monthly skew coefficients	MOMENT	NPU82	82	DISK1
	Time series distributions	FREQ	NPU83	83	DISK1
	Run length distributions above	RUNS	NPU84	84	DISK1
	Run length distributions below	RUNS	NPU85	85	DISK1
	Run length distributions about	RUNS	NPU86	86	DISK1
	Rippl deficit distributions	YIELD	NFU87, NFILE(1)	87	DISK1
	Rippl drought distributions	YIELD	NPU88, NFILE(2)	88	DISK1
	Rippl draw distributions	YIELD	NPU89, NFILE(3)	89	DISK1
	Rippl fill distributions	YIELD	NPU90, NFILE(4)	90	DISK1
	Rippl storages (not used)	YIELD	NPU91, NFILE(5)	91	DISK1
	Auto correlation functions	CORREL	NPU92	92	DISK1
	Cross correlation functions	CORREL	NPU93	93	DISK1
HISTORICAL STATISTIC SCRATCH FILES	Monthly moments and extrema	STATS	NPU1	1	DISK2[c]
	Annual moments	STATS	NPU2	2	DISK2
	Time series moments	STATS	NPU3	3	DISK2
	Time series frequencies	STATS	NPU4	4	DISK2
	Run length frequencies above and below	STATS	NPU9	9	DISK2
	Rippl storage, drought duration and analysis parameters	STATS	NPU10	10	DISK2
	Rippl deficit frequencies	STATS	NPU11	11	DISK2
	Rippl drought frequencies	STATS	NPU12	12	DISK2
	Rippl draw frequencies	STATS	NPU13	13	DISK2
	Rippl fill frequencies	STATS	NPU14	14	DISK2
	Auto and cross correlation functions	STATS	NPU15	15	DISK2

FILE TYPE	FILE DESCRIPTION	FILE REFERENCE BY PROGRAM(S)	FILE PROGRAM VARIABLE	ASSIGNED FILE NUMBER	ASSIGNED DEVICE TYPE
CROSS CORREL- ATION ANALYSIS SCRATCH FILES	Modified time series of station 1	STATS	NFILE(1)	50	DISK2
	Modified time series of station 2	STATS	NFILE(2)	51	DISK2
	. . .	. . .	. . .	. . .	. . .
	. . .	. . .	. . .	. . .	. . .
	. . .	. . .	. . .	. . .	. . .
	Modified time series of station 24	STATS	NFILE(24)	73	DISK2
	Modified time series of station 25	STATS	NFILE(25)	74	DISK2
CORREL- ATION COMP- ARISON SCRATCH FILES	Synthetic results vectors file No. 1	CORREL	NWF1	1	DISK2
	Synthetic results vectors file No. 2	CORREL	NWF2	2	DISK2
	Historical results vectors file	CORREL	NWF3	3	DISK2

(a): Private mountable tape

(b): Private mountable disk pack (IBM 2316).

(c): System scratch disk pack (IBM 3330).

TABLE 9.1 (CONT'D): SUMMARY OF PROGRAM SUITE FILES