UNITED STATES DEPARTMENT OF THE INTERIOR

GEOLOGICAL SURVEY

Brief descriptions of STATPAC and related statistical programs for the IBM Personal Computer

compiled by

W. D. Grundy[1] and A. T. Miesch[2]

Last update June 30, 1987

Open-File Report 87-411-A

DISCLAIMER

Although program tests have been made, no guarantee (expressed or implied) is made by the authors regarding program correctness, accuracy, or proper execution on all computer systems.

Any use of trade names in this report is for descriptive purposes only and does not imply endorsement by the U.S. Geological Survey. This report is preliminary and has not been reviewed for conformity with the U.S. Geological Survey editorial standards.

1 USGS, Lakewood, Colorado
2 Consultant, USGS, P.O. Box 1103, Grand Junction, CO 81502

The software described in this report is available on ten 5.25 inch double-sided, double-density, soft-sectored diskettes, formatted in Microsoft DOS version 2.x. Four diskettes contain program source code written in FORTRAN 77, IBM BASICA, 8086 assembly language, and Turbo Pascal. Five diskettes contain executable program files of several of the source programs, including all the kriging programs. A tenth diskette contains a library of object modules of FORTRAN and assembly language subprograms. This library saves users the trouble of recompiling these modules prior to linking with calling programs. See following index.

****

NOTE: ALL THE EXECUTABLE MODULES ON THE RELEASE DISKS REQUIRE AN INTEL 8087 NUMERIC COPROCESSOR (80287 FOR AN IBM PC AT). THE PROGRAMS WILL NOT RUN IF THE COPROCESSOR IS NOT PRESENT.

****

Users may order the diskettes in two different ways:
1) all ten diskettes -source, library, and executable modules,
2) five diskettes only -source and library.

The programs on the diskettes are subject to modification as bugs are discovered and corrected, and as improvments in program function are made. Users wishing to keep current can contact the Branch of Resource Analysis concerning the status of these updates.

Executable modules available on STATPAC disks

Program Name	Bytes	Date	Time	Disk No.
ADDORDEL EXE	40002	7-27-86	2:56p	7 of 10
BACKSCRL COM	6528	1-01-80	3:25a	1 of 10
BASTAT EXE	61291	5-13-87	9:21a	7 of 10
BROWSE COM	960	3-13-86	1:46p	1 of 10
CARD2STP EXE	46491	6-03-87	1:24p	7 of 10
CHARCONV COM	11181	11-21-86	3:35p	2 of 10
CHARSORT EXE	48152	7-25-86	10:42a	10 of 10
CHR2STP EXE	29472	5-30-86	7:02a	10 of 10
COOR EXE	35580	7-30-86	9:05a	9 of 10
CURVES EXE	51358	12-08-86	2:23p	10 of 10
FILTER EXE	47178	9-10-86	7:46a	6 of 10
FIXIT EXE	39910	12-08-86	2:26p	5 of 10
GAPMAP EXE	53868	4-07-86	3:43p	6 of 10
GSCUTM EXE	40456	7-29-86	7:48p	8 of 10
KBUF256 COM	42	9-26-86	8:04a	2 of 10
KRIGINFO COM	1084	12-30-85	11:28a	2 of 10
LISTER EXE	36983	7-28-87	7:52a	4 of 10
MAP---$ EXE	30810	4-23-86	8:25a	7 of 10
MAPPLT EXE	42277	5-11-87	10:46a	6 of 10
NEWNAM COM	96	10-22-86	11:57a	3 of 10
NORCHI EXE	39046	4-08-86	3:24p	6 of 10
NORMST EXE	47324	12-08-86	2:34p	5 of 10
NUMLINES COM	10508	11-24-86	6:40a	3 of 10
PERALALK EXE	35038	7-30-86	8:36a	5 of 10
PERMUTER EXE	33162	7-29-86	8:12p	6 of 10
POLYGRET EXE	44208	7-24-86	11:22a	10 of 10
PTILES1 EXE	36120	7-30-86	8:41a	5 of 10
PUBLST EXE	39712	4-04-86	10:20a	9 of 10
RANGERET EXE	38306	7-25-86	2:05p	10 of 10
REGRESS EXE	64240	3-10-87	7:45a	7 of 10
ROWIDRET EXE	34409	5-13-87	2:35p	7 of 10
ROWPATCH EXE	50791	5-13-87	1:09p	10 of 10
SS2DBLOK EXE	67275	4-16-87	2:47p	9 of 10
SS2DCONT EXE	106407	7-09-87	2:30p	8 of 10
SS2DDRES EXE	61191	4-16-87	2:58p	9 of 10
SS2DGAMH EXE	51771	5-15-87	9:10a	6 of 10
SS2DGRID EXE	64107	4-15-87	3:04p	9 of 10
SS2DMENU EXE	33215	7-14-87	10:16a	7 of 10
SS2DOUTL EXE	60622	7-25-86	9:03a	8 of 10
SS2DPREP EXE	44282	7-25-86	7:35a	8 of 10
SS2DVCTL EXE	38304	7-25-86	7:42a	8 of 10
SS2DXVAL EXE	64229	4-15-87	2:36p	8 of 10
STP2DAT EXE	46576	4-03-87	1:58p	7 of 10
STPSORT EXE	41968	7-24-86	9:26a	6 of 10
STRNGRET EXE	39786	7-25-86	2:23p	9 of 10
TABRAN COM	12100	12-08-86	6:56a	4 of 10
TERFIL EXE	40720	7-25-86	11:08a	10 of 10
TRANSF EXE	50198	4-16-87	11:19a	9 of 10
UNFILTER EXE	43094	9-10-86	12:20p	6 of 10
XYPLOT EXE	30760	4-07-86	1:16p	10 of 10

Contents

Brief descriptions of STATPAC and related statistical programs

for the IBM Personal Computer

Programs written for or modified for the PC

by W. D. Grundy and A. T. Miesch

Last update July 30, 1987

Introduction

This README file describes the USGS STATPAC (STATistical PACkage) computer programs as implemented on the IBM PC. The programs include such applications as computing basic statistics, data transformations, matrix operations, curve-fitting, regression analysis, factor analysis, discriminan analysis, analysis of variance, and analysis of spatial statistics (kriging) Data entry, data cleanup, file handling programs and a few elementary graphics programs are also included.

Program source codes and several compiled executable modules are included on this diskette and other diskettes of this report. Example STATPAC data files (indicated by the extension name ".STP") can be used to gain familiarity with the programs and their application. Two text files, STPINFO.ONE and STPINFO.TWO, demonstrate the usage of several of the STATPAC programs. STPINFO.ONE is intended for getting a user started using the STATPAC system, and STPINFO.TWO for demonstrating kriging and ancillary STATPAC programs. These information files assume that the user is familiar with PC- or MS-DOS and the correct usage of statistical techniques of data analysis. Terms such as "redirection of IO", "variogram", "mean", and "variance" are used without explanation in these files.

Programs described in this report require the following hardware:

IBM PC, PC/XT, PC/AT or compatible computer, running
under PC- or MS-DOS, Version 2.0 or later;
At least 320K bytes of memory;
At least one double-sided, double-density disk drive;
IBM color graphics adapter, enhanced graphics adapter, or equivalent;
IBM or Epson-equivalent dot-matrix printer (for 10-inch paper or wider);
An 8087 (80287 in PC/AT) numeric coprocessor.

A second floppy disk drive and a fixed disk drive are highly desirable hardware additions.

A FORTRAN 77 compiler (Microsoft, version 3.30 or later) is required as not all the FORTRAN programs present on these release disks are ready-to-run execution modules. (The user may want to modify some of the programs.) BASIC programs are written in IBM BASIC, which is a

modification of BASIC by Microsoft. Interpreted BASIC tends to run slowly, but compiled BASIC programs (run modules created by IBM's BASIC compiler) run much faster. A BASIC compiler is a highly desirable software addition.

Some subprograms are written in 8086 assembly language and will require the use of a macroassembler if any modifications of their coding are needed. Object codes from all the assembly language source programs on these release disks are contained in the STATPAC.LIB library. The contents of file STATPAC.LIB are given in file STATPAC.LST. Files STATPAC.LIB and STATPAC.LST are contained in one of the ten release disks which accompany this report. The user should be aware that the FORTRAN subprograms were compiled into object modules by Microsoft FORTRAN version 3.31. Users of Microsoft's version 4.0 compiler should either recompile these subprograms, or configure the version 4.0 libraries for version 3.31 compatibility. The assembly language object modules can be used "as-is" in either version 3.31 or 4.0.

Readers who are not familiar with the FORTRAN 77 language will find D. D. McCracken's (1984) text to be very useful in learning the language. Scanlon's (1985) book on assembly language programming for the PC and XT is also very informative.

NOTE: If these programs and their subroutines are to be compiled with Microsoft FORTRAN V4.0, the code fragment STATUS='NEW' found in FORTRAN source code OPEN statements, must be changed to STATUS='UNKNOWN' or the programs will halt execution immediately upon attempting to open an already existing file.

General remarks on the STATPAC system

The USGS STATPAC system of statistical programs was developed over more than 20 years by a number of mathematicians and programmers. Some of the system was originally written in ALGOL for the Burroughs 220 and 5000 systems, and later translated into FORTRAN for the IBM 360/370, DEC10, and Honeywell Multics systems. The programs are used chiefly in the fields of applied geochemistry and petrology, but may be used with any data that can be logically arranged into a matrix wherein the rows represent samples, or observations, and the columns represent variables, or measurements. Each data value can be qualified by any one of six different single-character codes which are used in the field of applied geochemistry, to have the following meanings:

N Constituent not detected

L Less than the associated value. Constituent present but concentration is too low to be measured by the analytical method being used

T Trace of constituent is present

G Greater than the associated value. Concentration is too large to be measured by the analytical method being used

H Concentration not measured because of analytical interference

B Blank. No value reported

Some of the STATPAC programs ignore the qualifying codes, some (such as those for multivariate analysis) reject the data set if the codes are present, and others use the codes as appropriate for the computation being performed. The qualifying codes can be eliminated from the data set if desired (see program REPLAC), but all aspects of treatment of the qualified data are the responsibility of the program user.

All programs are in FORTRAN 77 except where noted in the program descriptions. Programs with names ending in the character "$" are in IBM BASIC and either accept or write out standard .CHR files (see below).

Data are generally entered into unformatted STATPAC files from card-image files using the programs TABFIL or CARFIL (cards to STATPAC file), or small data sets can be entered into a STATPAC file with program TERFIL (terminal to STATPAC file). Program TERFIL prints out directions for its use. If a large data set is to be entered, the user should consider creating a response file with a text editor, and running program TERFIL using DOS' I/O redirection feature. For example, if all the replies to TERFIL's prompts are previously stored in disk file RESPONSE, TERFIL can be automatically run by merely typing:

```
TERFIL<RESPONSE
```

An output file from any STATPAC FORTRAN program can be used as input to any other STATPAC FORTRAN program. Once the data set has been entered on a STATPAC file, it can be transformed and modified in any manner with program FILFIL (STATPAC file to STATPAC file). Both programs CARFIL and FILFIL must be modified, re-compiled, and re-linked for each new application.

Three sample STATPAC files are included on the release disks -- CLARKFE.STP, DAVIS.STP, AND DEMO.STP. They are described later in this report. They are intended for use in demonstrating the STATPAC programs and for experimentation by the users of these programs.

The program naming convention used here commonly includes 3-character identifiers with the following meanings:

CAR a card-image or ASCII file

FIL
or
STP an unformatted STATPAC file

TER the terminal or keyboard

CHR a standardized ASCII file used for input to certain BASIC-language programs (i.e. those with a "$" at the end of their names).

CMN an ASCII file used for data communications (i.e., uploading or downloading).

TAB an ASCII data file that contains two rows for each observation. See program TABFIL.

Some of the FORTRAN programs and subroutines use the metacommand "$INCLUDE xxxxx".INC. The file xxxxx.INC must be present when the program is compiled.

Some programs that contain more than a few subroutine segments can be linked by typing LINK @xxxxx.LNK, where xxxxx is the program name. See for example program RRMODE and RRMODE.LNK or MATRIX and MATRIX.LNK.

The source codes for the programs are contained on 5-1/4 inch double-sided floppy diskettes maintained on file in U.S. Geological Survey Public Information Offices. Users supplying their own diskettes are invited to use the micro-computer facilities available in these offices to make duplicates. Comments or questions concerning program content should be directed to: STATPAC, Branch of Resource Analysis, Box 25046, MS 937, DFC, Denver, CO 80225, U.S.A.

STATPAC data files

STATPAC data files on USGS minicomputers can be converted to ASCII character files with Data General STATPAC program STP2CMN(resident on the Data General Computer), then down-loaded to the PC via telephone lines and converted back to unformatted STATPAC files with PC program CMN2STP.

The standard ASCII files needed for input to the BASIC programs can be created with program STP2CHR (and converted back with CHR2STP) or can be produced with program TERFIL$ by typing data in at the keyboard.

STATPAC programs in FORTRAN generally include declarator statements similar to:

```
CHARACTER*4 ID(2),IVID(m,2),NAME(4)
CHARACTER*1 IA(m)
INTEGER*4 LOC(2)
DIMENSION X(m)
```

where "m" is the maximum number of variables (columns) allowed in the input data file.

Unformatted STATPAC data files contain a "header" record followed by N data records. The header record is written by the FORTRAN statement:

```
WRITE(KI) ID,N,M,(IVID(I,1),IVID(I,2),I=1,M)
```

and read by:

```
READ(KI) ID,N,M,(IVID(I,1),IVID(I,2),I=1,M)
```

where KI is the device number, ID is an 8-character data set identifier, N is the number of rows in the data matrix and M is the number of columns. IVID is an 8-character identifier for each variable (column).

Data records are written to and read from STATPAC files with the following FORTRAN statements, respectively:

```
CALL JKPTLS(KI,I,NAME,LOC,X,IA,M)

CALL JKGTLS(KI,J,NAME,LOC,X,IA,M,IERR)
```

where I is a sequential row-number (generally the index for a DO loop), J is the row-number from the data file, NAME is a 16-character name for an observation (sample identifier), LOC contains two integers (generally latitude and longitude of the sampling site), X contains the numeric data values, and IA contains the qualifying codes. The first 8 characters of NAME are referred to as the primary row ID and the second 8 characters form the secondary row ID. IERR is an "unexpected end-of-file" error flag from subroutine JKGTLS and should be zero.

As a convention, unformatted STATPAC data file names usually end with the extension .STP and ASCII character data files with standardized format for the BASIC programs have extensions of .CHR. Files created for

transfer of data via phone lines have the extension .CMN. (See programs STP2CHR, CHR2STP, STP2CMN and CMN2STP for creating one type of data file from another.)

Classification of programs

Data Entry

CARD2STP
CARFIL
TABFIL
TERFIL *
TERFIL$

File conversion

CHARCONV
CHR2STP
CMN2STP
DIG2STP
PACR2STP
STP2CHR
STP2CMN
STP2DAT
UGR2FCO
UGR2GRD

Graphics

FVDGRM$
MAP---$
MAPPLT *
STDGRM$
TERPLT
XYDGRM$
XYPLOT *

Spatial Statistics (Kriging)

SS2DPREP *
SS2DVCTL *
SS2DGAMH *
SS2DXVAL *
SS2DBLOK *
SS2DGRID *
SS2DDRES *
SS2DCONT *
SS2DOUTL *
SS2DMENU
KRIGINFO *

Sorting

CHARSORT
STPSORT *

Retrieval

CHARSRCH
POLYGRET *
RANGERET *
ROWIDRET *
STRNGRET *
UCODERET

Mathematics & Statistics

ANOV-1
ANOV-2
BASTAT *
COHEN
CURVES
D0060
GAPMAP *
LNCLUST
MATRIX
NORCHI *
NORMST
PARCOR
PERALALK
PTILES1
RANDOM$
REGRESS *
STATCOR *
TABRAN
TRANSF *

Utility

ADDORDEL *
BACKSCRL
BROWSE
COOR
DUPFINDR
FDA
FILFIL
FILONE
FILTER *
FINDRG
FIXIT
GSCUTM
GXFIXX
KBUF256
LISTER
NEWIDS
NEWLOC
NEWNAM
NUMLINES
PERMUTER *
PRINTT
PUBLST *
REPLAC
ROWPATCH *
SELVAR
STPIDS
TABQ
UNFILTER *
WHATQ

Factor Analysis

EQFILE
EQBFIT
EQMODE
EQSPIN
EQTEST
RRMODE

* indicates program is accessible from program SS2DMENU

ADDORDEL
Performs any one of the following functions:
1) Merge two STATPAC data files columwise. The two files must have the same number of rows.
2) Merge two STATPAC data files rowwise. The two files must have the same number of columns.
3) Delete one or more sequences of rows from a STATPAC file.
4) Delete one or more sequences of columns from a STATPAC file.

See source code of ADDORDEL.FOR for names of required links.
After compiling ADDORDEL.FOR, link by typing:
LINK ADDORDEL /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN

ANOV-1 (formerly program D0038)
One-way nested (hierarchical) analysis of variance. Up to 10 levels may be nested and the design may be unbalanced. Variance component estimates may be written on a STATPAC output file. Non-significant mean-square estimates are pooled for revised significance tests. Output is directed to the printer which must provide 132 characters per line. The first NL variables in the STATPAC data file, where NL is the number of nested levels, must be indices that specify the position of the observation in the ANOV design. Program requires IBM- or Epson-compatible printer.
Link with: IOS

ANOV-2
Two-way analysis of variance with or without within-cell replications. Data on the STATPAC file must be ordered replications-columns-rows. That is, all replications for col.1,row1 -then all replications for col.1,row2 -then col.1,row3, etc. until col.1,row"n". The data for col.2,row1 then follow. There may be up to 35 variables on the STATPAC input file with up to 100 columns and 12 rows in the ANOV design. F-ratios are computed using the error term in the denominator and again using the pooled error and interaction terms. The latter F-ratio is to be used only where the interaction variance is not significant.
Link with: IOS

BACKSCRL.COM
A ram resident program which allows the user to review information which was scrolled off the screen. Type out file BACKSCRL.DOC for complete information.
NOTE: BACKSCRL may not function correctly if device ANSI.SYS is installed.

BASTAT
Computes basic statistics on a STATPAC data set and constructs horizontal bar-type histograms. Can also be used to create a new STATPAC file with the arrangement of the rows and variables changed according to the user's specifications.
After compiling BASTAT.FOR, link by typing:
LINK BASTAT /EXEPACK /SEGMENTS:256,,NUL,C:\COMPILER\STATPAC+
C:\COMPILER\8087+C:\COMPILER\FORTRAN

See source code of BASTAT.FOR for names of required links.

BROWSE

Program displays an ASCII file page-by-page or line-by-line, forward or backward. To invoke BROWSE, just type:

BROWSE filespec [/w]

The optional /w is used if the file to be scanned is a WORDSTAR file. Use PgDn and PgUp to move through the file page-wise or the up- and down-arrow keys to move a line at a time. The right- and left- arrow keys may be used for horizontal scrolling of the text. The Home and End keys move to the top or bottom of the file on display. To exit BROWSE, press the Esc key or type Ctrl-Break. Program was written by Charles Petzold (1986).

NOTE: BROWSE writes directly to video ram.

CARD2STP

CARD_image_TO_STatPac

Converts an ASCII data file with records of less than 81 characters into a STATPAC file. Program allows the user to build an input format by describing the location of the ASCII fields and their widths at run time. Read the comments in the prologue to the FORTRAN source code for program CARD2STP for more information. This program must be compiled with Microsoft FORTRAN version 4.0 as follows:

```
FL /c /FPi87 card2stp.for
link card2stp /exepack,,nul,c:\lib\statpac4+c:\lib\llibfor7
```

Where statpac4 is the version 4.0 STATPAC library.

CARFIL

Produces a STATPAC file from a card-image file. CARFIL must be modified and recompiled for each new card-image format.

Link with: IOS

CHARCONV

Converts an ASCII file from lower case characters (a - z) to upper case characters (A - Z), or from upper case characters to lower case characters. The source code for this program is written in Turbo Pascal. The compiled run module, CHARCONV.COM, is stored on one of the release disks.

CHARSORT

This program sorts a character file of 3000 lines or less in either ascending or descending order. No line can exceed a length of 254 characters. Lengths of text lines need not be uniform. The user can select from one to ten sort fields of varying widths in any desired sequence. The sum of the widths of the selected sort fields cannot exceed 48 characters. The sorted records are written on a character output file. Trailing blanks are stripped off each record. For files of less than 64k bytes and only one field to be sorted, the user may prefer to use the DOS-2 sort instead.

After compiling CHARSORT.FOR, link by typing:

```
LINK CHARSORT /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
```

See source code of CHARSORT.FOR for names of required links.

CHARSRCH
Searches a character file with lines of less than 255 characters each for the occurrence of a user-specified substring of less than 128 characters. Wild cards are permitted in the substring being sought. Retrieved records are written on a character output file. (Differs from DOS-2 "find" command in allowing user to specify the portion of the record to be searched). After compiling, link by typing:
LINK CHARSRCH /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of CHARSRCH.FOR for names of required links.

CHR2STP
To produce a binary STATPAC file from a standard-formatted ASCII file. See program STP2CHR for standard format.
Link with: IOS

CLARKFE.STP
This is a demonstration STATPAC data file derived from the iron data of Clark (1979, p. 95). It is used to demonstrate the kriging programs (see file STPINFO.TWO). The data set consists of fifty artificial sample points made up of an x-coordinate, a y-coordinate, and a percent iron.

CLARKTER.FIL
An ASCII "response" file which can be used by program TERFIL (see below) to create a STATPAC file identical to CLARKFE.STP (see above). Directions for using CLARKTER.FIL are given in the discussion of program TERFIL.

CMN2STP
To produce a binary STATPAC file from a .CMN (communications) file. See program STP2CMN.
Link with: IOS

COHEN
Used to derive means and standard deviations and/or geometric means and geometric deviations. If a variable is singly censored, the method of Cohen (1959) is used to estimate these values. Variables qualified with N, L, and/or T are censored at the low end of the distribution. Variables qualified with G are censored at the high end. No computations are performed if the variable is doubly censored (i.e., censored at both the low and high ends) or if the censoring points are variable (e.g. the presence of 10L and 20L) or if fewer than two values for a variable are unqualified. The validity of estimates by Cohen's method depends on the validity of the assumption that the data were drawn from normal or lognormal distributions. Frequently, in situations where the variable is severely censored, Cohen's method will give wild answers that are obviously incorrect.
See source program COHEN.FOR for names of required subprograms.
After compiling COHEN.FOR, link by typing:
LINK COHEN /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN

COOR
Transfers the values of latitude and longitude stored as integer values

in the 'LOC' array of a STATPAC row into corresponding STATPAC variables which are expressed in decimal degrees. The STATPAC rows containing the additional columns are written on a STATPAC output file. On option, program converts decimal degrees latitude and longitude stored as STATPAC variables into integer degrees, minutes, and seconds which are then stored in the 'LOC' array. A slight loss of precision is incurred when integer degrees, minutes, and seconds are converted into four-byte real variables containing decimal degrees.
After compiling COOR.FOR, link by typing:
LINK COOR /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\STATPAC
See source code of COOR.FOR for names of required links.

CURVES
To fit any of seven different functions to two variables by least squares. Observed, computed, and residual values may be printed, along with the 95% confidence intervals.
Link with: IOS, MATKLO

D0060
Performs a two group discriminant analysis. The two groups are the corresponding data matrices of two STATPAC data sets on a disk. The columns in the matrix contain the variables used in the analysis. An option is provided to allow deletion of variables estimated to be least significant, using a technique described by Davis and Sampson (1966). After deletion of the least significant variable, the discriminant function is recomputed. The stepwise action terminates when only one discriminator variable is left in the function.

The output consists of identifying information, means, product-cross product matrices for each data group, a combined sum product-cross product matrix, a constant vector, coefficients of the discriminant function, discriminant values for each row of the data set, average discriminant values for each data set, number of rows misclassified in each data set, critical value of the discriminant, Mahalanobis' D-squared, and F-value of D-squared, and estimates of the relative contribution of each variable to the discriminant. The basic method of computation is from Bennett and Franklin (1954).
Link with: ASCIIZ, BYTMOV, CHRKNT, CKSAFE, CONCAT, INT21H, IOFS, IOS, ISEG, SOUND

DAVIS.STP
A STATPAC demonstration file consisting of 52 sample points with xy coordinates and an elevation value. Derived from Davis (1973, pages 312-314) and intended for use in demonstrating universal kriging programs (see file STPINFO.TWO).

DEMO.STP
A STATPAC file consisting of fifty samples with ppm Se, Hg, and As values taken from the San Joaquin Valley, California. Intended for demonstrating STATPAC programs including those which obtain UTM coordinates from latitude-longitude values stored in the LOC array. (See file STPINFO.ONE).

DIG2STP
Converts an ASCII file created by program JKDIGIT into an unformatted STATPAC file of latitudes and longitudes and optional UTM coordinates. Program JKDIGIT was written by J. O. Kork (1986).
After compiling DIG2STP.FOR, link by typing:
LINK DIG2STP /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See comments in source code of DIG2STP.FOR for names of required links.

DUPFINDR
Writes to a character output file the row number and primary ID of all sets of data points which are within a user-specified distance of one another. This specified distance and the variable numbers of the XY coordinates are requested at run time.
Link with: IOS, FSORT, SOUND, TICKER

****

Programs in the "EQ" series are for "extended Q-mode factor analysis of compositional data" (see Miesch, 1976a, 1981). They are used principally for the development of petrologic mixing/unmixing (differentiation) models. It is required that the variables for each observation (row) in the STATPAC data matrix sum precisely to some constant, generally 100 percent. The principal advantages of "extended Q-mode factor analysis over conventional methods of factor analysis are that 1) the factor axes, rather than being determined solely by mathematics, can be selected to represent real or hypothetical geologic materials, 2) the factor scores, rather than being in units of transformed and normalized data, are in the same units as the compositional data, and 3) the factor loadings sum to unity for each sample and can be interpreted directly as mixing/unmixing proportions.

Program EQMODE is always run first. This program creates a binary MASTER or E97 file that can be inspected with program EQFILE. All of the other programs in the "EQ" series read this MASTER(E97) file. Seven of the "EQ" series program are described in Miesch (1976b).

EQBFIT
Used to derive "badness-of-fit" measures of samples to a factor model as represented in a MASTER(E97) file. Badness-of-fit, like the communality, depends only on the number of factors, and is computed as the residual/original value. The overall badness-of-fit measure for each sample is the quadratic mean of the measures for each variable.

EQFILE
Used to read a MASTER(E97) file created with program EQMODE and to compute varimax loadings and scores. Output may be directed to the screen or to a character output file. Output data may include any of the following: raw data, maximum and minimum values for each variable, principal component loadings, varimax loadings, communalities, principal component scores, varimax scores, or varimax transformation matrices.

EQMODE
All of the programs in the "EQ" series are for extended Q-mode factor analysis of compositional data that sum to a constant (generally 100%) for each observation (sample). Program EQMODE creates a file (named "E97") that contains all the basic information needed as input by the other programs in the series.
Link with: TRANS, EIGEN, VARMAX2, IOS

EQSPIN
Used to rotate a 3-dimensional vector system for display on a stereographic-type diagram. System may be rotated in a specified direction or so that the 1st principal component axis is vertical. May also be used to plot miscellaneous samples on the stereogram and to outline the area within the stereogram wherein all plotted and hypothetical vectors represent non-negative compositions. Compositions represented by 37 points along this boundary are written out on a character file named "COMP.LST". The 2nd and 3rd loadings for the rotated sample vectors and the 37 boundary points can be written on a STATPAC file. This may be converted to a .CHR file for screen display of the stereogram with program STDGRM$.
Link with: MATKLO, EXPAND, IOS

EQTEST
Used to test samples represented in a STATPAC data file for fit to a factor model represented in an existing MASTER(E97) file. The samples may or may not be those used to develop the model. Also derives coordinates of samples for plotting on a stereogram if a 3-factor model is being used. STATPAC output file contains the sample communalities and initial varimax loadings.
Link with: IOS

****

FDA.BAS
Program to create FDA.COM, which will search across subdirectories of both the hard disk and floppies for any files and report the path, size, and time of creation of any of its finds. Can be used with wildcard filespecs. Program was written by E. J. Muth (1986).

FINDRG
Used to Search a STATPAC data file looking for samples that contain values for selected variables within specified ranges.
Link with: IOS

FILONE
To read the header information (N, M, Data set ID, variable ID's) and the first observation in a STATPAC data set. This function is also performed by program FILTER (see below).
Link with: IOS

FILTER
To read the header information (N, M, Data set ID, variable ID's) and selected observations in a STATPAC data set. The ASCII representation of the STATPAC file is written on a disk file which can be printed or

edited. See program UNFILTER below. Output from FILTER can be directed to the printer or to the screen instead of to a disk file. See comments in source code for names of required links.
After compiling FILTER.FOR, link by typing:
LINK FILTER /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN

FILFIL
To read a STATPAC file, make data transformations, and write out a new STATPAC file. FILFIL must be modified and recompiled for each new application.
Link with: IOS

FIXIT
Changes selected values and/or qualifying codes in a STATPAC data set. This function is also performed by program ROWPATCH (See below).
Link with: IOS

FVDGRM$
A BASIC program for screen display of factor-variance diagrams from a .CHR file. The .CHR files are created with program STP2CHR or STP2DAT using output files from programs RRMODE or EQMODE.

GAPMAP
Searches a STATPAC file for statistically significant discontinuities in each variable. Data for each variable are first "normalized" using a 3-parameter log transformation, and then standardized to zero mean and unit standard deviation. The data are then ordered and the differences (gaps) between adjacent values are examined. The program prints the position of the largest gap and information regarding the 3-parameter lognormal transformation and its effectiveness. The STATPAC file must contain raw data - not logarithms. Qualified data may be present, but variables are skipped if the distribution is censored by more than 10% at both tails, or more than 50% at the low tail. Values qualified with N, L, T, or G are replaced by appropriate values for the computations. Values qualified with H or B are replaced by the average of the unqualified values for that variable. If xy-type coordinates are present in the data, the program will optionally divide the data into as many as five classes according to the four most significant gaps and print a map on the EPSON printer showing the relative geographic positions of the samples in each class. (Read comments in source code).
Link with: IOS

GSCUTM
Computes UTM north and east coordinates (expressed as kilometers rather than meters) using the STATPAC decimal degrees latitude and longitude variables. The UTM coordinates are written on a STATPAC output file along with the previously existing variables.
Link with: IOS

GXFIXX
To adjust each of the variables in a STATPAC data set to a common lower and/or upper cut-off value (i.e., lower and/or upper limits). This is to avoid the problem of having values of, say, 10L and 20L (or 100G and 200G) for the same variable. GXFIXX selects the lower and/or upper limit

that will require the fewest changes in order to make the adjust ments.
Writes a STATPAC output file containing the adjusted data.
Link with: IOS

JKGTLS
Subroutine to read data records from an unformatted STATPAC data file.
Contained in file IOS.FOR

JKPTLS
Subroutine to write data records to an unformatted STATPAC data file.
Contained in file IOS.FOR

KBUF256
Expands PC's keyboard buffer to 256 characters, giving a greater "type-ahead" capability. The new keyboard buffer is stored in that area of ram which is used for ram-resident programs, and is thus protected from overwrite. Do not invoke after Sidekick has been loaded.

LIBRARY.DOC
This file contains descriptions of the functions and usages of FORTRAN callable object modules contained in file STATPAC.LIB.

LISTER
Reads an ASCII text file and prints it onto a dot matrix printer at 55 lines per page. See comments in source code for directions on use.
After compiling lister.for, link by typing:
LINK LISTER /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\STATPAC
See scource code of LISTER.FOR for names or required links.

LNCLUST
Divides an ordered array of values for any selected STATPAC variable into K groups (K is specified by the user) so that the within-group variances are as small as possible. LNCLUST is from R.G. Garrett, Geological Survey of Canada. References to the mathematical techniques are given in the program comments in the source listing. Program requires an IBM- or Epson-compatible printer.
NOTE: PROGRAM EXECUTION TIME CAN BE EXCESSIVELY LONG FOR K GREATER THAN 3.
Link with: IOS, RSORT

MAP---$
Plots a "symbol" map on the IBM or EPSON printer. As many as seven symbols may be used for each map. User specifies the boundaries of the map so that the map may represent the entire area covered by the data set or only a portion of the area. The symbols may be defined by the user or left to the program.
Link with: IOS

MAPPLT
Reads in a STATPAC file and produces a file for printer plotting of a map of any selected variable. The x-y coordinates may have any origin. The map shows numbers ranging from 0-9 that specify the decile of the total range in which the mapped value occurs at

each x-y point on the map. If more than one observation occurs at any x-y location, the decile indicates their arithmetic average. On option, another map may be printed which shows the numbers of values averaged at each location.
Link with: IOS

MATRIX
Used to perform matrix computations on small matrices (up to 10x10). MATRIX is generally used for experimental work in matrix manipulations.
Link with: MATRIX1, EIGEN, MATKLO, IOS

NEWIDS
Writes a revised STATPAC file with modified data set ID, variable IDs, and/or sample IDs.
Link with: IOS

NEWLOC
Writes a revised STATPAC data file with new values for the 'LOC' array which are entered at the terminal. This function is also performed by program ROWPATCH (see below).
Link with: IOS

NEWNAM
A "COM" file which uses DOS function 86 to rename a file. This rename function can be used as a copy and delete command. For example the command: NEWNAM FILE1.DAT \SAVE\FILE1.DAT will place file1.dat in the directory SAVE and simultaneously remove it from the default directory. The syntax of the command is:

newnam [d][path]filespec [d][path]filespec

NEWNAM will not move a file from one disk drive to another, nor will it "clobber" an already existing file name.

NORCHI
Performs chi-square test for normality of a selected STATPAC variable. Program computes mean, variance, skewness, and kurtosis of the selected variable, a table of actual number of observations versus theoretical number of observations by class interval, the total chi-square value, the degrees of freedom (number of classes used minus three), and the probability of exceeding the computed value of chi-square. Program prints out to the PC's screen. As an option, the test can be performed on the distribution of natural logarithms of the STATPAC variable, provided it has no zero or negative values.
Link with: IOS

NORMST
Reads a STATPAC file of petrochemical data and creates a new STATPAC file of CIPW norms.
Link with: CIPW, IOS, UPRCASE.

NUMLINES
Counts the number of lines in an ASCII file and displays that count on the screen. The execution syntax of the program is:

numlines [d][path]filespec

where the drive and path are optional.
Non-ascii files yield spurious line counts.

The source program of NUMLINES is written in Turbo Pascal.

PACR2STP
Converts a specially formatted character file retrieved from a USGS Branch of Coal Geology PACER file into a STATPAC output file. Comments in the source code give instructions for its use.
Link with: IOS

PARCOR
Used to derive partial correlation coefficients. User supplies R12, R13, and R23. Program displays R12.3 (the correlation between variables 1 and 2 with variable 3 held constant.

PERALALK
Reads a STATPAC file of petrochemical data, computes the peraluminous and peralkaline indices, and writes these indices on a STATPAC output file. The peraluminous index is Al2O3/(Na2O+K2O+CaO) and the peralkaline index is (Na2O+K2O)/Al2O3, where all oxides for both indices are in molecular percents. After compiling PERALALK.FOR link by typing:
LINK PERALALK /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN.
See source code of PERALALK.FOR for names of required links.

PERMUTER
Rearranges (permutes) rows of a STATPAC input file in a random order. The random order will differ each time the program is run because the random number generator is initially seeded by the PC's system clock. The permuted rows are written on a STATPAC output file.
Link with: ASCIIZ, BYTMOV, CHRKNT, CKSAFE, CONCAT, FSORT, INT21H, IOFS, IOS, ISEG, RAND, SOUND, TICKER

POLYGRET
Retrieves from a STATPAC input file those samples whose coordinate values lie within a polygonal boundary. The user can select 1) only those samples which lie within or on the boundary, or 2) only those samples which lie outside the boundary. The retrieved sample-rows are written on a STATPAC output file. Nodes of the boundary are stored on a previously created character disk file which can be written by the EDLIN editor. Each node of the boundary is typed on a single line with the north coordinate (decimal degrees latitude or other vertical coordinate) typed first and the east coordinate typed second. The north and east coordinates must be separated by at least one blank space or by a comma. Embedded blanks in the north or east coordinates are not permitted. The boundary nodes must be entered in a clockwise direction.
Link with: AREA, ASCIIZ, BYTMOV, CHRKNT, CKSAFE, CONCAT, INT21H, IOFS, ISEG, PTLOC, SOUND

PRINTT
Prints the entire STATPAC file. A similar function is performed by program FILTER (q.v.).
Link with: IOS

PTILES1
Computes percentiles of all the variables in a STATPAC input file and tabulates them in a character output file.
Link with: ASCIIZ, BYTMOV, CHRKNT, CKSAFE, CONCAT, INT21H, IOFS, IOS, ISEG, RSORT, SOUND, TICKER

PUBLST
Program creates a fully labelled ASCII printer copy of a STATPAC input file. Output is suitable for use in manuscripts.
Link with: IOS

RANDOM$
A BASIC program for writing a character (.CHR) file of uncorrelated normal deviates. The file may contain up to 99 columns simulating uncorrelated variables and any number of rows.

RANGERET
Retrieves from a STATPAC file those rows in which a specified variable lies within limits which are also specified. Creates a STATPAC output file of the selected observations. After compiling RANGERET.FOR, link by typing:
LINK RANGERET /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of RANGERET.FOR for names of required links.

REGRESS (formerly program D0094)
Performs a stepwise multiple linear regression on a STATPAC file and computes related regression statistics. The basic method is that of Efroymson (1960).

The terms included in the derived regression equation are those which are significant at a prescribed confidence level. The observed, computed, and residual values of the dependent variable, and other selected variables, may be written to a STATPAC output file.

The user specifies a probability level, Q, as input for adding variables or deleting them from the regression equation. Q is usually entered as .05 or .01. If all selected variables are to be retained (i.e., stepwise regression is not desired), then Q is assigned a value of 1.0. Variables are added to the regression equation in the following manner: Each independent variable is tested for the proportion of the total sum of squares in the dependent variable that it might explain. The most significant variable in that respect is entered into the regression equation. Because the significance of an independent variable in the equation changes with the addition of each new variable, each variable in the equation is tested upon the addition of a new variable, and any variable shown to be no longer significant is deleted from the equation. If a variable is repeatedly added and deleted from the equation, that variable is permanently eliminated from the equation.

The variables to be treated can be controlled by selecting columns. The first selected column will represent the dependent variable. The independent variables are selected next and can be in any order. If a qualified value is encountered for any selected variable of any selected

row, the program deletes that row from the data set to be analyzed and reduces the total number of rows by one.

Regression residuals may be computed and printed and optionally written to a STATPAC file. Other printer output includes the following:

1. Verification of all input parameters.
2. Means and standard deviations of all selected variables
3. Correlation matrix for all selected variables - optional.
4. Record of sequence in which variables are added to and deleted from the regression equation.
5. Regression constant and coefficients.
6. Multiple correlation coefficient.
7. Sum of squares of the regression residuals.
8. Partial correlation coefficients between dependent variable and variables in the regression equation.
9. Regression weights (standardized partial regression coefficients).
10. Standard error of regression weights.
11. Standard errors of regression constant and coefficients.
12. Standard error of estimate of the dependent variable.
13. Percent of the total sum of squares of the dependent variable explained by the regression equation.
14. Listing of the observed, computed, and residual values of the dependent variable - optional.

This program requires an Epson- or IBM-compatible dot matrix printer.

Link with: ASCIIZ, BYTMOV, CHRKNT, CKSAFE, CONCAT, INT21H, IOFS, IOS, ISEG, SOUND

REPLAC

Used to replace (remove) qualifying codes in a STATPAC data file. Values of zero coded by N, L, or T are replaced by values supplied by the user. Then all values coded by N, L, or T are multiplied by 0.7 and the codes are removed. Values of zero coded by G are replaced by values supplied by the user. Then all values coded by G are divided by 0.7 and the G is removed. Values coded by H or B are replaced with the average uncoded value for the variable, and the codes are removed. It is recommended that the data be processed through program GXFIXX before REPLAC is used.
Link with: IOS

ROWIDRET

Retrieves from a STATPAC input file those rows whose primary or secondary IDs (either but not both) are stored on a character file which was previously created by the user. The user is given the option of retrieving 1) only those rows which are on this file, or 2) only those rows which are not. The retrieved rows are written on a STATPAC output file. The disk file containing the primary or secondary row IDs can be created by the EDLIN editor. The individual row IDs (either primary or secondary) are entered in FORTRAN format(a8) in columns 1-8, with only one row ID per line. Spelling of the row IDs must agree exactly with their spelling on the STATPAC input file. Leading and embedded blanks in the row IDs must be typed in exactly the same position

as on the STATPAC file. Trailing blanks, however, need not be typed. A retrieval based on both the primary and secondary row IDs will require two runs of the program with separate disk files of the respective row IDs.
Link with: ASCIIZ, BYTMOV, CHRKNT, CKSAFE, CONCAT, ICSRCH, IOFS, ISEG, SOUND

ROWPATCH
Edits the rowid, loc, x, and qualifying code arrays in the rows of a STATPAC file. Combines functions of programs NEWIDS, NEWLOC, and FIXIT. Displays instructions for use on the PC's screen. Can also be used to search through a STATPAC file without editing. Requires IBM color graphics adapter and is best run using a hard disk drive. After compiling ROWPATCH.FOR, link by typing:
LINK ROWPATCH /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of ROWPATCH.FOR for names of required links.

RRMODE
For R-mode factor analysis. Writes STATPAC output files of factor loadings and scores for four kinds of factor models (Principal Components, Varimax, Oblique model based on the extreme samples, and Oblique model based on the extreme variables). Output files of the recomputed data, the residuals, and/or badness-of-fit measures for the samples. Another output file contains the variable communalities. Coefficients used for computing the scores are also printed. RRMODE and nine of its subroutines may be compiled with batch file RRCOMP, by typing "RRCOMP". The file "RRMODE.INC" must be present when compiling in order to satisfy the INCLUDE: RRMODE.INC metacommand which is present in all ten RRxxxx files.
If IOS.OBJ is present, RRMODE can be linked after compilation by typing "LINK @RRMODE.LNK".
Link with: RRPCOM, RREIGN, RRFILE, RRVRMX, RRVROT, RROBLS, RROBLV, RRMATK, RRCOEF, and IOS

SELVAR
Rearranges the variables of a STATPAC file in a sequence specified by the user. This function is also performed by program BASTAT (See above).
Link with: IOS

****

Spatial Statistics in 2-Dimensions (SS2D)

Programs in the "SS2D" series implement the geostatistical method of estimation of geological variables that is known as kriging. STATPAC input data files must have cartesian coordinates in the x-y plane for each sample point, and no two sample points can have identical x-y coordinates. The variable to be estimated must have certain geological and statistical properties (Journel and Huijbregts, 1978, pp 196-207). The variable may have some qualified values, but qualified values will be bypassed and not included in the computations. Programs SS2DXVAL, SS2DGRID, and SS2DBLOK have options for using either ordinary or universa

kriging, the spherical, linear-root-parabolic, exponential, gaussian, and cubic semivariogram models, nested semivariograms, and geometric anisotropy. The error variance of the observed value of the geological

variable at each sample point, if known, can be included in the kriging. Kriging is done by the moving neighborhood technique, and the user specifies the search radius which defines the neighborhood size. Two include files, "PTKRIG.INC" and "BKKRIG.INC", must be present on disk when these programs and their subroutines are compiled.

The programs have restrictions on the use of extensions to the input and output file names. For details read the comments in the source code of these programs. This can be done for each program below by merely typing a line similar to the following examples:

```
FIND "C***" SS2DPREP.FOR|MORE
FIND "C***" SS2DVCTL.FOR|MORE
FIND "C***" SS2DGAMH.FOR|MORE
etc.
```

SS2DBLOK

Program computes the mean value and kriging variance of a geologic variable within a block of ground of any specified shape. Program produces a character output file containing the computed area of the block, the block variance, the kriged mean of the geologic variable, and the kriging variance of that variable, and optionally, a list of the sample values used in the estimation together with their kriging weights. Coordinates of the nodes defining the block can be entered from the terminal at run time, or can be stored on a character input file created by, say, the EDLIN editor. The north- and east-coordinates for each node are entered on a single line, and are separated by a space or comma. They must be entered in a clockwise direction.
After compiling SS2DBLOK.FOR, link by typing:

```
LINK SS2DBLOK /EXEPACK /SEGMENTS:256,,NUL,C:\COMPILER\STATPAC+
C:\COMPILER\8087+C:\COMPILER\FORTRAN
```

See source code of SS2DBLOK.FOR for names of required links.

SS2DDRES

Program computes drift residuals using method of generalized least squares. The user must enter the parameters of an assumed variogram and the exponents of the terms of an assumed form of drift. This program offers a method of developing a suitable variogram/drift model for non-stationary geological variables by trial-and-error iterations. Program produces a STATPAC output file of the computed drift residuals of the data points and their coordinates. This data file can be processe by program SS2DGAMH. The STATPAC input file for SS2DDRES must be pre-processed by SS2DPREP.
After compiling SS2DDRES.FOR, link by typing:

```
LINK SS2DDRES /EXEPACK /SEGMENTS:256,,NUL,C:\COMPILER\STATPAC+
C:\COMPILER\8087+C:\COMPILER\FORTRAN
```

See source code of SS2DDRES.FOR for names of required links.

SS2DGAMH

Computes semivariograms (gamma(h)) in one to five directions and produces a character file for printer plotting of the graphs. Program allows up to twenty distance class intervals, optional logarithmic

transformations of the geologic variable, and optionally creates a "punchout" character file for use by high-resolution screen plotting programs (such programs have not been implemented on the IBM PC, but the punchout file can be uploaded to the multics mainframe for producing these plots). SS2DGAMH requires a character control file (see SS2DVCTL below) and a preprocessed STATPAC input file (see SS2DPREP below). After compiling SS2DGAMH.FOR, link by typing:
LINK SS2DGAMH /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of SS2DGAMH.FOR for names of required links.

SS2DGRID
Program estimates values of a geological variable along a grid of points specified by the user. Program produces a STATPAC output file containing the north- and east-coordinates of each grid point, the kriged value, and the kriging variance. Program also produces a character output file containing the same information, but including a header displaying the user's variogram and grid specifications. STATPAC input file to this program must be preprocessed by SS2DPREP. After compiling SS2DGRID.FOR, link by typing:
LINK SS2DGRID /EXEPACK /SEGMENTS:256,,NUL,C:\COMPILER\STATPAC+
C:\COMPILER\8087+C:\COMPILER\FORTRAN
See source code of SS2DGRID.FOR for names of required links.

SS2DCONT
Creates a contour map of a kriged variable in high resolution on the PC's screen. Requires a STATPAC input file created by program SS2DGRID and the IBM color graphics adapter. Maximum number of grid points allowed is 10000. After compiling SS2DCONT, link by typing:
LINK SS2DCONT /EXEPACK /SEGMENTS:256,,NUL,C:\COMPILER\GRAFMS2+
C:\COMPILER\STATPAC+C:\COMPILER\8087+C:\COMPILER\FORTRAN.
Note: File GRAFMS2.LIB is a copyrighted file sold by Microcompatibles, Inc. This library is not included on these release disks. Instead, an already compiled and linked execution module of SS2DCONT is supplied on the release disks. See source code of SS2DCONT.FOR for names of required links.

KRIGINFO.COM
Program lists names of SS2D series programs on the screen together with a brief description of their function. The ordering of the list of programs is the usual sequence in which they will be executed during the course of a kriging study.

SS2DPREP
Preprocessing program which produces a STATPAC output file which is used as input to programs SS2DGAMH, SS2DGRID, and SS2DDRES. SS2DPREP sorts the rows of a STATPAC input file in ascending order of north- and east-coordinates, checks for duplicated data coordinate pairs, and produces the sorted STATPAC output file if no duplicates are found. The STATPAC input file must have the extension ".STP" and the preprocesse output file will have the root name given by the user with the program-assigned extension ".UPR" (Unformatted PReprocessed). The other "SS2D" programs require the extension .UPR and do not allow the user to assign any other extension name. After compiling SS2DPREP.FOR, link by typing:
LINK SS2DPREP /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+

C:\COMPILER\FORTRAN
See source code of SS2DPREP.FOR for names of required links.

SS2DMENU
A driver program which allows the user to select any one of 29 different programs for immediate execution. Requires KRIGINFO.COM. To execute, merely type SS2DMENU. SS2DMENU will first load and execute program KRIGINFO.COM, and then request the user to press the F9 key to display a menu of programs used in kriging (and useful in many other statistical applications). Pressing the F9 key causes the menu to appear on the screen. SS2DMENU, when the proper sequence of user key-presses is made, will load a selected program and cause it to execute. When execution of the loaded program terminates (either normally or abnormally), control reverts to SS2DMENU, and the user can make a different selection.
NOTE: A copy of DOS file COMMAND.COM must be present on the disk drive from which the computer was booted.
After compiling SS2DMENU.FOR, link by typing:
LINK SS2DMENU /EXEPACK,,NUL,C:\COMPILER\CEXEC+C:\COMPILER\STATPAC+
C:\COMPILER\8087+C:\COMPILER\FORTRAN
(CEXEC.LIB is supplied with Microsoft's V3.3 FORTRAN compiler.)
See source code of SS2DMENU.FOR for names of required links.

SS2DVCTL
Program creates a character control file for use by SS2DGAMH. The user is prompted for the distance class intervals for the variogram run, directions in which the variograms are to be computed, and the allowable angular deviations about the selected directions. The user may also select an option which will cause program SS2DGAMH to create a variogram plotfile suitable for high resolution plotting on the PC's screen. After compiling SS2DVCTL.FOR, link by typing:
LINK SS2DVCTL /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of SS2DVCTL.FOR for names of required links.

SS2DXVAL
Program is used to cross-validate semivariogram models to determine the suitability of the theoretical semivariogram model for use in programs SS2DBLOK, and SS2DGRID. Program computes kriged value and kriging standard deviation for each data point, using neighboring data points. Observed kriging errors are computed and statistically summarized. Program produces a STATPAC output file of the results for each data point and a character output file containing the same information, but including also a heading record of the users specifications of parameters for the cross-validation. After compiling SS2DXVAL.FOR, link by typing:
LINK SS2DXVAL /EXEPACK /SEGMENTS:256,,NUL,C:\COMPILER\STATPAC+
C:\COMPILER\8087+C:\COMPILER\FORTRAN
See source code of SS2DXVAL.FOR for names of required links.

SS2DOUTL
Searches for outliers in STATPAC data set using method of Krige and Magri (1982). Requires preliminary estimate of a variogram of the

variable.
After compiling SS2DOUTL.FOR, link by typing:
LINK SS2DOUTL /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of SS2DOUTL.FOR for names of required links.

****

STATCOR
Derives means, variances, ranges, and correlations for variables in a STATPAC data file. Qualified values are included in the computations. These statistics are also computed by program BASTAT (See above).
After compiling STATCOR.FOR, link by typing
LINK STATCOR /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See comments in source code of STATCOR.FOR for names of required links.

STATPAC.LIB
This file contains object modules of FORTRAN and assembly language subprograms which are required by many of the programs described in this report. The FORTRAN object modules were compiled using Microsoft's Version 3.31 FORTRAN compiler. Assembly language source code was assembled to FORTRAN callable object files using Microsoft's Version 3.0 macroassembler.

STATPAC.LST
This file contains a listing of the object modules contained in file STATPAC.LIB.

STDGRM$
A BASIC program for screen plotting a stereographic-type diagram from a .CHR file made from a .STP file (using program STP2CHR or STP2DAT) derived as output from program EQSPIN.

STP2CHR
Produces a standard-formatted ASCII file (to use as input to BASIC programs) from a unformatted STATPAC file. This function is also performed by program STP2DAT (see below).
Link with: IOS

STP2CMN
Produces an ASCII file for transfer via telephone lines. This function is also performed by program STP2DAT (see below).
Link with: IOS

STP2DAT
Produces an ASCII file from an unformatted STATPAC file.
Four different types of ASCII files can be created:

1) A ".DAT" file for input to Golden Software's QGRID87 and GRAPHER programs;
2) a ".CHR" file for input to A. T. Miesch's STATPAC programs written in BASIC (e.g., MAP---$);
3) a ".CMN" file for telecommunications with the Prime,

Data General, or VAX minicomputers;
4) a ".DIF" (data interchange format) file for use as input to the LOTUS 1-2-3 spreadsheet program.

The program will automatically assign an extension to the root name of the output file.
After compiling STP2DAT.FOR, link by typing:
LINK STP2DAT /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of STP2DAT.FOR for names of required links.

STPIDS
Writes an ASCII file containing the name of the STATPAC input file, the variable IDs, and all row IDs (primary and secondary, 3 IDs per line. After compiling STPIDS.FOR, link by typing:
LINK STPIDS /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN

STPINFO.ONE
An ASCII file which demonstrates the usage of many of the beginner's level STATPAC programs. This file may be typed out on the printer or examined with BROWSE.COM.

STPINFO.TWO
An ASCII file which demonstrates the usage of SS2D-series STATPAC programs as well as other programs necessary in a kriging study. This file may be typed out on the printer or examined with BROWSE.COM.

STPSORT
Sorts the rows of a STATPAC input file by primary row ID, by secondary row ID, or by one to three selected STATPAC variables. Sorting can be specified to be in either ascending or descending order. The STATPAC file cannot exceed 3000 rows. For STATPAC variables, the user may specify a sort on the absolute values of the variables rather than the signed values. The sorted rows are written on a STATPAC output file.
Link with: ASCIIZ, B2SORT, BYTMOV, CHRKNT, CKSAFE, CONCAT, FSORT, INT21H, IOFS, IOS, ISEG, ROSORT, SOUND, TICKER

STRNGRET
Searches each row id of a STATPAC input file for the occurrence of a character substring specified by the user. STATPAC rows in which the substring is found are written on a STATPAC output file. Wild cards are permitted in the search string. The user may also specify the retrieval to include only those STATPAC rows in which the sub-string is not found.
Link with: CKSAFE, REVRSE, FNDCUR, IOS, MATCH, MOVCUR, SOUND, TICKER

TABFIL
Used to convert an ASCII data file to a unformatted STATPAC file. Each observation in the ASCII file must be represented by two rows. The first row must contain the observation ID (up to 16 characters in columns 1-16). The second row must contain the data values separated by commas or blanks. If qualifying codes are to be entered, the codes are given after the corresponding numeric values and must be in single quotes. Every numeric value must have an associated code, even if the values is unqualified and the code, therefore, is a blank character. A row with

the ID "Sample 10ax" and the values of 0.1, 0.0B, 10.0L, and 100.0, for example, must be represented in the ASCII file as:

```
-------------
Sample 10ax
0.1 ' ' 0.0 'B' 10.0 'L' 100.0 ' '
-------------
```

although the blanks between numeric values and codes may also be commas. Helpful hint: In preparing an ASCII file with qualifying codes, make use of the global replacement capability of EDLIN or some other editor. After compiling TABFIL.FOR, link by typing:
LINK TABFIL /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of TABFIL.FOR for names of required links.

TABRAN
This program creates a table of numbers arranged in a random order. User specifies the starting numbers and ending numbers of the sequence. Numbers must be in the range 1-10000. The random numbers generated are displayed on the screen.
(TABRAN is coded in Turbo Pascal. The compiled Turbo Pascal program is TABRAN.COM and is included on the release disks.)

TABQ
Produces a summary table of the number and kinds of qualifying codes in a STATPAC data file. This function is also performed by program BASTAT (See above).
Link with: IOS

TERFIL
To enter data into a STATPAC file from the keyboard. Can accomodate entry of latitude-longitude and qualifying codes. An example of how the program can be used to enter data may be obtained by typing:
TERFIL <CLARKTER.FIL
The response file "clarkter.fil" is stored on the number 2 source disk.
After compiling TERFIL.FOR, link by typing:
LINK TERFIL /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of TERFIL.FOR for names of required links.

TERFIL$
To enter data into an ASCII file with format standardized for the BASIC programs.

TERPLT
Used to plot a crude ternary diagram on the printer. Any three variables can be selected from the STATPAC data file.
Link with: IOS

TRANSF
Used to read a STATPAC file, make data transformations or create new variables, and write a new STATPAC file. Eleven functions are provided, including addition, subtraction, multiplication, division, logarithms, and exponentiation. Qualifying codes are accepted in the data and are

translated as appropriate for the output file. For example, the expression "10.0/2.0L" (i.e., 10 divided by less than 2) will result in "5.0G" (greater than 5.0). Invalid expressions, such as "log(-2.0)" or "5.0/0.0" for example, will result in an answer of "0.0B" (i.e., a blank)
Link with: FIXQ, IOS

UCODERET
(User_CODEd_RETriever) retrieves from a STATPAC input file those rows which meet selection criteria specified by the user in the associated function IWANT. (Logical function IWANT begins in line 74 of the FORTRAN source program.) The user must 1) recode function IWANT so that it expresses the desired retrieval criteria; 2) compile the FORTRAN source program; and 3) link the resulting object module UCODERET.OBJ with IOS.OBJ. The user is given the option of retrieving only those rows which meet the user's selection criteria, or only those rows which do not. The retrieved rows are written on a STATPAC output file.
Link with: IOS

UGR2FCO
Converts an unformatted STATPAC grid file created by program SS2DGRID into a character grid file which can be used as input to the H-P 7550A and TEKTRONIX 4014 contouring programs developed by John Kork of the USGS. The output is also suitable for input to the Applicon color plotter routines of the Geophysics Branch. After compiling UGR2FCO, link by typing:
LINK UGR2FCO /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of UGR2FCO.FOR for names of required links.

UGR2GRD
Converts an unformatted STATPAC grid file created by program SS2DGRID into a binary grid file which can be used as input to GOLDEN SOFTWARE's TOPO87 and/or SURF87 programs. Read comments in source code of UGR2GRD for more information. After compiling UGR2GRD.FOR, link by typing:
LINK UGR2GRD /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN
See source code of UGR2GRD.FOR for names of required links.

UNFILTER
Converts a character file created by program FILTER back into a STATPAC file. The character file produced by FILTER can be readily edited using EDLIN. Using FILTER-EDLIN-UNFILTER in sequence offers an easy way of correcting errors in STATPAC files.
See comments in source code for names of required links.
After compiling UNFILTER.FOR, link by typing:
LINK UNFILTER /EXEPACK,,NUL,C:\COMPILER\STATPAC+C:\COMPILER\8087+
C:\COMPILER\FORTRAN

WHATQ
Creates a table showing the numeric values associated with each of the six qualifying codes that can be present in a STATPAC data file. If more than one numeric value is associated with any given code, the data file should be processed through GXFIXX or some similar routine before

any type of statistical analysis is performed.
Link with: IOS

XYDGRM$
A BASIC program for screen plotting points on an XY-type diagram using a .CHR file. The .CHR file is generally derived from a STATPAC (.STP) file using program STP2CHR.

XYPLOT
Writes for later printing a character output file containing a crude x versus y scatter diagram of any two selected STATPAC variables. The graph displays as numbers 1-9, the number of data pairs which occur at each point on the plot.
Link with: IOS

References cited

Bennett, C. A., and Franklin, N. S., 1954, Statistical analysis in chemistry and the chemical industry: New York, John Wiley and Sons, Inc., p. 288-295.

Cohen, A. C., Jr., 1959, Simplified estimators for normal distribution when samples are singly censored or truncated: Technometrics, Vol. 1, pp. 217-237.

Clark, Isobel, 1979, Practical Geostatistics, Applied Science Publishers, Ltd., London, UK (Reprinted 1982), 129 pages.

Davis, J. C., and Sampson, R. J., 1966, Computer contributions 4: State Geological Survey, the University of Kansas, Lawrence.

Davis, J. C., 1973, Statistics and data analysis in geology: John Wiley and Sons, New York, 500 pages.

Efroymson, M. A., 1960, Multiple regression analysis, in Ralston, A., and Wilf, H. S., eds., Mathematical methods for digital computers: John Wiley & Sons, New York, pages 191-203.

Journel, A. G., and Huijbregts, Ch. J., 1978, Mining geostatistics: Academic Press, New York, 600 pages.

Kork, J. O., 1986, JKDIGIT, A program to control a digitizing board for geologists, written in BASIC for an IBM personal computer: U. S. Geological Survey Bulletin 1616, 69 pages.

Krige, Daniel G., and Eduardo J. Magri, 1982, Studies of the effects of outliers and data transformations on variogram estimates for a base metal and a gold ore body: Mathematical Geology, Vol. 14, No. 6, pp. 557-564.

McCracken, Daniel D., 1984, Computing for engineers and scientists with FORTRAN 77: John Wiley and Sons, New York, 361 pages.

Miesch, A. T., 1976a, Q-mode factor analysis of geochemical and petrologic data matrices with constant row-sums: U.S. Geological Survey Professional Paper 574-G, 47 pages.

_____, 1976b, Interactive computer programs for petrologic modeling with extended Q-mode factor analysis: Computers & Geosciences, v. 2, no. 4, pages 439-492.

_____, 1981, Computer methods for geochemical and petrologic mixing problems: in Merriam, D. F., ed., Computer applications in the earth sciences: Plenum Press, New York, pages 243-265.

Muth, E. J., 1986, FDA.BAS: (in User-to-user, Best file finder yet), PC MAGAZINE, Vol. 5, No. 4, pages 249-250.

Petzold, Charles, 1986, Just looking, thank you: PC MAGAZINE, Vol. 5, No. 6, pages 253-262.

Scanlon, Leo J., 1985, IBM PC and XT assembly language: Brady Communications Company, New York, 395 pages.