DANIEL F. MERRIAM, Editor

FORTRAN IV PROGRAM FOR COMPUTATION AND DISPLAY OF PRINCIPAL COMPONENTS

By

WARREN C. WAHLSTEDT
Cities Service Oil Company

and

JOHN C. DAVIS
Kansas Geological Survey

COMPUTER CONTRIBUTION 21
State Geological Survey
The University of Kansas, Lawrence
1968

Editor's Remarks

Publication of the COMPUTER CONTRIBUTION series is now in its third year. From all indications programs in the series are being used and used successfully. We now have record of specific applications of our programs by almost one hundred individuals and organizations. About 170 program decks were purchased during the past year and more than 47,000 copies of the COMPUTER CONTRIBUTIONS and Special Distribution Series have been distributed. Scientists in 38 countries are regularly receiving copies of this series.

Application of different computer techniques are being reported in many articles in professional and trade journals, programs are being published, and several texts and reference books recently have become available. Training schools, lectures, seminars, colloquia and workshops are being conducted. Everywhere there is interest, and apparently success, in applying the "new" techniques.

Another "new" technique is reported here. As the authors state, principal components analysis may be used as a "search procedure" in examining large collections of information. Essentially it is a technique to find underlying structure or order in data. The technique utilizes simpler assumptions than factor techniques, and this in itself is a valid reason to try the method. Several examples of solutions to problems by the principal components method are given in the text and undoubtedly many others come immediately to mind.

For a limited time the Geological Survey will make available the program deck for the "FORTRAN IV program for computation and display of principal components" by W.C. Wahlstedt and J.C. Davis for \$10.00. Unless otherwise requested the deck will be sent on magnetic tape. A card deck is \$5.00 extra.

Although every precaution is taken to insure tested procedures and complete documentation, problems frequently are found in translating programs from one language to another or in adapting programs to different machines. If these problems are related to us, we can offer better programs in the future; therefore we solicit your help in identifying and correcting these problems.

We would like to thank Dr. John Imbrie of Brown University and Prof. R.A. Reyment of Uppsala University (Sweden) for serving on the editorial board. They have contributed much during their tenure and their help was most appreciated. At this time we welcome as an associate editor Dr. Walther Schwarzacher, who is a Reader in Geology at Queen's University in Belfast, Northern Ireland.

An up-to-date list of publications can be obtained by writing the Editor, COMPUTER CONTRIBUTION Series, Geological Survey, The University of Kansas, Lawrence, Kansas 66044.

FORTRAN IV PROGRAM FOR COMPUTATION AND DISPLAY OF PRINCIPAL COMPONENTS

By

WARREN C. WAHLSTEDT and JOHN C. DAVIS

INTRODUCTION

If several variables are measured on a set of samples, a linear transformation of these variables can be found that will result in new variables which are independent and account for, successively, as much of the total variation as possible. These new variables are called principal components, and the process of computing the transformation is principal components analysis. Many multivariate statistical procedures require data that are independent with respect to each other. This requisite is seldom met in geologic problems, where some degree of interdependency exists between almost all variables observed on a sample. By transforming raw data into principal components, independence is achieved and the new, transformed variates may be tested.

In addition, geologists may profitably use principal components analysis as a "search procedure" to examine a large collection of information in the hope of finding underlying structure or order in the data. Used in this way, principal components is much like factor analysis, and indeed, forms the initial step in many factor analytic procedures. Principal components analysis, however, is much simpler than factor analysis, and is not based on the somewhat elaborate set of assumptions involved in factor analysis. These assumptions have been criticized (Matalas and Reiher, 1967), and principal components analysis may constitute an acceptable substitute for researchers uneasy about the applicability of factor methods.

Use of principal components as a search technique is an expression of two hopes: firstly, that a few principal components will account for a large amount of the total variance in the data and secondly, that the nature of these important components can be deduced from an examination of the linear transformation. In many problems, these hopes will be realized and reasonable interpretations can be made of the components of the observations. However, it may not be possible to reduce the dimensionality of a given problem, or the method may produce components which cannot be meaningfully interpreted.

Principal components analysis consists of a series of operations on a covariance matrix which result in a series of transformed variables each accounting successively for the most possible variance. The same procedure may be applied to a matrix of correlations, resulting in transformed variables based on the maximum amount of intercorrelation. A correlation analysis is equivalent to that obtained by a covariance analysis of standardized variables. In such an analysis, all variables have equal weights and are independent of the magnitude of their measured units.

MATHEMATICAL DEVELOPMENT

Data are first read into the computer and converted into an m x m matrix of covariances or correlations, where m = number of variables. The covariance matrix is

$$A_{ij} = \left[\alpha_{ij}\right] = \frac{n\sum_{i=1}^{m}\sum_{j=1}^{m} X_i X_j - \left(\sum_{i=1}^{m} X_i \sum_{j=1}^{m} X_j\right)}{n(n-1)} .$$

The correlation matrix is formed by

$$A_{ij} = \left[r_{ij}\right] = \frac{n\sum_{i=1}^{m}\sum_{j=1}^{m} X_i X_j - \sum_{i=1}^{m} X_i \sum_{j=1}^{m} X}{\sqrt{\left(n\sum_{i=1}^{m} X_i^2 - \left(\sum_{i=1}^{m} X_i\right)^2\right)\left(n\sum_{j=1}^{m} X_j^2 - \left(\sum_{j=1}^{m} X_j\right)^2\right)}} .$$

Associated with every square matrix $[A]$ is a characteristic function

$$f(\lambda) = [A - \lambda I] = \begin{bmatrix} \alpha_{11}-\lambda & \alpha_{12} & \cdots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22}-\lambda & \cdots & \alpha_{2m} \\ --- & --- & --- & --- \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mm}-\lambda \end{bmatrix},$$

which has the property $f(\lambda) = 0$. From this matrix equation, m roots called eigenvalues or latent roots, may be extracted. Associated with each eigenvalue is an eigenvector or latent vector which is a column vector $[X_i]$ having the property $[A-\lambda I]\cdot[X_i] = 0$. If the eigenvalues of a matrix are distinct (that is, not identical), the associated eigenvectors are independent. These vectors are the transformations desired in principal components analysis. In this program the eigenvalues and eigenvectors are found by a subroutine modified from one published by Cooley and Lohnes (1962, p. 187-189). Their procedure was based in turn on the Jacobi algorithm as developed by Greenstadt (1960, p. 84-91).

The trace (or sum of diagonal elements) of the matrix and the sum of the eigenvalues are identical and represent either the total variance or the total correlation in the matrix. Eigenvalues appear in the order of their magnitude, and the percent of total variance or correlation accounted for by each can be calculated by dividing the eigenvalues by the trace. In most instances, the first few eigenvalues will account for more variance or correlation than any of the original variables, and the last eigenvalues will account for less. If the first few account for an acceptable percent of the total, the remaining eigenvalues may be discarded. The problem has then been reduced in dimensionality, from one of m variables to one of $y<m$ principal components.

The squared values of the terms of an eigenvector also may be summed, and the individual terms divided by the total. This yields the approximate percent contribution of the original variables to the principal component represented by that eigenvector. If only a few of the original variables account for most of a principal component, it may be interpreted by considering the nature of this combination, ignoring the contribution of other variables.

OPERATIONAL INSTRUCTIONS

This program is written in FORTRAN IV for the GE 625 Computer. Dimension statements are designed to allow up to 30 variables and 300 observations, although this may be altered by making simple adjustments within the program. The following control cards are necessary.

CONTROL CARD

Col. 1- 3 M, a right justified integer giving number of variables, up to 30.

Col. 4- 5 Blank

Col. 6 MANUAL, switch for covariance-correlation option.

1 = Use covariance matrix only.
2 = Use correlation matrix only.
3 = Use covariance and correlation matrices.

Col. 7 OPDAT, switch for data listing option.

T = No list of input data.
Blank = List of input data.

VARIABLE NAMES CARD

This is a group of M cards containing the names of variables, listed one name to a card. A variable name may occupy columns 1 to 18. Names should appear in the same sequence as variables in the data set to avoid mislabeling the output.

TITLE CARD

Col. 1-72 Any alphanumeric information.

VARIABLE FORMAT CARD

Col. 1-72 Format of data cards. The last field in the format should be L1, to read the last card signal.

DATA CARDS

Any number of data cards, up to 300, in the format specified above. The last data card must contain a "T" punched in the logic field defined in the variable format statement.

This completes one data set. If more than one set is to be processed, the sequence is repeated. A blank card should be placed after the final data set to provide normal termination of the program.

EXAMPLE PROBLEMS

Characterization of Pleistocene Tills by Clay Minerals

Data in the test example are taken from a study by Tien (in press) of clay minerals in tills and associated deposits of northeastern Kansas. Four types of deposits were recognized in the field: (1) a lower Kansan till, (2) lower Kansan fluvioglacial deposits, (3) upper Kansan till, and (4) associated upper Kansan fluvioglacial material. Seventy-four samples were analyzed by x-ray diffraction, yielding four variables on each sample. Variables are the relative intensities of the basal reflections of mixed-layer clay, illite, and montmorillonite, and the D.I. ratio (ratio between illite and kaolinite 001 peak heights).

Analysis by principal components is not necessary in this problem, because relationships between samples can be demonstrated on a triangular diagram (Fig. 1).

Two principal components, however, contain essentially all variation in the system, yielding an economical graphic display of the sample relations (Fig. 2). In essence, the analysis has reduced the number of variables to the true degrees of freedom in the system, as the D.I. ratio is a combination of two other variables and the remaining closed number system contains one less degree of freedom than the number of variables.

Lower Kansan tills can be distinguished easily from upper Kansan tills on the principal components display. Both units plot as distinct clusters separate from nontills. Upper and lower nontill samples form an elongate, diffuse band which does not separate into stratigraphic groups. Two till samples, taken from within a predominately nontill sequence, show characteristics of nontills. These are marked with a

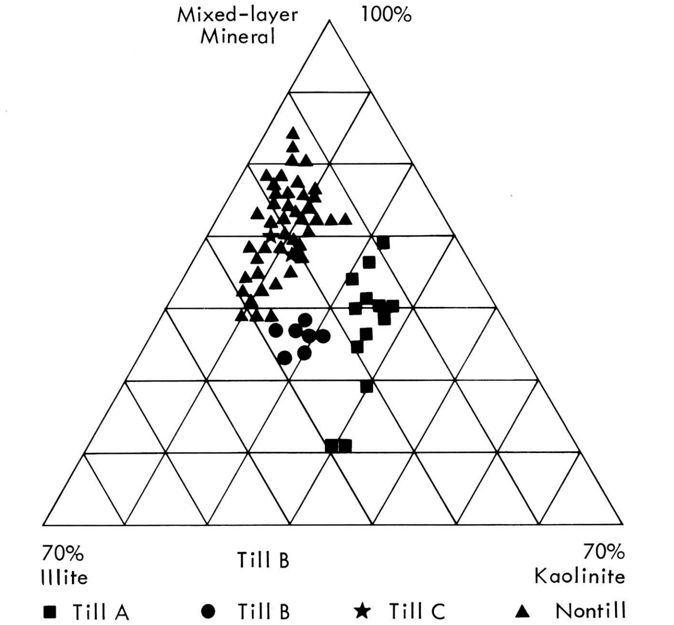


Figure 1.- Triangular diagram showing clay-mineral composition of Kansan tills (from Tien, in press).

special symbol (★) on Figures 1 and 2.

Data used in this problem are listed in the Appendix, with sample output from the principal components display program.

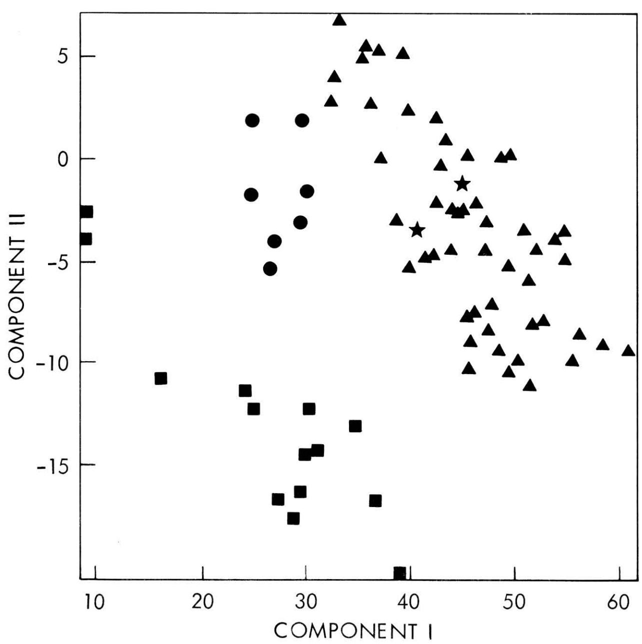


Figure 2.- Principal components plot of Component I versus Component II from clay-mineral compositions of Kansan tills. Symbols correspond to those on Figure 1.

Analysis of Income in Kansas by Counties.

Data have been compiled on personal income totals for Kansas counties during the years since instigation of income tax. These data are in the form of nine variables for each of the 105 counties; an example is shown in Table 1 which contains summaries of data for the year 1950. A principal components analysis was performed on the data to resolve the variance into independent sources. The first component accounts for 99 percent of the total variance, the second for 0.6 percent and the third for 0.4 percent. All other components account for less than 0.1 percent of the variance and were discarded. Eigenvectors corresponding to the first three components are listed in Table 2. The three sources of variation were interpreted from the eigenvector loadings as representing (1) commercial and industrial income, (2) government income, and (3) farm and ownership income.

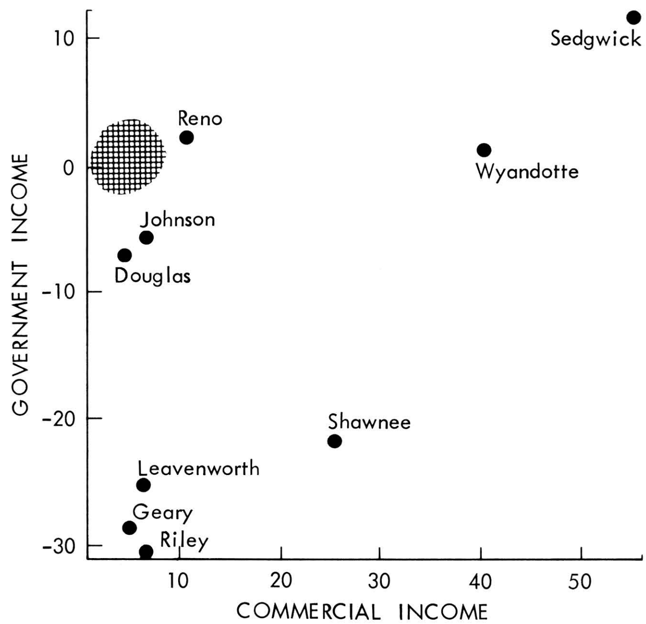


Figure 3.- Plot of commercial income (Component I) versus government income (Component II) for Kansas counties in 1950. Hatched area contains 96 counties.

Plots of the 105 counties on the coordinates defined by the eigenvectors are shown on Figures 3, 4, and 5. Only the distinctive counties are named. Sedgwick County contains the city of Wichita; Wyandotte and Johnson Counties are part of metropolitan Kansas City; Shawnee contains Topeka and a major Air Force base, Douglas contains The University of Kansas, and Riley, Geary, and Leavenworth Counties contain large military installations. All other counties in Kansas are uniform and show differences only in farm productivity. Although this sample analysis seems trivial and extracts only what

Table 1.- Summary statistics of Kansas county income data for 1950.

Variable	Mean*	Std. Dev.*	Covariances**								
Total personal income	25172.9	46374.6	0.22E10								
Farm income	5038.3	2027.2	0.24E8	0.41E7							
Total government payments	3723.1	6667.3	0.25E9	0.19E7	0.44E8						
Private nonfarm income	16688.5	41435.6	0.19E10	0.18E8	0.20E9	0.17E10					
Personal contributions	266.7	682.1	0.32E8	0.24E6	0.37E7	0.28E8	0.46E6				
Wage and salary income	13401.7	34303.1	0.16E10	0.12E8	0.18E9	0.14E10	0.23E8	0.12E10			
Proprietor income	7676.6	6491.7	0.28E9	0.77E7	0.30E8	0.24E9	0.39E7	0.19E9	0.42E8		
Property income	2894.2	4632.0	0.21E9	0.31E7	0.23E8	0.19E9	0.30E7	0.15E9	0.29E8	0.21E8	
Transfer payments	1467.9	2485.8	0.11E9	0.13E7	0.13E8	0.10E9	0.17E7	0.83E8	0.14E8	0.11E8	0.62E7

* in thousands of dollars.
** exponent notation.

Table 2.- First three eignevalues, percent contribution to total variance, and corresponding eigenvectors of Kansas county income data for 1950.

Eigenvalues	0.51E10	0.29E8	0.18E8
Eigenvalue Contribution	99.0%	0.6%	0.4%
Eigenvectors			
Total personal income	0.648	-0.181	0.369
Farm income	0.006	0.040	0.416
Total government payments	0.073	-0.761	0.094
Private nonfarm	0.578	0.533	-0.150
Personal contributions	0.010	-0.006	-0.009
Wages and salaries	0.478	-0.303	-0.471
Proprietor income	0.082	0.077	0.640
Property income	0.063	0.063	0.164
Transfer payments	0.034	-0.025	0.026

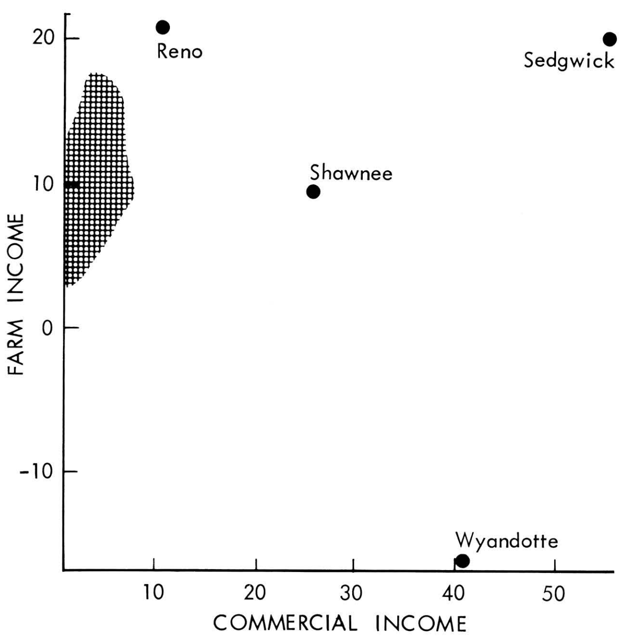


Figure 4.- Plot of commercial income versus farm income (Component III) for Kansas counties in 1950. Hatched area contains 101 counties.

is obvious from the political and geographic composition of Kansas, the method shows promise in dealing with more more complex economic systems, whether these are geographic regions or corporate entities.

Petrographic Variation in Sandstones

Okada (1966) has measured 12 variables on a suite of 38 samples of nongreywacke sandstones collected from British turbidite sequences. The data consist of measurements of six compositional variables and six physical attributes. A principal components analysis was performed on the data in an attempt to extract fundamental sources of variation. Means, standard deviations, and covariances are shown in Table 3. The first three principal components of the covariance matrix account for 97.5% of

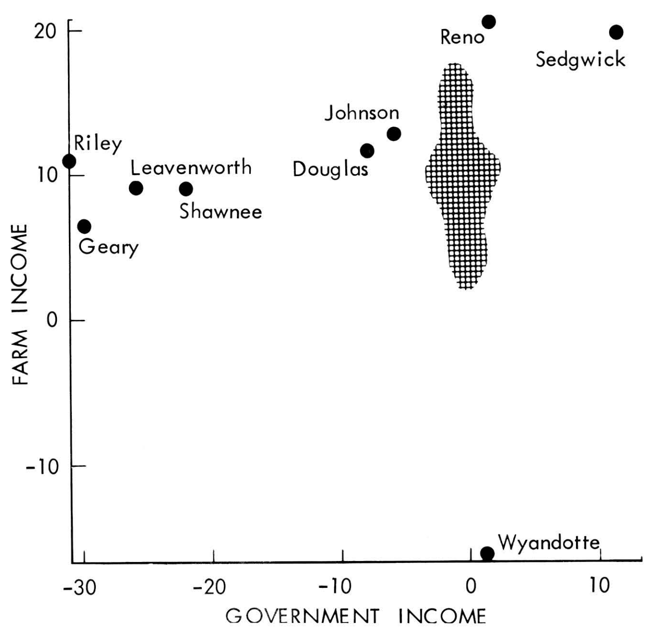


Figure 5.- Plot of government income versus farm income for Kansas counties in 1950. Hatched area contains 96 counties.

Table 3.- Means, standard deviations, and covariances of 12 variables measured on sandstones by Okada (1966).

Variable	Mean	Std. Dev.	Covariance Matrix											
Quartz	45.3	14.4	207.0											
Feldspar	10.4	5.2	23.5	26.7										
Rock fragments	7.9	9.5	-32.1	1.6	89.5									
Clay matrix	30.8	18.9	-177.9	-59.5	-81.7	356.0								
Cement	5.0	5.1	-14.5	8.2	22.4	-42.4	26.2							
Minor constituents	0.5	0.8	-6.3	-0.8	0.4	6.2	-0.1	0.6						
Median diameter	2.0	1.0	-1.1	-1.6	0.5	1.9	0.2	0.1	1.0					
Mean diameter	1.9	1.0	-0.8	-1.5	-0.1	2.4	**	0.1	1.0	1.0				
Standard deviation	1.0	0.4	1.6	0.9	1.9	-5.1	0.8	**	**	-0.1	0.2			
Skewness	-0.1	0.2	0.2	*	-0.4	0.4	-0.2	-0.1	**	**	**	*		
Kurtosis	0.6	0.2	1.4	**	0.2	-1.4	**		0.1	0.1	*	**	0.1	
Average roundness	0.3	0.1	0.7	0.2	0.3	-1.3	*	**	**	**	*	**	*	*

*	0.1	0.0
**	-0.1	0.0

variation in the data. Eigenvectors corresponding to these three are shown in Table 4. The first component, accounting for almost 64 percent of the variance, represents variation in the proportion of clay matrix. The second component accounts for an additional 28 percent of the variance and is produced by variation in the ratio quartz + clay: rock fragments + cement + feldspar. The third component accounts for about 5 percent of the variance and represents primarily the ratio between feldspar + cement and other constituents of the rock. Figures 6, 7, and 8 are plots of original samples on the orthogonal coordinates of the three major principal components.

Okada was able to distinguish Silurian from Cambrian sandstones on the basis of plots of rock-fragment index versus phi-mean. Essentially the same

Table 4.- First three eigenvalues, their percent contribution to total variance, and corresponding eigenvectors of Okada's (1966) sandstone data.

Eigenvalues	492.0	167.5	33.5
Eigenvalue Contribution	69.5%	23.6%	4.7%
Eigenvectors			
Quartz	0.514	-0.664	0.311
Feldspar	0.134	0.030	-0.614
Rock fragments	0.133	0.643	0.578
Clay matrix	-0.834	-0.281	0.151
Cement	0.069	0.256	-0.409
Minor constituents	-0.017	0.016	-0.008
Median diameter	-0.005	0.003	0.037
Mean diameter	-0.005	0.002	0.031
Standard deviation	0.011	0.011	-0.003
Skewness	-0.001	0.003	-0.002
Kurtosis	0.004	-0.002	0.012
Average roundness	0.003	0.000	0.001

distinction is apparent on plots containing the second principal component.

Physical parameters do not make a significant contribution to the first three eigenvectors and do not enter into the remaining components until the fifth, which represents variation in mean and median grain size. This component accounts for approximately 0.3 percent of the total variance.

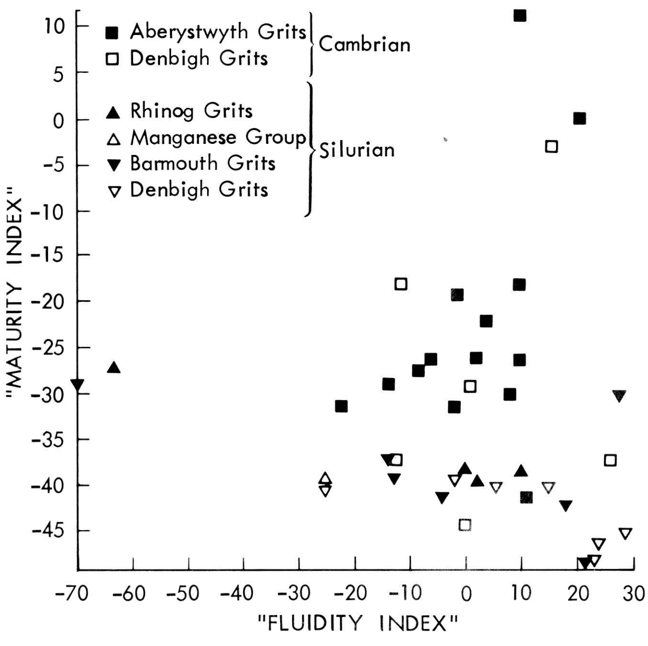
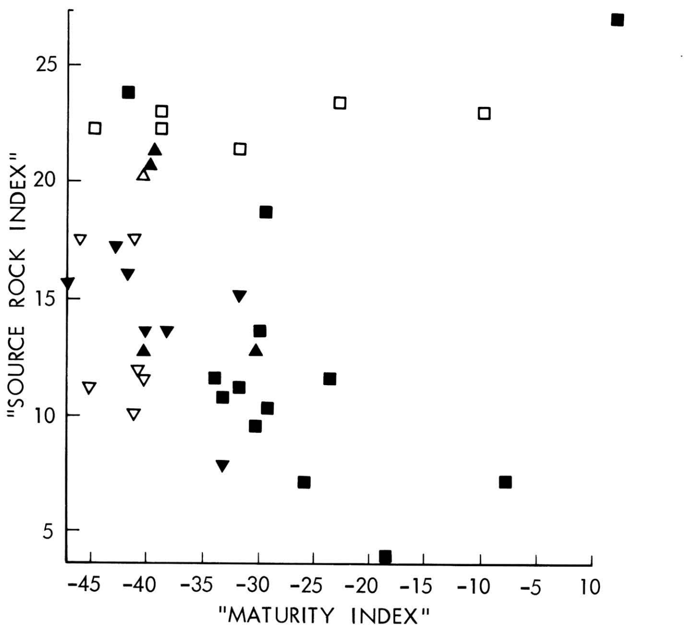


Figure 6.- Principal components plot of fluidity index (Component I) versus maturity index (Component II) for sandstone data from Okada (1966). Maturity index increases negatively; that is, greater proportions of quartz + clay are represented by negative values.

The three major principal components may be



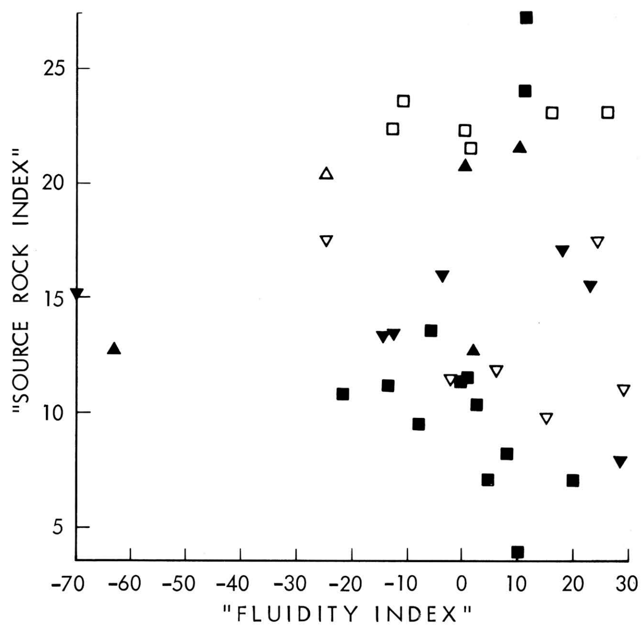


Figure 7.- Plot of fluidity index versus source rock index (Component III) for sandstone data from Okada (1966). Increasing values of source rock index indicate decreasing amounts of feldspar.

Figure 8.- Maturity index versus source rock index for sandstone data from Okada (1966).

equated to the sandstone classification parameters of Pettijohn (1957, p. 290). The first component corresponds to Pettijohn's fluidity index ("ratio of sand detritus to the interstitial detrial matrix."), the second to his maturity index (ratio of quartz...to feldspar plus rock fragments") and the third to his source rock index ("ratio of feldspar to rock fragments"). This coincidence may be fortuitous, or may reflect fundamental parameters of sandstones that were deduced by Pettijohn and can be extracted empirically by principal components analysis. This possibility merits further study utilizing more extensive data collections.

REFERENCES

Cooley, W.W., and Lohnes, P.R., 1962, Multivariate procedures for the behavioral sciences: John Wiley & Sons, New York, 211 p.

Greenstadt, J., 1960, The determination of the characteristic roots of a matrix by the Jacobi method, in A. Ralston and H.S. Wilf, eds., Mathematical methods for digital computers, v. 1: John Wiley & Sons, New York, 293 p.

Matalas, N.C., and Reiher, B.J., 1967, Some comments on the use of factor analysis: Water Resources Research, v. 3, no. 1, p. 213-223.

Okada, H., 1966, Non-greywacke "turbidite" sandstones in the Welsh geosyncline: Sedimentology, v. 7, no. 3, p. 211-232.

Pettijohn, F.P., 1957, Sedimentary rocks, 2nd ed., Harper & Bros., New York, 718 p.

Tien, P.L., in press, Differentiation of Pleistocene in northeastern Kansas by clay minerals: Clay and clay minerals, v. 1, no. 1, p. 1-8.

APPENDIX - Program Listing and Test Data

```
C               PRINCIPAL COMPONENTS PROGRAM
C
C               MAXIMUM OF 30 VARIABLES AND 300 SAMPLES
C
C        ORDER OF INPUT CARDS
C        1 - CONTROL CARD
C        2 - M VARIABLE DESCRIPTION CARDS
C        3 - TITLE CARD
C        4 - FORMAT CARD
C        5 - DATA CARDS
C
C        FORMAT  OF CONTROL CARD
C        COLUMN 1 - 3        M = NUMBER OF VARIABLES
C        COLUMN 6            IF MANUAL = 1  THE COVARIANCE OPTION IS TAKEN
C                            IF MANUAL = 2  THE CORRELATION OPTION IS TAKEN
C                            IF MANUAL = 3  BOTH OPTIONS ARE TAKEN
C        COLUMN 7            OPDAT
C                            IF COLUMN 7 IS LEFT BLANK THE ORIGINAL DATA
C                            WILL BE LISTED
C
      LOGICAL TEST ,OPDAT
      COMMON /WJ/ X(300,30),SSD(30,30),TITLE(20),M,N
      DIMENSION SX(30),SD(30),PERCON(30)
      DIMENSION SS(30,30),R(30,30),COVMAT(30,30)
      DIMENSION LABEL(30,4)
      DIMENSION FMT(12)
      EQUIVALENCE (SSD(1,1),SS(1,1)),(SX(1),PERCON(1))
  500 READ(5,5) M,MANUAL,OPDAT
      IF (M .LE. 0) CALL EXIT
      DO 103 L=1,M
  103 READ(5,10)(LABEL(L,IVX),IVX=1,4)
      DO 100 I=1,M
      SX(I)=0.0
      DO 100 J=1,M
      SS(I,J)=0.0
  100 CONTINUE
      READ(5,10) TITLE
      READ(5,10) FMT
      WRITE (6,1000) TITLE
      DO 16 N=1,300
      READ(5,FMT)(X(N,J),J=1,M),TEST
      IF (TEST) GO TO 26
      DO 20 I=1,M
      SX(I)=SX(I)+X(N,I)
      DO 20 J=1,M
   20 SS(I,J)=SS(I,J)+X(N,I)*X(N,J)
   16 CONTINUE
      WRITE(6,1130)N
      CALL EXIT
   26 N=N-1
      IF(OPDAT) GO TO 1643
      WRITE (6,1002)
      DO 101 I=1,N
      WRITE (6,1003) I
  101 WRITE (6,1004) (X(I,J),J=1,M)
 1643 EN=N
      DO 30 I=1,M
      DO 30 J=1,M
      SSD(I,J)=(SS(I,J)-SX(I)*SX(J)/EN)/(EN-1.0)
      COVMAT(I,J)=SSD(I,J)
   30 CONTINUE
      DO 35 I=1,M
      SD(I)=SQRT(SSD(I,I))
   35 SX(I)=SX(I)/EN
      WRITE (6,1000) TITLE
      WRITE (6,1005) M,N
      DO 102 I=1,M
  102 WRITE (6,1001) (LABEL(I,J),J=1,4),SX(I),SD(I)
      DO 60 I=1,M
      DO 60 J=1,M
      R(I,J)=SSD(I,J)/(SD(I)*SD(J))
   60 CONTINUE
      IF(MANUAL.EQ.2) GO TO 72
      WRITE (6,1000) TITLE
      WRITE (6,1006)
```

```
      CALL ARYOP(COVMAT,LABEL,M,3)
      CALL HDIAG(COVMAT,M,0,SSD,NR)
      ADITUP=0.0
      DO 9000 I=1,M
 9000 ADITUP=ADITUP+COVMAT(I,I)
      DO 9001 I=1,M
 9001 PERCON(I)=(COVMAT(I,I)/ADITUP)*100.0
      WRITE (6,1000) TITLE
      WRITE (6,1007)
      WRITE (6,1008) (COVMAT(I,I),I=1,M)
      WRITE (6,1009)
      WRITE (6,1008) (PERCON(I),I=1,M)
      WRITE (6,1010)
      CALL ARYOP2( SSD,LABEL,M,3)
      CALL JAZZUP
   72 IF(MANUAL.EQ.1) GO TO 500
      WRITE (6,1000) TITLE
      WRITE (6,1011)
      CALL ARYOP( R,LABEL,M,3)
      CALL HDIAG(R,M,0,SSD,NR)
      WRITE (6,1000) TITLE
      WRITE (6,1007)
      WRITE (6,1008) (R(I,I),I=1,M)
      ADITUP=0.0
      DO 9004I=1,M
 9004 ADITUP=ADITUP+R(I,I)
      DO 9005 I=1,M
 9005 PERCON(I)=((R(I,I)/ADITUP)*100.0     )
      WRITE (6,1009)
      WRITE (6,1008) (PERCON(I),I=1,M)
      WRITE (6,1010)
      CALL ARYOP2( SSD,LABEL,M,3)
      CALL JAZZUP
      WRITE (6,1000) TITLE
      TRACE=0.0
      DO 200 I=1,M
  200 TRACE=TRACE+R(I,I)
      WRITE(6,205)TRACE
      IF(M-IFIX(TRACE+.1)) 206,210,215
  206 WRITE(6,1220)
      GO TO 210
  215 WRITE(6,696)
  210 WRITE(6,220)
      PROD=R(1,1)
      DO 225 I=2,M
  225 PROD=PROD*R(I,I)
      WRITE(6,230)PROD
      GO TO 500
    5 FORMAT(2I3,L1)
   10 FORMAT(20A4)
  205 FORMAT(9H TRACE =  F13.5,/)
  220 FORMAT(24H0  CHECK ON DETERMINANTS)
  230 FORMAT(24H0  EIGENVALUE PRODUCT = F14.7)
  696 FORMAT(18H0  TRACE TOO SMALL)
 1000 FORMAT (1H1,20A4,//)
 1001 FORMAT(1X,4A4,5X,2F15.6)
 1002 FORMAT (11H INPUT DATA,//)
 1003 FORMAT (/,16H SAMPLE NUMBER =,I4)
 1004 FORMAT (1X,10F13.4)
 1005 FORMAT (22H NUMBER OF VARIABLES =,I3,//
     125H NUMBER OF OBSERVATIONS =,I4,///,
     29H VARIABLE,25X,25HMEAN   STANDARD DEVIATION,//)
 1006 FORMAT (18H COVARIANCE MATRIX,//)
 1007 FORMAT (12H EIGENVALUES)
 1008 FORMAT (/,1X,10G13.6)
 1009 FORMAT (//,45H PERCENT OF TOTAL CONTRIBUTION PER EIGENVALUE)
 1010 FORMAT (//,13H EIGENVECTORS)
 1011 FORMAT (19H CORRELATION MATRIX,//)
 1130 FORMAT(23H0 TOO MANY DATA POINTS ,I4)
 1220 FORMAT(16H0  TRACE TOO BIG)
      END
      SUBROUTINE HDIAG(H,N,IEGEN,U,NR)
C     THIS SUBROUTINE COMPUTES THE EIGENVALUES AND EIGENVECTORS OF A
C     REAL SYMMETRIC MATRIX H, OF ORDER N (WHERE N MUST BE LESS THAN 31)
C     AND PLACES THE EIGENVALUES IN THE DIAGONAL ELEMENTS OF THE MATRIX
C     H, AND PLACES THE EIGENVECTORS (NORMALIZED) IN THE COLUMNS OF U.
C     IEGEN IS SET AS 1 IF ONLY EIGENVALUES ARE DESIRED, AND IS SET TO 0
C     WHEN VECTORS ARE REQUIRED. NR CONTAINS THE NUMBER OF ROTATIONS DON
      DIMENSION H(30,30),U(30,30),X(30),IQ(30)
      IF(IEGEN)15,10,15
```

```
   10 DO 14 I=1,N
      DO 14 J=1,N
      IF(I-J)12,11,12
   11 U(I,J)=1.
      GO TO 14
   12 U(I,J)=0.
   14 CONTINUE
   15 NR=0
      IF(N-1)1000,1000,17
C     SCAN FOR LARGES OFF-DIAGONAL ELEMENT IN EACH ROW
C     X(I) CONTAINS LARGEST ELEMENT IN ITH ROW
C     IQ(I) HOLDS SECOND SUBSCRIPT DEFINING POSITION OF ELEMENT
   17 NMI1=N-1
      DO 30 I=1,NMI1
      X(I)=0.
      IPL1=I+1
      DO 30 J=IPL1,N
      IF(X(I)-ABS(H(I,J)))20,20,30
   20 X(I)=ABS(H(I,J))
      IQ(I)=J
   30 CONTINUE
C     SET INDICATOR FOR SHUT-OFF,RAP=2**-27,NR=NO. OF ROTATIONS
      RAP=7.450580596E-9
      HDTEST=1.0E38
C     FIND MAXIMUM OF X(I) S FOR PIVOT ELEMENT AND
C     TEST FOR END OF PROBLEM
   40 DO 70 I=1,NMI1
      IF(I-1)60,60,45
   45 IF(XMAX-X(I))60,70,70
   60 XMAX=X(I)
      IPIV=I
      JPIV=IQ(I)
   70 CONTINUE
C     IS MAX. X(I) EQUAL TO ZERO, IF LESS THAN HDTEST, REVISE HDTEST
      IF(XMAX)1000,1000,80
   80 IF(HDTEST)90,90,85
   85 IF(XMAX-HDTEST)90,90,148
   90 HDIMIN=ABS(H(1,1))
      DO 110 I=2,N
      IF(HDIMIN-ABS(H(I,I)))110,110,100
  100 HDIMIN=ABS(H(I,I))
  110 CONTINUE
      HDTEST=HDIMIN*RAP
C     RETURN IF MAX. H(I,J) LESS THAN (2**-27)ABS(H(K,K)-MIN)
      IF(HDTEST-XMAX)148,1000,1000
  148 NR = NR+1
C     COMPUTE TANGENT,SINE, AND COSINE, H(I,I),H(I,J)
  150 TANG=SIGN(2.0,(H(IPIV,IPIV)-H(JPIV,JPIV)))*H(IPIV,JPIV)/(ABS(H(IPI
 1501V,IPIV)-H(JPIV,JPIV))+SQRT((H(IPIV,IPIV)-H(JPIV,JPIV))**2+4.0*H(IP
 1502IV,JPIV)**2))
      COSINE=1.0/SQRT(1.0+TANG**2)
      SINE=TANG*COSINE
      HII=H(IPIV,IPIV)
      H(IPIV,IPIV)=COSINE**2*(HII+TANG*(2.*H(IPIV,JPIV)+TANG*H(JPIV,JPIV
     1)))
      H(JPIV,JPIV)=COSINE**2*(H(JPIV,JPIV)-TANG*(2.*H(IPIV,JPIV)-TANG*HI
     1I))
      H(IPIV,JPIV)=0.
C     PSEUDO RANK THE EIGNEVALUES
C     ADJUST SINE AND COS FOR COMPUTATION OF H(IK) AND U(IK)
      IF(H(IPIV,IPIV)-H(JPIV,JPIV))152,153,153
  152 HTEMP=H(IPIV,IPIV)
      H(IPIV,IPIV)=H(JPIV,JPIV)
      H(JPIV,JPIV)=HTEMP
C     RECOMPUTE SINE AND COSINE
      HTEMP=SIGN(1.0,-SINE)*COSINE
      COSINE=ABS(SINE)
      SINE=HTEMP
  153 CONTINUE
C     INSPECT THE IQS BETWEEN I+1 AND I-1 TO DETERMINE
C     WHETHER A NEW MAXIMUM VALUE SHOULD BE COMPUTED SINCE THE PRESENT
C     MAX IS IN THE I OR J ROW
      DO 350 I=1,NMI1
      IF(I-IPIV)210,350,200
  200 IF(I-JPIV)210,350,210
  210 IF(IQ(I)-IPIV)230,240,230
  230 IF(IQ(I)-JPIV)350,240,350
  240 K=IQ(I)
  250 HTEMP=H(I,K)
```

```
      H(I,K)=0.
      IPL1=I+1
C     SEARCH FOR DEPLETED ROW FOR NEW MAXIMUM
      DO 320 J=IPL1,N
      IF(X(I)-ABS(H(I,J)))300,300,320
  300 X(I)=ABS(H(I,J))
      IQ(I)=J
  320 CONTINUE
      H(I,K)=HTEMP
  350 CONTINUE
      X(IPIV)=0.
      X(JPIV)=0.
C     CHANGE THE OTHER ELEMENTS OF H
      DO 530 I=1,N
      IF(I-IPIV)370,530,420
  370 HTEMP=H(I,IPIV)
      H(I,IPIV)=COSINE*HTEMP+SINE*H(I,JPIV)
      IF(X(I)-ABS(H(I,IPIV)))380,390,390
  380 X(I)=ABS(H(I,IPIV))
      IQ(I)=IPIV
  390 H(I,JPIV)=-SINE*HTEMP+COSINE*H(I,JPIV)
      IF(X(I)-ABS(H(I,JPIV)))400,530,530
  400 X(I)=ABS(H(I,JPIV))
      IQ(I)=JPIV
      GO TO 530
  420 IF(I-JPIV)430,530,480
  430 HTEMP=H(IPIV,I)
      H(IPIV,I)=COSINE*HTEMP+SINE*H(I,JPIV)
      IF(X(IPIV)-ABS(H(IPIV,I)))440,450,450
  440 X(IPIV)=ABS(H(IPIV,I))
      IQ(IPIV)=I
  450 H(I,JPIV)=-SINE*HTEMP+COSINE*H(I,JPIV)
      IF(X(I)-ABS(H(I,JPIV)))400,530,530
  480 HTEMP=H(IPIV,I)
      H(IPIV,I)=COSINE*HTEMP+SINE*H(JPIV,I)
      IF(X(IPIV)-ABS(H(IPIV,I)))490,500,500
  490 X(IPIV)=ABS(H(IPIV,I))
      IQ(IPIV)=I
  500 H(JPIV,I)=-SINE*HTEMP+COSINE*H(JPIV,I)
      IF(X(JPIV)-ABS(H(JPIV,I)))510,530,530
  510 X(JPIV)=ABS(H(JPIV,I))
      IQ(JPIV)=I
  530 CONTINUE
C     TEST FOR COMPUTATION OF EIGENVECTORS
      IF(IEGEN)40,540,40
  540 DO 550 I=1,N
      HTEMP=U(I,IPIV)
      U(I,IPIV)=COSINE*HTEMP+SINE*U(I,JPIV)
  550 U(I,JPIV)=-SINE*HTEMP+COSINE*U(I,JPIV)
      GO TO 40
 1000 RETURN
      END
      SUBROUTINE JAZZUP
      INTEGER UP
      COMMON /WJ/ X(300,30),SSD(30,30),TITLE(20),M,N
      DIMENSION RESULT(300,3)
      DIMENSION XXMAX(3),XXMIN(3)
      DO 1 I=1,N
      DO 1 J=1,3
      RESULT(I,J)=0.0
      DO 1 K=1,M
    1 RESULT(I,J)=RESULT(I,J)+X(I,K)*SSD(K,J)
      WRITE (6,1000) TITLE
      WRITE (6,1001)
      LO=0
      UP=0
      LIMIT=(N/10)+1
      DO 888 IBX=1,LIMIT
      LO=UP+1
      UP=UP+10
      IF(UP.GT.N) UP=N
      WRITE (6,1003) (I,I=LO,UP)
      DO 100 J=1,3
  100 WRITE (6,1002) J,(RESULT(I,J),I=LO,UP)
      IF(UP.EQ.N) GO TO 889
  888 CONTINUE
  889 DO 20 I=1,3
      XXMAX(I)=RESULT(1,I)
      XXMIN(I)=RESULT(1,I)
      DO 20 J=1,N
```

```
      IF(XXMAX(I).LT.RESULT(J,I))XXMAX(I)=RESULT(J,I)
      IF(XXMIN(I).GT.RESULT(J,I))XXMIN(I)=RESULT(J,I)
   20 CONTINUE
      CALL PLOTER(RESULT,XXMAX,XXMIN,1,2)
      CALL PLOTER(RESULT,XXMAX,XXMIN,1,3)
      CALL PLOTER(RESULT,XXMAX,XXMIN,2,3)
      RETURN
 1000 FORMAT (1H1,20A4,//)
 1001 FORMAT (33H0ORIGINAL DATA TIMES EIGENVECTORS )
 1002 FORMAT(12H0EIGENVECTOR,I2, 7H * DATA,10F11.3)
 1003 FORMAT (//,21H0SAMPLE NUMBER        ,10I11)
      END
      SUBROUTINE ARYOP(ARAY,LABEL,M,N)
      DIMENSION ARAY(30,30),LABEL(30,4)
      DO 100 IZY=1,N
      J1=(10*(IZY-1))+1
      IX=10*IZY
      DO 10 I=J1,M
      II=I
      IF(II.GT.IX) II=IX
   10 WRITE(6,7)(LABEL(I,JV),JV=1,4),(ARAY(I,J),J=J1,II)
    7 FORMAT(1X,4A4,1X,10G11.4)
      WRITE(6,8)
    8 FORMAT(///)
      IF(M.LT.IX) GO TO 101
  100 CONTINUE
  101 RETURN
      END
      SUBROUTINE ARYOP2(ARAY,LABEL,M,N)
      DIMENSION ARAY(30,30),LABEL(30,4),NUMBER(2)
      DATA NUMBER/4HNUMB,2HER/
      DO 100 IZY = 1,N
      J1=(10*(IZY-1))+1
      IX=10*IZY
      IF(IX.GT.M) IX=M
      WRITE (6,1000) (NUMBER,IM1,IM1=J1,IX)
      DO 10  I=1,M
   10 WRITE (6,7)(LABEL(I,JV),JV=1,4),(ARAY(I,J),J=J1,IX)
    7 FORMAT(1X,4A4,1X10F11.4)
      IF(M.LE.IX) GO TO 101
  100 CONTINUE
  101 RETURN
 1000 FORMAT (//,20X,10(2X,A4,A2,I3))
      END
      SUBROUTINE PLOTER(RESULT,XMAX,XMIN,IXX,IYY)
      COMMON /WJ/ A(300,30),SSD(30,30),TITLE(20),M,N
      DIMENSION ISM1(11)
      DIMENSION RESULT(300,3),XMAX(3),XMIN(3)
      DIMENSION X(300),Y(300),IS(300)
      DIMENSION XC(300),YC(300),IOVP(300,5)
      DIMENSION MAP(110),ID(5),ISM(14)
      DATA ISM/1H ,1H*,1HA,1HB,1HC,1HD,1HE,1HF,1HG,1HH,1H+,1H-,1HI,
     16H100000/
      DATA ISM1/1H0,1H1,1H2,1H3,1H4,1H5,1H6,1H7,1H8,1H9,1H /
      DATA MINUS/1H-/
      WRITE (6,1004)
      DO 120 I=1,N
      X(I)=RESULT(I,IXX)
      Y(I)=RESULT(I,IYY)
  120 IS(I)=I
      DO 121 I=2,N
      IF (X(I)-X(I-1)) 121,121,20
   20 J=I
   21 XCH=X(J)
      X(J)=X(J-1)
      X(J-1)=XCH
      XCH=Y(J)
      Y(J)=Y(J-1)
      Y(J-1)=XCH
      ISC=IS(J)
      IS(J)=IS(J-1)
      IS(J-1)=ISC
      J=J-1
      IF (J .EQ. 1) GO TO 121
      IF (X(J)-X(J-1)) 121,121,21
  121 CONTINUE
      NUMOVP=0
      VINK=(XMAX(IYY)-XMIN(IYY))/59.0
      HINK=(XMAX(IXX)-XMIN(IXX))/99.0
      DO 100 IY=1,60
```

```
      DO 101 I=1,110
  101 MAP(I)=ISM(1)
      DO 102 I=1,101,10
  102 MAP(I)=ISM(13)
      DO 103 IX=1,N
      IYT=(XMAX(IYY)-Y(IX))/VINK+1.0
      IF (IYT .NE. IY) GO TO 103
      IXT=(X(IX)-XMIN(IXX))/HINK+1.0
      IF (IXT .LT. 1 .OR. IXT .GT. 100) GO TO 103
      ID(1)=ISM(1)
      ID(2)=ISM(11)
      ID(3)=IS(IX)/100
      ID(4)=(IS(IX)-ID(3)*100)/10
      ID(5)=IS(IX)-ID(3)*100-ID(4)*10
      NC=5
    1 IF (ID(3) .GT. 0) GO TO 2
      NC=NC-1
      ID(3)=ID(4)
      ID(4)=ID(5)
      ID(5)=ISM(1)
      GO TO 1
    2 DO 242 IWJ1=3,NC
      DO 242 IWJ=1,10
      IF(ID(IWJ1).EQ.IWJ-1)ID(IWJ1)=ISM1(IWJ)
  242 CONTINUE
      DO 105 I=2,NC
      IP=IXT+I-2
      IF (MAP(IP) .NE. ISM(1) .AND. MAP(IP) .NE. ISM(13)) GO TO 3
  105 CONTINUE
      DO 106 I=2,NC
      IP=IXT+I-2
  106 MAP(IP)=ID(I)
      GO TO 103
    3 NUMOVP=NUMOVP+1
      XC(NUMOVP)=X(IX)
      YC(NUMOVP)=Y(IX)
      DO 107 I=1,5
  107 IOVP(NUMOVP,I)=ID(I)
      IF (MAP(IP) .EQ. ISM(11) .OR. MAP(IP) .EQ. ISM(12)) IP=IP-1
    4 IF (IP .LE. 0) GO TO 103
      I=1
      IF (MAP(IP) .EQ. ISM(13)) GO TO 5
      DO 108 I=1,9
      IF (MAP(IP) .EQ. ISM(I)) GO TO 5
  108 CONTINUE
      GO TO 103
    5 MAP(IP)=ISM(I+1)
      IOVP(NUMOVP,1)=ISM(I+1)
  103 CONTINUE
      YP=XMAX(IYY)-FLOAT(IY-1)*VINK
      WRITE (6,1000) YP,MAP
  100 CONTINUE
      DO 130 I=1,101
  130 MAP(I)=MINUS
      DO 131 I=1,101,10
  131 MAP(I)=ISM(13)
      WRITE (6,1009) MAP
      DO 132 I=1,11
      X(I)=XMIN(IXX)+FLOAT(I-1)*10.0*HINK
  132 CONTINUE
      WRITE (6,1008) (X(I),I=1,11,2),(X(I),I=2,11,2)
      WRITE (6,1005) TITLE
      WRITE (6,1006) IXX
      WRITE (6,1007) IYY
      IF (NUMOVP .EQ. 0) RETURN
      WRITE (6,1001)
      WRITE (6,1002)
      DO 110 I=1,NUMOVP
      WRITE (6,1003) (IOVP(I,J),J=1,5),XC(I),YC(I)
  110 CONTINUE
      RETURN
 1000 FORMAT (1X,G15.5,2X,110A1)
 1001 FORMAT (////,1X,17H OVERPRINT VALUES,//)
 1002 FORMAT(37H POINT          XCOORD          YCOORD
 1003 FORMAT (5X,5A1,2G15.5)
 1004 FORMAT (1H1)
 1005 FORMAT(///,1X,20A4,///,18H IN THE MAP ABOVE,    )
 1006 FORMAT(27H0HORIZONTAL EIGENVECTOR =    I2   )
 1007 FORMAT(27H0   VERTICAL EIGENVECTOR =      I2)
 1008 FORMAT (3X,6F20.4,/,13X,5F20.4)
 1009 FORMAT (18X,110A1)
      END
```

PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

INPUT DATA

SAMPLE NUMBER = 1
60.0000 13.0000 27.0000 0.4800

SAMPLE NUMBER = 2
55.0000 19.0000 26.0000 0.7300

SAMPLE NUMBER = 3
69.0000 20.0000 11.0000 1.8200

SAMPLE NUMBER = 4
73.0000 17.0000 10.0000 1.7000

SAMPLE NUMBER = 5
72.0000 19.0000 9.0000 2.1100

SAMPLE NUMBER = 6
74.0000 18.0000 8.0000 2.2500

SAMPLE NUMBER = 7
63.0000 25.0000 12.0000 2.0800

SAMPLE NUMBER = 8
68.0000 22.0000 10.0000 2.2000

SAMPLE NUMBER = 9
68.0000 20.0000 12.0000 1.6700

SAMPLE NUMBER = 10
77.0000 13.0000 10.0000 1.3000

SAMPLE NUMBER = 11
72.0000 17.0000 11.0000 1.5500

SAMPLE NUMBER = 12
71.0000 17.0000 12.0000 1.4200

SAMPLE NUMBER = 13
75.0000 14.0000 11.0000 1.2800

SAMPLE NUMBER = 14
76.0000 14.0000 10.0000 1.4000

SAMPLE NUMBER = 15
73.0000 16.0000 11.0000 1.4500

SAMPLE NUMBER = 16
67.0000 21.0000 12.0000 1.7500

SAMPLE NUMBER = 17
70.0000 22.0000 8.0000 2.7500

SAMPLE NUMBER = 18
61.0000 29.0000 10.0000 2.9000

SAMPLE NUMBER = 19

62.0000 27.0000 11.0000 2.4500

SAMPLE NUMBER = 20
68.0000 20.0000 12.0000 1.6700

SAMPLE NUMBER = 21
59.0000 31.0000 10.0000 3.1000

SAMPLE NUMBER = 22
65.0000 26.0000 9.0000 2.8900

SAMPLE NUMBER = 23
62.0000 29.0000 9.0000 3.2200

SAMPLE NUMBER = 24
67.0000 20.0000 13.0000 1.5400

SAMPLE NUMBER = 25
60.0000 17.0000 23.0000 0.7400

SAMPLE NUMBER = 26
58.0000 24.0000 18.0000 1.3300

SAMPLE NUMBER = 27
77.0000 15.0000 8.0000 1.8800

SAMPLE NUMBER = 28
77.0000 15.0000 8.0000 1.8800

SAMPLE NUMBER = 29
71.0000 21.0000 8.0000 2.6300

SAMPLE NUMBER = 30
80.0000 13.0000 7.0000 1.8600

SAMPLE NUMBER = 31
71.0000 21.0000 8.0000 2.6300

SAMPLE NUMBER = 32
75.0000 19.0000 6.0000 3.1700

SAMPLE NUMBER = 33
76.0000 17.0000 7.0000 2.4300

SAMPLE NUMBER = 34
72.0000 20.0000 8.0000 2.5000

SAMPLE NUMBER = 35
84.0000 12.0000 4.0000 3.0000

SAMPLE NUMBER = 36
82.0000 13.0000 5.0000 2.6000

SAMPLE NUMBER = 37
78.0000 18.0000 4.0000 4.5000

SAMPLE NUMBER = 38
57.0000 28.0000 15.0000 1.8700

SAMPLE NUMBER = 39
54.0000 26.0000 20.0000 1.3000

SAMPLE NUMBER = 40

72.0000	16.0000	12.0000	1.3300

SAMPLE NUMBER = 41

74.0000	15.0000	11.0000	1.3600

SAMPLE NUMBER = 42

70.0000	23.0000	7.0000	3.2900

SAMPLE NUMBER = 43

68.0000	24.0000	8.0000	3.0000

SAMPLE NUMBER = 44

66.0000	12.0000	22.0000	0.5500

SAMPLE NUMBER = 45

60.0000	14.0000	26.0000	0.5400

SAMPLE NUMBER = 46

49.0000	21.0000	30.0000	0.7000

SAMPLE NUMBER = 47

56.0000	23.0000	21.0000	1.1000

SAMPLE NUMBER = 48

53.0000	29.0000	18.0000	1.6100

SAMPLE NUMBER = 49

73.0000	22.0000	5.0000	4.4000

SAMPLE NUMBER = 50

59.0000	29.0000	12.0000	2.4200

SAMPLE NUMBER = 51

65.0000	22.0000	13.0000	1.6900

SAMPLE NUMBER = 52

56.0000	18.0000	26.0000	0.6900

SAMPLE NUMBER = 53

56.0000	24.0000	20.0000	1.2000

SAMPLE NUMBER = 54

70.0000	21.0000	9.0000	2.3300

SAMPLE NUMBER = 55

70.0000	21.0000	9.0000	2.3300

SAMPLE NUMBER = 56

67.0000	25.0000	8.0000	3.1300

SAMPLE NUMBER = 57

69.0000	9.0000	22.0000	0.4100

SAMPLE NUMBER = 58

64.0000	15.0000	21.0000	0.7100

SAMPLE NUMBER = 59

59.0000	14.0000	27.0000	0.5200

SAMPLE NUMBER = 60

68.0000	23.0000	9.0000	2.5600

SAMPLE NUMBER = 61

70.0000	21.0000	9.0000	2.3400

SAMPLE NUMBER = 62

76.0000	18.0000	6.0000	1.1300

SAMPLE NUMBER = 63

73.0000	22.0000	5.0000	4.4000

SAMPLE NUMBER = 64

77.0000	18.0000	5.0000	3.6000

SAMPLE NUMBER = 65

78.0000	17.0000	5.0000	3.4000

SAMPLE NUMBER = 66

72.0000	15.0000	13.0000	1.1500

SAMPLE NUMBER = 67

80.0000	14.0000	6.0000	2.3300

SAMPLE NUMBER = 68

61.0000	15.0000	24.0000	0.6300

SAMPLE NUMBER = 69

61.0000	15.0000	24.0000	0.6300

SAMPLE NUMBER = 70

41.0000	28.0000	31.0000	0.9000

SAMPLE NUMBER = 71

41.0000	29.0000	30.0000	0.9700

SAMPLE NUMBER = 72

58.0000	25.0000	17.0000	1.4700

SAMPLE NUMBER = 73

64.0000	28.0000	8.0000	3.5000

SAMPLE NUMBER = 74

59.0000	28.0000	13.0000	2.1500

PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

NUMBER OF VARIABLES = 4

NUMBER OF OBSERVATIONS = 74

VARIABLE	MEAN	STANDARD DEVIATION
MIXED-LAYER CLAY	66.945946	8.876461
ILLITE	20.013513	5.219807
KAOLINITE	13.040540	7.200531
D.I. RATIO	1.952702	0.993941

PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

COVARIANCE MATRIX

MIXED-LAYER CLAY	78.79			
ILLITE	-27.10	27.25		
KAOLINITE	-51.70	-0.1512	51.85	
D.I. RATIO	3.830	1.842	-5.671	0.9879

PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

EIGENVALUES

123.877 34.7552 0.241466-0.823334E-06

PERCENT OF TOTAL CONTRIBUTION PER EIGENVALUE

77.9720 21.8760 0.151986-0.518233E-06

EIGENVECTORS

	NUMBER 1	NUMBER 2	NUMBER 3	NUMBER 4
MIXED-LAYER CLAY	0.790	-0.206	0.014	0.577
ILLITE	-0.220	0.782	0.081	0.577
KAOLINITE	-0.570	-0.577	-0.095	0.577
D.I. RATIO	0.048	0.116	-0.992	0.000

PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

ORIGINAL DATA TIMES EIGENVECTORS

SAMPLE NUMBER	1	2	3	4	5	6	7	8	9	10
EIGENVECTOR 1 * DATA	29.171	24.484	43.932	48.316	47.676	50.053	37.535	43.291	42.565	52.336
EIGENVECTOR 2 * DATA	-17.681	-11.354	-4.677	-7.283	-4.849	-5.490	-0.078	-2.286	-5.065	-11.281
EIGENVECTOR 3 * DATA	-1.152	-0.888	-0.271	-0.245	-0.409	-0.506	-0.302	-0.405	-0.231	-0.117

SAMPLE NUMBER	11	12	13	14	15	16	17	18	19	20
EIGENVECTOR 1 * DATA	46.949	45.582	49.965	51.331	47.954	41.559	46.038	36.256	36.894	42.565
EIGENVECTOR 2 * DATA	-7.672	-8.058	-10.667	-10.282	-8.671	-4.068	-1.480	4.710	2.311	-5.065
EIGENVECTOR 3 * DATA	-0.205	-0.185	-0.139	-0.149	-0.173	-0.243	-0.733	-0.629	-0.426	-0.231

SAMPLE NUMBER	21	22	23	24	25	26	27	28	29	30
EIGENVECTOR 1 * DATA	34.246	40.645	37.632	41.198	30.586	30.347	53.065	53.065	47.042	56.444
EIGENVECTOR 2 * DATA	6.709	2.116	5.118	-5.451	-12.216	-3.378	-8.496	-8.496	-2.482	-10.103
EIGENVECTOR 3 * DATA	-0.693	-0.712	-0.838	-0.211	-0.706	-0.278	-0.340	-0.340	-0.681	-0.346

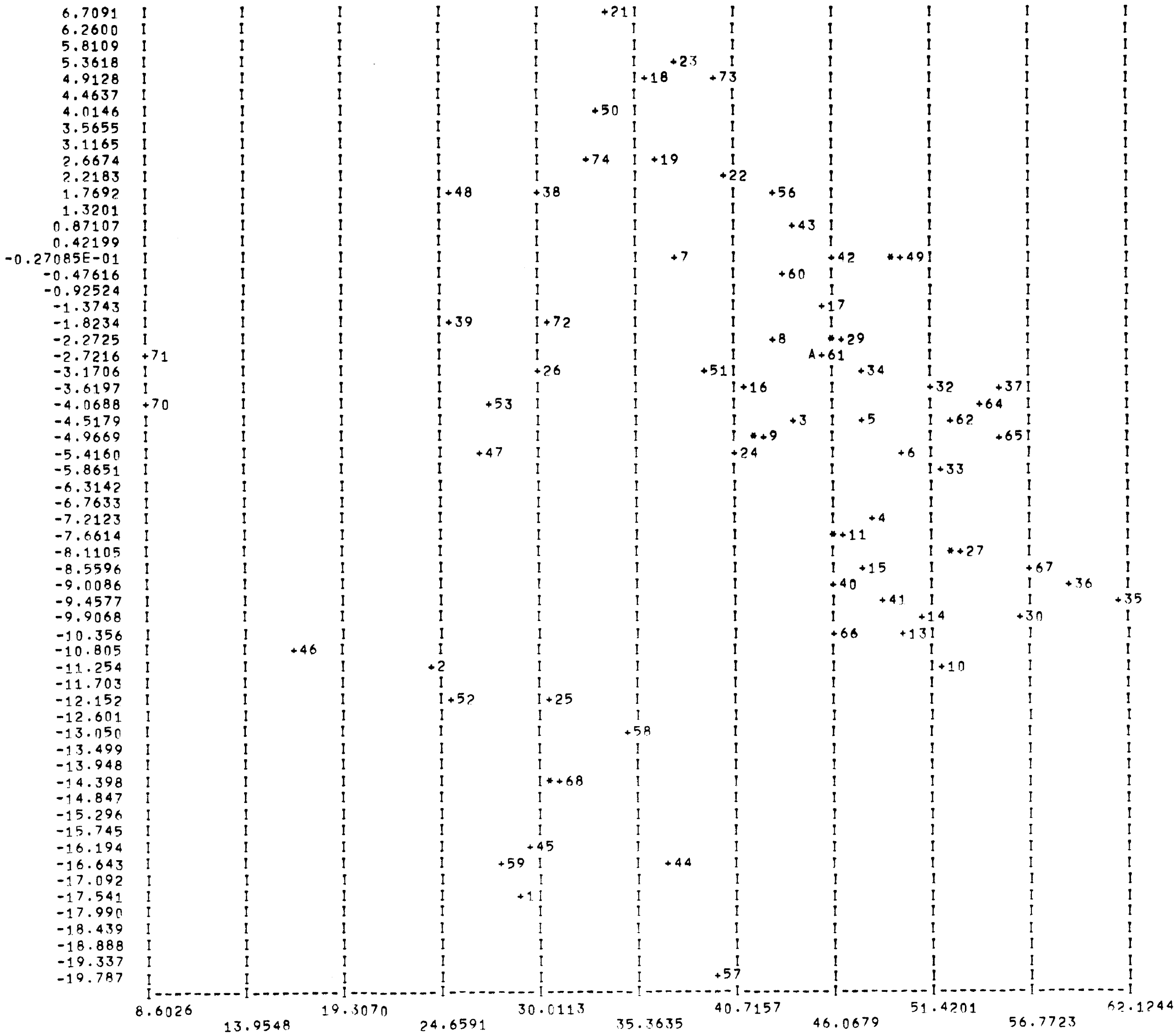
SAMPLE NUMBER	31	32	33	34	35	36	37	38	39	40
EIGENVECTOR 1 * DATA	47.042	51.808	52.432	48.046	61.589	59.200	55.602	30.415	25.605	46.588
EIGENVECTOR 2 * DATA	-2.482	-3.653	-6.086	-3.485	-9.846	-9.276	-3.745	1.748	-2.148	-9.056
EIGENVECTOR 3 * DATA	-0.681	-1.134	-0.643	-0.620	-1.218	-0.863	-2.303	-0.219	-0.331	-0.163

SAMPLE NUMBER	41	42	43	44	45	46	47	48	49	50
EIGENVECTOR 1 * DATA	48.959	46.414	44.030	36.986	29.524	17.022	27.265	25.311	50.198	33.512
EIGENVECTOR 2 * DATA	-9.669	-0.059	0.524	-16.807	-16.316	-10.866	-5.506	1.593	-0.176	3.913
EIGENVECTOR 3 * DATA	-0.151	-1.093	-0.847	-0.745	-1.036	-1.159	-0.443	-0.220	-2.044	-0.371

SAMPLE NUMBER	51	52	53	54	55	56	57	58	59	60
EIGENVECTOR 1 * DATA	39.186	25.492	27.620	45.667	45.667	43.027	40.008	35.324	28.163	43.659
EIGENVECTOR 2 * DATA	-3.458	-12.347	-4.135	-2.888	-2.888	1.527	-19.787	-13.454	-16.689	-0.885
EIGENVECTOR 3 * DATA	-0.225	-0.916	-0.366	-0.492	-0.492	-0.909	-0.808	-0.594	-1.125	-0.586

SAMPLE NUMBER	61	62	63	64	65	66	67	68	69	70
EIGENVECTOR 1 * DATA	45.668	52.721	50.198	54.199	55.199	46.228	56.817	31.239	31.239	8.603
EIGENVECTOR 2 * DATA	-2.887	-4.878	-0.176	-4.220	-5.231	-10.436	-8.690	-14.576	-14.576	-4.299
EIGENVECTOR 3 * DATA	-0.502	0.823	-2.044	-1.519	-1.388	-0.160	-0.636	-0.840	-0.840	-0.996

SAMPLE NUMBER	71	72	73	74
EIGENVECTOR 1 * DATA	8.957	30.704	40.015	33.149
EIGENVECTOR 2 * DATA	-2.932	-2.003	4.534	2.523
EIGENVECTOR 3 * DATA	-0.890	-0.241	-1.074	-0.279



PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

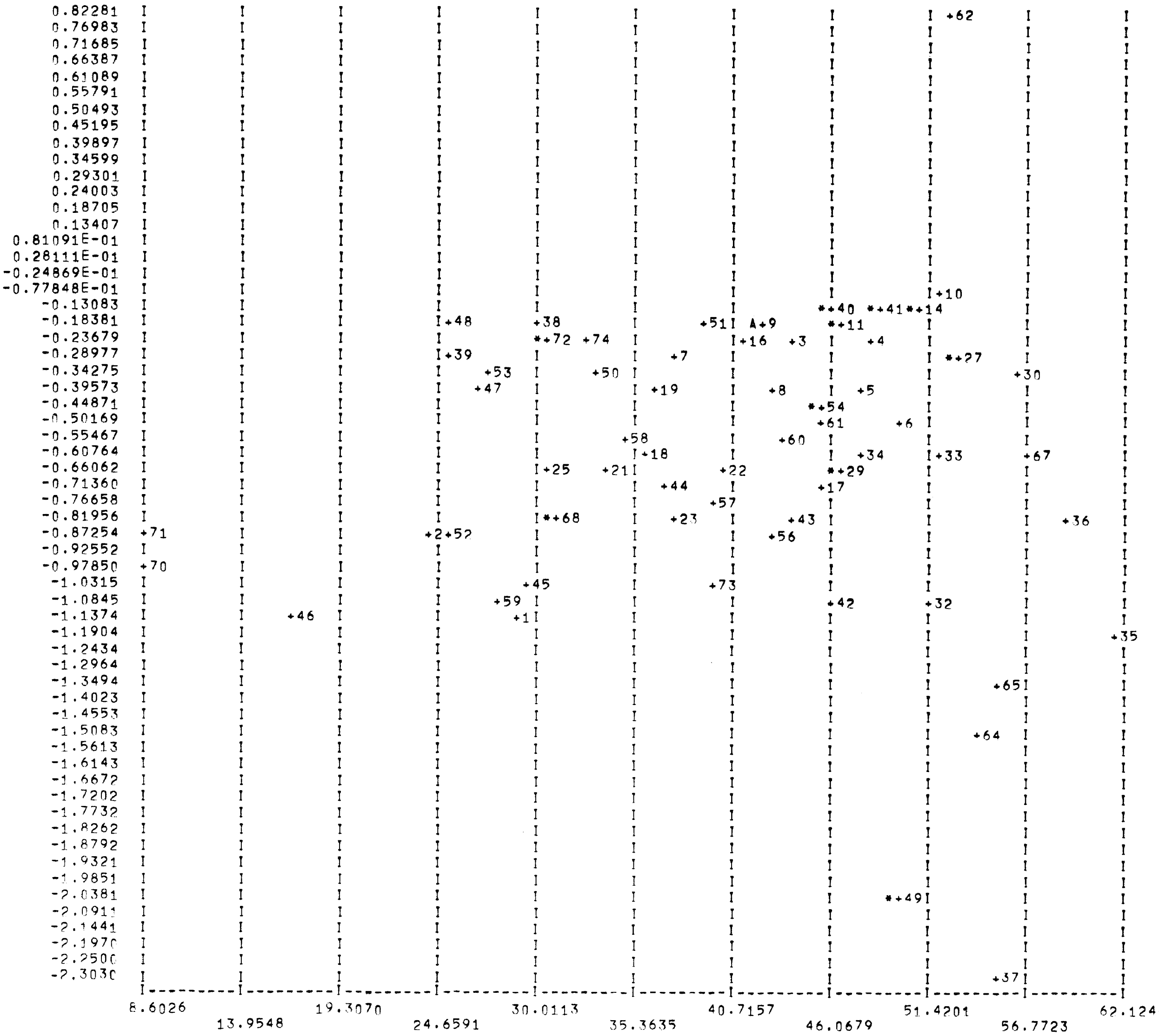
IN THE MAP ABOVE,

HORIZONTAL EIGENVECTOR = 1

VERTICAL EIGENVECTOR = 2

OVERPRINT VALUES

POINT	XCOORD	YCOORD
*+63	50.198	-0.17606
*+31	47.042	-2.4822
*+54	45.667	-2.8879
A+55	45.667	-2.8879
*+20	42.565	-5.0651
*+12	45.582	-8.0576
*+28	53.065	-8.4963
*+69	31.239	-14.576



PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

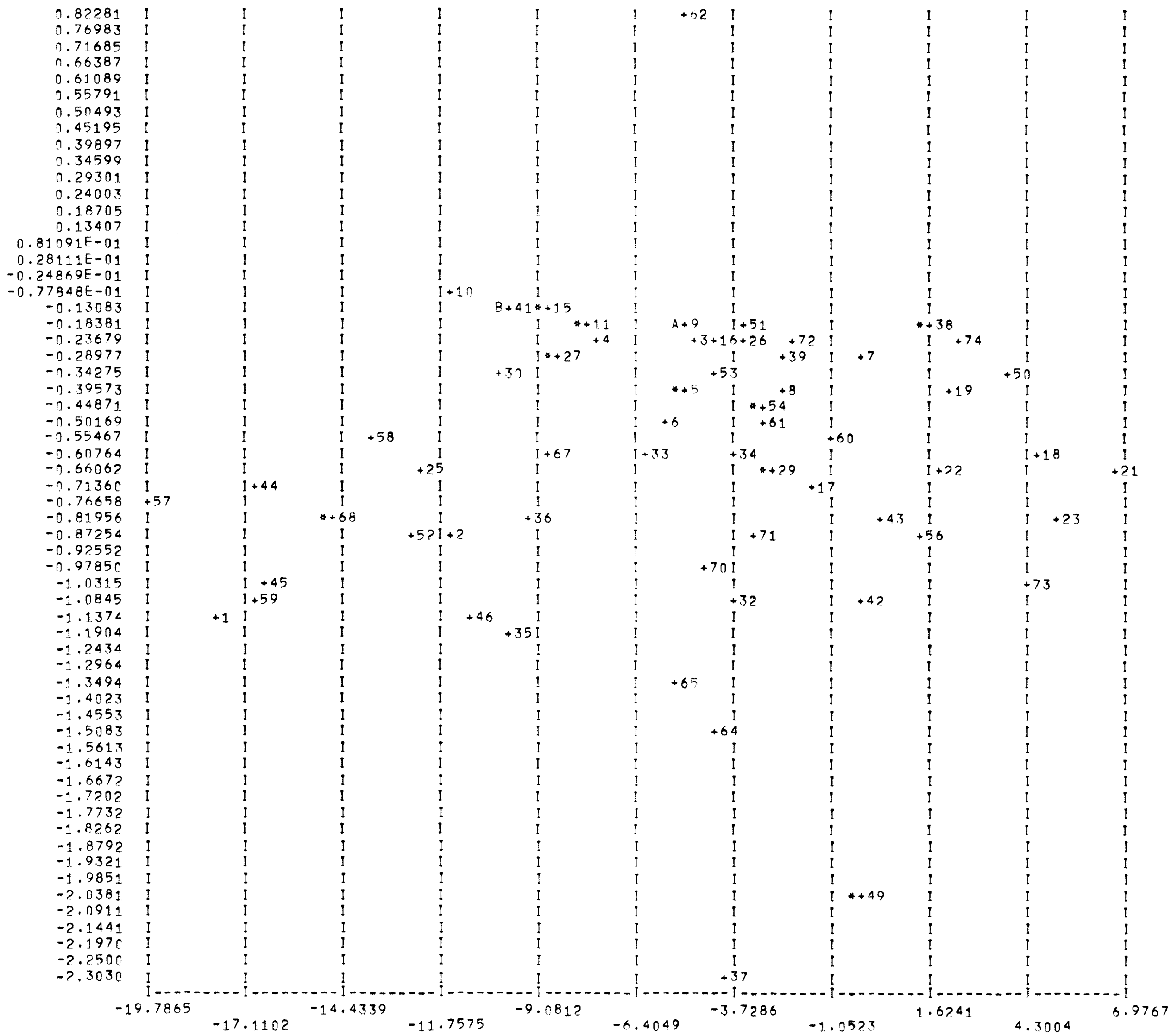
IN THE MAP ABOVE,

HORIZONTAL EIGENVECTOR = 1

VERTICAL EIGENVECTOR = 3

OVERPRINT VALUES

POINT	XCOORD	YCOORD
*+13	49.965	-0.13851
*+15	47.954	-0.17283
*+66	46.228	-0.15990
*+12	45.582	-0.18466
*+20	42.565	-0.23117
A+24	41.198	-0.21096
*+26	30.347	-0.27777
*+28	53.065	-0.34030
*+55	45.667	-0.49248
*+31	47.042	-0.68134
*+69	31.239	-0.84043
*+63	50.198	-2.0439



PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

IN THE MAP ABOVE,

HORIZONTAL EIGENVECTOR = 2

VERTICAL EIGENVECTOR = 3

OVERPRINT VALUES

POINT	XCOORD	YCOORD
*+40	-9.0559	-0.16254
*+14	-10.282	-0.14880
A+66	-10.436	-0.15990
B+13	-10.667	-0.13851
*+48	1.5932	-0.21970
*+20	-5.0651	-0.23117
A+24	-5.4510	-0.21096
*+12	-8.0576	-0.18466
*+28	-8.4963	-0.34030
*+47	-5.5057	-0.44306
*+55	-2.8879	-0.49248
*+31	-2.4822	-0.68134
*+69	-14.576	-0.84043
*+63	-0.17606	-2.0439

PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

CORRELATION MATRIX

MIXED-LAYER CLAY	1.000			
ILLITE	-0.5848	1.000		
KAOLINITE	-0.8088	-0.4024E-02	1.000	
D.I. RATIO	0.4341	0.3550	-0.7924	1.000

PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

EIGENVALUES

2.38378 1.47468 0.141542-0.189600E-07

PERCENT OF TOTAL CONTRIBUTION PER EIGENVALUE

59.5945 36.8669 3.53855-0.474001E-06

EIGENVECTORS

	NUMBER 1	NUMBER 2	NUMBER 3	NUMBER 4
MIXED-LAYER CLAY	0.575	-0.376	-0.171	-0.706
ILLITE	-0.110	0.801	-0.416	-0.415
KAOLINITE	-0.629	-0.117	0.512	-0.573
D.I. RATIO	0.512	0.450	0.731	-0.000

PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

ORIGINAL DATA TIMES EIGENVECTORS

SAMPLE NUMBER	1	2	3	4	5	6	7	8	9	10
EIGENVECTOR 1 * DATA	16.316	13.543	31.468	34.663	34.708	36.667	26.977	31.498	30.188	37.195
EIGENVECTOR 2 * DATA	-15.106	-8.187	-10.411	-14.258	-11.978	-13.352	-4.147	-8.145	-10.220	-19.149
EIGENVECTOR 3 * DATA	-1.474	-3.448	-13.143	-13.177	-14.052	-14.387	-13.498	-14.039	-12.570	-12.487

SAMPLE NUMBER	11	12	13	14	15	16	17	18	19	20
EIGENVECTOR 1 * DATA	33.383	32.113	35.297	36.562	34.015	29.545	34.187	27.068	27.002	30.188
EIGENVECTOR 2 * DATA	-14.066	-13.865	-17.721	-17.926	-15.289	-9.006	-8.416	0.414	-1.884	-10.220
EIGENVECTOR 3 * DATA	-12.604	-12.016	-12.064	-12.660	-12.431	-12.757	-15.003	-15.247	-14.401	-12.570

SAMPLE NUMBER	21	22	23	24	25	26	27	28	29	30
EIGENVECTOR 1 * DATA	25.802	30.318	28.435	28.918	18.526	20.057	38.530	38.530	34.809	41.092
EIGENVECTOR 2 * DATA	2.860	-3.383	0.299	-10.019	-11.317	-4.105	-17.051	-17.051	-9.648	-19.675
EIGENVECTOR 3 * DATA	-15.592	-15.200	-15.696	-11.982	-4.999	-9.702	-13.920	-13.920	-14.845	-14.127

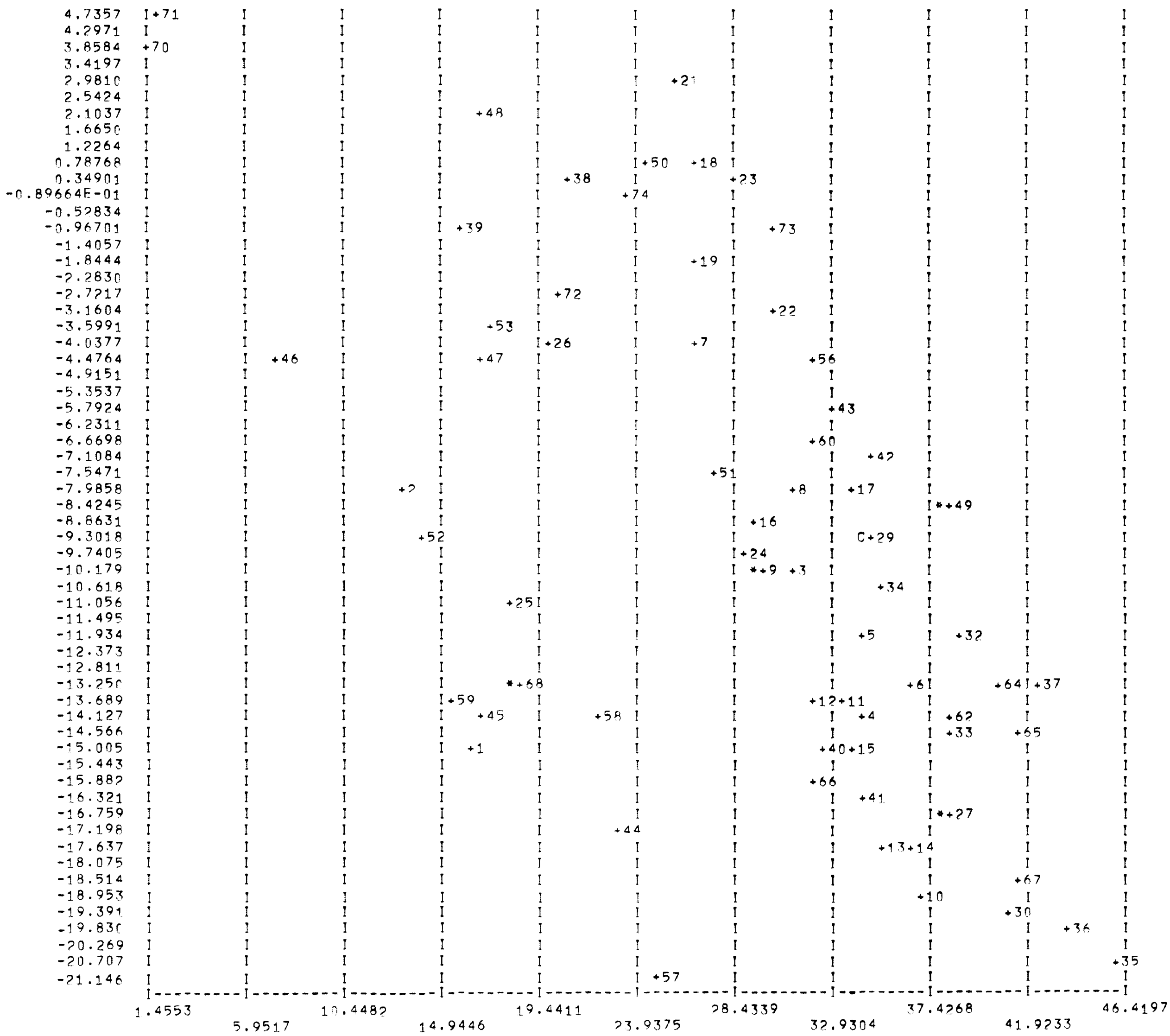
SAMPLE NUMBER	31	32	33	34	35	36	37	38	39	40
EIGENVECTOR 1 * DATA	34.809	38.861	38.647	35.427	45.970	43.878	42.634	21.207	16.266	32.751
EIGENVECTOR 2 * DATA	-9.648	-12.279	-14.708	-10.884	-21.118	-19.861	-13.377	0.070	-1.244	-15.083
EIGENVECTOR 3 * DATA	-14.845	-15.326	-14.692	-14.695	-15.097	-14.952	-15.473	-12.339	-8.849	-11.836

SAMPLE NUMBER	41	42	43	44	45	46	47	48	49	50
EIGENVECTOR 1 * DATA	34.654	34.983	32.947	23.053	16.866	7.345	17.013	16.779	38.642	24.415
EIGENVECTOR 2 * DATA	-16.507	-7.255	-5.948	-17.550	-14.161	-4.807	-4.607	1.910	-8.452	0.717
EIGENVECTOR 3 * DATA	-12.252	-15.537	-15.312	-4.593	-2.359	-1.229	-7.576	-10.725	-15.846	-14.231

SAMPLE NUMBER	51	52	53	54	55	56	57	58	59	60
EIGENVECTOR 1 * DATA	27.627	14.206	17.583	33.452	33.452	32.329	25.034	22.286	15.653	32.202
EIGENVECTOR 2 * DATA	-7.596	-9.383	-3.644	-9.523	-9.523	-4.712	-21.146	-14.204	-13.911	-7.064
EIGENVECTOR 3 * DATA	-12.363	-3.232	-8.431	-14.382	-14.382	-15.462	-3.959	-5.896	-1.691	-14.705

SAMPLE NUMBER	61	62	63	64	65	66	67	68	69	70
EIGENVECTOR 1 * DATA	33.457	38.500	38.642	40.969	41.551	32.139	41.852	18.635	18.635	1.455
EIGENVECTOR 2 * DATA	-9.519	-14.375	-8.452	-13.523	-14.790	-16.082	-18.546	-13.462	-13.462	3.786
EIGENVECTOR 3 * DATA	-14.375	-16.572	-15.846	-15.448	-15.349	-11.039	-14.712	-3.905	-3.905	-2.119

SAMPLE NUMBER	71	72	73	74
EIGENVECTOR 1 * DATA	2.010	20.648	30.466	23.757
EIGENVECTOR 2 * DATA	4.736	-3.124	-1.012	-0.323
EIGENVECTOR 3 * DATA	-2.996	-10.528	-15.928	-13.500



PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

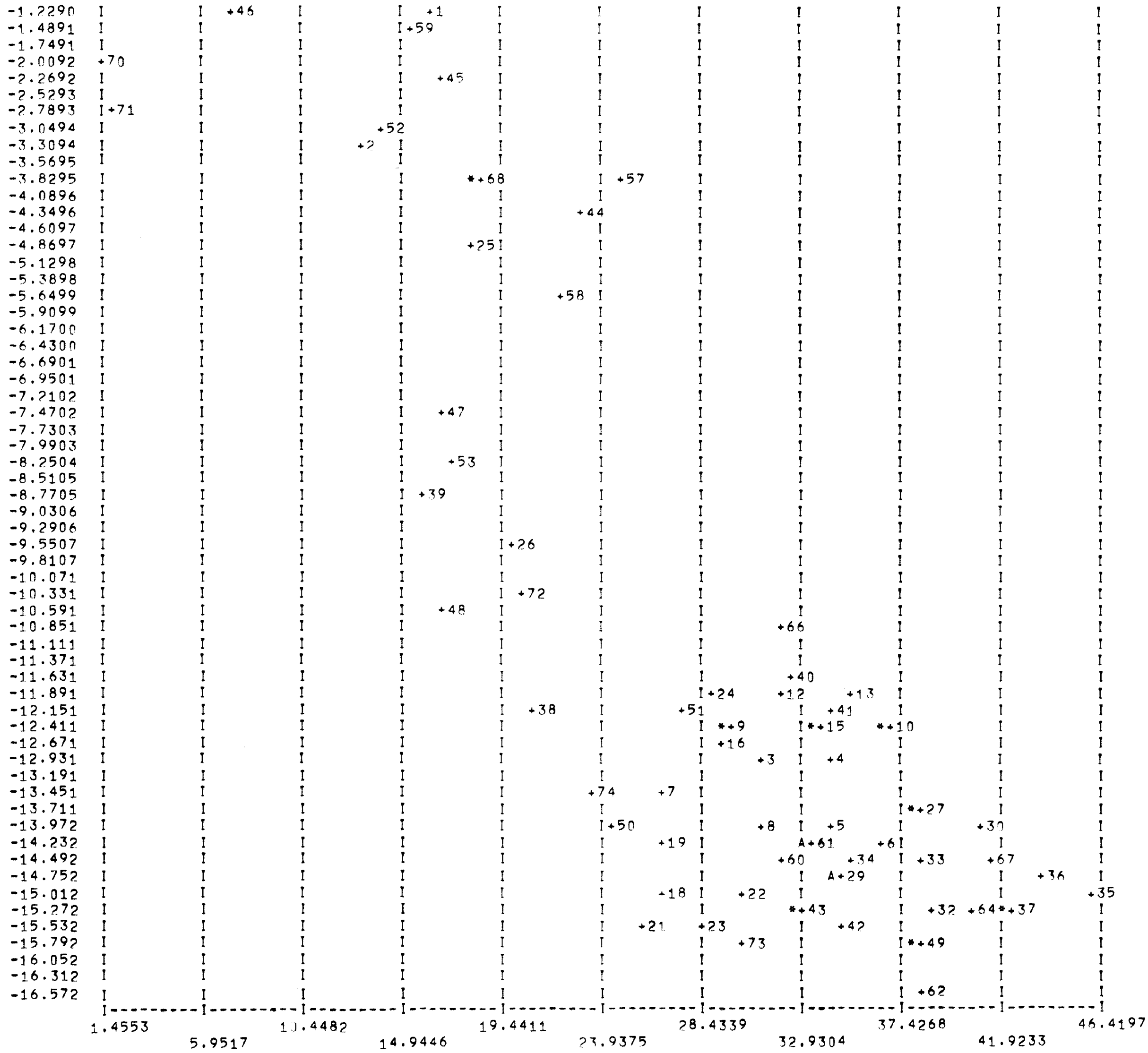
IN THE MAP ABOVE,

HORIZONTAL EIGENVECTOR = 1

VERTICAL EIGENVECTOR = 2

OVERPRINT VALUES

POINT	XCOORD	YCOORD
*+63	38.642	-8.4519
*+31	34.809	-9.6478
A+61	33.457	-9.5188
B+54	33.452	-9.5233
C+55	33.452	-9.5233
*+20	30.188	-10.220
*+69	18.635	-13.462
*+28	38.530	-17.051



PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

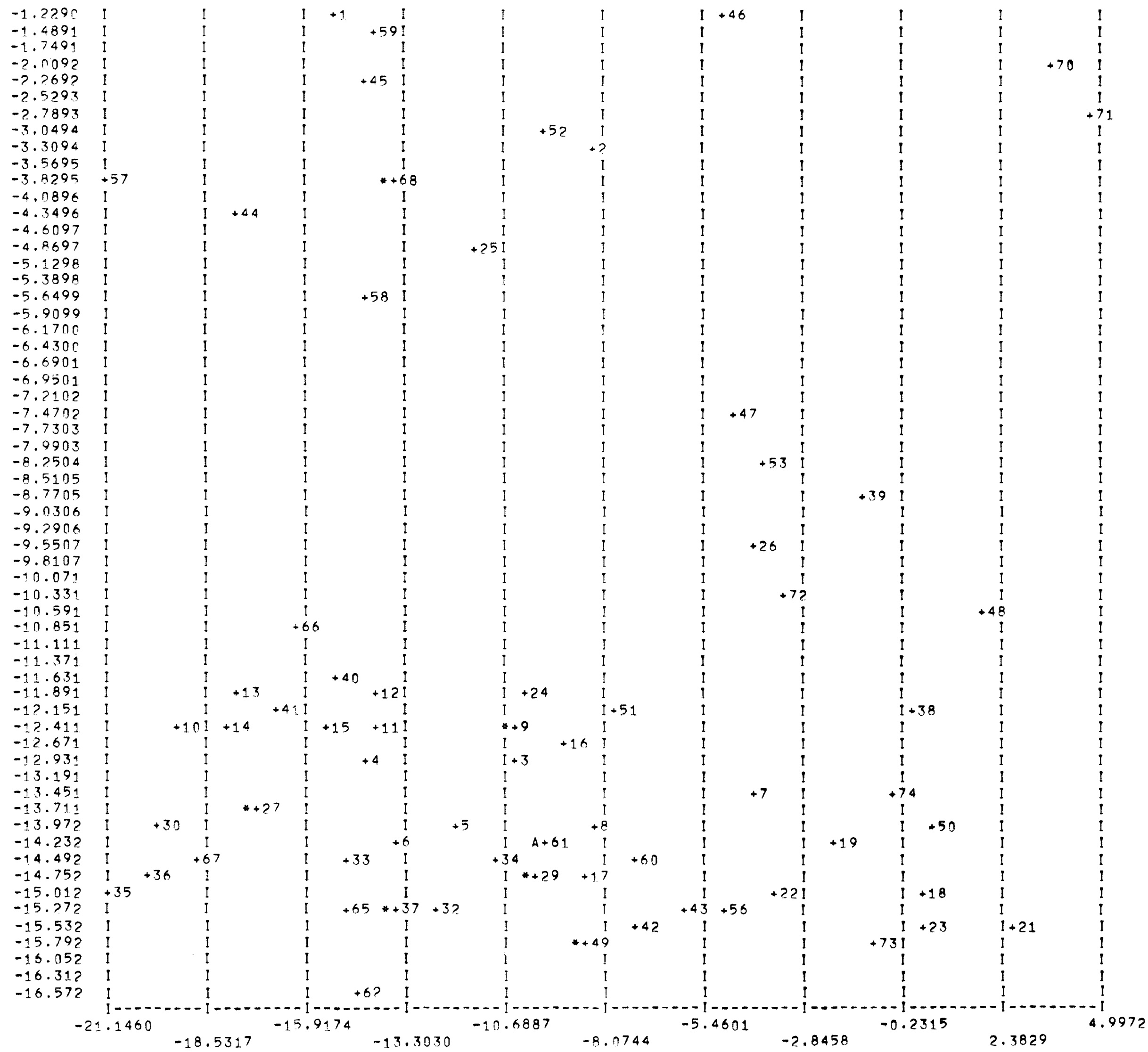
IN THE MAP ABOVE,

HORIZONTAL EIGENVECTOR = 1

VERTICAL EIGENVECTOR = 3

OVERPRINT VALUES

POINT	XCOORD	YCOORD
*+69	18.635	-3.9052
*+14	36.562	-12.660
*+11	33.383	-12.604
*+20	30.188	-12.570
*+28	38.530	-13.920
*+54	33.452	-14.382
A+55	33.452	-14.382
*+31	34.809	-14.845
A+17	34.187	-15.003
*+65	41.551	-15.349
*+56	32.329	-15.462
*+63	38.642	-15.846



PRINCIPAL COMPONENTS ANALYSIS OF PLEISTOCENE CLAYS OF KANSAS, BY TIEN.

IN THE MAP ABOVE,

HORIZONTAL EIGENVECTOR = 2

VERTICAL EIGENVECTOR = 3

OVERPRINT VALUES

POINT	XCOORD	YCOORD
*+69	-13.462	-3.9052
*+20	-10.220	-12.570
*+28	-17.051	-13.920
*+54	-9.5233	-14.382
A+55	-9.5233	-14.382
*+31	-9.6478	-14.845
*+64	-13.523	-15.448
*+63	-8.4519	-15.846

KANSAS GEOLOGICAL SURVEY COMPUTER PROGRAM
THE UNIVERSITY OF KANSAS, LAWRENCE

PROGRAM ABSTRACT

Title (If subroutine state in title):

FORTRAN IV program for computation and display of principal components

Date:

Author, organization: Warren J. Wahlstedt, Cities Service Oil Company, Cities Service Building, Bartlesville, Oklahoma

Direct inquiries to: W.J. Wahlstedt, or

Name: John C. Davis Address: Kansas Geological Survey
Lawrence, Kansas 66044

Purpose/description: Computes principal components of covariance or correlation matrix and plots original data in the form of the first three principal components.

Mathematical method: Eigenvalues and eigenvectors are extracted from the symmetrical covariance or correlation matrix by a modified Jacobi method.

Restrictions, range: Up to 30 variables and 300 observations may be used.

Computer manufacturer: General Electric Model: 625

Programming language: FORTRAN IV (G)

Memory required: 25 K Approximate running time: $\frac{1}{2}$ min. for 100 samples, 10 variables

Special peripheral equipment required: none

Remarks (special compilers or operating systems, required word lengths, number of successful runs, other machine versions, additional information useful for operation or modification of program)

Program has been run successfully many times on the GE 625 and IBM/360 Model 65

COMPUTER CONTRIBUTIONS

Kansas Geological Survey
University of Kansas
Lawrence, Kansas

Computer Contribution

1. Mathematical simulation of marine sedimentation with IBM 7090/7094 computers, by J.W. Harbaugh, 1966 . . . . $1.00
2. A generalized two-dimensional regression procedure, by J.R. Dempsey, 1966 . . . . . $0.50
3. FORTRAN IV and MAP program for computation and plotting of trend surfaces for degrees 1 through 6, by Mont O'Leary, R.H. Lippert, and O.T. Spitz, 1966 . . . . . . $0.75
4. FORTRAN II program for multivariate discriminant analysis using an IBM 1620 computer, by J.C. Davis and R.J. Sampson, 1966 . . . . $0.50
5. FORTRAN IV program using double Fourier series for surface fitting of irregularly spaced data, by W.R. James, 1966. . . . . $0.75
6. FORTRAN IV program for estimation of cladistic relationships using the IBM 7040, by R.L. Bartcher, 1966 . . . . $1.00
7. Computer applications in the earth sciences: Colloquium on classification procedures, edited by D.F. Merriam, 1966 . . . . $1.00
8. Prediction of the performance of a solution gas drive reservoir by Muskat's equation, by Apolonio Baca, 1967 . . . . $1.00
9. FORTRAN IV program for mathematical simulation of marine sedimentation with IBM 7040 or 7094 computers, by J.W. Harbaugh and W.J. Wahlstedt, 1967 . . . . . . $1.00
10. Three-dimensional response surface program in FORTRAN II for the IBM 1620 computer, by R.J. Sampson and J.C. Davis, 1967 . . . . $0.75
11. FORTRAN IV program for vector trend analyses of directional data, by W.T. Fox, 1967 . . $1.00
12. Computer applications in the earth sciences: Colloquium on trend analysis, edited by D.F. Merriam and N.C. Cocke, 1967 . . . . $1.00
13. FORTRAN IV computer programs for Markov chain experiments in geology, by W.C. Krumbein, 1967 . . . . $1.00
14. FORTRAN IV programs to determine surface roughness in topography for the CDC 3400 computer, by R.D. Hobson, 1967 . . . . $1.00
15. FORTRAN II program for progressive linear fit of surfaces on a quadratic base using an IBM 1620 computer, by A.J. Cole, C. Jordan, and D.F. Merriam, 1967 . . . . . $1.00
16. FORTRAN IV program for the GE 625 to compute the power spectrum of geological surfaces, by J.E. Esler and F.W. Preston, 1967 . . . . $0.75
17. FORTRAN IV program for Q-mode cluster analysis of nonquantitative data using IBM 7090/7094 computers, by G.F. Bonham-Carter, 1967. . . . . $1.00
18. Computer applications in the earth sciences: Colloquium on time-series analysis, D.F. Merriam, editor, 1967 . . . . $1.00
19. FORTRAN II time-trend package for the IBM 1620 computer, by J.C. Davis and R.J. Sampson, 1967 . . . . $1.00
20. Computer programs for multivariate analysis in geology, edited by D.F. Merriam, 1968 . . $1.00
21. FORTRAN IV program for computation and display of principal components, by W.C. Wahlstedt and J.C. Davis, 1968 . . . . $1.00