DEV1, A FORTRAN PROGRAM FOR ONE-WAY ANALYSIS OF VARIANCE OF ANALYTICAL DATA

by

Jean S. Kane
U.S. Geological Survey
Reston, Virginia 22092

DEV1, A FORTRAN PROGRAM FOR ONE-WAY ANALYSIS OF VARIANCE OF ANALYTICAL DATA

The USGS standard rock program provides reference materials for geochemical investigations, principally for our own laboratories but also for other geological surveys, other governmental agencies, universities, and research institutes. The utility of these reference materials requires the existance of a large body of reliable analytical data and the assurance that any one material is homogeneous for most, if not all, elements over the bottles of issue.

An experimental design with a single variable of classification to provide these data involves analysis of these reference materials for the major and minor elements, and many trace elements. The variable of classification is the bottles of issue. Using two or more bottles randomly selected from the total stock, the analyst makes replicate determinations of each of the constituents of interest in random order. The calculations of the analysis of variance result in the mean concentration of the element, the mean sum of squares within bottles, which we equate to the analytical variance (precision), a mean sum of squares between bottles which permits us to evaluate bottle variance, and an F ratio of the mean sum of squares between bottles to the mean sum of squares within bottles. This ratio tells us whether there is a bottle variance (when $F > 1$) and a comparison of this ratio with values of the F distribution at a specified probability level with appropriate degrees of freedom will tell us if the F ratio, and therefore the bottle variance, is significant. If it is not significant, homogeniety of the element from bottle to bottle within the total stock may be assumed.

This design is useful also to study day to day variations in a given analytical method , with days rather than bottles as the variable of classification. With several methods for determining a single method, this same design can give

valuable comparisons between methods. In this case the F ratio is the ratio of mean sum of squares between methods to mean sum of squares within methods.

One way analysis of variance is also useful in assessing new analytical methods. These are tested by analyzing a randomized collection of two to four bottles each of the reference rock materials. The new method is considered acceptable if the mean value obtained for each rock is in reasonable agreement with previously accepted literature values and if the analytical variance for the new method is not significantly greater than those of currently employed procedures.

Fortran Program DEV1

Fortran program DEV1 is available on Multics to calculate a one way analysis of variance on data from multiple determinations of a single element. The experimental design generally used by the Chemical Analysis Project (Reston) for testing a new analytical procedure involves four replications on each of four splits. DEV1 is written to handle symmetrical data arrays of up to this typical 4x4 size. Larger arrays and ones with an unequal number of observations per bottle will require some program modification which is currently under development. This modification, when completed, will be available as DEV2.

To produce a table suitable, after photoreduction, for publication, subroutines ROUND and CHAR insure that the computer prints all data and results with the desired number of significant figures. From the input rock name, the subroutine ROCKTYPE identifies the type of rock, eg. granite or dunite, for inclusion in the table.

Trivial changes in table headings, in placement of the "read, rock" and "read, element" statements, and in the "write (10, 150)" statement, statement number 65, by the user will adapt this program to report the determination of several elements on a single rock. In this form DEV1 can be used in reporting data for the standard rock program. An examination of the compilations of data for reference materials [1-4] demonstrates the problem of measuring objectively the accuracy of any analytical mean concentration. Within the same laboratory mean concentrations for an element arrived at by different analytical methods often show less agreement than the precision reported for the individual methods would suggest.

A major aim of our method development is to minimize discrepancies in mean value determined by a variety of methods. Experimental designs in which an analytical method is the variable of classification provide a measure of success

in achieving that aim. Availability of DEV1 on Multics to all laboratory personnel greatly facilitates the analysis of variance for this design.

Availability of DEV1 on Multics will also facilitate quality control checks on the bulk of the analytical data which the laboratories supplies daily to geologists and others. It makes it quite simple for each analyst to check his/her skill in using a given procedure against that of other analysts using the same procedure. It facilitates comparisons of duplicates repeated at weekly or monthly intervals for the duration of long term projects. It can be a tool in defining limitations of a given method, both for the analyst striving to improve the method, and for the geologist attempting to interpret the data produced by use of the method.

A user's guide for Multics execution of these programs follows. Appendices A, B, C, D list the fortran language programs DEV1, ROCKTYPE, ROUND and CHAR, and present the logic flow diagrams for each. Table I is a sample set of data, while Table II illustrates the computer output.

Multics User's Guide for DEV1

Read and execute access to DEV1 and its' called subroutines is available to all users of Multics. Execution of the program requires the user to

1) log in

2) link to DEV1, ROCKTYPE, ROUND, and CHAR:

```
lk >udd>ChemanalR>JKane>DEV1
lk >udd>ChemanalR>JKane>ROCKTYPE
lk >udd>ChemanalR>JKane>ROUND
lk >udd>ChemanalR>JKane>CHAR
```

3) Set carriage control for file 10 and attach that file for the collection of output:

```
set_cc file10 -on
io attach file10 vfile_ zz -append
```

4) initiate program execution on the first data set:

```
DEV1
```

5) enter first data set as requested by the computer with a series of prompts:

"enter chemical symbol for element"
"enter rock name"
"enter analytical data for groupl"

All data entries are discussed in detail in the next section of this report.

6) cycle the program through the second and subsequent data sets with a keyboard entry of the digit (1) in response to the cue "if more data enter 1". Again enter data as requested.

7) terminate calculations with the keyboard entry of the digit (0) after data have been entered in response to the cue "if you wish to terminate...."

8) close data files and secure a hard copy of computer output:

```
cf -a
dp zz
```

After the command cf -a the command pr zz can be used to examine output immediately on the computer screen.

9) logout

Input Data Needed for DEV1 Execution

Fortran name	Discussion
elemnt	chemical symbol for the analyte
	It is necessary to use the symbol and not the name of the analyte due to a restriction on the number of characters in the variable name and to assure alignment of output column headings with data columns.
conc	concentration units for the analyte
	This input value is restricted to a six character representation
	e.g.,ng/gm,ppm,ug/gm,%.
	Since this variable appears in the title of the table, all analyses must be in the same concentration units. The format for the computer output allows for a range from 0.001 to 9999. Thus the full range from 1 ng/gm to 0.999% can be represented as ppm when the analytical data for a single element spans such a range.

maxtt — the maximum number of groups

This appears in the output table heading to identify values in data array.

rock — the rock name

Format is critical to identification of type by the ROCKTYPE subroutine, for example, GSP-1 is a correct name; GSP, gsp-1, Gsp etc. will all result in erroneous type identification, The program listing of ROCKTYPE in appendix B includes all accepted rock names. New standard rocks can be added if desired by the user.

sn — sample size - the total number of analyses included in the mean

tt (and n) — the number of groups

st (and m) — the number of analyses per group

DEV1 assumes "st" to be equal for all groups.

X(1,1)...X(st,1) — st analytical values for the first group

X(1,n)...X(st,n) — st values for the nth group

X(1,tt)...X(st,tt) — st analytical values for the last group

f1, f2 — f ratio below which among group variation is insignificant at the 95% and 99% level of confidence, respectively.

nsig — number of significant figures supported by the precision of the data

While trace analysis data will support only one to three significant figures, minor and major element analyses may justify more. The subroutine ROUND can handle any number. CHAR maps output into an eight character format, with the decimal point as the fifth character. Therefore only numbers equal to or greater than one can be written with more than three significant numbers. Select concentration units accordingly. Alternatively edit CHAR for a 9 or more digit array.

Statistical Calculations

Fortran Name		
pmean	$\bar{X}$ (parent mean)	$= \frac{1}{sn} \sum_{i=1}^{st} \sum_{j=1}^{tt} Xij$
qw	MSS within (within - groups sum of squares)	$= \sum_{i=1}^{st} \sum_{j=1}^{tt} (Xij-\bar{X}j)^2$
qa	MSS among (among - groups sum of squares)	$= st \sum_{j=1}^{tt} (\bar{X}j - \bar{X})^2$
ssw	within - groups variance	$= \frac{qw}{sn-tt}$
ssa	among - means variance	$= \frac{qa}{tt-1}$
f	F-ratio	$= \frac{ssa}{ssw}$
analvar	V within (Analytical variance)	$= \sqrt{ssw}$
relvar	relative anal. variance (%)	$= \left(\frac{analvar}{pmean}\right) 100$

Table I

Sample Input Data

element	Cl
concentration	ppm
total no. samples	16
no. groups, no. samples per group	4, 4
rock	BCR-1
analytical values, group 1	107, 105, 107, 101
group 2	107, 98.9, 101, 109
group 3	111, 109, 99.6, 107
group 4	106, 110, 110, 111
f1, f2	3.49, 5.95
no. of significant figures	3

Table II

Cl [illegible] content in ppm of USGS standard rocks

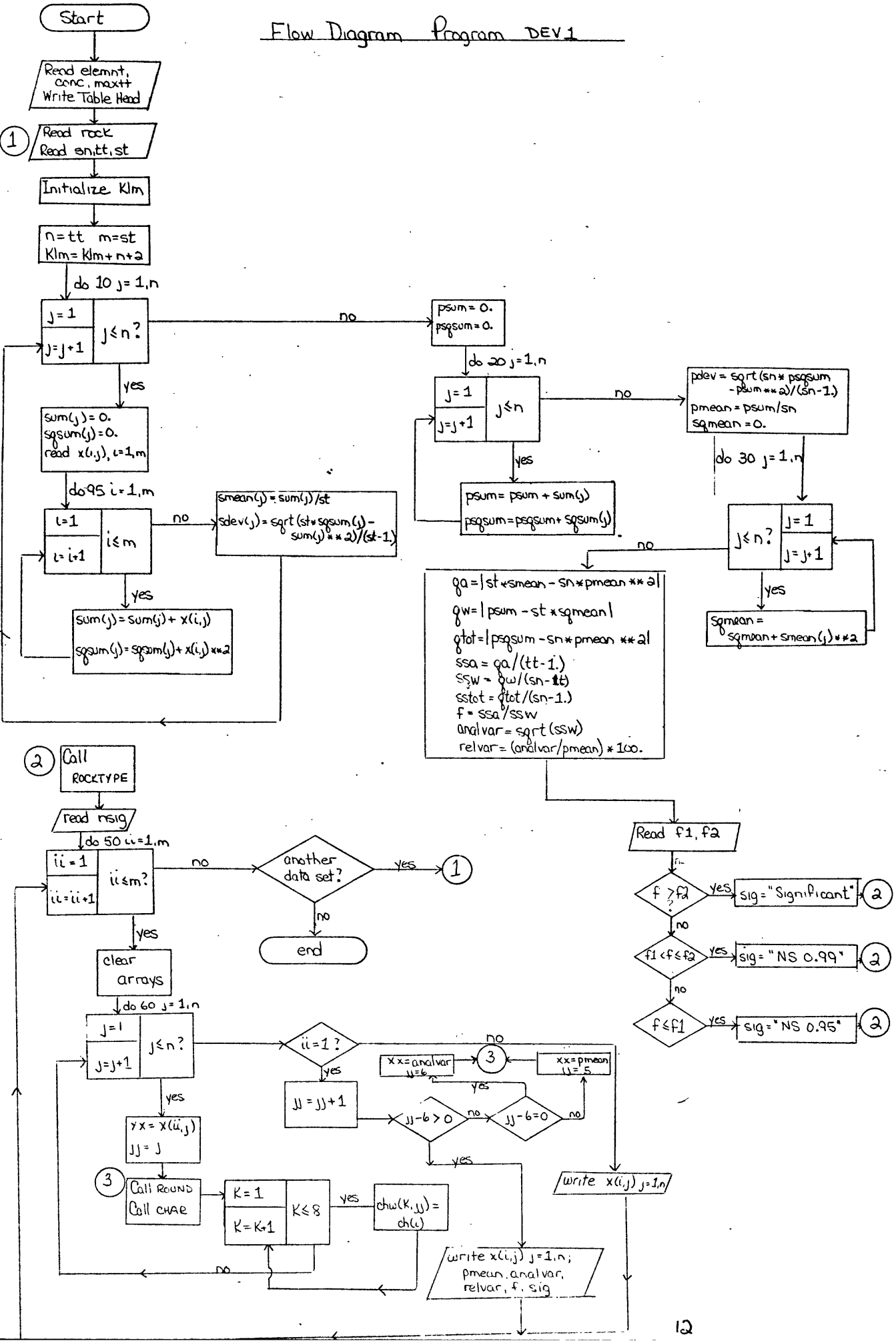
Rock	Type	Cl II in ppm				Mean	Variance within bottle	Rel. Var. in percent	F Ratio	
	group	1	2	3	4					
SCo-1	shale	108.	106.	118.	119.	114.	7.45	7.	0.333	NS 0.95
		114.	107.	120.	106.					
		118.	124.	120.	122.					
		107.	125.	109.	109.					
DTS-1	dunite	14.1	19.2	14.7	12.3	14.7	2.04	14.	1.084	NS 0.95
		16.9	16.5	14.2	14.2					
		13.5	16.1	12.2	14.1					
		14.9	12.2	17.8	13.0					
DNC-1	diabase	15.2	13.8	14.3	13.9	13.9	1.20	9.	0.458	NS 0.95
		14.5	13.9	15.7	14.7					
		12.5	12.7	13.5	12.5					
		12.2	15.9	14.1	13.1					
G-2	granite	65.4	49.0	53.1	54.8	55.0	3.19	6.	2.375	NS 0.95
		54.7	52.7	57.2	54.3					
		56.9	55.5	52.9	58.4					
		56.3	52.7	53.6	52.4					
GSP-1	granodiorite	323.	334.	346.	344.	339.	8.19	2.	0.322	NS 0.95
		347.	342.	333.	332.					
		330.	346.	344.	334.					
		347.	346.	331.	340.					
AGV-1	andesite	122.	120.	125.	107.	116.	9.21	8.	0.299	NS 0.95
		105.	107.	105.	125.					
		124.	105.	120.	124.					
		107.	125.	120.	122.					

```
com fortran program DEV1 to calculate standard deviation and f ratio
  dimension t(15),nx(4)
  dimension x(4,4),sum(4),sqsum(4),smean(4),sdev(4)
   character *6 elemnt,conc
  character *6 rock
  character *18 t,type
  character *1 ch(8)
  character *1 chw(8,6)
  character *14 sig
com Print chart headings
  write (6,1)
1 format(5x,"enter chemical symbol for element")
  read (5,2)elemnt
2 format (a6)
  write (6,3)
3 format (5x,"enter concentration units")
  read (5,2) conc
  write (6,17)
17 format (5x,"enter maximum number of groups")
  read,maxtt
  write (10,4) elemnt,conc
4 format (10x,a6," content in ",a6," of USGS standard rocks"///)
  write (10,5)elemnt,conc
5 format (2x,"Rock",7x,"Type",18x,a6," in ",a6,13x,"Mean",3x,"Variance",7x,"Rel. Dev.",3x,"F Ratio")
  write (10,6)(i,i=1,maxtt)
6 format (22x,"group",5x,i1,3(7x,i1),14x,"within bottle",2x,"in percent")
  klm=7
70 write (6,7)
7 format(5x,"enter rock name")
  read (5,8)  rock
8 format (a6)
  write (6,9)
9 format (5x,"enter total number of determinations,number of groups,",25x,"number of determinations per group")
  read,sn,tt,st
com calculate mean and deviations for individual groups
  n=tt
  klm=  klm+n+2
  m=st
  do 10 j=1,n
 sum(j)=0.0
  sqsum(j)=0.0
  write (6,13)j
13 format (5x, "enter analytical values for group number",1x,i1)
  read,(x(i,j),i=1,m)
  do 95 i=1,m
  sum(j)=sum(j)+x(i,j)
95 sqsum(j)=sqsum(j)+x(i,j)**2
  smean(j)=sum(j)/st
10 sdev(j)=sqrt(st*sqsum(j)-sum(j)**2)/(st-1.)
com  Calculate parent mean
  psum=0.0
  psqsum=0.0
  do 20 j=1,n
  psum=psum+sum(j)
20 psqsum=psqsum+sqsum(j)
  pdev=sqrt(sn*psqsum-psum**2)/(sn-1.)
  pmean=psum/sn
  sqmean=0.0
  do 30 j=1,n
30 sqmean=sqmean+smean(j)**2
```

```
com Calculate analytical variance, f ratio
  qa=abs(st*sqmean-sn*pmean**2)
  qw=abs(psqsum-st*sqmean)
  qtot=abs(psqsum-sn*pmean**2)
  ssa=qa/(tt-1.)
  ssw=qw/(sn-tt)
  sstot=qtot/(sn-1.)
  f=ssa/ssw
  analvar=sqrt(ssw)
  relvar=(analvar/pmean)*100.
com Determine significance of among bottle variation
  write (6,11)
11 format (5x,"enter maximum f ratio for 95% confidence")
  read ,f1
  write (6,12)
12 format (5x,"enter maximum f ratio for 99% confidence")
 read ,f2
  if (klm .gt.60) write (10,650)
650 format ("1")
if (klm .gt.60)klm=0
com prepare output file
com  subroutine will match rock name and rock type
  call ROCKTYPE(rock,type)
   if (f .gt. f2)sig="  Significant"
   if (f .gt. f1 .and. f .le. f2)sig="  NS 0.99"
   if (f .le. f1)sig="  NS 0.95"
   write (6,16)
16 format(5x,"enter number of significant figures")
   read, nsig
com subroutines round numbers and put in format for output table
   do 50 ii=1,m
   do 80 k=1,8
   do 80 j=1,6
80 chw(k,j)=" "
   do 60 j=1,n
   xx=x(ii,j)
   jj=j
55 call ROUND (xx,nsig,np)
   call CHAR(xx,np,nsig,ch)
   do 60 k=1,8
60 chw(k,jj)=ch(k)
   if (ii .gt. 1) go to 25
   jj=jj+1
   if (jj-6)45,35,65
45 xx=pmean
  jj=5
   go to 55
35 xx=analvar
  jj=6
   go to 55
65 write (10,150)rock,type,(chw(i,1),i=1,8),(chw(i,2),i=1,8),(chw(i,3),i=1,8),(chw(i,4),i=1,8),(chw(i,5),i=1,8),(chw(i,6),i=1,6),relvar,
f,sig
   go to 50
25 write(10,160)(chw(i,1),i=1,8),(chw(i,2),i=1,8),(chw(i,3),i=1,8),(chw(i,4),i=1,8)
150 format(//,2x,a6,2x,a18,4(8a1),5x,2(8a1),12x,f3.0,9x,f5.3,a14)
160 format(28x,4(8a1))
50 continue
com repeat calculation on additional data set or terminate
  write (6,14)
14 format(5x,"if you wish to terminate calculations enter 0",35x,"if more data enter 1")


  read ,kk
  if (kk .eq. 1) go to 70
  write (6,15)
15 format(5x,"calculations are complete. Use command cf -a to close output",25x,"files. Follow with dp zz command for hard copy printout
",25x,"and pr zz for writeout on terminal screen")
end
```

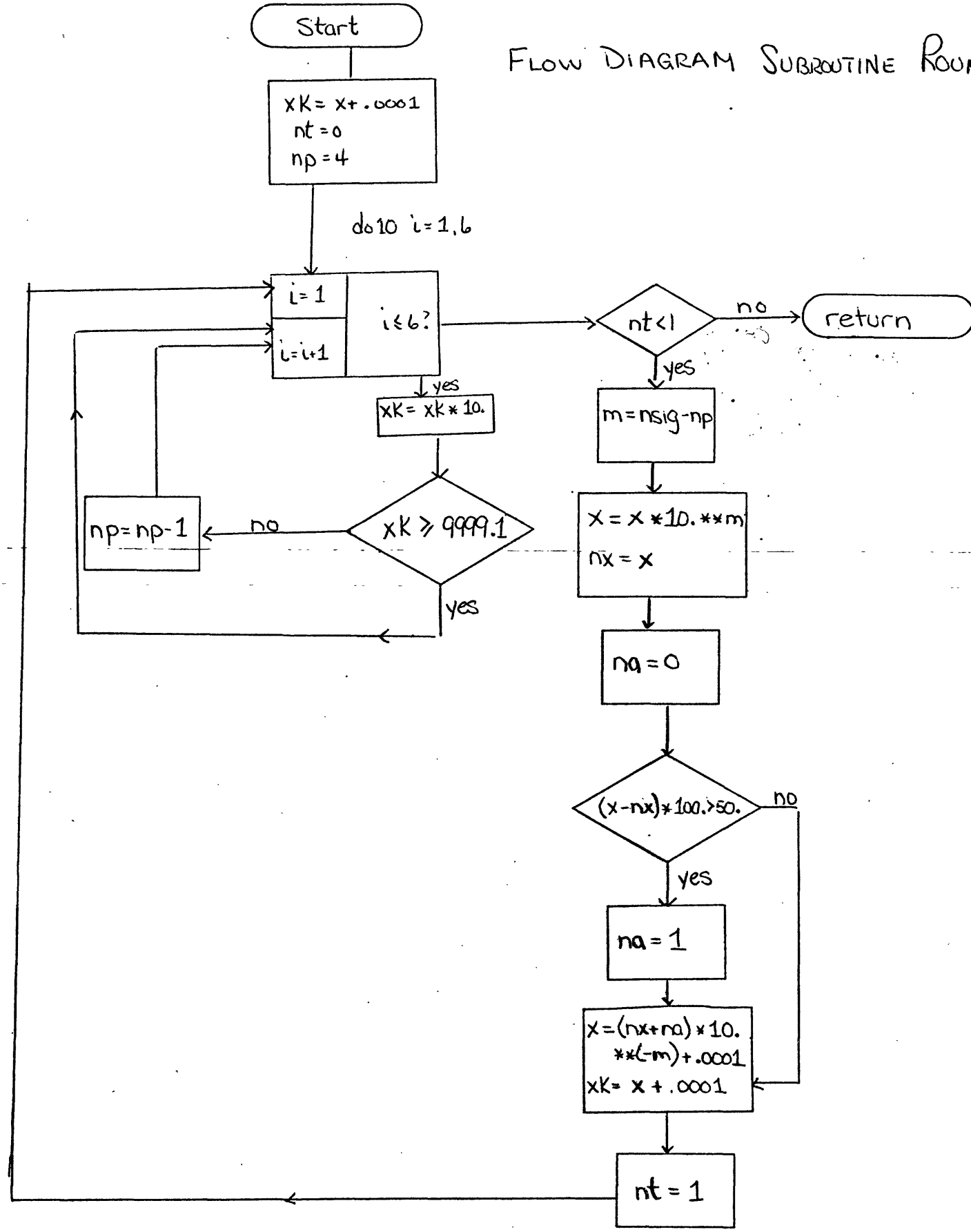
Flow Diagram Program DEV1



APPENDIX B

```
c prepare output file
c subroutine defines rock type,given rock name
subroutine ROCKTYPE(rock,type)
dimension t(15)
  character *6 rock
  character *18 t,type
t(1)="basalt"
t(2)="granite"
t(3)="granodiorite"
t(4)="nepheline syenite"
t(5)="marine mud"
t(6)="dunite"
t(7)="peridotite"
t(8)="shale"
t(9)="andesite"
t(10)="quartz latite"
t(11)="rhyolite"
t(12)="mica schist"
t(13)="diabase"
t(14)="rhyolitic pumice"
if (rock .ne. "BCR-1")go to 11
type=t(1)
go to 20
11 if (rock .ne. "G-2")go to 12
type =t(2)
go to 20
12 if (rock .ne. "GSP-1") go to 13
type = t(3)
go to 20
13 if (rock .ne. "STM-1")go to 14
type=t(4)
go to 20
14 if (rock .ne. "MAG-1")go to 15
type=t(5)
go to 20
15 if (rock .ne. "DTS-1")go to 16
type=t(6)
go to 20
16 if (rock .ne. "PCC-1") go to 17
type =t(7)
go to 20
17 if (rock .ne. "SCo-1") go to 18
type=t(8)
go to 20
18 if (rock .ne. "SGR-1") go to 19
type=t(8)
go to 20
19 if (rock .ne. "AGV-1") go to 21
type=t(9)
go to 20
21 if (rock .ne. "QLO-1")go to 22
type=t(10)
go to 20
22 if (rock .ne. "RGM-1") go to 23
type=t(11)
go to 20
23 if (rock .ne. "SDC-1") go to 24
type =t(12)
go to 20
24 if (rock .ne. "DNC-1") go to 25
type=t(13)
go to 20
25 if (rock .ne. "BHVO-1") go to 26
type=t(1)
go to 20
26 if (rock .ne. "BIR-1") go to 27
type=t(1)
go to 20
27 if (rock .ne. "W-2") go to 28
type=t(13)
go to 20
28 if (rock .eq. "RJM-22")type=t(14)
20 continue
return
end
```

Flow Diagram Subroutine Round

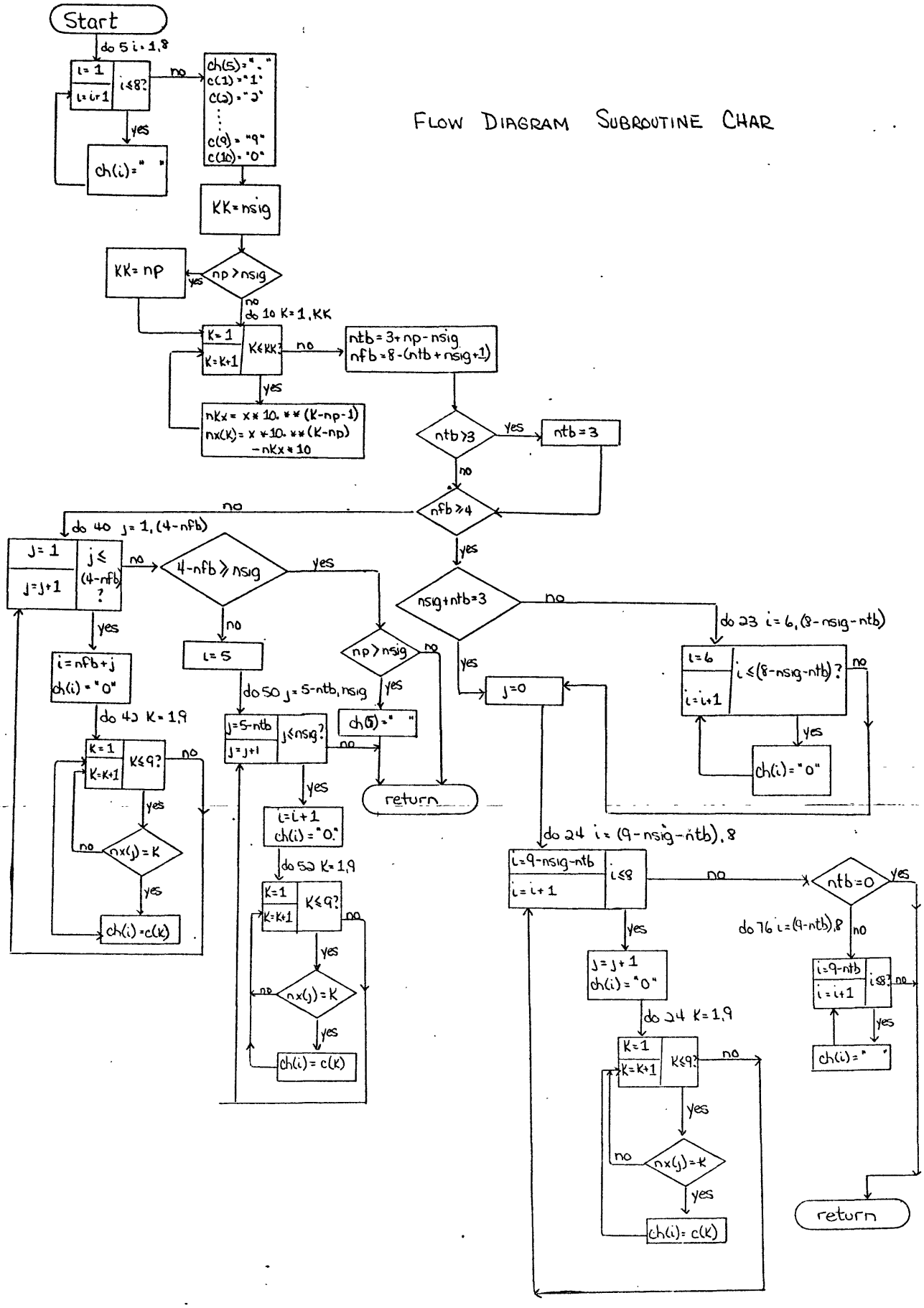


```
      subroutine ROUND(xx,nsig,np)

      xk=xx+0.00001
      nt=0
com   determine no. of places in front of decimal
      np=4
10    xk=xk*10.
      if (xk .ge. 9999.9)go to 5
      np=np-1
      go to 10
5     m=nsig-np
com     convert xx to integer with nsig significant figures
      x=xx*10.0**m
      nx=x
com   determine whether nx needs to be rounded up
      na=0
      if(((x-nx)*100.0) .gt. 50.)na=1
com   convert nx to real with np sig figures and original decimal
      xx=(nx+na)*10.0**(-m)+0.00001
com   re-determine np in case rounding added another place
      if(nx .eq. 9 .and. nsig .eq. 1)np=np+1
      if(nx .eq. 99 .and. nsig .eq. 2)np=np+1
      if(nx .eq. 999 .and. nsig .eq. 3)np=np+1
      if(nx .eq. 9999 .and. nsig .eq. 4)np=np+1
99    return
      end
```

```
      subroutine CHAR(xx,np,nsig,ch)
      dimension nx(4)
      character*1 c(10),ch(8)
com initialize character strings
      do 5 i=1,8
      ch(i)=" "
5     continue
      ch(5)="."
      c(1)="1"
      c(2)="2"
      c(3)="3"
      c(4)="4"
      c(5)="5"
      c(6)="6"
      c(7)="7"
      c(8)="8"
      c(9)="9"
com put indivdual digits of xx in nx
      kk=nsig
      if(np .gt. nsig)kk=np
      do 10 k=1,kk
      nkx=xx*10.0**(k-np-1)
10    nx(k)=xx*10.0**(k-np)-nkx*10
com determine no. of frontal(nfb) and trailing(ntb) blanks in ch
      ntb=3+np-nsig
      nfb=8-(ntb+nsig+1)
      if(ntb .gt. 3)ntb=3
com  transfer nx to ch; branch to 2nd routine if xx<1.0
      if(nfb .ge. 4)go to 20
      do 42 j=1,(4-nfb)
      i=nfb+j
      ch(i)="0"
      do 42 k=1,9
42    if(nx(j) .eq. k)ch(i)=c(k)
com  nos. in front of decimal entered; branch if no nos. behind decimal
      if((4-nfb) .ge. nsig)go to 99
      i=5
      do 52 j=(5-nfb),nsig
      i=i+1
      ch(i)="0"
      do 52 k=1,9
52    if(nx(j) .eq. k)ch(i)=c(k)
      go to 99
com  enter routine for xx<1.0
20    if((nsig+ntb) .eq. 3)go to 22
      do 23 i=6,(8-nsig-ntb)
23    ch(i)="0"
22    j=0
      do 24 i=(9-nsig-ntb),8
      j=j+1
      ch(i)="0"
      do 24 k=1,9
24    if(nx(j) .eq. k)ch(i)=c(k)
      if(ntb .eq. 0)go to 99
      do 76 i=(9-ntb),8
76    ch(i)=" "
99    continue
      if(np .gt. nsig)ch(5)=" "
 return
```

FLOW DIAGRAM SUBROUTINE CHAR



References

1. Flanagan, F.J., Geol. Survey Prof. Paper 840, 1976
2. Flanagan, F.J., Geochim. & Cosmochim. Acta. 33, 81-120, 1969.
3. Flanagan, F.J., Geochim. & Cosmochim. Acta. 34, 121-125, 1970
4. Abbey, Sidney, Geol. Survey of Canada Paper 77-34, 1977.