National Uranium Resource Evaluation

STATISTICAL TECHNIQUES APPLIED TO AERIAL RADIOMETRIC SURVEYS (STAARS):

PRINCIPAL COMPONENTS ANALYSIS
USER’S MANUAL

Bendix
Field Engineering Corporation
Grand Junction Operations
Grand Junction, CO 81502

January 1981

PREPARED FOR THE U.S. DEPARTMENT OF ENERGY
Assistant Secretary for Resource Applications
Grand Junction Office, Colorado
This report is a result of work performed by Bendix Field Engineering Corporation, Operating Contractor for the U.S. Department of Energy, as part of the National Uranium Resource Evaluation. NURE is a program of the U.S. Department of Energy's Grand Junction, Colorado, Office to acquire and compile geologic and other information with which to assess the magnitude and distribution of uranium resources and to determine areas favorable for the occurrence of uranium in the United States.

Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed in this report, or represents that its use would not infringe privately owned rights. Reference therein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
STATISTICAL TECHNIQUES APPLIED TO AERIAL RADIOMETRIC SURVEYS (STAARS):

PRINCIPAL COMPONENTS ANALYSIS
USER'S MANUAL

C. D. Koch
F. L. Pirkle
J. S. Schmidt

BENDIX FIELD ENGINEERING CORPORATION
Grand Junction Operations
P.O. Box 1569
Grand Junction, Colorado 81502

January 1981

PREPARED FOR THE U.S. DEPARTMENT OF ENERGY
Assistant Secretary for Resource Applications
Grand Junction Office, Colorado
Under Contract No. DE-AC13-76GJ01664
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>5</td>
</tr>
<tr>
<td>Introduction</td>
<td>7</td>
</tr>
<tr>
<td>Principal Components Analysis Theory</td>
<td>8</td>
</tr>
<tr>
<td>Principal Components Analysis Versus Factor Analysis</td>
<td>20</td>
</tr>
<tr>
<td>Program Operation</td>
<td>21</td>
</tr>
<tr>
<td>An Example from Aerial Radiometric Data</td>
<td>23</td>
</tr>
</tbody>
</table>

## FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Covariance Plot</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>Eigenvector Plot</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>Processing Flow Chart for Principal Components Analysis Program</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>Output 1. Statistics Derived from a Principal Components Analysis</td>
<td>29</td>
</tr>
<tr>
<td>5</td>
<td>Output 2. Frequency of Distribution for the Principal Components</td>
<td>29</td>
</tr>
<tr>
<td>6a</td>
<td>Output 3. Histograms of the Principal Component 1</td>
<td>30</td>
</tr>
<tr>
<td>6b</td>
<td>Output 3. Histograms of the Principal Component 2</td>
<td>30</td>
</tr>
<tr>
<td>6c</td>
<td>Output 4. Histograms of the Principal Component 3</td>
<td>31</td>
</tr>
<tr>
<td>7</td>
<td>Output 4. Plot of Points of Interest to Interpreter for All Three Components</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>Generalized Geology of Detail Area C.</td>
<td>33</td>
</tr>
</tbody>
</table>

## APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Program Listing</td>
<td>37</td>
</tr>
<tr>
<td>B</td>
<td>Sample Output</td>
<td>69</td>
</tr>
<tr>
<td>C</td>
<td>PCA Documentation File</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Suggested Reading</td>
<td>83</td>
</tr>
</tbody>
</table>
ABSTRACT

A Principal Components Analysis (PCA) program has been written to aid in the interpretation of multivariate aerial radiometric data collected by the U.S. Department of Energy (DOE) under the National Uranium Resource Evaluation (NURE) program. The variations exhibited by these data have been reduced and classified into a number of linear combinations by using the PCA program. The PCA program then generates histograms and outlier maps of the individual variates. Black and white plots can be made on a Calcomp plotter by the application of follow-up programs. All programs referred to in this guide were written for a DEC-10.

From this analysis a geologist may begin to interpret the data structure. Insight into geological processes underlying the data may be obtained.
INTRODUCTION

The U.S. Department of Energy (DOE), under the National Uranium Resource Evaluation (NURE) program, has been conducting aerial radiometric surveys over various portions of the United States (Bendix Field Engineering Corporation, 1979). The data collected in these surveys include multichannel observations of gamma-ray emissions from which the uranium (from $^{214}\text{Bi}$), thorium (from $^{208}\text{Tl}$), and potassium (from $^{40}\text{K}$) contributions to the total activity can be determined. The multivariate structure of the data suggests that statistical techniques of multivariate nature are appropriate.

Multivariate exploratory techniques can be used to suggest revealing projections and transformations of the data. Principal components analysis is one such technique. Its application is straightforward, providing a suitable starting point for such exploration.

Principal components analysis is a technique which provides insight into the structure of multivariate data. Many analysts consider it "strictly descriptive, (having) an interesting deterministic geometrical interpretation..." (Bargmann, 1969, p. 573). They view the principal components as possibly useful but strictly artificial constructs. However, it has been shown that when principal components analysis is applied to aerial radiometric data it may shed light on the geologic history of an area and may aid in the identification of areas favorable for uranium deposition (Pirkle et al., 1980).
Other workers have also found principal components analysis useful in analyzing aerial radiometric data. For instance, Los Alamos Scientific Laboratory (1975), as part of a DOE/Grand Junction sponsored study, applied principal components analysis to the three variables Bi, Ti, and K and the ratios $\frac{Bi}{K}$, $\frac{Bi}{Ti}$, and $\frac{Ti}{K}$. This study indicated that 98 percent of the variation could be explained using only Bi, Ti, and K data. Duval (1976, 1977) applied factor analysis, a technique similar to principal components analysis (see section on Principal Components Analysis Versus Factor Analysis), to aerial data and concluded that the technique was useful in mapping geology. Pirkle et al. (1980) used principal components analysis to develop an interpretation of aerial radiometric data suggestive of depositional processes.

In addition to the Bi, Ti, and K data, the flight-line number, record number, latitude, longitude, and geology for each observation are recorded. This information enables principal components to be calculated on subsets of the data or the total data set. Subset calculations may be accomplished as a function of the flight-line numbers or geologic labels, depending on the desires of the user. Maps of the locations of the desired component scores are made possible by using latitude and longitude information and a Mercator projection routine. The user may set these values so that the points plotted are outliers within the interpreter's context.

**PRINCIPAL COMPONENTS ANALYSIS THEORY**

The mathematical objective of principal components analysis is to determine $N$ uncorrelated components that express linear combinations of the original data.
To illustrate the principal components analysis theory, a three variable data set comprised of potassium (K), bismuth (Bi), and thallium (Ti) counts will be used. The determination of uncorrelated components is accomplished by first calculating the means, variances, and covariances of these data and setting these values into matrix form. The standard covariance equation is:

\[
CV_{ik} = \frac{1}{M} \sum_{j=1}^{M} [(X_{ij} - \bar{X}_i)(X_{kj} - \bar{X}_k)]
\]  

(1)

where M is the total number of data points and \( \bar{X} \) is the mean of the respective variable. The integers \( i \) and \( k \) represent the K, Bi, or Ti values in the jth sample (Davis, 1973).

Equation (2) is derived from equation (1). This version was used initially within the program because the covariance calculation is accomplished with only one pass through the data. This equation proved to be slightly inaccurate as well as unstable. These two difficulties were overcome by implementing the West-Howell Algorithm (Chan and Lewis, 1979; Howell, 1980) which calculates the covariance by means of equation (3):

\[
CV_{ik} = [(1/M) \sum_{j=1}^{M} X_{ij} X_{kj}] - [(1/M) \sum_{j=1}^{M} (X_{ij} \sum_{i=1}^{M} X_{ij}) (X_{kj} \sum_{j=1}^{M} X_{kj})]
\]  

(2)

\[
CV_{ik} = \frac{1}{(M-1)} \sum_{j=2}^{M} (j-1) (A_{i,j} A_k)
\]  

(3)
where the subscripts are the same as in equations (1) and (2) and the values for variable A are:

$$A_i = X_{i1}$$

and where the integer $i$ represents the variate, $K$, $B_i$, or $T_1$, and 1 is the first value for each. After A is initialized it then takes on the values given in equation (4):

$$A_i = \sum_{j=2}^{M} \frac{1}{j} \left[ X_{ij} - (A_{i-1}) \right]$$

(4)

The West-Howell Algorithm is an optimized version of the one-pass Hanson Algorithm (Hanson, 1975). Basically, the West-Howell Algorithm calculates the covariance matrix in a single pass through the data. The values of $A$ tend to level out as the value of $j$ approaches $M$ which in turn forces CV to approach the respective variances and covariances of the data. This method is not as accurate as equation (1), but equation (1) requires two passes. Equation (2) requires only one pass through the data but is less accurate than is the West. The West Algorithm is a good blend of speed and accuracy (West, 1979).

The matrix CV is symmetrical and square with the respective variances on the diagonal. This satisfies the definition of a covariance matrix. (See Subroutine COVAR in Appendix A program listing.)
The next step in obtaining principal components is to calculate the eigenvalues and eigenvectors of the covariance matrix \( CV \). Before the actual calculations of these values are discussed, the concept of eigenvalues and eigenvectors will be explored.

At this point variances and covariances of the \( K \), \( Bi \), and \( Tl \) data have been calculated and set into a matrix labeled \( CV \). The elements of \( CV \) or any other \( N \times N \) matrix may be thought of as defining coordinates in \( N \)-dimensional space. If the \( K \) and \( Bi \) variances and covariance are considered then these values may be expressed in two dimensional space on paper. If the \( K \) variance is plotted on the \( X \) axis, then the \( Bi \) variance may be plotted on the \( Y \) axis and the \( K, Bi \) covariance may be plotted as shown in Figure 1.

![Figure 1. A Plot Illustrating a Two-Dimensional Representation of a Covariance Diagram.](image-url)
Figure 1 also shows two additional lines or vectors, $V_1$ and $V_2$. These two vectors show the structure or relationship of the $K$ and $Bi$ variances with respect to each other. The magnitudes of $V_1$ and $V_2$ define points in two dimensional space and may be regarded as lying on the boundary of an ellipse. To define this ellipse the principal axes must first be determined. This is accomplished by calculating the eigenvalues and eigenvectors. The eigenvectors are the direction cosines of the principal axes. The eigenvalues are the magnitudes of these axes. Figure 2 shows these axes superimposed over the covariance plot shown in Figure 1.
Figure 2 shows the principal axes as being perpendicular or orthogonally positioned to each other. It may also be seen that one axis is of a greater magnitude than the other. Both effects are indicative of eigenvalues and eigenvectors.

The eigenvalues, $\lambda$, are defined as a set of values that when multiplied by the identity matrix, $I$, and subtracted from the covariance matrix $CV$ leaves the determinant of the covariance matrix equal to zero.

Let $CV^*$ be a new matrix which accounts for the original matrix $CV$ and the eigenvalues, $\lambda$, and has a zero determinant. For example, in our three-parameter case of $Bi$, $K$, and $Tl$:

$$CV^* = CV - \lambda I$$

$$\begin{bmatrix}
CV^*_{11} & CV^*_{12} & CV^*_{13} \\
CV^*_{21} & CV^*_{22} & CV^*_{23} \\
CV^*_{31} & CV^*_{32} & CV^*_{33}
\end{bmatrix} =
\begin{bmatrix}
(CV_{11} - \lambda) & CV_{12} & CV_{13} \\
CV_{21} & (CV_{22} - \lambda) & CV_{23} \\
CV_{31} & CV_{32} & (CV_{33} - \lambda)
\end{bmatrix}$$

(5)

Where:

$\lambda$ is the unknown eigenvalue and $I$ is the identity matrix. The identity matrix has ones on the diagonal elements and zeroes on all of the off-diagonal elements.
The next step is to calculate the determinant of the matrix $CV^*$. A determinant is a single number extracted from a square matrix by a series of operations of the form represented in equation (6). The determinant is represented symbolically by $\det CV^*$ or $|CV^*|$ and is defined as the sum of N terms of the form:

$$|CV^*| = \sum_{i=1}^{k} (-1)^k CV^*_{i1} \cdot CV^*_{i2} \cdot \ldots \cdot CV^*_{in}, \quad (6)$$

where $n$ is the number of rows (or columns), the subscripts $i_1, i_2, \ldots, i_n$ taken in order, and $k$ is the number of exchanges of two elements necessary to place the $i$th subscripts in the order of $1, 2, \ldots, n$. Each term contains one and only one element from each row and each column. In the case of the matrix $CV^*$, the determinant may be calculated in the following way:

$$\text{TEMP1} = (-1)^k CV^*_{11} \cdot CV^*_{22} \cdot CV^*_{33} : \text{No exchanges are necessary so } k \text{ is even and the sign of TEMP1 remains unchanged.}$$

$$\text{TEMP2} = (-1)^k CV^*_{11} \cdot CV^*_{23} \cdot CV^*_{32} : \text{One exchange is necessary so } k \text{ is odd and the sign of TEMP2 is changed.}$$

$$\text{TEMP3} = (-1)^k CV^*_{12} \cdot CV^*_{23} \cdot CV^*_{31}$$
$$= (-1)^k CV^*_{12} \cdot CV^*_{31} \cdot CV^*_{23}$$
$$= (-1)^k CV^*_{31} \cdot CV^*_{12} \cdot CV^*_{23}$$

where TEMP1, TEMP2, and TEMP3 are being used to demonstrate the sign changes necessary in calculating the determinant of $CV^*$. 

14
There will be 3! (factorial) or 6 of these calculations for the matrix CV*.

Once all the temporary values have been determined they must be summed, resulting in the determinant of CV*. But because CV* has the unknown value of \( \lambda \) included in the temporary variables, the determinant of CV* will take the form as shown in equation (7).

\[
|CV*| = a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4
\]  

(7)

The values \( a_1, a_2, a_3, \) and \( a_4 \) are known, leaving only the \( \lambda \) to be determined. The values of \( \lambda \) are equal to the roots of equation (7). This satisfies the condition that the determinant of CV* be equal to zero. Also, because equation (7) is a cubic, there may be three real solutions of \( \lambda \). The roots extracted from equation (7) are:

\[
(\lambda + b_1) \cdot (\lambda + b_2) \cdot (\lambda + b_3) = 0
\]

\[
\lambda_1 = -b_1, \quad \lambda_2 = -b_2, \quad \lambda_3 = -b_3
\]  

(8)

The eigenvectors are the principal components of the data set and they are calculated next.

The eigenvectors are defined as a set of values that form a matrix \( \mu \) such that when the matrix CV* is multiplied by \( \mu \) the result is zero. In matrix form:

\[
CV* \cdot \mu = 0 \text{ or } (CV - \lambda I) \mu = 0
\]  

(9)
The determination of \( \mu \) must be done in columns. To calculate the first column of \( \mu \), substitute the first value of \( \lambda \) into equation (9). Then using a temporary vector, \( T \), equation (9) becomes:

\[
\begin{bmatrix}
CV_{11} - \lambda_1 & CV_{12} & CV_{13} \\
CV_{21} & CV_{22} - \lambda_1 & CV_{23} \\
CV_{31} & CV_{32} & CV_{33} - \lambda_1
\end{bmatrix}
\begin{bmatrix}
T_1 \\
T_2 \\
T_3
\end{bmatrix}
= 0
\]  

(10)

Multiplying equation (10) gives:

\[
T_1 (CV_{11} - \lambda_1) + T_2 CV_{12} + T_3 CV_{13} = 0 
\]  

(11)

\[
T_1 CV_{21} + T_2 (CV_{22} - \lambda_1) + T_3 CV_{23} = 0 
\]  

(12)

\[
T_1 CV_{31} + T_2 CV_{32} + T_3 (CV_{33} - \lambda_1) = 0 
\]  

(13)

Solving for \( T_1 \) in equations (11) and (12) gives:

\[
T_1 = \frac{-T_2 CV_{12} - T_3 CV_{13}}{(CV_{11} - \lambda_1)} 
\]  

(14)

\[
T_1 = \frac{-T_2 (CV_{22} - \lambda_1) - T_3 CV_{23}}{CV_{21}} 
\]  

(15)
Because equations (14) and (15) are both equal to $T_1$ it follows that they are equal to each other. Using this equality and solving for $T_2$ gives:

$$T_2 = \frac{T_3[CV_{13}CV_{21} - CV_{23}(CV_{11} - \lambda_1)]}{(CV_{11} - \lambda_1)(CV_{12} - \lambda_1) - CV_{12}CV_{21}}$$  \hspace{1cm} (16)$$

Looking at equation (16) all values are known except $T_2$ and $T_3$. Substituting a constant "A" for the known values, equation (16) becomes:

$$T_2 = T_3A$$  \hspace{1cm} (17)$$

If the value "1" is assumed for $T_2$ then $T_3$ becomes:

$$T_3 = 1/A$$  \hspace{1cm} (18)$$

Substituting the values of $T_2$ and $T_3$ into equation (14) gives:

$$T_1 = \frac{-CV_{22} + \lambda_1 - CV_{23}/A}{CV_{21}}$$  \hspace{1cm} (19)$$
All values are known so $T_1$ is also known. At this point the first column of $u$ may be calculated by:

$$u_{11} = \frac{T_1}{(T_1^2 + T_2^2 + T_3^2)^{1/2}} \quad (20)$$

$$u_{21} = \frac{T_2}{(T_1^2 + T_2^2 + T_3^2)^{1/2}} \quad (21)$$

$$u_{31} = \frac{T_3}{(T_1^2 + T_2^2 + T_3^2)^{1/2}} \quad (22)$$

The calculations of the second and third columns of $u$ are accomplished in the same manner with the exception that $\lambda_2$ is substituted into equation (10) when solving for the second column of $u$ and $\lambda_3$ is substituted into equation (10) when solving for the third column of $u$.

The eigenvectors and principal components are the same and will now be referred to as PC.

At this point a new, orthonormalized data set will be generated from the original data using equation (23). The normalization permits each of the three resulting orthogonalized distributions to be compared on a uniform basis. The orthonormalization is given by:

$$Y_{ij} = \sum_{j=1}^{3} \text{PC}_{jk} \cdot \frac{(X_{ji} - \bar{X}_j) / \lambda_k^{1/2}} \quad (23)$$

Where: $i$ is the $i$-th sample, $j$ represents the K, Bi, or Tl data values, and $k$ represents the principal component associated with K, Bi, and Tl respectively. This new data set represents the variation of the variate explained by each principal component. The $j$-th variable of the new data will be uncorrelated, orthogonally positioned, to the other variables of the data set.
The reader should understand that there are several methods in existence that calculate the eigenvalue and eigenvectors from the covariance matrix. This paper has dealt with one of the more basic methods in an effort to make clear the concepts of eigenvalue and eigenvector calculations. The PCA program, on the other hand, utilizes the Jacobi Method of eigenvalue and eigenvector extraction. This method is much more efficient from a standpoint of program size and speed of execution than would be a program designed to rigidly follow the method put forth in this paper. The Jacobi Method (Van Nostrands, 1967 and Carnahan et al., 1973) begins with the premise that eigenvectors are orthogonally positioned with respect to one another and, therefore, may be calculated using combinations of sine and cosine functions. It also assumes that the eigenvalues be considered when calculating the eigenvectors (equation 9). These assumptions enable the program to calculate the eigenvalues and eigenvectors simultaneously. This accounts for a significant savings in computer time and reduces the program size.

At the start of the program the identity matrix is initialized, the object matrix is input (the matrix upon which the eigenvalues and eigenvectors are to be calculated), and a threshold value is determined. This threshold value (determined by the degree of accuracy desired, usually $10^{-9}$) is used to decide which elements in the object matrix are to be used in the next set of calculations and which are not. Should an element value be less than this threshold value it is bypassed and the next element is considered. This decision is made at the beginning of a series of iterations which, if the element value is greater than the threshold, normalize that element and calculate the sine and cosine of it with respect to another factor. This factor is created from its neighboring diagonal elements. The sine and cosine values are then used to calculate a replacement value for that element in the object matrix as well.
as a new value for the parallel element in the identity matrix. Once this iterative process is complete, all the elements of both matrices have been replaced. The matrix that had initially been the identity matrix contains the eigenvectors and the matrix that originally had been the object matrix now has the eigenvalues as its diagonal elements. The program then ranks the eigenvalues in descending order and ends execution.

PRINCIPAL COMPONENTS ANALYSIS VERSUS FACTOR ANALYSIS

Principal components analysis and factor analysis are often misunderstood in that the differences between them are not fully appreciated. The following discussion leans heavily on the work of K. Campbell (written communication) and Jöreskog et al. (1976). The fundamental differences between principal components analysis and factor analysis are (1) which factors are defined, and (2) the assumptions made in regards to the residuals. (Residuals are defined as the differences between actual structure of the data and the structure predicted by the calculated factors.)

The principal components analysis determines factors that account for all the variance of all the observed variables. The factors determined in factor analysis account maximally for the intercorrelation of all the variables. Hence, it can be said principal components analysis is variance oriented whereas factor analysis is correlation oriented.

The residual terms are assumed to be small in principal components analysis, however, the converse is true in factor analysis. This implies that a large part of the total variance in a variable is important in the principal components analysis. On the other hand, factor analysis uses only that part of a
variable that correlates with the other variables. In both methods the re-
siduals are assumed to be uncorrelated with the factors that are determined.

In the principal components analysis there are no assumptions made about the
residuals with respect to themselves, whereas in factor analysis the residuals
are assumed to be uncorrelated among themselves.

Finally, the goal of principal components analysis is to reproduce, or fit
both the diagonal and off-diagonal elements of the covariance matrix. The
goal of factor analysis is to fit the off-diagonal elements only, that is, the
covariance elements and not the variance elements.

PROGRAM OPERATION

The PCA program is comprised of three sections. The first is named PCOMP
and is the executive section where data file names are input, the covariance
or correlation matrix is calculated, and where eigenvalues and eigenvectors
are determined. The second section, HISPC, is the histogram routine that
normalizes the original data (equation 23), determines the histograms to be
printed and/or plotted, and creates the outlier data files. The third section
called PIC3 creates the outlier maps.

Outlier data selection is one function of the histogram routine, HISPC.
Selection has been accomplished in two ways: 1) by stripping out the
normalized data values that are equal to or exceed a user determined sensitivity,
and, 2) by sorting out normalized data values that fall in a user defined window.
The window method proved to be more desirable because of its nature. The user
is able to map a specific range of data. For example, if the histogram of a
data set shows it to be bimodal, the user could map only one of the modes without having to consider the data that fell above the upper bound of the mode.

Another option of the PCA program takes the square roots of the data before any other calculations are performed. This method minimized the tendency of the variances to increase in proportion to the means of the data. The principal components experienced small changes but tended to better represent inhomogeneities in ground concentrations.

AN EXAMPLE FROM AERIAL RADIOMETRIC DATA

An example of a principal components analysis is presented using the three radiometric observations (Bi, T1, and K) obtained from 20,051 data points. The data for this example were collected over a portion of the Texas Coastal Plain by Geodata International, Inc. in 1974 (Geodata International, Inc., 1975). In the Geodata report the area under investigation is referred to as Detail Area C. Figure 3 illustrates the flow of processing data with the principal components program presented in this user's guide.

The first analytical decision facing the user is whether to use the covariance matrix or correlation matrix. As a general rule, a covariance matrix should be used when all variables are expressed in the same physical unit, for example, concentrations in parts per million (ppm). However, when the input variables are of different units, e.g., ppm and percentage, then the analyst is advised to use the correlation matrix. This choice of calculation techniques is required whether or not aerial radiometric data is under consideration. In the present example all of the variables are expressed in terms of counts per second; thus, the covariance matrix is utilized.
Figure 3. Processing Flow Chart for Principal Components Analysis Program.
Figure 3. (con't.)
Figure 3. (con't.)
The user next decides whether or not to take the square root of each individual observation. If the variance of observations of an element has a tendency to increase in proportion to the mean, a phenomenon characteristic of true count data, the square root of each observation should be taken to reduce the artificial trend. Since aerial radiometric data, even after considerable reduction, exhibit this characteristic of count rate data, the square root of each observation was taken prior to the principal components analysis.

The remaining options depend on the end product desired by the user. One can analyze all available data, or in the case of the aerial data, only that data on certain flight lines. All geologies can be analyzed together or an individual geology may be analyzed. In this example all available data points were analyzed without considering individual formations.

There are four output sections to each analysis. The output sections for this example are listed at the end of this section. Output 1 (Figure 5) lists pertinent statistics needed by the geologist performing the analysis. The mean vector, with the values in the order of parameter input, is listed. The covariance matrix appears next. The diagonal of this matrix equals the variance of the parameter. Examination of Output 1 reveals that the variance does not increase in proportion to the mean; thus, the square root transformation has aided in correcting the count rate statistics problem.

The eigenvalues are used to determine the amount of variation explained by each principal component. The percent of variation explained is obtained by dividing each eigenvalue by the sum of the eigenvalues.
Next are the principal components. The columnar loadings are in the same order as the parameters were input. These loadings are used to help the geologist interpret the structure of the data and are used in creating the histograms in Output 2 (Figure 6) and Output 3 (Figure 7). They are also instrumental in creating the displays of Output 4 (Figure 8).

Finally the correlations between the principal components and the original variables are listed.

With the given information the geologist may begin to interpret the data structure. There may be times when only the structure of a data set may be obtained. If geologic significance cannot be assigned to the analysis, the interpreter may always take the results obtained and compare them to results from similar analyses. Geometrics, Inc. uses this approach in their principal components analysis of aerial data. (For an example see Geometrics, Inc., 1978.) In their approach they perform a principal components analysis on a favorable formation for uranium resources in a quadrangle and then do the same for other formations and look for results that are similar to the results obtained from favorable formation analysis.

The histograms provided in Output 2 allow the interpreter to see the frequency of observations represented in class intervals of Output 3. Output 2 and Output 3 are generated for each of the components, and as a program option either output can be displayed without the other. The histograms aid in revealing a data structure that may have geologic significance. For example, a bimodal histogram as seen in Figure 6 for principal component 1 may be indicating the mixing of geologies when an analysis is performed without regard to separating formations or it may reflect facies of mineralogical differences.
when a single formation is being analyzed. A bimodal second principal component histogram may be representing various degrees of radioelement mobility while the third principal component histogram may be depicting non-mobilized radioelement distributions (Pirkle et al., 1980; Pirkle in preparation). Depending on the loadings of elements on the principal components other interpretations may be considered for the histograms.

Output 4 (Figure 7) allows a spatial viewing of points comprising portions of the histograms that are of interest to the interpreter. All points greater than or less than some critical value (sensitivity parameters) may be plotted to give an outlier map (Figure 7), or all points between two values may be plotted to help determine if a significant spatial clustering is observed. If significant clusterings occur than the investigator has a clue as to where additional data may be needed and the type of data needed to verify an interpretation.

The following will illustrate how one may use Output 3 (Figure 6) and Output 4 (Figure 7). Looking at the histogram PC-1 in Figure 6 an investigator may be interested in the interval 2-3. There appears to be a slight increase in the number of observations found in this interval. A plot of the geographic position of the points in this region (Figure 7, Principal Component 1) reveals two distinct groupings of data. If a generalized geologic map (Figure 8) is overlain, it appears that these points occur in the alluvium and part of the undivided Jackson. Thus, it seems geology has influenced this histogram. Pirkle (in preparation) provides more details as to how the geologist may develop an interpretation of aerial radiometric data from a principal components analysis.
Figure 4. Output 1. Statistics Derived from a Principal Components Analysis. *

*Order of variable output is the same as the order of the variable input. In this analysis the order is K, Bi, TI.

Figure 5. Output 2. Frequency Distribution for the Principal Components.
Figure 6a. Output 3. Histograms of the Principal Component 1.*
*Horizontal axis is standard deviation. Vertical axis is frequency.

Figure 6b. Output 3. Histograms of the Principal Component 2.*
*Horizontal axis is standard deviation. Vertical axis is frequency.
Figure 6c. Output 4. Histograms of the Principal Component 3.*
*Horizontal axis is standard deviation. Vertical axis is frequency.
Figure 7. Output 4. Plot of Points of Interest to Interpreter for All Three Components.*

*These plots show points that fall between 2.00-3.00 standard deviations. Figures 7a, 7b, and 7c show the region begin plotted.

(A geology overlay of Figure 8 may be found in the jacket at the end of this report.)
Figure 7. (con't.)

TEXAS GULF COAST DATA
PRINC COMP 2 SENSITIVITY 2.00 - 3.00 MNL PMTS PLOTTED 413
TEXAS GULF COAST DATA
PRINC COMP 3 SENSITIVITY 2.00 - 3.00 MM PTS PLOTTED 522

Figure 7. (con't.)
EXPLANATION

1. Alluvium
2. Fluviatile terrace deposits (higher gravels)
3. Beaumont and Lissie Fms
4. Willis Fm
5. Fleming Fm
6. Oakville Fm
7. Catahoula Fm
8. Jackson Group (undivided)
9. Yegua Fm

Figure 8. Generalized Geology of Detail Area C. *

*A clear film of this figure that may be overlain onto Figure 7 may be found in the jacket at the end of this report.
APPENDIX A

Program Listing
### APPENDIX A

#### Program Listing

<table>
<thead>
<tr>
<th>Program Name</th>
<th>Function</th>
</tr>
</thead>
</table>
| NEWPC        | Receives input and output data file names  
Calculates data means, variances, and covariances  
Prints output of calculated values  
calls other subroutines |
| INST         | Prints program execution instructions |
| HEAD         | Receives outlier map information |
| READIT       | Inputs data from input data file |
| EIGEN        | Calculates eigenvalues and eigenvectors |
| HISPC1       | Calculates histogram printouts and bar graphs utilizing  
AG II (Advanced Graphics 2) plot functions |
| PIC3         | Plots outlier data location on a Mercator projection  
Outputs map title information |
| BASG         | Determines map dimensions in latitude and longitude  
Initializes Mercator projection subroutine |
| MRCMAP       | Plots outside border  
calls NMERC, NMERCP, NMERCC, and ANOTE, all utilizing  
Plot 10 functions |
| NMERC        | Receives initialization information for Mercator projection  
Selects calculation constants for Mercator projection |
| NMERCP       | Initializes the north–south and east–west physical scale factors |
| NMERCC       | Receives latitude and longitude information and returns  
X–Y information (Mercator projection subroutine) |
| TAN          | A function which returns the tangent of an argument in radians |
| SCNF         | A function which returns a plus, minus, or zero determined for a cotangent calculation in the Mercator projection subroutine |
| ANOTE        | Plots annotation of map border |
**Program NEWPC**

**Array**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORR</td>
<td>Correlation with original data</td>
</tr>
<tr>
<td>NSAVE</td>
<td>Number of outliers for PC1-PC3</td>
</tr>
<tr>
<td>TSXY</td>
<td>Weighted mean values</td>
</tr>
<tr>
<td>CVARA</td>
<td>Working array for the subroutine EIGEN containing covariance initially and eigenvalues finally</td>
</tr>
<tr>
<td>S</td>
<td>Standard deviations for the three variates</td>
</tr>
<tr>
<td>ITITL</td>
<td>Passes to the subroutine HEAD and returns with the title printed across the top of each outlier map</td>
</tr>
<tr>
<td>IFMT</td>
<td>The input data format</td>
</tr>
<tr>
<td>IFILE</td>
<td>Device numbers for the outlier map data</td>
</tr>
</tbody>
</table>

**Labeled Common**

<table>
<thead>
<tr>
<th>Block</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLK1</td>
<td>XBAR</td>
<td>Variate mean values</td>
</tr>
<tr>
<td></td>
<td>CVAR</td>
<td>Variate covariance values</td>
</tr>
<tr>
<td></td>
<td>XLAM</td>
<td>Eigenvalues</td>
</tr>
<tr>
<td></td>
<td>PC</td>
<td>Eigenvectors or principal components</td>
</tr>
<tr>
<td>BLK2</td>
<td>XLONG</td>
<td>Data longitude values</td>
</tr>
<tr>
<td></td>
<td>XLAT</td>
<td>Data latitude values</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>Data Values</td>
</tr>
<tr>
<td></td>
<td>CRPA</td>
<td>ICRT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Plot 10 internal flag for outputting to a tektronix terminal</td>
</tr>
<tr>
<td></td>
<td>IPAPER</td>
<td>Plot 10 internal flag for outputting to a Calcomp plotter</td>
</tr>
</tbody>
</table>

**COVAR**

West Algorithm for covariance calculation
Subroutine Eigen

Array

A
Initially covariance values and
eigenvalues at completion

B
Initially zeroes and eigenvectors or
principal components at completion

Subroutine HISPCI

Array

TH1
The histogram for PC1

TH2
The histogram for PC2

TH3
The histogram for PC3

Z
The base or X axis values for TH1,
TH2, and TH3 or Y axis values

ITL1
ITL2
ITL3
Contains the ADE characters for the
PC1, PC2, and PC3 titles at the top
of each histogram

All other arrays were previously
described.

Subroutine PIC3

Array

MTIT
Title printed across each outlier map

ITIT
Heading containing encoded
information as to which PC was
mapped, the sensitivity used, and the
number of normalized data points
exceeding the sensitivity

Labeled Common

MPORG XO
X value in inches of the longitude
from the central meridian

YO
Y value in inches of the latitude
from the central meridian

XORG
X value in inches of the origin
(lower left corner of map)

YORG
Y value in inches of the origin
(lower left corner of map)
Subroutine BASG

Array

Contents

**XEW**  
X measurements in inches of the actual east-west map dimension

**YEW**  
Y measurements in inches of the actual east-west map dimension (usually zero)

**XNS**  
X measurements in inches of the actual north-south map dimension (usually zero)

**YNS**  
Y measurements in inches of the actual north-south map dimension

Labeled Common

**DATA CMERID**  
Central meridian longitude in decimal degrees

**CLAT**  
Latitude of the southwest corner of the map

**CLON**  
Longitude of the southwest corner of the map

**SPA**  
Spheroid constants selection flag

**SCAL**  
Map scale inch/inches

**PSCEW**  
Physical scale factor for the theoretical east-west distance produced by the mapping equation (actual distance/theoretical distance)

**PSCNS**  
Physical scale factor for the theoretical north-south distance produced by the mapping equation (actual distance/theoretical distance)

**DEGEW**  
Map size east-west in decimal degrees

**DEGNS**  
Map size north-south in decimal degrees

**TICSZ**  
North-south annotation interval in decimal degrees
Subroutine MRCMAP

All labeled common blocks have previously been described.

Subroutine NMERC

<table>
<thead>
<tr>
<th>Array</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>TANPHI</td>
<td>Stores intermediate values for each point plotted</td>
</tr>
<tr>
<td>COTPHI</td>
<td>Reciprocal of TANPHI</td>
</tr>
<tr>
<td>A and B</td>
<td>Stores intermediate values used in the calculations of TANPHI</td>
</tr>
<tr>
<td>AZ</td>
<td>Stores the first of two UTM constants</td>
</tr>
<tr>
<td>BZ</td>
<td>Stores the second UTM constant</td>
</tr>
</tbody>
</table>

Subroutine READIT

All arrays have been previously described.

Subroutine HEAD

<table>
<thead>
<tr>
<th>Array</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>JTITL</td>
<td>Used in subroutine BASG under the name of MTITL</td>
</tr>
<tr>
<td>ITITL</td>
<td>Described previously</td>
</tr>
</tbody>
</table>
Subroutine INST

Array

IT

Contents

Receives documentation file instructions

Subroutine COVAR

Array

X

CVAR

AM and AMl

Contents

Described previously

Described previously

Stores intermediate values for covariance calculations
PROGRAM PCOMP
C**********************************************************************
C PROGRAM PCOMP
C MODEL NUMBER: 3.0
C PURPOSE: TO CALCULATE THE PRINCIPAL COMPONENTS OR EIGENVALUES AND EIGENVECTORS FROM A MULTIVARIATE DATA SET
C PROGRAMMER: CHUCK KOCH
C**********************************************************************

DIMENSION CORR(3,3),NSAVE(3),CVARA(3,3),SENSE(2)
DIMENSION S(3),ITITL(10),IFMT(16)
DIMENSION XLONG(1000),XLAT(1000),X(3,1000),IFILE(3)
COMMON /BLK1/XBAR(3),CVAR(3,3),XLAM(3),PC(3,3)
COMMON /CRPA/ICRT,IPAPER

DATA IFILE/23,24,25/

DATA IFILE1/20/

DATA IFILE2/21/

DATA IFILE3/22/

DATA ICRT/1/
DATA IPAPER/0/

ITE = 5
C INITIALIZE THE GRAPHICS LIBRARY PLOT-10
C
CALL INITT(120)
WRITE(ITE,981)
READ(ITE,982) IANS
IF (.IANS.EQ.'NO') GOTO 983
C
THE SUBROUTINE INST PRINTS TO THE TERMINAL
A DOCUMENTATION FILE STATING THE NECESSARY INPUTS
TO SUCCESSFULLY RUN THIS PROGRAM.
C
CALL INST
CALL ERASE
983 CONTINUE
KOUNT = 0
NPTS = 0
DO 5 K=1,3
XBAR(K) = 0.
DO 5 J=1,3
CVR(J,K) = 0.
WRITE(ITE,1001)
OPEN(UNIT=IFILE1,DIALOG)
WRITE(ITE,1002)
OPEN(UNIT=IFILE2,DIALOG)
WRITE(ITE,1013)
READ(ITE,982) IFLG
WRITE(ITE,2011)
READ(ITE,2012) IFMT
DO 1112 JT=1,3
WRITE(ITE,1020) JT
OPEN(UNIT=IFILE(JT),DIALOG)
1112 CONTINUE
CALL ERASE
C
THE HEAD SUBROUTINE ENABLES THE INPUTS OF
ALL THE INFORMATION NEEDED TO GENERATE THE
HISTOGRAMS AND OUTLIER MAPS.
C
CALL HEAD(IFILE3,NCELLSHMINHMAXKPLOTITITL)
C
INPUT THE SENSE VALUE. THIS VALUE IS USED IN THE HISPC
SUBROUTINE AS THE DECIDING FACTOR AS TO WHICH DATA
POINTS ARE TO BE USED IN THE OUTLIER MAPS.
C
1118 CONTINUE
WRITE(ITE,1113)
READ(ITE,1114) SENSE(1)
WRITE(ITE,1115)
READ(ITE,1114) SENSE(2)
C
CHECK TO MAKE SURE THAT THE MAX SENSE IS
GREATER THAN THE MIN SENSE
C
IF(SENSE(1).LT.SENSE(2)) GOTO 1117
WRITE(ITE,1116)
GOTO 1118
1117 CONTINUE
C
READ THE FLAG TO CALCULATE COVARIANCE OR CORRELATION
C
118 CONTINUE
WRITE(ITE,2006)
READ(ITE,2007) ICORR
IF (ICORR.EQ.'SHCORRE').OR.(ICORR.EQ.'SHCOVAR')) GOTO 117
WRITE(ITE,2008)
GOTO 118
117 CONTINUE
C
READ LINE NUMBER, LAT, LON, GEO, K, U AND TH DATA
C
10 CALL READIT(XLAT,XLONG,X,NPTS,IFILE1,IFMT)
IF PRINCIPAL COMPONENTS HAVE ALREADY BEEN CALCULATED DO NOT REREAD NOR RECALCULATE THEM. HOWEVER ONE CALL TO READIT MUST BE MADE IN ORDER TO SET THE OPTIONS.

IF(IFLG.NE.2HNO) GOTO 1021
IF(NPTS.EQ.0) GOTO 100

COUNT THE TOTAL NUMBER OF POINTS READ AND STORE IN KOUNT.
KOUNT=KOUNT+NPTS

SUM THE K, U AND T DATA AND STORE IN XBAR.
CALCULATE THE COVARIANCES OF THE K, U AND T DATA AND STORE IN THE MATRIX CUAR ALSO LOAD THE WORKING ARRAY CVARA FROM CUAR.

DO 50 I=1,NPTS
   DO 40 K=1,3
      XBAR(K)=XBAR(K)+X(K,I)
   40 CONTINUE
50 CONTINUE
CALL COVAR(XNPTS,CVAR)
GO TO 10

XN=FLOAT(KOUNT)

XBAR CONTAINS CALCULATED DATA MEANS

DO 120 K=1,3
   XBAR(K)=XBAR(K)/XN
120 CONTINUE

DO 115 J=K,3
   CUAR(K,J)=CUAR(K,J)/(XN-1.)
115 CVAR(J,K)=CVAR(K,J)

S(K)=SQRT(CUAR(K,K))

IF COVARIANCE MATRIX IS TO BE USED GO TO 116

IF(ICORR.EQ.5HCOVAR) GOTO 116
DO 211 I=1,3
   DO 211 J=I,3
      CUAR(I,J)=CUAR(I,J)/(S(I)*S(J))
211 CONTINUE

DO 110 I=1,3
   DO 110 J=I,3
      CVARA(I,J)=CUAR(I,J)
110 CONTINUE

STORE THE K, U AND T STANDARD DEVIATIONS IN THE VECTOR S

M = 3
INITIALIZE M TO 3
CALL EIGEN SUBROUTINE TO EXTRACT THE EIGENVALUES AND EIGENVECTORS FROM THE WORKING ARRAY CUARA. CUARA AND PC ARE 3X3 MATRICES
50 PASS M,M TO EIGEN.
CALL EIGEN(CVAR,PC,M,M)
LOAD THE VECTOR XLAM FROM THE DIAGONALS OF CUARA WHICH CONTAINS THE EIGENVALUES.
DO 1331 L=1,3
   PRINT *,XLAM(L),XLAM(L)
1331
WRITE THE MEAN VECTOR, COVARIANCE ARRAY, EIGENVALUE VECTOR, PRINCIPAL COMPONENT ARRAY AND K, U AND T STANDARD DEVIATIONS TO A FILE THROUGH UNIT IFIL2.
WRITE(IFILE2,1000) XBAR
WRITE(IFILE2,1000) CVAR
WRITE(IFILE2,1000) XLAM
WRITE(IFILE2,1000) PC
WRITE(IFILE2,1000) S
GOTO 1031

C IF THE PRINCIPAL COMPONENTS ETC. WERE PREVIOUSLY
C CALCULATED THEN READ THEM FROM A FILE THROUGH UNIT
C IFILE2.

1021 CALL ERASE
READ(IFILE2,1000) XBAR
READ(IFILE2,1000) CVAR
READ(IFILE2,1000) XLAM
READ(IFILE2,1000) PC
READ(IFILE2,1000) S
GOTO 1031

1031 CONTINUE

C PRINT OUT TO THE TERMINAL ALL CALCULATED VALUES
C AND MAKE A HARD COPY OF THEM.
C
WRITE(ITE,2000) XBAR
IF(ICORR.EQ.5HCORRE) GOTO 1090
WRITE(ITE,2001) CVAR
GOTO 1091

1090 WRITE(ITE,2009) CVAR
1091 CONTINUE
WRITE(ITE,2002) XLAM
WRITE(ITE,2003)((PC(JK),K-1,3),J-1,3)

C COMPUTE CORRELATION OF EACH PRINCIPAL COMPONENT
C WITH EACH ORIGINAL VARIATE.
C
DO 160 K-1,3
DO 160 J-1,3
160 CORR(J,K)=PC(J,K)*SQRT(XLAM(K))/S(J)
WRITE(ITE,2004)((CORR(JK),K-1,3),J-1,3)
CALL HDCOPY

C HISPC IS THE SUBROUTINE WHICH GENERATES THE
C HISTOGRAMS OF PC1,PC2 AND PC3 AND IS RESPONSIBLE FOR
C CREATING THE 3 FILES WHICH CONTAIN THE OUTLIER
C MAP INFORMATION.
C
CALL HISPC(NCELLS,IFILE1,SENSE,MSAVE,ITITL
1,DMIN,DMAX,KPLOT,IFILE)
DO 1111 KAT-1,3
IPRINC-KAT
JFILE-IFILE(KAT)
MSAVE-NSAVE(KAT)

C PIC3 IS THE SUBROUTINE WHICH CREATES THE OUTLIER MAPS
C BY USING A MERCADOR PROJECTION CALCULATION.
C
CALL PIC3(JFILE,IFILE3,IPRINC,SENSE,MSAVE,ITITL)
1111 CONTINUE
CLOSE(UNIT-IFILE1)
CLOSE(UNIT-IFILE2)
CLOSE(UNIT-IFILE3)
CLOSE(UNIT-IFILE(1))
CLOSE(UNIT-IFILE(2))
CLOSE(UNIT-IFILE(3))

981 FORMAT(42H DOES THE USER REQUIRE INSTRUCTIONS TO RUN
1 28H THIS PROGRAM (YES OR NO) ? ,)
982 FORMAT(A2)
1000 FORMAT(3E15.8)
1001 FORMAT(34H INPUT DATA FILE TO BE READ FROM ?)
1002 FORMAT(34H INPUT DATA FILE TO HOLD PRINCIPAL
1 14H COMPONENTS ? )
1013 FORMAT(40H HAVE PC S BEEN CALCULATED FOR THIS DATA
1 23H SET BEFORE (YES,NO) ? ,)
SUBROUTINE INST

C THIS SUBROUTINE PRINTS THE DOCUMENTATION
C FILE CALLED PCOMP.DOC DURING PROGRAM EXECUTION
C TO PRECLUDE HAVING TO STOP EXECUTION TO READ THE
C DOCUMENTATION. IT ALSO MAKES A HARD COPY AS IT PRINTS.
C
DIMENSION IT(16)
C
C SET DEVICE 10 TO FILENAME PCOMP.DOC
C
OPEN (UNIT=10,FILE='PCOMP.DOC')
CALL ERASE
DO 1 I=1,60
   READ(10,2,END=3) IT
   WRITE(5,2) IT
   CALL HDCOPY
   CALL ERASE
GOTO 4
3   CALL HDCOPY
   CLOSE(UNIT=10)
2   FORMAT(1X,16A5)
   RETURN
1   FORMAT(12H MEAN VECTOR//3F12.3)
   FORMAT(12H COVARIANCE MATRIX//(3F12.3))
   FORMAT(12H EIGENVALUES//3F12.3)
   FORMAT(37H THE MAX SENSITIVITY IS LESS THAN THE
   1 41H MINIMUM SENSITIVITY RELOAD THESE VALUES )
   FORMAT(/12H PRINCIPAL COMPONENTS//7X,SHFIRST,6X,6HSECOND,
   1 7X,5HTHIRD//(3F12.5))
   FORMAT(/35H CORRELATION WITH ORIGINAL ELEMENTS//(3F12.5))
   FORMAT(2006 FORMAT(36H DOES THE USER WISH TO CALCULATE THE ,/
   1 53H CORRELATION OR COVARIANCE MATRIX (CORRE OR COVAR) ? ,$)
   FORMAT(2007 FORMAT(16H INCORRECT INPUT )
   2008 FORMAT(16H INCOMPLETE INPUT )
   2009 FORMAT(16H CORRELATION MATRIX//(3F12.3))
   2011 FORMAT(16H INPUT THE FORMAT OF THE DATA (80 CHARS) ? )
   2012 FORMAT(16A5)
   STOP
END

SUBROUTINE READIT (XLAT,XLONG,X,NPTSIFILEIIFMT)

DIMENSION XLAT(1000),XLONG(1000),X(3,1000),IFMT(16)

DATA IFLAG/0/,ITE/5/
C
C THIS ROUTINE READS DATA FROM THE FILE ASSIGNED TO UNIT
C IFILEI. IFMT IS THE FORMAT OF THE DATA TO BE READ.
C THE DATA MUST BE ORDERED AS FOLLOWS:
C 1) FLIGHT LINE NUMBER - LINE
C 2) LATITUDE - XLATT
C 3) LONGITUDE - XLONG
C 4) GEOLOGY - NGE0
C 5) DATA VALUE 1 - X1
C 6) DATA VALUE 2 - X2
C 7) DATA VALUE 3 - X3
C
C THIS ROUTINE ALSO HAS PROVISIONS TO SORT THE DATA ON A
C FLIGHT LINE AND/OR GEOLOGIC LABEL BASIS AND MAY OR MAY NOT
C TAKE THE SQUARE ROOTS OF THE DATA. ANY COMBINATION OF THESE
C PROVISIONS MAY BE USED.
C
C INITIALIZE THE POINT COUNTER
C NPTS=0
C IFLAG ALLOWS ONLY ONE READING OF OPTIONS
C IF(IFLAG.EQ.0) GOTO 10
WRITE(ITE,100)
C READ THE FLAG FOR FLIGHT LINE SORT
C READ(ITE,100) IFL
IF(IFL.EQ.2HYE) GOTO 200
WRITE(ITE,101)
C READ THE BEGINNING FLIGHT LINE NUMBER
C READ(ITE,1001) IFL1
WRITE(ITE,102)
C READ THE ENDING FLIGHT LINE NUMBER
C READ(ITE,1001) IFL2
WRITE(ITE,103)
C READ THE FLAG FOR GEOLOGICAL SORT
C READ(ITE,100) IGL
IF(IGL.EQ.2HYE) GOTO 300
WRITE(ITE,104)
C READ THE DESIRED GEOLOGIC CODE
C READ(ITE,1002) IGSRT
WRITE(ITE,105)
C READ THE FLAG TO TAKE SQUARE ROOTS
C READ(ITE,100) ISOR
IFLAG=1
CALL ERASE
10 CONTINUE
C READ INPUT FILE AND TEST FOR END OF FILE.
C READ(IFILE,IFMT,END=965) LIME,XLATT,XLONGG,NGEO,X1,X2,X3
IF(IFL.EQ.2HYE) GOTO 400
C ASSIGN INPUT FILE DATA TO APPROPRIATE MATRICES
C RETURN DATA TO CALLING ROUTINE WHEN MATRICES ARE FULL
C IF(LIME.LT.IFL1) GOTO 10
IF(LIME.GT.IFL2) GOTO 965
400 IF(NGEO.EQ.IGSRT) GOTO 10
410 IF(ISOR.EQ.2HYE) GOTO 420
C CHECK TO SEE IF X1,X2 OR X3 ARE NEGATIVE. IF SO ZERO
C IT BEFORE TAKING THE SQUARE ROOT
C IF(X1.LT.0) X1=0.
IF(X2.LT.0) X2=0.
IF(X3.LT.0) X3=0.
X1=SQRT(X1)
X2=SQRT(X2)
X3=SQRT(X3)
420 CONTINUE
NPTS=NPTS+1
XLAT(NPTS)=XLATT
XLONGG(NPTS)=XLONGG
X(1,NPTS)=X1
X(2,NPTS)=X2
X(3,NPTS)=X3
C LIMIT THE NUMBER OF DATA POINTS READ AT ONE TIME TO 1000
C IF(NPTS.EQ.1000) GO TO 965
GO TO 10
C 965 CONTINUE
100 FORMAT(36H DOES THE USER WISH TO SORT THE DATA,/,1
36H ON A FLIGHT LINE BASIS (YES OR NO) ?,S)
101 FORMAT(37H INPUT BEGINING FLIGHT LINE NUMBER ?,S)
102 FORMAT(35H INPUT ENDING FLIGHT LINE NUMBER ?,S)
103 FORMAT(32H DOES THE USER WISH TO SORT ON A GEOLOGICAL
1 21H BASIS (YES OR NO) ?,S)
104 FORMAT(32H INPUT THE GEOLOGY TO BE USED ?,S)
105 FORMAT(39H DOESS THE USER WISH TO TAKE THE
1 28H SQUARE ROOTS OF THE DATA ?,S)
1000 FORMAT(A2)
1001 FORMAT(I5)
1002 FORMAT(AS)
RETURN
END

SUBROUTINE EIGEN(A,B,N,N1)

C THIS ROUTINE CALCULATES THE EIGENVALUES AND
C EIGENVECTORS OF THE OBJECT MATRIX
C WHICH IS AN N X N1 SYMMETRIC MATRIX.
C UPON COMPLETION THE EIGENVALUES ARE STORED IN THE
C DIAGONAL ELEMENTS OF A AND THE EIGENVECTORS ARE
C STORED IN THE MATRIX B BY COLUMNS.
C DIMENSION A(N1,N1),B(N1,N1)
TS=1.E-09
AN=0.
C CALCULATE THE INITIAL AND FINAL NORMS
C AND SET MATRIX B TO THE IDENTITY MATRIX.
C DO 10 I=1,N
  DO 10 J=1,N
    IF(I-J).NE.2,1,2
      B(I,J)=1.
  GOTO 10
2    B(I,J)=0.
    AN=AN+A(I,J)**2
10   CONTINUE
C INITIALIZE THE INDICATORS AND COMPUTE THE THRESHOLD.
C AN=SORT(AN)
C CALCULATE THE THRESHOLD VALUE
C FN=(AN*TS)/FLOAT(N)
THR=AN
23 THR=THR/FLOAT(N)
3 IND=0
C BEGIN THE ITERATIVE PROCESS TO FIND THE EIGENVALUES
C AND EIGENVECTORS
C DO 20 I=2,N
  II=I-1
  DO 30 J=1,II
    IF (ABS(A(J,I))-THR) 30,4,4
COMPUTE THE SINES AND THE COSINES AND SQUARES EACH WHICH WILL SET THE OFF-DIAGONALS AT 90 DEGREES TO EACH OTHER.

NORMALIZE THE DATA VALUE

\[ \begin{align*}
AL &= A(J,I) \\
AM &= (A(J,J) - A(I,I))/2. \\
AO &= AL/\sqrt{AL^2 + AM^2} \\
TM &= \sqrt{1 - AO^2}
\end{align*} \]

CALCULATE THE SINE AND ITS SQUARE

\[ \begin{align*}
\text{SINX} &= AO/\sqrt{2 \times (1 + TM)} \\
\text{SINX}^2 &= \text{SINX}^2
\end{align*} \]

CALCULATE THE COSINE AND ITS SQUARE

\[ \begin{align*}
\text{COSX} &= \sqrt{1 - \text{SINX}^2} \\
\text{COSX}^2 &= \text{COSX}^2
\end{align*} \]

ROTATE COLUMNS I AND J

\[ \begin{align*}
&\text{DO 40 K=1,N} \\
&\quad \text{IF(K-J)} \ 7,8,7 \\
&\quad \text{IF(K-I)} \ 11,11,11 \\
&\quad \text{ATT} = A(K,J)
\end{align*} \]

CALCULATE NEW DATA VALUES IN THE OBJECT MATRIX USING THE SINES AND COSINES COMPUTED BEFORE

\[ \begin{align*}
A(K,J) &= \text{ATT} \times \text{COSX} - A(K,I) \times \text{SINX} \\
A(K,I) &= \text{ATT} \times \text{SINX} + A(K,I) \times \text{COSX} \\
B(K,J) &= B(K,J)
\end{align*} \]

CALCULATE NEW VALUES IN THE IDENTITY MATRIX USING THE SINES AND COSINES COMPUTED BEFORE

\[ \begin{align*}
B(K,J) &= B(K,J) \times \text{COSX} - B(K,I) \times \text{SINX} \\
B(K,I) &= B(K,I) \times \text{SINX} + B(K,I) \times \text{COSX}
\end{align*} \]

CONTINUE

CALCULATE THE EIGENVALUES AND SET INTO THE DIAGONAL ELEMENTS OF THE OBJECT MATRIX

\[ \begin{align*}
\text{XTT} &= 2 \times A(J,I) \times \text{SINX} \times \text{COSX} \\
\text{ATT} &= A(J,J) \\
\text{BTT} &= A(I,I) \\
A(J,J) &= \text{ATT} \times \text{COSX}^2 + BTT \times \text{SINX}^2 - \text{XTT} \\
A(I,I) &= \text{ATT} \times \text{SINX}^2 + BTT \times \text{COSX}^2 - \text{XTT} \\
A(J,I) &= (\text{ATT} - \text{BTT}) \times \text{SINX} \times \text{COSX} + A(J,I) \times (\text{COSX}^2 - \text{SINX}^2) \\
A(I,J) &= A(J,I)
\end{align*} \]

SET OPPOSITE OFFDIAGONAL ELEMENTS EQUAL (SQUARE MATRIX)

\[ \begin{align*}
&\text{DO 50 K=1,N} \\
&\quad A(J,K) = A(K,J) \\
&\quad A(I,K) = A(K,I)
\end{align*} \]

CONTINUE

IF(IND) \ 60,60,3

CONTINUE

IF(THR-FN) \ 70,70,23

SORT THE EIGENVALUES AND EIGENVECTORS

\[ \begin{align*}
&\text{DO 80 I=2,N} \\
&\quad J-I
\end{align*} \]
29 IF(A(J-1,J-1)-A(J,J)) 90,80,80
90 ATT-A(J-1,J-1)
A(J-1,J-1)*ATT
DO 100 K-1,N
ATT-B(K,J-1)
B(K,J-1)-B(K,J)
100 B(K,J)-ATT
J+J-1
IF(J-1) 80,80,29
80 CONTINUE
RETURN
END

SUBROUTINE HEAD(IFILE3,NCELLS,HMIN,HMAX,KPLOT,ITITL)

C THIS SUBROUTINE GENERATES THE INFORMATION NECESSARY
C FOR THE HISTOGRAMS AND OUTLIER MAPS
C
DIMENSION JTITL(16),ITITL(10)
C
ITE IS THE DEVICE NUMBER OF THE TERMINAL
C
DATA ITE/5/
WRITE(ITE,21)
C
ASSIGN A FILENAME TO UNIT IFILE3
C
OPEN(UNIT=IFILE3,DIALOG)
C
DECIDE WHAT HISTOGRAM OUTPUTS WILL BE DONE
C
WRITE(ITE,25)
READ(ITE,126) KPLOT
WRITE(ITE,23)
C
INPUT THE MINIMUM VALUE FOR THE HISTOGRAMS TO
C BE GENERATED FROM
C
READ(ITE,3) HMIN
WRITE(ITE,41)
C
INPUT THE MAXIMUM VALUE FOR THE HISTOGRAMS TO
C BE GENERATED FROM
C
READ(ITE,3) HMAX
C
CALCULATE THE NUMBER OF CELLS TO BE USED IN THE
C HISTOGRAMS BUT LIMIT TO A MAXIMUM OF 401
C
NCELLS=(HMAX-HMIN)*10.
NCELLS=NCELLS+1
IF(NCELLS.GT.401) NCELLS=401
C
OUTPUT THE NUMBER OF CELLS
C
WRITE(ITE,31) NCELLS
WRITE(ITE,30)
C
INPUT INFORMATION ABOUT THE MAPS
C
READ(ITE,32) ITITL
WRITE(ITE,42)
READ(ITE,43) ITL
C
IF OUTLIER MAP INFORMATION HAS BEEN STORED IN
C IFILE3 PREVIOUSLY DO NOT REENTER IT
C
IF(ITL.NE.2HNO) GOTO 24
C
INPUT NECESSARY MAP INFORMATION
WRITE(ITE,10)
READ(ITE,61) XLAT
WRITE(ITE,12)
READ(ITE,62) XLONG
WRITE(ITE,26)
READ(ITE,63) CMER

MAKE SURE CENTRAL MERIDIAN IS NEGATIVE FOR THE U.S.
CMER=-ABS(CMER)
WRITE(ITE,13)
READ(ITE,64) EUSZ
WRITE(ITE,66)
READ(ITE,64) MSSZ
WRITE(ITE,14)
READ(ITE,65) EUTC
WRITE(ITE,15)
READ(ITE,65) NSTC
WRITE(ITE,16)
READ(ITE,18) JTITL
RITE ALL INFORMATION PLUS 17.46 AND 28.94 TO IFILE3.
THE 17.46 AND 28.94 ARE THE MAP SIZE IN INCHES NECESSARY FOR THE ROUTINE PIC3
WRITE(IFILE3,22)XLAT,XLONG,MSSZ,EUSZ,CMER,EUTC,
1 HSTC,JTITL
21 FORMAT(40H FILE NAME HOLDING HEADER INFORMATION ? )
25 FORMAT(42H SELECT HISTOGRAM PLOTS, PRINTOUT, OR BOTH,/, 1 43H PLOTS ONLY=+1, BOTH=0, PRINTOUT ONLY=-1 , / ,$)
126 FORMAT(I2)
23 FORMAT(45H INPUT THE LOWER LIMIT TO BE USED IN THE PLOTS ?,$)
31 FORMAT(51H THE NUMBER OF CELLS TO BE USED IN THE HISTOGRAMS * 1 ,14,/)
30 FORMAT(35H GIVE SPECIAL DATA TITLE FOR MAP ? )
32 FORMAT(10A5)
12 FORMAT(42H HAS HEADER FILE BEEN WRITTEN (YES,NO) ?,$)
43 FORMAT(A2)
10 FORMAT(32H ENTER THE NORTHWEST CORNER LATITUDE ?,$)
61 FORMAT(F12.5)
12 FORMAT(32H NORTHWEST CORNER LONGITUDE ?,$)
62 FORMAT(F9.3)
26 FORMAT(20H CENTRAL MERIDIAN ?,$)
63 FORMAT(F7.2)
13 FORMAT(32H E-U MAP SIZE (DECIMAL DEGS) ?,$)
64 FORMAT(F7.2)
66 FORMAT(32H N-S MAP SIZE (DECIMAL DEGS) ?,$)
14 FORMAT(26H E-U TIC MARK INCREMENT ?,$)
65 FORMAT(F6.3)
15 FORMAT(26H N-S TIC MARK INCREMENT ?,$)
66 FORMAT(29H MAP TITLE (80 CHARACTERS) ?)
18 FORMAT(16A5)
22 FORMAT(F12.3,F9.3,/,3F7.2,2F6.3,/,16A5,/, 1 10H 17.46,/,10H 28.94)
24 RETURN
END

SUBROUTINE COVAR(X,NPTS,CVAR)
DIMENSION X(3,1000),CVAR(3,3),TEM(3),TEM1(3)
WITH IFLAG=0/

THIS IS THE U.S. ALGORITHM FOR CALCULATING THE COVARIANCE MATRIX
C IFLAG IS USED TO INITIALIZE THE TEMPORARY VARIABLES
C
IF (IFLAG.NE.0) GOTO 100
IFLAG=1
C
ZERO THE COVARIANCE MATRIX
C
DO 50 I=1,3
DO 50 J=1,3
50 CVAR(I,J)=0.
C
INITIALIZE THE TEMPORARY STORAGE VECTOR
C
TEM1(1)=X(1,1)
TEM1(2)=X(2,1)
TEM1(3)=X(3,1)
L=2
LT=0
GOTO 200
100 L=1
200 CONTINUE
C
BEGIN THE CALCULATIONS
C
DO 10 I=1,3
DO 20 J=1,3
C
IF I AND J ARE THE SAME VALUE THEN A VARIANCE
C CALCULATION IS UNDERWAY OTHERWISE ITS A COVARIANCE
C CALCULATION
C
TX=TEM1(I)
TY=TEM1(J)
DO 30 K=L,NPTS
C
CALCULATE THE TOTAL NUMBER OF POINTS UNTIL NOW
A=FLOAT(K+LT)
C
CALCULATE THE DIFFERENCE BETWEEN THE DATA VALUE AND
C THE AVERAGE SUMMED DIFFERENCE
OX=X(I,K)-TX
C
AVERAGE THE DIFFERENCE
RX=OX/A
C
SUMMATE THE AVERAGE DIFFERENCE
TX-TX+RX
C
DO THE SAME CALCULATIONS ON THE SECOND DATA VALUE
QY=X(J,K)-TY
RY-QY/A
TY+TY+RY
C
CALCULATE THE VARIANCE - COVARIANCE MATRIX
C
CVAR(I,J)=CVAR(I,J)+(A-1.)*OX*RY
30 CONTINUE
30 CONTINUE
C
STORE TEMPORARY VALUES
C
TEM(I)=TX
10 CONTINUE
DO 40 I=1,3
40 TEM1(I)=TEM(I)
C
SUM TOTAL NUMBER OF POINTS UNIT NOW
SUBROUTINE HISPC(NCELLS,IFILE1,SEN,NSAVE,IFMT
1,HMIN,HMAX,KPLOT,IFILE)

C
C COMPUTE HISTOGRAMS FOR EACH OF THE THREE PRINCIPAL
C COMPONENTS, AND PRINT THEM OUT AND/OR PLOT THEM
C USING PLOT10 AND AGII.
C
C VECTORS TH1,TH2 AND TH3 CONTAIN THE HISTOGRAMS
C FOR THE THREE PRINCIPAL COMPONENTS. IFILE IS
C A VECTOR CONTAINING THE UNIT NUMBERS OF THE FILES
C WHICH WILL RECEIVE THE OUTLIER DATA.
C
COMMON /BLKI/XBAR(3),CVAR(3,3),XLAM(3),PC(3,3)
DIMENSION TH1(402),TH2(402),TH3(402),Z(402)
DIMENSION IFILE(3),NSAVE(3),IFMT(16),SEN(2)
DIMENSION XLONG(1000),XLAT(1000),X(3,1000)
DIMENSION ITL1(4),ITL2(4),ITL3(4)

C
C SET LABELS TO BE PLACED AT THE TOP OF EACH PLOT
C
DATA ITL1/80,67,45,49/
DATA ITL2/80,67,45,50/
DATA ITL3/80,67,45,51/

C
C REWIND THE FILES TO BE USED FOR THE OUTLIERS
C
REWIND IFILE1
REWIND IFILE(1)
REWIND IFILE(2)
REWIND IFILE(3)

C
C INITIALIZE THE PLOT10 LIBRARY
C
CALL RESET
CALL TERM(3,4096)
CALL CHRSIZ(4)
10 CONTINUE
20 CONTINUE

C
C SET THE CELL WIDTH
C
DH=.1

C
C TAKE THE SQUARE ROOTS OF THE EIGENVALUES IN
C PREPARATION FOR DATA NORMALIZATION
C
DO 30 K=1,3
30 XLAM(K)=1./SQRT(XLAM(K))
Z(2)=HMIN

C
C ZERO THE CELL COUNTERS TH1,TH2 AND TH3 ALSO LOAD
C THE X AXIS VECTOR
C
DO 31 J=2,NCELLS
31 TH1(J)=0.
TH2(J)=0.
TH3(J)=0.

C
C INITIALIZE THE TOTAL POINT COUNTER AND THE OUTLIER
C COUNTERS
C
NTOTAL=0
DO 17 I=1,3
17 NSAVE(I)=0
SORT DATA INTO THREE HISTOGRAMS IN ONE PASS THROUGH IFILE1.

CALL READIT(XLAT,XLONG,X,NPTS,IFILE1,IFMT)
IF(NPTS.EQ.0) GOTO 100

CALCULATE THE TOTAL NUMBER OF POINTS
NTOTAL=NTOTAL+NPTS
DO 90 I=1,NPTS
DO 80 K=1,3
TEMP=0.
90 CONTINUE

CALCULATE THE NORMALIZED PC VALUE AND STORE IT IN TEMP
DO 60 J=1,3
60 TEMP=TEMP+PC(J,K)*(X(J,I)-XBAR(J))
TEMP=TEMP*XLAM(K)

DETERMINE THE CELL THAT THE TEMP VALUE BELONGS IN
NH=IFIX((TEMP-HMIN)/DH)+2
IF(NH.GT.NCELLS+1) NH=NCELLS+1
IF(NH.LT.1) NH=1
GOTO (76,77,78),K

ADD 1 TO THE NUMBER STORED IN THE CORRECT TH VECTOR
76 TH1(NH)=TH1(NH)+1.
GOTO 79
77 TH2(NH)=TH2(NH)+1.
GOTO 79
78 TH3(NH)=TH3(NH)+1.
79 CONTINUE

SUBTRACT THE SENSITIVITY FROM THE NORMALIZED VALUE
AND CHECK TO SEE IF IT IS GREATER THAN THE ZERO, IF SO WRITE LAT
, LONG AND NORMALIZED VALUE TO THE CORRECT FILE.
IF((TEMP.LT.SEN(1)).OR.(TEMP.GT.SEN(2))) GOTO 80

INCREMENT THE OUTLIER COUNTER BY 1
NSAVE(K)=NSAVE(K)+1

OUTPUT TO THE APPROPRIATE FILE
JFILE=IFILE(K)
WRITE (JFILE,3000) XLAT(I),XLONG(I),TEMP
80 CONTINUE
90 CONTINUE
GO TO 50

PRINT OUT HISTOGRAMS
KOUT=10
MOUT=10.IDH
DO 150 K=1,3
CALL ERASE
WRITE(5,2001) K,DH,NCELLS,NTOTAL
150 CONTINUE
C PRINT OUT THE COUNTS IN THE RESPECTIVE VECTORS

120  GOTO (117,118,119),K
117  WRITE(5,2002) I1,XP,(TH1(I),I-I1+1,I2+1)
     GOTO 116
118  WRITE(5,2002) I1,XP,(TH2(I),I-I1+1,I2+1)
     GOTO 116
119  WRITE(5,2002) I1,XP,(TH3(I),I-I1+1,I2+1)
     CONTINUE
116 CONTINUE

I1=I1+1
IF (I1.GT.NCELLS) GO TO 160
I2=MIN0(I2+NROUNCELLS)
XP=XP+DROU
GOTO 120

160 CONTINUE
CALL HDCOPY

C DRAW THE PLOTS

200 IF (KPLOT.LT.0) GO TO 300
DO 250 K=1,3

C INITIALIZE THE ADVANCED GRAPHICS II LIBRARY

CALL BINITT
CALL ERASE
CALL CHRSTZ(4)
CALL PLACE(3HSTD)
X1=600.
X2=3600.
ISET=2*(X2-X1)/(NCELLS-1)
IF((ISET-ISET).GE..5) ISET=ISET+1
CALL XDEN(9)
CALL XFRM(2)
CALL XLEN(40)

C PLOT THE RESPECTIVE HISTOGRAM

12 CONTINUE
CALL UBARST(0,ISET,0)
CALL XFRM(1)
CALL CHECK(2,TH1)
CALL DISPLAY(Z,TH1)
CALL MOVABS(2020,3120)
CALL HLABEL(4,ITL1)
     GOTO 15
13 CONTINUE
CALL UBARST(0,ISET,0)
CALL XFRM(1)
CALL CHECK(2,TH2)
CALL DISPLAY(Z,TH2)
CALL MOVABS(2020,3120)
CALL HLABEL(4,ITL2)
     GOTO 15
14 CONTINUE
CALL UBARST(0,ISET,0)
CALL XFRM(1)
CALL CHECK(2,TH3)
CALL DISPLAY(Z,TH3)
CALL MOVABS(2020,3120)
CALL HLABEL(4,ITL3)
     GOTO 15
15 CONTINUE

IYT=450
IXT=600-(IEST/2)
ITT=ISET10
NT = (NCELLS - 1) / 10
CALL MOVABS(600, 450)
DO 251 I = 1, NT + 1
CALL DRUREL(0, 50)
   IXT + IYT
CALL MOVABS(IXT, IYT)
   CALL CONTINUE
250 CALL HDCOPY
   CALL CONTINUE
C
300 CALL CONTINUE
C
1000 FORMAT(3E15.8)
1001 FORMAT(20H HISTOGRAM FOR P.C., I2)
3000 FORMAT(3F10.4)
2001 FORMAT(/, /, 25H HISTOGRAM FOR COMPONENT, I2, /, 14H CELL WIDTH =
1 FB.3, /, 3X, 10H J START, 4X, 16H H(J), ..., H(J+9), /,
2 17H NUMBER OF CELLS, I4, /, 28H THE TOTAL NUMBER OF POINTS, I6)
2002 FORMAT(I4, FB.2, 3X, 10F6.0)
RETURN
END

SUBROUTINE PIC3(JFILE, IFILE3, IPRINC, SENSE, 1 MSAVE, MTIT)
C
C BENDIX FIELD ENGINEERING CORPORATION
C UNDER U.S. DEPT. OF ENERGY CONTRACT
C EY-76-C-13-1664
C
C*******************************************************************************
C PROGRAM: PIC3
C MODEL NUMBER: 1.0
C PURPOSE: PLOT QUADE OUTLINE AND LOCATIONS OF OBSERVATIONS
C
C*******************************************************************************
C
C RV DATE SRO $ BY REASON FOR CHANGE
C
C
C
C*******************************************************************************
C
C*******************************************************************************
C
C DIMENSION MTIT(16), ITIT(13), SENSE(2)
C COMMON/MPORG/XO,YO,XORG,YORG
C COMMON/FILE3,JFILE
C MAP INITIALIZATION ROUTINE
C CALL BASG(IFILE3)
C READ THE THREE VALUES TO BE PLOTTED
C CALL READ(JFILE, 10, END=99) RLAT, RLON, XVAL
C GET THE X-Y COORDINATES OF THE POINT READ
C AND PLOT IT
12 CALL NMERCC( RLAT, -ABS(RLON), X, Y)
   CALL MOVEA((X-XO)+XORG, (Y-YO)+YORG)
   CALL ANCHO(43)
   GOTO 50
99 CONTINUE
C SET THE PLOT TO ALPHA MODE
C CALL ANMODE
C
C ENCODE PERTINENT INFO INTO ITIT
C ENCODE(65, 30, ITIT) IPRINC, SENSE(1), SENSE(2), MSAVE
C CALL CHRSLZ(3)
C MOVE TO LOWER LEFT CORNER
CALL MOVABS(50,100)

C OUTPUT MAP TITLE
CALL AOUTST(80,MTIT)

C MOVE UNDER THE MAP TITLE
CALL MOVABS(10,10)

C OUTPUT PERTINENT INFO
CALL AOUTST(65,ITIT)
CALL HDCOPY
30 FORMAT(3F10.3)

RETURN
END

C*****************************************************************
C PROGRAM: BASG
C PURPOSE: TO PLOT A 1 BY 2 DEGREE QUADRANGLE OUTLINE
C*****************************************************************

BASE MAPS SHOWING THE GEOGRAPHIC GRATICULE OF SPECIFIED
1 TO 250,000 SCALE MAPS PUBLISHED BY THE US GEOLOGICAL
SURVEY WERE REQUIRED BY BENDIX FIELD ENGINEERING GEOLOGY
DIVISION. THE BASG PROGRAM WAS WRITTEN TO PLOT
THOSE GRATICULES TO OVERLAY NEGATIVES OF THE USGS PLATES
OF THE CULTURE AND HYDROGRAPHY FOR EACH QUADRANGLE.
MAPS FOR AREAS NORTH OF 60 DEGREES LATITUDE WILL AUTOMATICALLY
BE CHANGED TO 1 BY 3 DEGREES.

SUBROUTINE BASG(IFILE3)

C PROGRAM TO GENERATE GEOGRAPHIC BASE MAP OVERLAYS FOR ENTIRE UNITED STATES

COMMON /DATA/ CMERID,CLAT,CLON,SPH,SCAL,
1 PSCEW,PSCHS,DEGEU,DEGNs,TICSZ,TICUZ,
2 MTITL(16),MAP(6),JDMA(3)
3 ,JDELS(6),JOU0(5),DOEV
 COMMON /crpa/ icrt,paper

C--------- INPUT DATA FORMAT 5-CARDS/QUAD
C 1 VERSION ( DOE OR GEOL )
C 2 LOCATION, MAPNAME CARD
C 3 SIZE, TICMARKS, CENTRAL MERIDIAN
C 4 TITLE (FOR UPPER LEFT OVER MAP)
C 5 LINE AB MEASUREMENTS
C 6 LINE CD MEASUREMENTS

COMPUTE THE THEORETICAL CENTRAL MERIDIAN AND CENTER E-U
DISTANCES OF THE NURE QUADRANGLES

C-------- THIS Initializes THE PLOT BLOCK COUNTER
paper=0
icrt=1

C OVERRIDE OPTION CARD 1
DOEV=1.

C INPUT THE NORTHWEST LAT, LON NTMS IDENT AND MAP NAME
C DEGREES N-S AND E-W, CENTRAL MERIDIAN AND TIC MARK
C INTERVAL N-S AND E-W.

READ(IFILE3,10,END=99)RNULA,RNULO,JDMA,MAP,
1 DEGNs,DEGEU,CMERID,TICUZ,TICSZ

C-------- BEGIN TO DETERMINE IF DEFAULT MAP SIZE AND TIC MARK
C INTERVALS WILL BE APPLIED OR IF THIS MAP IS ABOVE 60
C DEGREES NORTH LATITUDE.
C
C------- READ TITLE CARD
READ(14,END=99) MTITL
C
C SET DEFAULT TIC MARK INTERVAL AND MAP SIZE IF NOT
C DEFINED PREVIOUSLY.
C
IF(DEGNS.EQ.0.)DEGNS=1.0
IF(DEGEW.EQ.0.)DEGEW=2.0
C
IF(DEGEU.EQ.0..AND.RNULA.GE.60.)DEGEU=3.0
C
IF(TICSZ.EQ.0.)TICSZ=0.25
C
IF(TICWZ.EQ.0..AND.RNULA.LT.60.)TICWZ=0.25
C
C SET PHYSICAL SCALE FACTORS TO 1.
PSNS=1.
PSEU=1.
C
C INITIALIZE THE PLOTTING ROUTINES
C
CALL RESET
CALL ERASE
C TELL PLOT ROUTINES THAT YOU ARE AN ENHANCED GRAPHICS TERMINAL
CALL TERM(3,2096)
C
CALL TERM(3,4096)
C
SET CHARACTER SIZE TO THE SMALLEST
CALL CHRSIZ(4)
C
SET SCREEN WINDOW LIMITS
CALL SWINDO(1,4000,1,3000)
C
DETERMIN X AND Y MAX VALUES
XRAN=(53.*DEGEU)/2.
YRAN=31.*DEGNS
C
SET X AND Y MINIMUM VALUES
XMIN=-8.
YMIN=1.
C
SET DATA VALUE WINDOW
CALL DWINDO(XMIN,XRAN,YMIN,YRAN)
C
DRAW A BORDER AROUND THE AREA WHERE THE MAP WILL BE
CALL MOVEA(XMIN,YRAN)
CALL DRAWA(XRAN,YRAN)
CALL DRAWA(XRAN,YMIN)
CALL DRAWA(XMIN,YRAN)
CALL MOVEA(XMIN,YMIN)
C
MOVE TO THE TOP OF THE BORDER AND PRINTOUT THE TITLE
C
RETURN CHARACTER SIZE TO 3
C
CALL CHRSIZ(3)
CALL MOVABS(800,2000)
CALL AOUTST(80,MTITL)
C
RETURN CHARACTER SIZE TO 4
C
CALL CHRSIZ(4)
C
SET MAP SCALE TO 250000.
C
SCAL=250000.
RCAL=SCAL/39.37008
C
SPHEROID NUMBER 2 (CLARK 1866)
C
SPH=2.
C
BE SURE LONGITUDES ARE NEGATIVE FOR U.S.
C
RNULO=ABS(RNULO)
CLAT=RNULA-DEGNS
CLON=RNULO
C
FIND CENTRAL MERIDIAN
AO=RNULO+(DEGEU/2.)
C--------- USE GEOGRAPHICAL CENTER OF MAP AREA FOR CENTRAL MERIDIAN  
C--------- UNLESS IT WAS SPECIFICALLY REPORTED IN THE INPUT DATA SET  
   IF(CMERID.EQ.0.)CMERID=A0  
C   INITIALIZE MERCATOR PROJECTION TO GO THROUGH THE CENTER OF  
C   EACH MAP.  
   CALL NMERC(CMERID,RCAL,SPH,Z)  
   CALL NMERC(0.,0.,PSNS,PSEU)  
   CALL CHRSIZ(4)  
C   CALL MRCMAP(XRAN,YRAN)  
C 99 CONTINUE  
C 10 FORMAT(6X,F6.3,F9.3,1X,2A5,A2,1X,5A5,A4,/,
   1F7.2,F7.2,F7.2,1X,F5.3,1X,F5.3)  
14 FORMAT(16AS)  
12 FORMAT(2F10.I)  
RETURN  
END  
SUBROUTINE MRCMAP(XRAN,YRAN)  
C
C
C.............. BEGIN TO FIND ORIGIN OFFSET TO CENTER OF NEU MAP FROM  
C THE TRIM LINE ORIGIN. DO THIS BY FINDING SEPARATION  
C OF NORTHWEST CORNER FROM CENTRAL MERIDIAN  
C
C   CALL NMERCC(CLAT,CMERID,X0,Y0)  
C   CALL NMERCC(CHULA,CHULO,XNU,YNU)  
C............. FIND INCHES FROM NW CORNER TO CM  
   YNU=ABS(YNU-Y0)  
   XNU=ABS(XNU-X0)  
C............. THE OFFSET TO CENTER OF MAP  
   XORG=(XRAN-.8)/2.  
   YORG=.2*YRAN  
   NTCHS=DEGNS/TICSZ  
   NTCEU=DEGEU/TICUZ  
C--------- XTIC IS THE SIZE OF THE EDGE TIC MARKS  
   XTIC=.5  
C
C............. WE START PLOTTING AT THE SOUTHWEST CORNER OF THE MAP  
   BLAT=CLAT  
   BLON=CLON  
   CALL NMERCC(CLAT,CLON,X,Y)  
   XXO=X-X0+XORG  
   YVO=Y-Y0+YORG  
   CALL MOVEA(XXO,YVO)  
C   XYU=XORG*.02  
   HGT=XYU*.2.  
   IF(NOTAT.EQ.0)CALL ANOTE(CLON,XXO,YVO,0)  
C
C............. LOWER EDGE OF MAP  
C   DO 10 J=1,NTCEU  
   BLRJ=BLON+FLOAT(J)*TICUZ  
   CALL NMERCC(BLAT,BLRJ,X,Y)  

62
XXO-X-XO+XORG
YYO-Y-YO+YORG
CALL DRAUA(XXO,YYO)
YZ+YYO+XTIC
CALL DRAUA(XXO,YZ)
CALL MOVEA(XXO,YYO)

C------ANNOTE THE LOW EDGE OF MAP
IF(NOTAT.EQ.0)CALL ANOTE(BLRJ,XXO,YYO,1)
10 CONTINUE
C
C..........RIGHT SIDE OF MAP
C
IF(NOTAT.EQ.0)CALL ANOTE(BLON,XXO,YYO,5)
BLON+BLON+DEGEW
DO 20 J=1,NTCNS
BLRJ-BLAT+FLOAT(J)*TICSZ
CALL NMERCC(BLRJ,BLON,X,Y)
XXO=X-XO+XORG
YYO=Y-YO+YORG
CALL DRAUA(XXO,YYO)
XZ-XXO+XTIC
CALL DRAUA(XZ,YYO)
CALL MOVEA(XXO,YYO)
IF(NOTAT.EQ.0)CALL ANOTE(BLRJ,XXO,YYO,2)
20 CONTINUE
C
C..........TOP OF MAP
C
IF(NOTAT.EQ.0)CALL ANOTE(BLON,XXO,YYO,6)
BLAT-BLAT+DEGS
DO 30 J=1,NTCEU
BLRJ-BLAT-FLOAT(J)*TICUZ
CALL NMERCC(BLAT,BLRJ,X,Y)
XXO=X-XO+XORG
YYO=Y-YO+YORG
CALL DRAUA(XXO,YYO)
YZ+YYO+XTIC
CALL DRAUA(XXO,YZ)
CALL MOVEA(XXO,YYO)
IF(NOTAT.EQ.0)CALL ANOTE(BLRJ,XXO,YYO,3)
30 CONTINUE
C
C..........LEFT SIDE OF MAP
C
IF(NOTAT.EQ.0)CALL ANOTE((BLAT+DEGS),XXO,YYO,7)
BLON+BLON-DEGEW
DO 40 J=1,NTCNS
BLRJ-BLAT-FLOAT(J)*TICSZ
CALL NMERCC(BLAT,BLRJ,X,Y)
XXO=X-XO+XORG
YYO=Y-YO+YORG
CALL DRAUA(XXO,YYO)
XZ-XXO+XTIC
CALL DRAUA(XZ,YYO)
CALL MOVEA(XXO,YYO)
IF(NOTAT.EQ.0)CALL ANOTE(BLRJ,XXO,YYO,4)
40 CONTINUE
C
C........CENTRAL TIC MARKS INSIDE MAP AREA
C
NCNS+NTCNS-1
NCEU+NTCEU-1
DO 50 J=1,NCEW
DO 50 K=1,NCNS
XZ-CLAT+FLOAT(K)*TICSZ
YZ-CLON+FLOAT(J)*TICUZ
CALL NMERCC(XZ,YZ,X,Y)
XXO=X-XO+XORG
YYO=Y-YO+YORG
CALL MOVEA((XXO+XYU),YYO)
CALL DRAUA((XXO+XYU),YYO)
CALL MOVEA((XXO-XYU),YYO)
CALL DRAUA((XXO-XYU),YYO)

63
XZ = XORG + XMAX
CALL MOVEA(XZ, -YORG)

C
RETURN
END

SUBROUTINE NMERC(AL, BL, X, Y)

C
C------ TRANSVERSE MERCATOR PROJECTION
C------ 5 SPHEROIDS
C------ YEAR EQUATORIAL POLAR COMP-PRINCIPAL
C------ NAME AXIS METERS AXIS METERS RESSION USER
C
C 1 1909
C HAYFORD 6378388. 6356912. 297.0 U.S.A. --INTERNATIONAL
C 2 1866
C CLARK 6378206.4 6356583.8 295.0 U.S.A.
C 3 1880
C CLARK 6378249.145 6356514.8696 293.5 SOUTH AFRICA
C 4 1830
C EVEREST 6377276.345 6356075.4134 300.8 INDIA
C 5 1841
C BESSEL 6377397.155 6356978.0629 299.15 GERMANY, INDONESIA, NETHERLANDS
C 6 1858
C CLARK 6377563. 6356257. 299.3 GREAT BRITAIN
C 7 1948
C KRASSOVSKY6378245. 6356863. 298.3 RUSSIA, EASTERN COUNTRIES
C 8 1967
C I.U.G.G. 6378160. 6356775. 298.25 INTERNATIONALLY ADOPTED
C
C DIMENSION TANPHI(5), COTPHI(5), A(5), B(5), AZ(5), BZ(5)
EXTERNAL TAN
REAL KO

C------ SPHEROID DEFINITIONS
C
C------- INITIAL ENTRY DEFINES PROJECTION
ENTRY NMERCPI(AL, BL, X, Y)
PSCEU = X
PSCH5 = Y
RETURN

C
ENTRY NMERC(AL, BL, X, Y)
C AL = CENTRAL MERIDIAN BL = SCALE
C X = SPHEROID NUMBER Y = (IGNORED)

IF (PSCEU.LE.0.) PSCEU = 1.
IF (PSCH5.LE.0.) PSCH5 = 1.
AZ(1) = 6378388.
AZ(2) = 6378206.4
AZ(3) = 6378249.145
AZ(4) = 6377276.345
AZ(5) = 6377397.155

BZ(1) = 6356911.946
BZ(2) = 6356583.8
BZ(3) = 6356514.896
BZ(4) = 6356075.413
BZ(5) = 6356078.96

KO = 0.996
PI = 3.1415926
ADJ = 0.0174533
ALT = AL
SC = BL
IX = X
IF (IX.GT.5.OR.IX.LT.1) GOTO 100
GOTO 101

C------ CHOOSE THE CLARK 1866 SPHEROID BY DEFAULT
100 IX = 2
101 AA = AZ(IX)
BB = BZ(IX)
NORMAL ENTRY FOR CONVERTING (LAT, LON) TO (X, Y)

ENTRY  NMBERRC(AL, BL, X, Y)
AL = LATITUDE
BL = LONGITUDE
X = X(METERS)
Y = Y(METERS)

ORIGIN (0., 0.) IS AT (EQUATOR, CENT. MER.)
SOUTH LATS ARE - WEST LONS ARE -
NORTH LATS ARE + EAST LONS ARE +

DL = BL - ALT
IF (BL.GE.0.) GOTO 2
DL = ALT - BL
2
P = DL * 36
C1 = P**5
C2 = C1**2
C3 = SIN1**2
C4 = C3 * SIN1
C5 = C4 * SIN1
C6 = C5 * SIN1
C7 = C6 * SIN1
AS = AA**2
BS = BB**2
AMB = AS - BS
E = AM * BS
E-SORT(E)
EPS = AMB / BS
ALL = ADJ * AL
SL = SIN(AL)
S2L = SL**2
C9 = S2L * E
C10 = S2L * EPS
CL = COS(AL)
C11 = CL**2
C12 = C11**2
TL = TAN(AL)
T2L = TL**2
RHO = AA * (1. - E) / (1. - C9)**1.5
U = RHO * (1. + C11)
C13 = U * CL
C14 = C13 * KO
C15 = C14 * SL
INU = 1
IF (ALL) 7, 9, 8
9
ONE = 0.
GOTO 40
7
INU = -1
8
TANPHI(1) = TAN(ABS(AL))
A(1) = 1
B(1) = SQRT(1. - E**2)
DO 10 I = 2, 5
IM1 = I - 1
AJ1 = A(IM1)
BJ1 = B(IM1)
A(I) = (AJ1 + BJ1) / 2.
B(I) = SQRT(AJ1 * BJ1)
TANPHI(I) = (AJ1 + BJ1) * TANPHI(IM1) / (AJ1 - BJ1) * TANPHI(IM1)**2
COTPHI(I) = 1. / TANPHI(I)
10 CONTINUE
COTPHI(1) = 1. / TANPHI(1)
PART1 = ATAN(TANPHI(5))
PART2 = PART1 * PI / 2
DO 20 I = 1, 5
PART2 = PART2 + 2. * (I - 1) / (1. - SQNF(COTPHI(I)))
20
phi4 = phi4 + 2. * (I - 1) / (1. - SQNF(COTPHI(I)))
PART1=0.
PART2=0.
DO 30 I=1,4
RA2:A(I)*I
RB2:B(I)*I
PART2=PART2+(RA2-RB2)*((1./SQRT(1.+COTPHI(I+1)*IX))/MX(I+1))
PART1=PART1+2.*IX(RA2-RB2)
ONE=ONE+(IX-PI)/2.*PHI4/(A(3)+B(3))+PART2/4.
30

40
TUO=C15*E9/2.*C3
THREE=C15*E15/24.*CSL*C5*(5.-T2L+9.*C11+4.*C12)
FOUR=C14*E4*W
FIVE=C14*CSL*1.1264.*C4*(1.-T2L+C11)
A6=C21*E24*7/20.*C15*CSL**2*61.-58.*T2L+
T4L*240.*CSL*330.*C10
Z26=C11*E26*CSL*2/120.
B5=(15.-T2L+T4L+14.*C11-58.*C10)
B5=B5*X
X=-(FOUR+TWO*P**3+THREE*P**4+AG)/SC
Y=(ONE+TUO*P**2+THREE*P**4+AG)/SC
RETURN
END

FUNCTION TAN(ARG)
C..THIS JUST IN CASE THERE IS NO TAN FUNCTION ON COMPILER
C..TAN=SIN(ARG)/COS(ARG)
RETURN
END

FUNCTION SGNF(ARG)
C.C..SET THE SIGN OF ARG
C.
C..IF(ARG.GT.0.) SGNF=1.
C..IF(ARG.LT.0.) SGNF=-1.
C..IF(ARG.EQ.0.) SGNF=0.
RETURN
END

SUBROUTINE ANOTE(ANGLXX,YY,KIND)
C..ANOTE WILL PLACE EITHER THE MINUTES VALUE OR THE WHOLE
C..DEGREE VALUE IN AN APPROPRIATE POSITION NEXT TO A TICK MARK
C..LOCATED AT (X,Y).
C..KIND TELLS WHICH EDGE OF THE MAP THE ANNOTATION IS TO BE MADE.
C..1 - LOWER EDGE
C..2 - RIGHT EDGE
C..3 - TOP EDGE
C..4 - LEFT EDGE
C.
C-------------------------------
DIMENSION XNUM(8)
C..CHECK FOR A VALID EDGE DESIGNATOR
IF(KIND.GT.8.OR.KIND.LT.1)RETURN

X=X*XX
Y=YY
AN=ABS(ANGL*10.)
IAN=AN
REMM=FLOAT(IAN)/10.

10 CONTINUE
ENCODE(8,111,XNUM) Remm
CALL MOVREL(-100,-50)
CALL AOUTST(8,XNUM)
GOTO 100

20 CONTINUE
ENCODE(8,111,XNUM) REMM
CALL MOUREL(50,0)
CALL AOUTST(6,XNUM)
GOTO 100

C-------  TOP OF MAP
C
30 CONTINUE
ENCODE(8,111,XNUM) REMM
CALL MOUREL(-100,50)
CALL AOUTST(8,XNUM)
GOTO 100

C-------  LEFT EDGE OF MAP
C
40 CONTINUE
ENCODE(8,111,XNUM) REMM
CALL MOUREL(-200,0)
CALL AOUTST(8,XNUM)
GOTO 100

C

C-------  SOUTHEAST CORNER, LATITUDE
50 CONTINUE
ENCODE(8,111,XNUM) REMM
CALL MOUREL(50,0)
CALL AOUTST(8,XNUM)
GOTO 100

C

C-------  NORTHEAST CORNER, LONGITUDE
60 CONTINUE
ENCODE(8,111,XNUM) REMM
CALL MOUREL(-100,50)
CALL AOUTST(8,XNUM)
GOTO 100

C

C-------  NORTHWEST CORNER, LATITUDE
70 CONTINUE
ENCODE(8,111,XNUM) REMM
CALL MOUREL(-200,0)
CALL AOUTST(8,XNUM)
GOTO 100

C

C-------  SOUTHWEST CORNER, LONGITUDE
80 CONTINUE
ENCODE(8,111,XNUM) REMM
CALL MOUREL(-100,-50)
CALL AOUTST(8,XNUM)
GOTO 100

C

100 CONTINUE
CALL MOVE(A,XX,YY)
111 FORMAT(F5.1)
RETURN
END
### Mean Vector

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>47.570</td>
<td>71.600</td>
<td>37.020</td>
</tr>
</tbody>
</table>

### Covariance Matrix

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>198.147</td>
<td>38.473</td>
<td>-7.880</td>
</tr>
<tr>
<td>38.473</td>
<td>324.384</td>
<td>63.160</td>
</tr>
<tr>
<td>-7.880</td>
<td>63.160</td>
<td>65.131</td>
</tr>
</tbody>
</table>

### Eigenvalues

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>347.543</td>
<td>191.394</td>
<td>48.725</td>
</tr>
</tbody>
</table>

### Principal Components

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRST</td>
<td>SECOND</td>
<td>THIRD</td>
</tr>
<tr>
<td>0.23383</td>
<td>0.95021</td>
<td>0.20598</td>
</tr>
<tr>
<td>0.95021</td>
<td>-0.20264</td>
<td>-0.16104</td>
</tr>
<tr>
<td>0.20598</td>
<td>-0.16104</td>
<td>0.96511</td>
</tr>
</tbody>
</table>

### Correlation with Original Elements

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.30968</td>
<td>0.94922</td>
<td>0.05547</td>
</tr>
<tr>
<td>0.98354</td>
<td>-0.09175</td>
<td>0.03476</td>
</tr>
</tbody>
</table>

### Histogram for Component 1

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>START</td>
<td>H(J),...H(J+9)</td>
</tr>
<tr>
<td>NUMBER OF CELLS</td>
<td>101</td>
<td>THE TOTAL NUMBER OF POINTS</td>
</tr>
<tr>
<td>1</td>
<td>-5.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>11</td>
<td>-4.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>21</td>
<td>-3.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>31</td>
<td>-2.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>41</td>
<td>-1.00</td>
<td>4. 4. 4. 4. 4. 4. 4. 4. 4.</td>
</tr>
<tr>
<td>51</td>
<td>0.00</td>
<td>4. 4. 4. 4. 4. 4. 4. 4. 4.</td>
</tr>
<tr>
<td>61</td>
<td>1.00</td>
<td>4. 4. 4. 4. 4. 4. 4. 4. 4.</td>
</tr>
<tr>
<td>71</td>
<td>2.00</td>
<td>1. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>81</td>
<td>3.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>91</td>
<td>4.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>101</td>
<td>5.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
</tbody>
</table>

### Histogram for Component 2

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>START</td>
<td>H(J),...H(J+9)</td>
</tr>
<tr>
<td>NUMBER OF CELLS</td>
<td>101</td>
<td>THE TOTAL NUMBER OF POINTS</td>
</tr>
<tr>
<td>1</td>
<td>-5.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>11</td>
<td>-4.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>21</td>
<td>-3.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>31</td>
<td>-2.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>41</td>
<td>-1.00</td>
<td>4. 4. 4. 4. 4. 4. 4. 4. 4.</td>
</tr>
<tr>
<td>51</td>
<td>0.00</td>
<td>4. 4. 4. 4. 4. 4. 4. 4. 4.</td>
</tr>
<tr>
<td>61</td>
<td>1.00</td>
<td>4. 4. 4. 4. 4. 4. 4. 4. 4.</td>
</tr>
<tr>
<td>71</td>
<td>2.00</td>
<td>1. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>81</td>
<td>3.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>91</td>
<td>4.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>101</td>
<td>5.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
</tbody>
</table>

### Histogram for Component 3

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>START</td>
<td>H(J),...H(J+9)</td>
</tr>
<tr>
<td>NUMBER OF CELLS</td>
<td>101</td>
<td>THE TOTAL NUMBER OF POINTS</td>
</tr>
<tr>
<td>1</td>
<td>-5.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>11</td>
<td>-4.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>21</td>
<td>-3.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>31</td>
<td>-2.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>41</td>
<td>-1.00</td>
<td>4. 4. 4. 4. 4. 4. 4. 4. 4.</td>
</tr>
<tr>
<td>51</td>
<td>0.00</td>
<td>4. 4. 4. 4. 4. 4. 4. 4. 4.</td>
</tr>
<tr>
<td>61</td>
<td>1.00</td>
<td>4. 4. 4. 4. 4. 4. 4. 4. 4.</td>
</tr>
<tr>
<td>71</td>
<td>2.00</td>
<td>1. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>81</td>
<td>3.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>91</td>
<td>4.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
<tr>
<td>101</td>
<td>5.00</td>
<td>0. 0. 0. 0. 0. 0. 0. 0. 0.</td>
</tr>
</tbody>
</table>

72
TEXAS GULF COAST 1 DETAILED AREA C

TEST DATA SET
PRINC COMP 1 SENSITIVITY 3.89 - 3.89 NM PTS PLOTTED 4
DOES THE USER REQUIRE INSTRUCTIONS TO RUN THIS PROGRAM (YES OR NO)? NO

INPUT DATA FILE TO BE READ FROM?
Unit*20 : /ACCESS=SEGINOU/MODE=ASCII

Enter new file specs. End with an $(ALT):
*TEXASC.DAT$

INPUT DATA FILE TO HOLD PRINCIPAL COMPONENTS?
Unit*21 : /ACCESS=SEGINOU/MODE=ASCII

Enter new file specs. End with an $(ALT):
*TEXASC.PCS$

HAVE PC S BEEN CALCULATED FOR THIS DATA SET BEFORE (YES, NO)? NO

INPUT THE FORMAT OF THE DATA (80 CHAR)?
(14,7X,2F10.5,A5,3F6.1)

INPUT DATA FILE FOR PC 1 OUTLIERS USED IN THE MAP?
Unit*23 : /ACCESS=SEGINOU/MODE=ASCII

Enter new file specs. End with an $(ALT):
*TEXASC.PC1$

INPUT DATA FILE FOR PC 2 OUTLIERS USED IN THE MAP?
Unit*24 : /ACCESS=SEGINOU/MODE=ASCII

Enter new file specs. End with an $(ALT):
*TEXASC.PC2$

INPUT DATA FILE FOR PC 3 OUTLIERS USED IN THE MAP?
Unit*25 : /ACCESS=SEGINOU/MODE=ASCII

Enter new file specs. End with an $(ALT):
*TEXASC.PC3$

FILE NAME HOLDING HEADER INFORMATION?
Unit*22 : /ACCESS=SEGINOU/MODE=ASCII

Enter new file specs. End with an $(ALT):
*TEXASC.HED$

SELECT HISTOGRAM PLOTS, PRINTOUT, OR BOTH
PLOTS ONLY=+1, BOTH=0, PRINTOUT ONLY=-1 ? 0
INPUT THE LOWER LIMIT TO BE USED IN THE PLOTS ? -5.
INPUT THE UPPER LIMIT TO BE USED IN THE PLOTS ? 5.

THE NUMBER OF CELLS TO BE USED IN THE HISTOGRAMS = 101

GIVE SPECIAL DATA TITLE FOR MAP?
TEXAS GULF COAST DATA
HAS HEADER FILE BEEN WRITTEN (YES, NO)? YES
INPUT THE LOWER LIMIT SENSITIVITY FOR THE OUTLIER MAPS ? 2.
INPUT THE UPPER LIMIT SENSITIVITY FOR THE OUTLIER MAPS ? 3.

DOES THE USER WISH TO CALCULATE THE CORRELATION OR COVARIANCE MATRIX (CORRE OR COVAR)? COVAR

DOES THE USER WISH TO SORT THE DATA ON A FLIGHT LINE BASIS (YES OR NO)? NO
DOES THE USER WISH TO SORT ON A GEOLOGICAL BASIS (YES OR NO)? NO
DOES THE USER WISH TO WISH TO TAKE THE SQUARE ROOTS OF THE DATA? YES

76
APPENDIX C

PCA Documentation File
THIS PROGRAM WAS DESIGNED TO USE THE PLOTIO AND AG2 GRAPHIC ROUTINES AND A TEKTRONIX 4015-1 TERMINAL WITH 4096 ADDRESSABLE POINTS AND A HARD COPY UNIT.

THE INPUT PARAMETERS NECESSARY TO EXECUTE THE PRINCIPAL COMPONENT PROGRAM ARE AS FOLLOWS:

1) THE NAME OF THE DATA FILE TO BE PROCESSED.

NOTE: THIS DATA FILE MAY BE IN ANY FORMAT WITH THE LATITUDE FIRST, THE LONGITUDE SECOND AND THE K, U & T DATA LAST. THE K, U & T DATA MAY BE IN ANY ORDER.


3) THE PROGRAM WILL NOW ASK IF THE PRINCIPAL COMPONENTS HAVE BEEN CALCULATED BEFORE. IF THE ANSWER IS YES THEN THE FILENAME SUPPLIED IN STEP 2 MUST CORRESPOND TO THE FILENAME USED BEFORE. THIS CAN RESULT IN SIGNIFICANT SAVINGS IN TIME DEPENDING ON THE SIZE OF THE DATA SET.

4) INPUT THE DATA FORMAT. THIS INPUT IS LIMITED TO 80 CHARACTERS AND MUST BEGIN AND END WITH PARENTHESES.

5) THREE FILENAMES MUST BE ENTERED NEXT. THESE FILES WILL RECEIVE THE DATA USED IN THE MAPS. THE CONVENTION FOR NAMING THESE FILES IS:

   XXXXXXX.PC1
   XXXXXXX.PC2
   XXXXXXX.PC3

6) THE PROGRAM WILL ERASE THE SCREEN AND ASK FOR THE FILENAME CONTAINING THE HEADER INFORMATION IF THE HEADER HAS BEEN CREATED PREVIOUSLY. IF NOT, THEN NAME THE FILE XXXXXXX.HD.

7) AT THIS POINT THE USER IS ASKED TO SELECT HISTOGRAM PLOTS, PRINTOUTS OR BOTH BY INPUTING A (+1) FOR PLOTS ONLY, A (0) FOR BOTH, OR A (-1) FOR PRINTOUTS ONLY.

8) THE PROGRAM NOW ASKS FOR THE MINIMUM VALUE TO BE USED IN THE HISTOGRAMS. THIS VALUE MAY BE + OR - SO LONG AS IT IS LESS THAN THE MAXIMUM VALUE FOR THE HISTOGRAMS. THIS NUMBER IS INPUT IN FLOATING POINT.

9) THE MAXIMUM VALUE FOR THE HISTOGRAMS IS NOW LOADED. THIS NUMBER IS INPUT IN FLOATING POINT.

NOTE: THE USER MUST BEAR IN MIND THAT THE DATA PLOTTED IN THE HISTOGRAMS IS NORMALIZED DATA AND THE MIN & MAX VALUES INPUT IN STEPS 8 & 9 WILL BE REPRESENTATIVE OF THE STANDARD DEVIATION OF THAT DATA.


11) THE PROGRAM ASKS FOR A SPECIAL DATA TITLE FOR THE MAPS. ANY PERTINENT INFORMATION ABOUT THE DATA MAY BE ENTERED HERE.


NOTE: IF THE ANSWER TO 83 IS YES, THEN PROCEED WITH 821.
13) IT NOW ASKS FOR THE NORTHWEST CORNER LATITUDE IN DECIMAL DEGREES.

14) IT NOW ASKS FOR THE NORTHWEST CORNER LONGITUDE IN DECIMAL DEGREES.

15) INPUT THE CENTRAL MERIDIAN IN DECIMAL DEGREES.

16) INPUT THE N-S MAP SIZE IN DECIMAL DEGREES.

17) INPUT THE E-U MAP SIZE IN DECIMAL DEGREES.

18) INPUT THE E-U TIC INTERVAL IN DECIMAL DEGREES.

19) INPUT THE N-S TIC INTERVAL IN DECIMAL DEGREES.

20) INPUT THE TITLE FOR THE MAPS (80 CHARACTERS MAX).

21) THE NEXT TWO INPUTS ARE USED TO SET THE BOUNDS FOR OUTLIER SELECTION. INPUT THE MINIMUM SENSITIVITY.

22) INPUT MAXIMUM SENSITIVITY.

NOTE: THESE SENSITIVITIES ARE THE BASIS FOR GENERATING THE MAPS OF THE OUTLIERS. THEY MAY BE ANY NUMBER + OR - SO LONG AS THE MAXIMUM VALUE IS GREATER THAN THE MINIMUM VALUE. ALL THE NORMALIZED DATA POINTS THAT FALL WITHIN THIS RANGE WILL BE PLOTTED.

23) THE PROGRAM NOW ASKS THE USER TO SELECT EITHER THE COVARIANCE MATRIX OR THE CORRELATION MATRIX. THE CORRELATION MATRIX IS USED WHEN CALCULATING PRINCIPAL COMPONENTS ON DATA OF DIFFERENT UNITS. INPUT (COVAR) FOR A COVARIANCE MATRIX OR (CORRE) FOR THE CORRELATION MATRIX.

24) THE PROGRAM NOW ASKS IF THE USER WISHES TO SORT THE DATA ACCORDING TO FLIGHT LINES. A YES OR NO ANSWER IS REQUIRED.

25) IF THE ANSWER TO QUESTION 24 IS YES THEN INPUT THE STARTING FLIGHT LINE NUMBER.

26) IF THE ANSWER TO QUESTION 24 IS YES THEN INPUT THE ENDING FLIGHT LINE NUMBER.

27) THE PROGRAM ASKS WHETHER OR NOT THE USER WISHES TO SORT THE DATA ACCORDING TO A CERTAIN GEOLOGY. A YES OR NO ANSWER IS REQUIRED.

28) IF THE ANSWER TO QUESTION 27 IS YES THEN INPUT THE DESIRED GEOLOGIC CODE (FORMAT A5).

29) THE PROGRAM NOW ASKS IF THE USER WISHES TO TAKE THE SQUARE ROOTS OF THE DATA. TAKING THE SQUARE ROOTS OF THE DATA IS A GOOD FIRST ORDER APPROXIMATION OF POISON COUNTING STATISTICS. A YES OR NO ANSWER IS REQUIRED.

30) AT THIS POINT ALL USER INTERACTION IS COMPLETED. THE PROGRAM WILL MAKE PRINTOUTS AND PLOTS AND MAKE COPIES OF EACH AUTOMATICALLY.
REFERENCES


Campbell, Katherine, Los Alamos Scientific Laboratories, Principal components computational analysis and principal components versus factor analysis, (Written communication).


Hanson, R. J., 1975, Stably updating mean and standard deviation of data, in Communications of the ACM, v. 18, no. 1, p. 57-58.


Los Alamos Scientific Laboratory, 1975, Quarterly progress report of work on the geostatistics project of the Jackson/Goliad Formation in Texas as carried out by the Los Alamos Scientific Laboratory: July-September, 1975: U.S. Energy Research and Development Administration, LASL C5-75-10, unpublished report.


Additional documents in this series, "Statistical Techniques Applied to Aerial Radiometric Surveys (STAARS)", are currently in preparation and scheduled for release in the near future. Tentative titles and/or subject matter covered are as follows:

Introduction to STAARS Series

Principal Components Analysis as Applied to Aerial Data

One document has been released at this time - "Statistical Techniques Applied to Aerial Radiometric Surveys (STAARS): Percentile Estimation with the Normal and Lognormal Distributions, GJBX-123(80)."
SUGGESTED READING


Kane, V. K., Baer, T., and Begovich, C. L., 1977, Principal component testing for outliers: U.S. Energy Research and Development Administration, GJBX-71(77), Open-file report, 72 p.


