DEPARTMENT OF THE INTERIOR
U.S. GEOLOGICAL SURVEY

Computer Program for Geochemical Correlation Studies

by

John B. Rapp

U.S. Geological Survey
Menlo Park, California 94025

OPEN FILE REPORT 84-415

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

This report discusses one of the computer programs which were written for our study of oil and rock correlations of samples from the North Slope, Alaska (Kvenvolden and others). The program produces tables of calculated data which are used for correlations. The program can be used for all combinations (oil/oil, oil/rock, rock/rock) of correlations by simply selecting which samples are inputted. Any group of chemical compounds may be used in the program but we used different classes of hydrocarbons (normal alkanes, isoprenoids, steranes and hopanes) in our study. The report is oriented toward users as opposed to programmers. If the reader needs more detail, the listing of the program is given in the Appendix and contains remark statements to assist the reader. The program is interactive and all input requests are in the form of questions and non-numeric input requests are usually followed by a list of the possible choices. Examples of the output produced by the program are presented in the tables. Choices and limitations of input and output are discussed.

The program is written in BASIC and run on a Honeywell 61/80 computer with the Multics operating system. The reasons for using BASIC are that it is an interactive language and that it is widespread in today's microcomputers. The program is designed to be run on a wide variety of computers including 8 bit microcomputers. Many variables and matrices are re-defined several times in order to conserve memory (for running on microcomputers with small memories).

OVERVIEW

The name of the program is OIL/ROCK.BASIC. The objective of the program is to assist the user in grouping samples (oils and/or rocks) according to their hydrocarbon composition. The data base consists of several files which are constructed before the program is run. Some files contain names of samples or names of hydrocarbon compounds. Other files contain analyses of hydrocarbons in the samples (either concentrations (ppb) or peak heights). The program begins with an initialization section in which the user determines which data file (analyses of which hydrocarbon class—e.g. HOPANE) to be used in the calculations. This data is placed in one matrix, the corresponding names of the hydrocarbon compounds are placed in another matrix and the sample names are placed into a third matrix. The program now branches to one of the subprograms, to the selection of a new data matrix or to the end of the program. This is the major branching point to which all subprograms return. The names of the four subprograms are Ratio, Carbon Preference Index, Odd-Even Predominance, and Dissimilarity Index. The subprogram Ratio calculates and outputs simple ratios (single numerator and single denominator) or multiple ratios (summations in the numerator and/or summations in the denominator). It can be used with compounds from any of the hydrocarbon classes (e.g. STERANES). The subprograms Carbon Preference Index and Odd-Even Predominance are two different methods of calculating a ratio between odd-carbon-numbered normal alkanes and even-carbon-numbered normal alkanes. The subprogram Dissimilarity Index (DI) compares a number of variables in two samples and results in a number which is a measure of the differences between the two samples. The technique is useful in comparing the overall composition (as opposed to only one or two variables) of samples and in grouping samples which have low dissimilarity indices (the output table has been sorted in ascending order and samples with similar DIs are adjacent to each other). This
subprogram may be used with compounds from any of the hydrocarbons.

***MATRICES***

Many matrices are common to all the subprograms. All the matrices are stored within the computer as files and the requested matrices are read by the programs at run time.

One set of matrices contains data (in two dimensions - sample number versus compound number). Matrix ALKANE contains normal alkanes and the isoprenoids pristane and phytane, matrix HOPANE contains tricyclic and pentacyclic terpanes and matrix STERANE contains steranes and diasteranes. The data is quantitated in two ways: 1. concentration (ppb) - where a quantitative standard is available (e.g. matrix ALKANE - ppb contains concentrations of normal alkanes and the isoprenoids pristane and phytane) and 2. peak height - where only a qualitative standard (peaks identified but not quantitated in concentration) is available (e.g. the matrices STERANE, HOPANE and ALKANE-HT).

Another set of matrices contains names. One matrix contains sample names of oils (No. 1-8, 24) and rocks (No. 9-23) from the North Slope of Alaska and was called NS-NAME. Table 1 contains sample numbers and sample names in the Matrix NS-NAME. Other matrices contained compound names and were called the concatenation of the data matrix name and the string "-str" (e.g. STERANE-STR and HOPANE-STR). Table 7 contains the compound numbers and abbreviated names of the compounds in the matrices HOPANE and STERANE (contents of matrices HOPANE-STR and STERANE-STR). Table 8 contains compound numbers and compound names of the terpanes and steranes used in the study. The alkane matrices consisted of n-C-12 to n-C-40 (compounds 1 to 29), pris (pristane:compound 30) and phyt (phytane:compound 31).

Several matrices and simple variables are re-defined several times in a single subprogram. This is done for several operations where memory is needed to store intermediate results of calculations. Examples of such operations are sort routines and summations. Normally, the practice of re-defining variables should be avoided in programming but is used here to conserve memory so that these programs might be run on microcomputers with limited memories.

Some of the limitations imposed on the program are the following maximums: 25 samples, and 40 compounds in each data matrix, 20 letters in a sample name, 8 letters in a compound name in all subprograms and 9 ratios, 9 numerators and 9 denominators in the ratio subprogram. Most of these limitations are imposed to fit this one set of data and can easily be expanded to larger values for larger data sets. This entails changing the dimension statement and a few lines of program. Some of the limitations are imposed by the printer width (132 characters). The name lengths are limited in order to get more columns of data per page. Most of the time the number of characters in the column heading name is larger than the number of characters in the data.

***OIL/ROCK.BASIC***

As stated previously, the objective of the program is to assist the user in grouping samples (oil and rocks) according to their hydrocarbon composition. The output from this program are tables of data ordered by (1.) requested samples and (2.) sorted data. Observing sample similarities, differences and groupings is much easier when sorted data are used. The program is quite versatile in that there is a very large number of possible
combinations of samples and compounds. All the samples (oils and rocks) or
only selected ones may be compared and thus all combinations
(oil/oil,oil/rock,rock/rock) of correlations may be produced. With the
present program, there is a limitation on the combinations of hydrocarbon
compounds which may be compared. Only compounds within a single matrix may be
compared (e.g. only compounds within HOPANE with each other and not with
compounds in STERANE). The program can be easily modified (by combining the
matrices) so that any combination of compounds can be compared. The program
listing is given in the Appendix. Remark statements (rem) are used in the
listing to delineate subprograms and subroutines and to clarify certain
statements. This program is composed of an initialization section and four
subprograms: Ratio, Carbon Preference Index, Odd-Even Predominance and
Dissimilarity Index.

Initialization. After dimensioning the matrices, the program requests
the name of the data file to be used. The program then loads that data file
into matrix m and defines file No. 2 as the concatenation of the data file
name and the string"-str" (this file contains the names of the compounds).
File No. 3 is defined as "NS-NAME": the names of the samples. The program
then requests whether the user wants to calculate a ratio of compounds, a
dissimilarity index, a Carbon Preference Index, or an Odd/Even Predominance or
to select a new data matrix or to stop the program. This is the major
branching point to which all sub-programs return.

Ratio. The sub-program Ratio calculates and outputs simple ratios
(single numerator and single denominator) or multiple ratios (summations in
the numerator and/or summations in the denominator). The program first asks
which samples to include in the output table. It asks the total number of
samples and if the sample numbers are contiguous or not (the rock sample
numbers are contiguous and the oil sample numbers are contiguous except number
24 which was added later). This is to make sample number entry easier - its
easier to enter a range than each individual sample. An example would be if
all rock samples were requested - it would be easier to input the two number
range (9,23) than to input the fifteen individual sample numbers
(9,10,11,...,22,23). It then asks which ratios to print in the table. There
is a maximum number of ratios of 9 because the table would get too crowded
with more. The next question is "Multiple Numerator/Denominators (y/n)? An
answer of "n" implies a simple ratio and asks a different set of questions
about numerators and denominators.

The output table contains the run date, sample number, sample name, and
requested ratios. The heading for simple ratios is the abbreviations of the
numerator and denominators above the appropriate columns whereas the heading
for multiple ratios is just RATIO 1, RATIO 2, etc. because of space
limitations. The initial table is printed in the order that the sample
numbers were inputted. The program then asks if the user would like a table
with the ratios sorted. If the answer is yes then the question is asked "sort
according to which ratio." The table is then printed in ascending order of
the ratio selected with the other ratios following the same sample order. If
the answer is no to the sort question then no further table is produced. If a
denominator is zero for a sample, asterisks are printed for the value of the
ratio for that sample. Table 2 is an example of the output for simple
ratios. The first part is the output in the order of requested samples while
the second part is sorted according to the value of pristane/phytane. Note
the asterisks for sample eight: none of the requested alkanes was present in
the sample. It is easy to observe in the sorted table that samples 1,2,3, and
4 form a group with similar ratios. Table 3 is an example of the output for
multiple ratios. The numerator is the sum of four diasteranes and the
denominator is the sum of four steranes. The second part of the table
contains the ratios sorted in ascending order. The group formed by samples
1, 2, 3, 4, and 7 is much more easily delineated in the sorted table than in the
unsorted table. This grouping is also supported by information reported
earlier (Kvenvolden and others).

Carbon Preference Index. This subprogram calculates the carbon
preference index (CPI) of the normal alkanes of samples according to an
extension of the method of Bray and Evans (1961). CPI is a "ratio" of odd-
carbon-numbered to even-carbon-numbered normal alkanes and is calculated as
follows:

\[
\text{CPI} = \frac{\sum_{i=j}^{k} \frac{C_{2i+1}}{2}}{\sum_{i=j}^{k} \frac{1}{C_{2i}}} + \frac{C_{2i}}{2} + \frac{C_{2i}}{2}
\]

Where \( C_{2i} \) = the concentration measure (weight percent, mole
percent, concentration, peak height, etc) of a
normal alkane with carbon number \( 2i \).

\( j \) = one half the lowest carbon numbered alkane in the
range.

\( k \) = (one half the highest carbon numbered alkane in
the range) minus 1.

The range of the original method was limited to \( n-C_{24} \) to \( n-C_{34} \) (\( j=12, k=16 \))
whereas the range of this subprogram is user defined. CPI is useful as a
maturity indicator (i.e. recent sediments have high values whereas ancient
sediments have values approaching 1).

The program first checks which matrix is being used. It can only be used
with the "ALKANE-" matrices, otherwise an error is printed. Next, the range
of hydrocarbons to be used in the calculation is requested. If the carbon
numbers entered aren't both even then an error is noted and the question is
repeated. Because of the structure of the alkane matrix, the CPI range is
limited to \( n-C_{12} \) to \( n-C_{40} \). The structure can easily be modified to extend the
range.

The subprogram's output is a table containing the run date, name of the
matrix used in the calculations, a heading and a sorted list of CPIs. If all
the compounds in the selected range are present in the sample, the CPI is
printed. If any of the first three (lowest carbon numbered) compounds in the
selected range are missing, CPI is not printed but the lowest missing
compound's carbon number is printed in the "missing alkanes" column. If any
compound beyond the first three is missing, the lowest missing compound's
carbon number is printed in the "missing alkanes" column and a "pseudo CPI" is
calculated using summations accumulated to that point and is printed in
parenthesis in the CPI column. The CPIs are sorted and printed in ascending
order in the table. The "pseudo CPI" should not be used with or compared to
the real CPIs in the table. It is only printed in order to give the user an
estimate of CPI in range from the lowest carbon numbered compound in the
selected range to the compound just before the missing compound. Table 4 is
an example of various CPI outputs. The range is \( n-C_{16} \) to \( n-C_{28} \). The values
for oils are printed first. Sample 8 doesn't contain \( n-C_{16} \) so no CPI is
printed. Sample 6 doesn't contain \( n-C_{20} \) so a pseudo CPI is printed in
parenthesis. The values for rocks are printed next. Note that sample 10 is higher than the rest, indicating less maturity.

**Odd-Even Predominance.** This subprogram calculates the odd-even predominance (OEP) of the normal alkanes of samples according to the method of Scalan and Smith (1969). The calculation uses five consecutive members of the n-alkane homologous series. OEP is a ratio of odd-carbon-numbered to even-carbon-numbered normal alkanes and is calculated as follows:

\[
OEP = \frac{C_i + 6C_{i+2} + C_{i+4}}{4C_{i+1} + 4C_{i+3}} (-1)^{i+1}
\]

where \(C_i\) = the concentration measure (weight percent, mole percent, concentration, peak height, etc) of a normal alkane with carbon number \(i\).

The OEP value is assigned to the alkane in the center of the group (alkane containing \(i + 2\) carbon atoms). It is used to relate samples and as a maturity indicator (young samples have high values and mature samples have values around 1.0). It is similar to CPI but with more filtering - large variations in the raw data have less effect on OEP than on CPI. The program first checks that the current data matrix is "ALKANE-." It then asks which carbon number the user wishes the OEP to center around and which samples are to be used.

The program's output is a table containing the run date, name of the matrix used, a heading containing the carbon number about which the OEP was calculated and a sorted list of OEPs. If any of the values of the five consecutive alkanes in a sample is missing then no OEP is printed but the sample number, sample name, and the carbon number of the lowest missing alkane are printed at the top of the table. The rest of the table contains OEP values in ascending order. This ordering is useful in grouping samples since samples with similar OEPs are adjacent to each other. Table 5 is on OEP output centering at \(n-C_{25}\). Oil and rock OEPs were printed separately for clarity but could easily have been printed together. Samples 6 and 8 didn't contain \(n-C_{23}\) so their OEPs are not printed. Because the rock OEP listed is sorted, it is easy to see (1.) that generally the Pebble and Torok samples have higher OEPs than the rest of the samples and (2.) that sample 10 is very different from the rest of the samples (1.57 versus 1.29 for the adjacent sample).

**Dissimilarity Index.** Dissimilarity Index (DI) (Montalvo, 1978) is the result of a mathematical calculation which compares a number of selected variables in two samples (somewhat analogous to correlation coefficient which compares many samples for two variables - comparing samples versus comparing variables). The technique is useful in comparing the overall composition (as opposed to only one or two variables) of two samples and in grouping samples which have low dissimilarity indices. The variables in each sample are first normalized such that the summation of all the selected variables in that sample equals 1.0.
\[ V_{xy} = \frac{C_{xy}}{t} \sum_{i=1}^{t} C_{xi} \]

where \( x \) = sample number
\( y \) = compound number
\( V_{xy} \) = normalized variable for sample \( x \) and compound \( y \)
\( C_{xy} \) = concentration measure (weight \%, mole \%, concentration (ppb), peak height, etc) of compound \( y \) in sample \( x \)

The absolute difference between the normalized variable in the two samples is summed for all the compounds. The summation is divided by 2.0 to give the dissimilarity index for the sample pair:

\[ DI_{x_1x_2} = \frac{1}{2} \sum_{i=1}^{t} |V_{x_1i} - V_{x_2i}| \]

where \( DI_{x_1x_2} \) = dissimilarity index for the sample pair \( x_1 \) and \( x_2 \)
\( t \) = total number of compounds used
\( V_{x_1i} \) = normalized variable for compound \( i \) in sample \( x_1 \)
\( V_{x_2i} \) = normalized variable for compound \( i \) in sample \( x_2 \)

The limits for dissimilarity index are 0 for a perfect match (identical composition) and 1.0 for a complete mismatch (all compounds present in one sample are absent in the other sample and vice-versa). The program first asks about the samples to be run. The total number of samples to be printed in the table is requested. Next, sample numbers are requested - contiguous ranges of sample numbers should be inputted before sample numbers which aren't contiguous (ranges are easier to input than entering every individual sample number). Information on which compounds to use is inputted next. If all compounds in a matrix are to be used then a matrix \( c \) is filled in ascending order with all the compound numbers. If all compounds in a matrix aren't to be used then the ranges and compound numbers are inputted, sorted in ascending order and placed into matrix \( c \). All the requested samples are checked to see if at least one of the requested compounds is present and if not, a message is printed stating none of the compounds are present in that sample (the sample is then excluded from further calculations and printouts).

The program outputs a table of dissimilarity indices. A reference sample number is requested (can be any sample - isn't limited to those requested above). The DIs printed are all comparisons between this reference sample and the above requested (comparison) samples. The reference may be an oil and the comparison samples may be oils (oil/oil comparisons) or rocks (oil/rock comparisons). The reference may be a rock sample and the comparison samples may be rocks (rock/rock comparisons) or oils (oil/rock comparisons). The current date, compound matrix name, reference sample number and name are printed at the beginning of the table. The table consists of a sorted list of
DIs in ascending order together with the corresponding sample number and name. The difference between adjacent DIs is printed under the heading of DELTA. Printing the DIs in ascending order is useful because members of the same group as the reference will generally be together at the top of the list. Printing the delta values is sometimes helpful in determining boundaries between groups. Table 6 contains the DIs of 39 hopanes in all the rock samples compared to oil sample 2. Note the large delta value (.249) between samples 14 and 17. It shows the division between the Shublik/Kingak group, which is more closely associated with oil sample 2, and the Pebble/Torok group. The two exceptions (samples 18 and 23) have been discussed previously (Kvenvolden and others),
REFERENCES


TABLE CAPTIONS

Table 1. Contents of matrix NS-NAME (sample numbers and sample names). (numbers following rock names are depths (in feet) below the surface)

Table 2. Input and output of (simple) Ratio subprogram pris/n-C_{17}, phyt/n-C_{18} and pris/phyt in oil samples: 1. in order of requested samples and 2. sorted according to ascending value of pris/phyt.

Table 3. Input and output of (multiple) Ratio subprogram. Diasteranes/ Steranes in oil samples: 1. in order of requested samples and 2. sorted in ascending order.

Table 4. Input and output of Carbon Preference Index subprogram. Carbon Preference Index (in the n-C_{16} to n-C_{28} range) of oil and rock samples.

Table 5. Input and output of Odd-Even Predominance subprogram. Odd-Even Predominance centering at n-C_{25} of oil and rock samples.

Table 6. Input and output of Dissimilarity Index subprogram. Dissimilarity Indices of 39 hopanes in rock samples compared to oil sample 2.

Table 7. Contents of matrices HOPANE-STR and STERANE-STR (compound numbers and abbreviated compound names).

Table 8. Compound numbers and names of the terpanes and steranes used in the study.
## SAMPLES

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Prudhoe - Put River</td>
</tr>
<tr>
<td>2</td>
<td>South Barrow 20</td>
</tr>
<tr>
<td>3</td>
<td>South Barrow 19</td>
</tr>
<tr>
<td>4</td>
<td>Fish Creek 1</td>
</tr>
<tr>
<td>5</td>
<td>Simpson Core Test 4</td>
</tr>
<tr>
<td>6</td>
<td>Seabee 1 DST 3</td>
</tr>
<tr>
<td>7</td>
<td>Dalton 1 DST 2</td>
</tr>
<tr>
<td>8</td>
<td>Cape Simpson SL</td>
</tr>
<tr>
<td>24</td>
<td>Umiat 4</td>
</tr>
</tbody>
</table>

## OILS

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Pebble shale 6361</td>
</tr>
<tr>
<td>10</td>
<td>Pebble shale 1977</td>
</tr>
<tr>
<td>11</td>
<td>Kingsak Shale 14577</td>
</tr>
<tr>
<td>12</td>
<td>Fort. Mtn. Fm 12014</td>
</tr>
<tr>
<td>13</td>
<td>Sadlerochit Gp 13831</td>
</tr>
<tr>
<td>14</td>
<td>Shublik Fm 12273</td>
</tr>
<tr>
<td>15</td>
<td>Kingsak Shale 11704</td>
</tr>
<tr>
<td>16</td>
<td>Torok Fm 7054</td>
</tr>
<tr>
<td>17</td>
<td>Pebble shale 6405</td>
</tr>
<tr>
<td>18</td>
<td>Shublik Fm 10273</td>
</tr>
<tr>
<td>19</td>
<td>Pebble shale 7371</td>
</tr>
<tr>
<td>20</td>
<td>Torok Fm 3792</td>
</tr>
<tr>
<td>21</td>
<td>Shublik Fm 9042</td>
</tr>
<tr>
<td>22</td>
<td>Torok Fm 4994</td>
</tr>
<tr>
<td>23</td>
<td>Kingsak Shale 7390</td>
</tr>
</tbody>
</table>

## ROCKS

Table 1
Name of matrix file = alkane-h
t
Ratio(r), Dissimilarity Index(d), Cpi/Dep(c), New Matrix(m), or Stop(s)? r
Input contiguous ranges before non-contiguous samples
Are samples contiguous(c) or not(n)? c
What range (low, high)? 1, 8
Are samples contiguous(c) or not(n)? n
Sample Number =? 24
Number of ratios in table (9 max.) =? 3
Multiple Numerator/Denominators(y/n)? n
Ratio 1 numerator compound number =? 30
Ratio 1 denominator compound number =? 6
Ratio 2 numerator compound number =? 31
Ratio 2 denominator compound number =? 7
Ratio 3 numerator compound number =? 30
Ratio 3 denominator compound number =? 31
date = 01/16/84 matrix is alkane-h
t
<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Sample Name</th>
<th>prist</th>
<th>phyt</th>
<th>prist</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n-C17</td>
<td>n-C18</td>
<td>phyt</td>
</tr>
<tr>
<td>1</td>
<td>Prudhoe - Put River</td>
<td>0.388</td>
<td>0.360</td>
<td>1.444</td>
</tr>
<tr>
<td>2</td>
<td>South Barrow 20</td>
<td>0.419</td>
<td>0.349</td>
<td>1.556</td>
</tr>
<tr>
<td>3</td>
<td>South Barrow 19</td>
<td>0.273</td>
<td>0.223</td>
<td>1.565</td>
</tr>
<tr>
<td>4</td>
<td>Fish Creek 1</td>
<td>0.382</td>
<td>0.267</td>
<td>1.750</td>
</tr>
<tr>
<td>5</td>
<td>Simpson Core Test 4</td>
<td>0.453</td>
<td>0.180</td>
<td>3.545</td>
</tr>
<tr>
<td>6</td>
<td>Seabee 1 DST 3</td>
<td>0.438</td>
<td>0.200</td>
<td>14.000</td>
</tr>
<tr>
<td>7</td>
<td>Dalton 1 DST 2</td>
<td>0.343</td>
<td>0.400</td>
<td>1.000</td>
</tr>
<tr>
<td>8</td>
<td>Cape Simpson SL</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>24</td>
<td>Umiat 4</td>
<td>0.458</td>
<td>0.248</td>
<td>2.200</td>
</tr>
</tbody>
</table>

t
Print table with ratios sorted(y/n)? y
Sort according to which ratio? 3
date = 01/16/84 matrix is alkane-h
t
<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Sample Name</th>
<th>prist</th>
<th>phyt</th>
<th>prist</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n-C17</td>
<td>n-C18</td>
<td>phyt</td>
</tr>
<tr>
<td>7</td>
<td>Dalton 1 DST 2</td>
<td>0.343</td>
<td>0.400</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>Prudhoe - Put River</td>
<td>0.388</td>
<td>0.360</td>
<td>1.444</td>
</tr>
<tr>
<td>2</td>
<td>South Barrow 20</td>
<td>0.419</td>
<td>0.349</td>
<td>1.556</td>
</tr>
<tr>
<td>3</td>
<td>South Barrow 19</td>
<td>0.273</td>
<td>0.223</td>
<td>1.565</td>
</tr>
<tr>
<td>4</td>
<td>Fish Creek 1</td>
<td>0.382</td>
<td>0.267</td>
<td>1.750</td>
</tr>
<tr>
<td>24</td>
<td>Umiat 4</td>
<td>0.458</td>
<td>0.248</td>
<td>2.200</td>
</tr>
<tr>
<td>5</td>
<td>Simpson Core Test 4</td>
<td>0.453</td>
<td>0.180</td>
<td>3.545</td>
</tr>
<tr>
<td>6</td>
<td>Seabee 1 DST 3</td>
<td>0.438</td>
<td>0.200</td>
<td>14.000</td>
</tr>
<tr>
<td>8</td>
<td>Cape Simpson SL</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>

Another sort(y/n)? n
Name of matrix file = sterane
Ratio(r), Dissimilarity Index(d), Cpi/Dep(c), New Matrix(m), or Stop(s)? r
Number of samples in table = 9
Input contiguous ranges before non-contiguous samples
Are samples contiguous (c) or not (n)? c
What range (low, high)? 1, 8
Are samples contiguous (c) or not (n)? n
Sample Number = 24
Number of ratios in table (9 max.) = 1
Multiple Numerator/Denominators (y/n)? y
Maximum Numerator or Denominators is 9 each.
Number of numerators in ratio 1 = 4
  Numerator 1 = 1
  Numerator 2 = 2
  Numerator 3 = 10
  Numerator 4 = 13
Number of denominators in ratio 1 = 4
  Denominator 1 = 9
  Denominator 2 = 12
  Denominator 3 = 20
  Denominator 4 = 23
Ratio 1
  Numerators = 27baSD + 27baRD + 29baSD + 29baRD
  Denominators = 27aaaSS + 27aaaRS + 29aaaSS + 29aaaRS

date = 01/16/84 matrix is sterane

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prudhoe - Put River</td>
<td>1.736</td>
</tr>
<tr>
<td>South Barrow 20</td>
<td>1.389</td>
</tr>
<tr>
<td>South Barrow 19</td>
<td>1.518</td>
</tr>
<tr>
<td>Fish Creek 1</td>
<td>1.842</td>
</tr>
<tr>
<td>Simpson Core Test 4</td>
<td>2.230</td>
</tr>
<tr>
<td>Seabee 1 DST 3</td>
<td>**</td>
</tr>
<tr>
<td>Dalton 1 DST 2</td>
<td>1.495</td>
</tr>
<tr>
<td>Cape Simpson SL</td>
<td>2.349</td>
</tr>
<tr>
<td>Umiat 4</td>
<td>3.325</td>
</tr>
</tbody>
</table>

Print table with ratios sorted (y/n)? y
Sort according to which ratio? 1

date = 01/16/84 matrix is sterane

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Barrow 20</td>
<td>1.389</td>
</tr>
<tr>
<td>Dalton 1 DST 2</td>
<td>1.495</td>
</tr>
<tr>
<td>South Barrow 19</td>
<td>1.518</td>
</tr>
<tr>
<td>Prudhoe - Put River</td>
<td>1.736</td>
</tr>
<tr>
<td>Fish Creek 1</td>
<td>1.842</td>
</tr>
<tr>
<td>Simpson Core Test 4</td>
<td>2.230</td>
</tr>
<tr>
<td>Cape Simpson SL</td>
<td>2.349</td>
</tr>
<tr>
<td>Umiat 4</td>
<td>3.325</td>
</tr>
<tr>
<td>Seabee 1 DST 3</td>
<td>**</td>
</tr>
</tbody>
</table>

Another sort (y/n)? n

Table 3
Name of matrix file: alkane-ht
Ratio(r), Dissimilarity Index(d), CPI/OEP(c), New Matrix(m), or Stop(s)?

What range? C(even) to C(even)? 16, 28
Number of samples in table = ? 9
Input contiguous ranges before non-contiguous samples
Are samples contiguous (c) or not (n)? c
What range? low, high)? 1, 8
Are samples contiguous (c) or not (n)? n
Sample Number = ? 24

date = 01/16/84 matrix is alkane-ht

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Sample Name</th>
<th>CPI</th>
<th>Missing Alkanes</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Cape Simpson SL</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Umiat 4</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Fish Creek 1</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>South Barrow 19</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Dalton 1 DST 2</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Prudhoe - Put River</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Simpson Core Test 4</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>South Barrow 20</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Seabee 1 DST 3</td>
<td>1.13</td>
<td></td>
</tr>
</tbody>
</table>

Other samples with same range (y/n)? y

Number of samples in table = ? 15
Input contiguous ranges before non-contiguous samples
Are samples contiguous (c) or not (n)? c
What range? low, high)? 9, 23

date = 01/16/84 matrix is alkane-ht

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Sample Name</th>
<th>CPI</th>
<th>Missing Alkanes</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>Shublik Fm 9042</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Kingsak Shale 14577</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Fort. Mtn. Fm 12014</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Shublik Fm 12273</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Sadlerochit Gr 13831</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Shublik Fm 10273</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Kingsak Shale 7390</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Pebble shale 7371</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Pebble shale 6405</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Kingsak Shale 11704</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Pebble shale 6361</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Torok Fm 3792</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Torok Fm 7054</td>
<td>1.14</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Torok Fm 4994</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Pebble shale 1977</td>
<td>1.27</td>
<td></td>
</tr>
</tbody>
</table>

Other samples with same range (y/n)? n
Date = 01/16/84
Matrix is alkane-ht

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Sample Name</th>
<th>OEP</th>
<th>Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Cape Simpson SL</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>Seabee 1 DST 3</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>7</td>
<td>Dalton 1 DST 2</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>South Barrow 19</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Prudhoe - Put River</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>South Barrow 20</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Umiat 4</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Simpson Core Test 4</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Fish Creek 1</td>
<td>1.06</td>
<td></td>
</tr>
</tbody>
</table>

Other samples with same range (y/n)? y
Number of samples in table =? 15
Input contiguous ranges before non-contiguous samples
Are samples contiguous (c) or not (n)? c
What range (low, high)? 9, 23

Date = 01/16/84
Matrix is alkane-ht

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Sample Name</th>
<th>OEP</th>
<th>Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Shublik Fm 10273</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Fort. Mtn. Fm 12014</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Shublik Fm 9042</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Kingsak Shale 14577</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Sadlerochit Gr 13831</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Shublik Fm 12273</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Kingsak Shale 7390</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Pebble shale 6405</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Pebble shale 7371</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Kingsak Shale 11704</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Pebble shale 6361</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Torok Fm 7054</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Torok Fm 4994</td>
<td>1.23</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Torok Fm 3792</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Pebble shale 1977</td>
<td>1.57</td>
<td></td>
</tr>
</tbody>
</table>

Other samples with same range (y/n)? n

Table 5
Name of matrix file = hopane
Number of samples in table = 15
Are samples contiguous (c) or not (n)? c
What range (low, high)? 9-23
Use all (a) compounds in matrix or not (n)? a
Total number of compounds to be used = 39

Date = 01/16/84 matrix is hopane

Reference Sample Number (may be other than inputted samples) = 2

**DISSIMILARITY INDEX OF SAMPLE 2**  
South Barrow 20

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Sample Name</th>
<th>Dissimilarity Index</th>
<th>Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>Shublik Fm 9042</td>
<td>0.094</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Kinsak Shale 11704</td>
<td>0.111</td>
<td>0.016</td>
</tr>
<tr>
<td>11</td>
<td>Kinsak Shale 14577</td>
<td>0.111</td>
<td>0.001</td>
</tr>
<tr>
<td>12</td>
<td>Fort. Mtn. Fm 12014</td>
<td>0.152</td>
<td>0.041</td>
</tr>
<tr>
<td>14</td>
<td>Shublik Fm 12273</td>
<td>0.197</td>
<td>0.045</td>
</tr>
<tr>
<td>17</td>
<td>Pebble shale 6405</td>
<td>0.445</td>
<td>0.249</td>
</tr>
<tr>
<td>23</td>
<td>Kinsak Shale 7390</td>
<td>0.451</td>
<td>0.006</td>
</tr>
<tr>
<td>13</td>
<td>Sadlerochit Gr 13831</td>
<td>0.470</td>
<td>0.018</td>
</tr>
<tr>
<td>19</td>
<td>Pebble shale 7371</td>
<td>0.521</td>
<td>0.051</td>
</tr>
<tr>
<td>22</td>
<td>Torok Fm 4994</td>
<td>0.536</td>
<td>0.015</td>
</tr>
<tr>
<td>9</td>
<td>Pebble shale 6361</td>
<td>0.542</td>
<td>0.006</td>
</tr>
<tr>
<td>18</td>
<td>Shublik Fm 10273</td>
<td>0.550</td>
<td>0.008</td>
</tr>
<tr>
<td>20</td>
<td>Torok Fm 3792</td>
<td>0.556</td>
<td>0.006</td>
</tr>
<tr>
<td>16</td>
<td>Torok Fm 7054</td>
<td>0.558</td>
<td>0.003</td>
</tr>
<tr>
<td>10</td>
<td>Pebble shale 1977</td>
<td>0.658</td>
<td>0.100</td>
</tr>
</tbody>
</table>

Another Reference Sample for DI. (y/n)? n

Table 6
<table>
<thead>
<tr>
<th>TERPANES</th>
<th>STERANES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C19 TC</td>
<td>27baSD</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>C20 TC</td>
<td>27baRD</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>C21 TC</td>
<td>27abSD</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>C22 TC</td>
<td>27abRD</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>C23 TC</td>
<td>28baSD</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>C24 TC</td>
<td>A</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>C25 TC</td>
<td>B</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>A</td>
<td>28baRD</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>C26TC(S)</td>
<td>27aaaSS</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>C26TC(R)</td>
<td>29baSD</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>B'</td>
<td>C</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>27aaaRS</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>C28TC(S)</td>
<td>29baRD</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>C28TC(R)</td>
<td>D</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>C29TC(S)</td>
<td>E</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>C29TC(R)</td>
<td>28aaaSS</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Ts</td>
<td>28abbRS</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Tm'</td>
<td>28abbSS</td>
</tr>
<tr>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Tm</td>
<td>29aaaRS</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>C</td>
<td>29abbRS</td>
</tr>
<tr>
<td></td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>23</td>
</tr>
</tbody>
</table>

Table 7
### Terpanes (m/z 191)

<table>
<thead>
<tr>
<th>Peak No.</th>
<th>Compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>C19 tricyclic terpane</td>
</tr>
<tr>
<td>2.</td>
<td>C20 tricyclic terpane</td>
</tr>
<tr>
<td>3.</td>
<td>C21 tricyclic terpane</td>
</tr>
<tr>
<td>4.</td>
<td>C22 tricyclic terpane</td>
</tr>
<tr>
<td>5.</td>
<td>C23 tricyclic terpane</td>
</tr>
<tr>
<td>6.</td>
<td>C24 tricyclic terpane</td>
</tr>
<tr>
<td>7.</td>
<td>C25 tricyclic terpane</td>
</tr>
<tr>
<td>8.</td>
<td>A C24 tetracyclic terpane</td>
</tr>
<tr>
<td>9.</td>
<td>C26 tricyclic terpane (S7)</td>
</tr>
<tr>
<td>10.</td>
<td>C26 tricyclic terpane? (R?)</td>
</tr>
<tr>
<td>11.</td>
<td>B' tricyclic terpane?</td>
</tr>
<tr>
<td>12.</td>
<td>B' tricyclic terpane?</td>
</tr>
<tr>
<td>13.</td>
<td>C28 tricyclic terpane (S7)</td>
</tr>
<tr>
<td>14.</td>
<td>C28 tricyclic terpane (R?)</td>
</tr>
<tr>
<td>15.</td>
<td>C29 tricyclic terpane</td>
</tr>
<tr>
<td>16.</td>
<td>C29 tricyclic terpane (R7)</td>
</tr>
<tr>
<td>17.</td>
<td>18a(4),21B(H)-22,29,30-trisnorhopane (Ts)</td>
</tr>
<tr>
<td>18.</td>
<td>Tw' (C30) triterpane?</td>
</tr>
<tr>
<td>19.</td>
<td>17a(H),21B(H)-22,29,30-trisnorhopane (Ts)</td>
</tr>
<tr>
<td>20.</td>
<td>C (C28) triterpane?</td>
</tr>
<tr>
<td>21.</td>
<td>C' triterpane?</td>
</tr>
<tr>
<td>22.</td>
<td>D triterpane?</td>
</tr>
<tr>
<td>23.</td>
<td>D' (C29) triterpane?</td>
</tr>
<tr>
<td>24.</td>
<td>17a(H),21B(H)-30-normorexane</td>
</tr>
<tr>
<td>25.</td>
<td>E triterpane?</td>
</tr>
<tr>
<td>26.</td>
<td>17B(H),21a(H)-30-normoretane</td>
</tr>
<tr>
<td>27.</td>
<td>17a(H),21B(H)-hopenane</td>
</tr>
<tr>
<td>28.</td>
<td>17B(H),21a(H)-morexane</td>
</tr>
<tr>
<td>29.</td>
<td>17a(H),21B(H)-30-homohopane (22S)</td>
</tr>
<tr>
<td>30.</td>
<td>17a(H),21a(H)-30-homohopane (22R)</td>
</tr>
<tr>
<td>31.</td>
<td>Gammacerane?</td>
</tr>
<tr>
<td>32.</td>
<td>17B(H),21a(H)-30-homomoretane</td>
</tr>
<tr>
<td>33.</td>
<td>17a(H),21B(H)-30,31-bishomohopane (22S)</td>
</tr>
<tr>
<td>34.</td>
<td>17a(H),21B(H)-30,31-bishomohopane (22R)</td>
</tr>
<tr>
<td>35.</td>
<td>17B(H),21a(H)-30,31-bishomoretane</td>
</tr>
<tr>
<td>36.</td>
<td>17B(H),21B(H)-30,31,32-trishomohopane (22S)</td>
</tr>
<tr>
<td>37.</td>
<td>17B(H),21B(H)-30,31,32-trishomohopane (22R)</td>
</tr>
<tr>
<td>38.</td>
<td>17B(H),21B(H)-30,31,32,337-tetrakishomohopane (22S)</td>
</tr>
<tr>
<td>39.</td>
<td>17B(H),21B(H)-30,31,32,337-tetrakishomohopane (22R)</td>
</tr>
</tbody>
</table>

### Steranes (m/z 217)

<table>
<thead>
<tr>
<th>Peak No.</th>
<th>Compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>13B(H),17a(H) diacholestane (20S)</td>
</tr>
<tr>
<td>2.</td>
<td>13B(H),17a(H) diacholestane (20R)</td>
</tr>
<tr>
<td>3.</td>
<td>13a(H),17B(H) diacholestane (20S)</td>
</tr>
<tr>
<td>4.</td>
<td>13a(H),17B(H) diacholestane (20R)</td>
</tr>
<tr>
<td>5.</td>
<td>13B(H),17a(H) diasterane (20S)</td>
</tr>
<tr>
<td>6.</td>
<td>13B(H),17a(H) diasterane (20R)</td>
</tr>
<tr>
<td>7.</td>
<td>13a(H),14a(H),17a(H) cholestane (20S)</td>
</tr>
<tr>
<td>8.</td>
<td>13a(H),14a(H),17a(H) cholestane (20R)</td>
</tr>
<tr>
<td>9.</td>
<td>13B(H),17a(H) diastigastane (20S)</td>
</tr>
<tr>
<td>10.</td>
<td>13B(H),17a(H) diastigastane (20R)</td>
</tr>
<tr>
<td>11.</td>
<td>C Sterane?</td>
</tr>
<tr>
<td>12.</td>
<td>5a(H),14a(H),17a(H) cholestane (20R)</td>
</tr>
<tr>
<td>13.</td>
<td>13B(H),17a(H) diastigastane (20R)</td>
</tr>
<tr>
<td>14.</td>
<td>D Sterane?</td>
</tr>
<tr>
<td>15.</td>
<td>E Sterane?</td>
</tr>
<tr>
<td>16.</td>
<td>5a(H),14a(H),17a(H) ergostane (20R)</td>
</tr>
<tr>
<td>17.</td>
<td>5a(H),14B(H),17B(H) ergostane (20R)</td>
</tr>
<tr>
<td>18.</td>
<td>5a(H),14B(H),17B(H) ergostane (20S)</td>
</tr>
<tr>
<td>19.</td>
<td>5a(H),14a(H),17a(H) ergostane (20R)</td>
</tr>
<tr>
<td>20.</td>
<td>5a(H),14B(H),17B(H) stigmastane (20S)</td>
</tr>
<tr>
<td>21.</td>
<td>5a(H),14B(H),17B(H) stigmastane (20R)</td>
</tr>
<tr>
<td>22.</td>
<td>5a(H),14B(H),17B(H) stigmastane (20S)</td>
</tr>
<tr>
<td>23.</td>
<td>5a(H),14B(H),17B(H) stigmastane (20R)</td>
</tr>
</tbody>
</table>

Table 8
APPENDIX

Listing of Computer Program OIL/ROCK.BASIC
10  DIM c(40), d(9), f(25*2), m(25, 40), n(9), r(25, 9), s(25)
20  PRINT "Name of matrix file=?"
30  INPUT f$
40  INPUT *file $1:f$
50  reset $1:0
60  MAT READ #1*m
70  LET c* = f*$"-str"
80  File *2 'c$
90  File *3: 'ns-name$
100 reset *3:0
110  PRINT "Ratio(r), Dissimilarity Index(d), Cpi/Dep(c), New Matrix(m), or Stop(s)?"
120  input p$
130  IF SEG*(p$[1] = "r" then 460
140  IF SEG*(p$[1] = "c" then 1390
150  IF SEG*(p$[1] = "d" then 2410
160  IF SEG*(p$[1] = "s" then 4820
170  IF SEG*(p$[1] = "m" then 30
180  GO TO 120
190  PRINT "Number of samples in table =";
200  INPUT s
210  PRINT "Input contiguous ranges before non-contiguous samples"
220  LET x = 1
230  PRINT "Are samples contiguous(c) or not(n)?"
240  INPUT m$
250  IF m$ = "c" then 290
260  IF m$ = "n" then 400
270  GO TO 240
280  REM contiguous samples
290  PRINT "What range(low, high)?"
300  INPUT l, h1
310  IF l < h1 then 340
320  GO TO 300
330  FOR y = l to h1
340    LET s(x) = y
350  LET x = x + 1
360  NEXT y
370  IF x <= s then 240
380  GO TO 450
390  REM non-contiguous samples
400  PRINT "Sample Number =";
410  INPUT s(x)
420  LET x = x + 1
430  IF x <= s then 410
440  GO TO 450
450  RETURN
460  REM Ratio Program
470  GOSUB 200
480  PRINT "Number of ratios in table(9 max.) =";
490  INPUT r
500  IF r > 9 then 480
510  MAT r = zer
520  PRINT "Multiple Numerator/Denominators(y/n)?"
530  INPUT p$
540  IF p$ = "y" then 3810
550  IF p$ = "n" then 570
560  GO TO 520
570  FOR x = 1 to r
580  PRINT "Ratio ;" x ;" numerator compound number =";
590  INPUT n(x)
600  PRINT "Ratio ;" x ;" denominator compound number =";
610  NEXT x
610 INPUT d(x)
620 NEXT x
630 FOR x = 1 TO r
640 FOR y = 1 TO s
650 IF m(s(y), d(x)) <> 0 THEN 680
660 LET r(s(y), x) = 9999
670 GO TO 690
680 LET r(s(y), x) = m(s(y), n(x)) / m(s(y), d(x))
690 NEXT y
700 NEXT x
710 Gosub 910
720 Gosub 1240
730 PRINT "Print table with ratios sorted(y/n)"
740 INPUT p$
750 IF p$ = "n" THEN 120
760 Gosub 1160
770 Gosub 3390
780 Gosub 910
790 Gosub 1240
800 PRINT "Another sort(y/n)"
810 INPUT p$
820 IF p$ = "y" THEN 760
830 IF p$ = "n" THEN 800
840 PRINT "Use same ratios with different samples(y/n)"
850 INPUT p$
860 IF p$ = "n" THEN 120
870 IF p$ = "y" THEN 890
880 GO TO 840
890 Gosub 200
900 GO TO 630
910 REM Print SIMPLE RATIO heading
920 PRINT
930 PRINT "date=", d3t$, "matrix is "
940 PRINT
950 PRINT "Sample   Sample   ";
960 FOR x = 1 TO r
970 reset #2:n(x) - 1
980 READ #2:n$
990 PRINT using", "n$;";
1000 NEXT x
1010 PRINT
1020 PRINT "tab(30);
1030 FOR x = 1 TO r
1040 PRINT "------- ";
1050 NEXT x
1060 PRINT
1070 PRINT "Number   Name ";
1080 FOR x = 1 TO r
1090 reset #2:d(x) - 1
1100 READ #2:n$
1110 PRINT using", "n$;";
1120 NEXT x
1130 PRINT
1140 PRINT
1150 RETURN
1160 PRINT "Sort according to which ratio";
1170 INPUT m
1180 IF m <= r THEN 1200
1190 GO TO 1160
1200 FOR x = 1 TO s
1210 \text{let } f(x;2) = r(s(x),m) \\
1220 \text{next } x \\
1230 \text{return} \\
1240 \text{for } n=1 \text{ to } s \\
1250 \text{reset } \#3:s(n)-1 \\
1260 \text{read } \#3:s$ \\
1270 \text{print using"} -r s(n),s$; \\
1280 \text{for } x=1 \text{ to } r \\
1290 \text{if } r(s(n),x)<>9999 \text{ then } 1320 \\
1300 \text{print"} ** "; \\
1310 \text{go to } 1330 \\
1320 \text{print using"} -### *** ";r(s(n),x); \\
1330 \text{next } x \\
1340 \text{print} \\
1350 \text{next } n \\
1360 \text{print} \\
1370 \text{print} \\
1380 \text{return} \\
1390 \text{rem CPI/OEP Calculations} \\
1400 \text{if } \text{ses$(f$,1,6)}="alkane" \text{ then } 1430 \\
1410 \text{print } "WRONG MATRIX" \\
1420 \text{go to } 30 \\
1430 \text{print"CPI(c) or OEP(p)";} \\
1440 \text{input } p$ \\
1450 \text{if } p$="P" \text{ then } 2040 \\
1460 \text{if } p$="C" \text{ then } 1480 \\
1470 \text{go to } 1430 \\
1480 \text{print"What range?C(even) to C(even)";} \\
1490 \text{input } l,h \\
1500 \text{if } 2*\text{int}(1/2)<1 \text{ then } 1530 \\
1510 \text{if } 2*\text{int}(h/2)<h \text{ then } 1530 \\
1520 \text{go to } 1550 \\
1530 \text{print"Input even numbers only"} \\
1540 \text{go to } 1480 \\
1550 \text{mat } f=zer \\
1560 \text{gosub } 200 \\
1570 \text{print} \\
1580 \text{print"date";}dat$;"matrix is ";f$ \\
1590 \text{print} \\
1600 \text{print"Sample";}tab(10);"Sample";tab(33);"CPI" \\
1610 \text{print tab(31);} \\
1620 \text{print using"n_C-#### Missing",l,h} \\
1630 \text{print"Number Name";}tab(30); \\
1640 \text{print using"Center=-# Alkanes";}(l+h)/2 \\
1650 \text{print} \\
1660 \text{for } n=1 \text{ to } s \\
1670 \text{for } i=1 \text{ to } l+2 \\
1680 \text{if } m(s(n),i-11)<0 \text{ then } 1710 \\
1690 \text{let } f(s(n),i)=i \\
1700 \text{go to } 1830 \\
1710 \text{next } i \\
1720 \text{let } a=b=c=0 \\
1730 \text{for } i=1 \text{ to } h-2 \text{ step } 2 \\
1740 \text{if } m(s(n),i-9)=0 \text{ then } 1810 \\
1750 \text{let } a=a+m(s(n),i-10) \\
1760 \text{let } b=b+m(s(n),i-11) \\
1770 \text{let } c=c+m(s(n),i-9) \\
1780 \text{next } i \\
1790 \text{let } f(n,2)=a/2*(1/b+1/c) \\
1800 \text{go to } 1830
1810 'let f(s(n),1)=i+2
1820 let f(n,2)=a/2*(1/b+1/c)
1830 next n
1840 dosub 3390
1850 for n=1 to s
1860 reset $3:s(n)-1
1870 read $3:s$
1880 print using"* -#### <######################################################## ",’s(n),s$;
1890 if f(s(n),1)=0 then 1930
1900 if f(n,2)=0 then 1950
1910 print using"-##",f(n,2),f(s(n),1)
1920 go to 1790
1930 print using"-##",f(n,2)
1940 go to 1790
1950 print tab(44);
1960 print using"-##",f(s(n),1)
1970 next n
1980 print
1990 print "Other samples with same range (y/n)";
2000 input $p$
2010 if $p=$"y" then 1550
2020 if $p=$"n" then 120
2030 go to 1990
2040 print"DEF centering around what carbon number";
2050 input 1
2060 let 1=1-2
2070 mat f=zer
2080 sosub 200
2090 print
2100 print"date=",dat$,"matrix is ";$f$
2110 print
2120 print"Sample";tab(11);"Sample";tab(33);"OEP";tab(42);"Missing"
2130 print"Number";tab(30);
2140 print using"Center=--# Alkanes",l+2
2150 print
2160 for n=1 to s
2170 for i=1 to 1+4
2180 if m(s(n),i-11)<>0 then 2210
2190 let f(s(n),1)=i
2200 go to 2230
2210 next i
2220 let f(n,2)=((m(s(n),l-11)+6*m(s(n),l-9)+m(s(n),1-7))/(4*m(s(n),l-10)+4*m(s\c(n),1-8)))^(-1"(1+1))
2230 next n
2240 sosub 3390
2250 for n=1 to s
2260 reset $3:s(n)-1
2270 read $3:s$
2280 print using"* -#### <######################################################## ",’s(n),s$;
2290 if f(s(n),1)=0 then 2330
2300 print tab(44);
2310 print using"-##",f(s(n),1)
2320 go to 2340
2330 print using"-##",f(n,2)
2340 next n
2350 print
2360 print"Other samples with same range(y/n)";
2370 input $p$
2380 if $p=$"y" then 2070
2390 if $p=$"n" then 120
2400 go to 2360
2410. rem dissimilarity index
2420gosub 200
2430print"Use all (a) compounds in matrix or not (n)?";
2440input p$
2450 print"Total number of compounds to be used=";
2460 input n
2470if p$="a" then 2500
2480if p$="n" then 2540
2490 go to 2430
2500for x=1 to n
2510let c(x)=x
2520next x
2530 go to 2850
2540let x=1
2550print"Input contiguous ranges before non-contig. compounds;"
2560print"Contiguous Compounds (y/n)?";
2570input p$
2580if p$="y" then 2660
2590if p$="n" then 2610
2600 go to 2560
2610print"Compound Number=";
2620input c(x)
2630let x=x+1
2640if x>n then 2750
2650 go to 2610
2660print"What range (low, high)?";
2670 input l,h
2680if l<h then 2700
2690 go to 2660
2700for y=1 to h
2710let c(x)=y
2720let x=x+1
2730next y
2740if x<=n then 2560
2750 rem sort in increasing order
2760for x=n to 2 step -1
2770for y=x-1 to 1 step -1
2780if c(y)<c(x) then 2820
2790let u=c(y)
2800let c(y)=c(x)
2810let c(x)=u
2820next y
2830next x
2840print
2850rem create normalized matrix of all samples
2860for y=1 to 25
2870let t=0
2880for x=1 to n
2890let t=t+m(y,c(x))
2900next x
2910if t<>0 then 3090
2920let m(y,0)=1
2930for x=1 to n
2940let m(y,x)=0
2950next x
2960let x=1
2970if s(x)=y then 3010
2980let x=x+1
2990if x<s then 2970
3000 go to 3120
3010 reset $3:3(y)-1
3020 read $3:n$
3030 gosub 3050
3040 go to 3120
3050 print
3060 print"NOTE-NONE OF THE INDICATED COMPOUNDS ARE PRESENT IN ";n$
3070 print
3080 return
3090 for x=1 to n
3100 let m(y,x)=m(y,c(x))/t
3110 next x
3120 next y
3130 print
3140 print"date";dat$;"matrix is ";f$
3150 print
3160 print"Reference Sample Number(May be other than inputted samples)=";
3170 input a
3180 reset $3:a-1
3190 read $3:n$
3200 if m(a,0)=0 then 3230
3210 gosub 3050
3220 go to 3710
3230 print
3240 print'DISSIMILARITY INDEX OF SAMPLE ";a;" ";n$
3250 print
3260 for y=1 to s
3270 let t=0
3280 for x=1 to n
3290 let t=t+abs(m(s(y),x)-m(a,x))
3300 next x
3310 let f(y,2)=t/2
3320 next y
3330 print"Sample Number Name Diss. Delta"
3340 print"Sample Sample Diss. Delta"
3350 print
3360 gosub 3390
3370 let f=s=0
3380 go to 3520
3390 rem sort in increasing order
3400 for x=1 to s-1
3410 for y=x+1 to s
3420 if f(x,2)<f(y,2) then 3490
3430 let t=f(x,2)
3440 let u=s(x)
3450 let f(x,2)=f(y,2)
3460 let s(x)=s(y)
3470 let f(y,2)=t
3480 let s(y)=u
3490 next y
3500 next x
3510 return
3520 for x=1 to s
3530 if s(x)=a then 3680
3540 if m(s(x),0)=0 then 3570
3550 let g=g+1
3560 go to 3680
3570 if m(s(x-1),0)=1 then 3600
3580 let d=f(x,2)-f(x-1,2)
3590 go to 3610
3600 let d=f(x,2)-f(x-1-g,2)
3610 *reset *3:s(x)-1
3620 read *3:n$
3630 if f=1 then 3670
3640 let f=1
3650 print using" -# "**"************** -***",s(x),n$,f(x,2)
3660 go to 3680
3670 *print using" -# "**"************** -***",s(x),n$,f(x,2),d
3680 next x
3690 *print
3700 *print
3710 *printAnother Reference Sample for D.I.(y/n)*;
3720 *input p$
3730 if p$="y" then 3160
3740 if p$="n" then 3760
3750 go to 3710
3760 rem must recreate matrix m
3770 for x=1 to 25
3780 let m(x,0)=0
3790 next x
3800 go to 60
3810 rem Multiple numerators/denominators Ratios
3820 rem Fill numerator/denominator matrix
3830 *print"Maximum Numerators or Denominators is 9 each."
3840 for x=1 to r
3850 *print"Number of numerators in ratio";x;"=";
3860 input n
3870 if n>9 then 3850
3880 let r(10,x)=n
3890 for y=1 to n
3900 *print" Numerator";y;"=";
3910 *input r(y,x)
3920 next y
3930 *print"Number of denominators in ratio";x;"=";
3940 input d
3950 if d>9 then 3930
3960 let r(20,x)=d
3970 for y=1 to d
3980 *print" Denominator";y;"=";
3990 *input r(y+10,x)
4000 next y
4010 next x
4020 for x=1 to r
4030 *print"Ratio";x
4040 print* Numerators=";
4050 for y=1 to r(10,x)
4060 *reset *2:r(y,x)-1
4070 *read *2:n$
4080 if y=r(10,x) then 4110
4090 *print n$;"+";
4100 go to 4120
4110 *print n$
4120 next y
4130 *print Denominators=";
4140 for y=1 to r(20,x)
4150 *reset *2:r(y+10,x)-1
4160 *read *2:n$
4170 if y=r(20,x) then 4200
4180 *print n$;"+";
4190 go to 4210
4200 *print n$
4210
4200 \textbf{end}