DEVELOPMENT OF A MICROCOMPUTER PROGRAM TO CURVE-FIT PETROLEUM DATA

Clifton L. Adams II

Texas Tech University
Lubbock, Texas

Abstract:

This paper reviews basic techniques required to determine an equation suitable to describe a given set of data points. In general, the data is matched against ten of the more common curves encountered by the petroleum engineer. The conversion of each of the ten curves into a psuedo-linear equation is discussed. Coefficients for each equation are determined by the least squares method and the method of averages.

A microcomputer program that incorporates these concepts is presented. The program is written in BASIC for the Apple II microcomputer.

Introduction:

With the advent of programmable calculators and, more recently, personal computers, it is not surprising that there is a growing interest in the ability to determine an equation suitable to describe a given curve. For those with limited exposure to the higher levels of mathematics, an approach which attempts to correlate any set of ordered pairs (x,y) to the most basic of all curves, the straight line, would seem most appreciated.

I am entirely indebted to Dr. R. E. Carlile, Chairman of the Petroleum Engineering Department, Texas Tech University, and his book, "FORTRAN and Computer Mathematics for the Engineer and Scientist,"[1] for the format and mathematical content presented in this paper. My contribution consists of the development of the BASIC computer program and documentation that embodies these concepts.

Overview:

The equation for a straight line should readily be recognized as:

$$Y = A*X + B \tag{Eq. 1}$$

where A is the slope of the line and B is the point where the line intersects the y axis. Thus, given any two points on the line (X1,Y1) and (X2,Y2), the equation for this curve can be determined calculated as:

$$A = \frac{Y2 - Y1}{X2 - X1} \tag{Eq. 2}$$

and therefore:

$$B = Yi - A*Xi \qquad \text{(Eq. 3)}$$

The first discussion will develop two methods of determining the coefficients A and B by an averaging process. Each of the these methods will provide its own coefficients which are then tested with each linearized function for goodness of fit. For each of the ten linearized functions, two equations will be developed that can generate a dependent variable (Yi') which will be compared to the observed variable (Yi). The equation with the lowest overall deviation factor is selected for a detailed print-out which includes the equation, complete with coefficients, a listing of the actual input data (X,Y), the calculated Y, and the percent deviation for each point as well as the overall deviation factor.

The BASIC program that accomplishes this task was written in APPLESOFT II BASIC and 3.3 DOS on a 64K Apple IIe microcomputer with one disk drive and an optional Epson RX-80 dot-matrix printer.

Determination of Coefficients A and B:

As discussed earlier, once we have our desired generic linear function, an accounting must be made for either the error inherent in collecting field data or having to deal with nonconforming data. In the simpliest terms, the collected points will not usually fall on any single straight line drawn through them. Two methods will now be discussed to deal with this problem. This section will also serve as the discussion of the test for simple linearity.

Method of Averages:

This method, hereafter referred to as MOA, basically utilizes a simple averaging technique. We have two unknown values to find, A and B, and so given the two following equations:

$$Y1 = A*X1 + B \qquad \text{(Eq. 3)}$$
$$Y2 = A*X2 + B \qquad \text{(Eq. 4)}$$

If X1, X2, Y1, and Y2 are known, Eqs. 3 and 4 can be solved simultaneously for A and B.

We can let X1 and Y1 equal the average value of the first half of the actual X and Y values respectively while X2 and Y2 equal the average value of the last half of the actual X and Y values respectively. The effect of stray points is thus somewhat minimized. With these values substituted into Eqs. 3 and 4 they can now be solved simultaneously for A and B.

As an example, temperature in the earth in most cases is known to be a simpler linear function of depth and surface temperature. Referring to Table 1, column 1 contains a hypothetical series of depths at which the temperatures found in column 2 were recorded.

Applying the MOA technique:

Depth	Temp
0	72
1000	87
2000	108
3000	128
+ 4000	+ 141
10000	536
5000	164
6000	181
7000	199
8000	214
9000	235
+ 10000	+ 251
45000	1244

X1 = 10000/5 = 2000

Y1 = 536/5 = 107.2

X2 = 45000/6 = 7500

Y2 = 1244/6 = 207.33

Therefore Eqs. 3 and 4 become:

107.2 = A + 2000*B
207.33 = A + 7500*B

which solved simultaneously yield:

A = 70.8
B = 0.0182

which substituted into Eq. 1 yields: Yi = 70.8 + 0.0182*Xi (Eq. 4a)

Column 3 of Table 1 is the calculated temperature at depth Xi using this equation. Column 4 of Table 1 shows the deviation of the calculated temperature from the observed temperature. Percent deviation here and throughout this paper is calculated as follows:

$$\% \text{ Dev} = \frac{Yi - Yi'}{Yi} * 100 \qquad \text{(Eq. 5)}$$

Figure 1 shows observed temperature values along with temperature calculated with Eq. 4a.

Least Squares Method:

The Least Squares Method (LS) is a more sophisticated and in most cases more accurate technique than the Method of Averages for determining the coefficients of a linear function. The basic concept involves the assumption that there exists a polynomial such that the sum of squares of the distance between the dependent variable generated by this polynomial (Yi') and the actual dependent variable (Yi) is a minimum.

If we let: $$Yi' = A + B*Xi \qquad \text{(Eq. 6)}$$

then: $$\sum_{i=1}^{n} (Yi = Yi')^2 = \sum_{i=1}^{n} (Yi - A - B*Xi)^2 \qquad \text{(Eq. 7)}$$

To find the minimum, we take the partial derivatives of Eq. 7 and set each equal to 0.

Thus: $$\frac{d(\text{Eq. 7})}{d(A)} = -2*\sum_{i=1}^{n} (Yi - A - B*Xi) = 0 \qquad \text{(Eq. 8)}$$

$$\frac{d(\text{Eq. 7})}{d(B)} = -2*\sum_{i=1}^{n} (Yi - A - B*Xi) * (Xi) = 0 \qquad \text{(Eq. 9)}$$

where n is the number of data points (Xi, Yi).

Eqs. 8 and 9 can be rewritten:

$$(\text{Eq. 8}) = \sum_{i=1}^{n} (Yi) = (A*n) + B*\sum_{i=1}^{n} Xi \qquad \text{(Eq. 10)}$$

$$(\text{Eq. 9}) = \sum_{i=1}^{n} (Xi*Yi) = (A*Xi) + B*\sum_{i=1}^{n} (Xi)^2 \qquad \text{(Eq. 11)}$$

As Xi and Yi are given, there are two equations (Eqs. 10 and 11) and two unknowns (A and B) from which A and B can easily be calculated.

To illustrate this method, refer again to Table 1.

$$\sum_{i=1}^{n} (Xi) = 55000$$

$$\sum_{i=1}^{n} (Yi) = 1780$$

$$\sum_{i=1}^{n} (Xi) = 3.85 * 10^8$$

$$\sum_{i=1}^{n} (Xi * Yi) = 1.0887 * 10^7$$

Substituting these values into Eqs. 10 and 11:

$$1780 = 11*A + 5500*B \qquad \text{(Eq. 10)}$$

$$1.0887*10^7 = 5500*A + 3.85*10^8*B \qquad \text{(Eq. 11)}$$

Solving these two equations simultaneously:

$$A = 71.3181$$

$$B = 0.0181$$

With these values substituted into Eq. 1, the values found in column 3a of Table 1 were generated. The percent average deviation found in column 4a was calculated with Eq. 5.

In this case, the overall deviation using the least squares derived coefficient (1.04%) is slightly lower than that using the MOA derived coefficient (1.19%). The computer program will evaluate cases with both methods as there are situations were the MOA will yield better results.

Semilog:

Consider the following equation:

$$\ln Yi = A + B*Xi \qquad \text{(Eq. 12)}$$

Let $Yi' = \ln Yi$.

Then

$$Yi' = A + B*Xi \qquad \text{(Eq. 13)}$$

It should be apparent to the reader that Eq. 13 will yield a straight line with slope B and y-intercept A. The previous discussion of coefficient determination will apply directly to Eq. 13. Once these coefficients have been found they can be substituted back into Eq. 2 to serve as the general equation.

To illustrate, Srini-Vasan[2] has shown that, in general, the log of the ratio of either the plastic viscosity or apparent viscosity of a drilling fluid to the viscosity of water at the same temperature is equal to a linear function of that temperature. That is:

$$\ln(FV/WV) = A + B*T$$

where FV = Viscosity of drilling fluid (cp) @ T ('F)
and WV = Viscosity of water (cp) @ T ('F)

Given the data in columns 1, 2 and 3 of Table 2a, an equation can be developed to predict an unobserved viscosity ratio at various temperatures. Column 4 is the ratio of column 2/column 3. Column 5 is the natural log of column 4. Column 5 values correspond to the Yi' variable in Eq. 13. These values, with the values in column 1 as the Xi values, allow the A and B coefficients to be calculated by the MOA and LS techniques discussed in the previous section.

By the MOA method:

$$A = 1.8689$$
$$B = 0.0056$$

With these values substituted into Eq. 12,

$$\ln Yi = 1.8689 + 0.0056 \qquad \text{(Eq. 14)}$$

This equation generated the calculated viscosity ratios in column 6. Using column 6, the calculated ratios, with column 4, the observed ratios, by Eq. 5, gives the percent deviation found in column 7.

By the LS method:

$$A = 1.8686$$
$$B = 0.0056$$

which produces:

$$\ln Yi = 1.8686 + 0.0056*Xi \qquad \text{(Eq. 15).}$$

Table 2b summarizes the results using Eq. 15. Figure 2 plots the results of both methods.

Semilog with Curvature:

Sometimes the plot of the log of the dependent variable (Yi) against the dependent variable (Xi) will yield a straight line for part of the data range and a curve over another section of the range. Figure 3[3] is a linear plot of such a situation. One method of linearizing this curve is to determine a shift factor (s) such that:

$$\ln (Yi - s) = A + B*Xi \qquad \text{(Eq. 16)}$$

yields a straight line over the entire data set. The shift factor (s) can be determined as follows:

$$s = \frac{Yn*Y1 - Ym^2}{Yn + Y1 - 2*Ym} \qquad \text{(Eq. 17)}$$

where: Y1 = First Yi in data set (Xi,Yi)
Yn = Last Yi in data set (Xi,Yi)
Ym = Yi corresponding to Xm, the arithmetic mean of Xi
$Xm = \frac{X1 + Xn}{2}$ (X1 = First Xi;Xn = Last Xi)

Columns 1 and 2 of Table 3 are values taken from Figure 2. Using this data to determine the shift factor (s):

Y1 = 1.56
Yn = 0.31
Xm = (40 + 200)/2 = 120

At Xi = 120, Yi = 0.56; therefore Ym = 0.56.
(Note: If Xm is equal to an unobserved value of Xi, Ym can be obtained by linear interpolation).

Substituting these values into Eq. 17 gives:

$$s = \frac{(0.31 * 1.56) - (0.56^2)}{(0.31 + 1.56) - (2*0.56)} = 0.2267$$

Using the least squares method, the coefficients A and B are found to be:

A = 0.9037
B = -0.0164

Substituting these values into Eq. 16, we have:

$$\ln(Yi - 0.2267) = 0.9037 - 0.0164*Xi \qquad \text{(Eq. 18)}$$

This equation, using column 1 values for Xi, generate the water viscosity values shown in column 3. Column 4 is the percent deviation of column 3 from column 2. The worst individual point fit (4.26%) is less than the width of the curve in Figure 2, making Eq. 18 a good substitute for Figure 2.

Log-Log:

When data pairs have a linear log-log relationship, the general equation can be stated:

$$\ln Yi = \ln A + B*(\ln Xi) \qquad \text{(Eq. 19)}$$

Letting: $Yi' = \ln Yi$
$A' = \ln A$
$Xi' = \ln Xi$

Eq. 19 can be written: $$Yi' = A' + B*Xi' \qquad \text{(Eq. 20)}$$

which is the form of Eq. 1. The coefficients A' and B can be determined as before by either the MOA or LS method and substituted into Eq. 19 for the general equation of the curve.

As an example, the pressure drop in a laminar pipe flow has a log-log relationship to the flow rate, all other factors remaining constant. Table 4, columns 1 and 2 are actual experimental data reported by Melton and Saunders[4] for a water suspension of calcium carbonate (73.5 lbm/cu ft.) through a 6.92 ft. horizontal steel pipe of 0.278-in. ID.

The least squares method of coefficient determination yields:

$$A' = 0.7202$$
$$B = 0.2754$$

With substitution in Eq. 19, the general equation becomes:

$$\ln Yi = 0.7202 + 0.2754 \qquad \text{(Eq. 21)}$$

Column 3 of Table 6 is the calculated pressure drop using Eq. 21. The over-all fit is a very good 0.98%.

Parabolic:

The standard equation for a parabola is:

$$Yi = C + A*Xi + B*X^2 \qquad \text{(Eq. 22)}$$

In order to linearize this function, consider two points selected from the data set, (X1,Y1) and (Xi,Yi). With each of these pairs of values substituted into Eq. 22, we have:

$$Yi = C + A*Xi + B*Xi^2$$
$$Y1 = C + A*X1 + B*X1^2$$

which solved simultaneously produces:

$$Yi - Y1 = A*(Xi - X1) + B*(Xi^2 - X1^2)$$

This equation can be manipulated to read:

$$(Yi - Y1)/(Xi - X1) = A + B*(Xi + X1)$$

Letting: $(Yi - Y1)/(Xi - X1) = Yi'$
and $Xi + X1 = Xi'$

we have: $$Yi' = A + B*Xi' \quad \text{(Eq. 23)}$$

This is the familiar form for which the coefficients can be derived as usual by either the least squares or MOA method. Once A and B have been determined, they can be substituted into Eq. 22 with any known values (Xi, Yi) from the data set to find the value of C. When all three coefficients (A,B,C) are substituted into Eq. 22, you have the general equation for the curve.

To illustrate the parabolic correlation, refer to Table 5. Columns 1 and 2 were taken from an SPE paper on phase equilibria in carbon dioxide/hydrocarbon systems presented by Turek, et al[5] in 1979. Column 1 is the mole percent CO_2 in a mixture with a particular composition of synthetic oil. Column 2 is the corresponding saturation pressure at a temperature of 322 deg. K. Applying the least squares routine, the following coefficients were determined:

$$A = 0.0891$$
$$B = -0.001$$
$$C = 10.197$$

Substitution into Eq. 22 yields the general equation:

$$Yi = 10.197 + 0.0891*Xi = 0.001*(Xi^2)$$

The calculated values of saturation pressure using this equation are recorded in column 3 with the deviation percent in column 4. The overall deviation percent was 0.48%, indicating a very good fit.

Hyperbolic:

A general equation for a hyperbola can be written in the following manner:

$$\frac{Xi - X1}{Yi - Y1} = A + B*Xi \qquad \text{(Eq. 24)}$$

where Xi and Y1 are equal to the first ordered pair in the data set. Solving for Yi:

$$Yi = \frac{Xi - X1}{A + B*Yi} + Y1 \qquad \text{(Eq. 25)}$$

Letting:

$$Yi' = \frac{Xi - X1}{Yi - Y1} \qquad \text{(Eq. 26)}$$

which substituted in Eq. 24 yields:

$$Yi' = A + B*Xi$$

which is the general form of a linear equation. The coefficients A and B are determined as before and substituted back into Eq. 25 to produce a general equation for the data set.

To illustrate, Black[6] determined that in reservoirs whose temperature ranged from 140-240 degrees F, the oil viscosity at reservoir conditions could be related to the API gravity of the stock tank oil flashed at 0 psig separator pressure by the curve shown in Figure 6. Table 6 lists selected values taken from this chart. Column 1 contains the oil viscosities (Xi), while column 2 contains the corresponding API gravity values (Yi). In column 3 are the Yi' values determined by Eq. 26 with X1 = 0.325 cp and Y1 = 45.0 API. When these values run through the least squares and MOA routines, the MOA method gives the best coefficents as:

$$A = 0.002$$
$$B = -0.0433$$

With these values substituted into Eq. 25, the general equation for the curve becomes:

$$Yi = 45 + (Xi - 0.325)/(0.002 - 0.0433*Xi)$$

Using this equation and the Xi values in column 1, the calculated Yi's in column 4 were determined. Column 5 lists the percent deviation of each calculated Yi from the actual Yi of column 2. The overall fit for this curve is a good 1.19%.

Sigmoidal:

A sigmoidal function is easily recognized by the characteristic S-shaped curve it produces. Petroleum engineers see this relationship in the case of permeability vs. saturation curves, fractional flow vs. water saturation curves as well as others.

To correlate this function, we will depart from the usual pattern of variable manipulation into some psuedo-linear relationship from which coefficients can be derived. Instead, a technique using the modified Gompertz equation will be developed.

For this correlation to work, the independent variables (Xi) must be separated by a constant amount; they must be an equal distance apart. If this is not the case, a plot of the data can be made from which values of Xi can be taken at equal intervals. The modified Gompertz equation takes the form:

$$Yi = S + A*B^{C^{Xi/D-1}} \qquad \text{(Eq. 27)}$$

The coefficients S, A, B and D are determined in the following manner:

1) S = Y(0)
   S is equal to the value of Y corresponding to X = 0.
   If there is no X = 0 then S = 0.

2) D = X(2)-X(1)
   D is equal to the equalized interval distance between all Xi values.

3) $C = ((S2 - S3)/(S1 - S2))^{-N}$ (Eq. 27a)
   where: S1 = sum of first third ln(Yi-S) values
   S2 = sum of second third ln(Yi-S) values
   S3 = sum of last third ln(Yi-S) values
   N = INT (# of data points (Xi,Yi))/3
   (Integer value of nearest multiple of 3)

4) $$A = \ln^{-1}\left\{(1/N) * \left\{S1 - \frac{(S1-S2)}{(1-C)^N}\right\}\right\} \qquad \text{(Eq. 29)}$$

5) $$B = \ln^{-1}\left\{\frac{(S1-S2)*(1-C)}{(1-C^N)^2}\right\} \qquad \text{(Eq. 30)}$$

The following illustration should serve to clarify this procedure.

The specific heat of air (Cp) at 0-1 atmospheres is related to the temperature ('F) in a sigmoidal manner.[7] Column 2 of Table 7 lists the specific heat values for air (Yi) at the temperature found in column 1 (Xi). Column 3 is the intermediate ln (Yi-S) values with:

S = Y(0) = 0.2420
D = X(2) - X(1) = 100 - 0 = 100
N = 12/3 = 4
S1 = sum of first third terms of column 3 = -26.1959
S2 = sum of second third terms of column 3 = -17.8481
S3 = sum of third third terms of column 3 = -14.8744

By Eq. 28:

$$C = \left\{ \frac{(-17.8481) - (-14.8744)}{(-26.1959) - (-17.8481)} \right\}^4 = 0.7726$$

By Eq. 29:

$$A = \ln^{-1} \left\{ (1/4) * \left\{ (-26.1959) - \frac{(26.1959) - (-17.8481)}{(1 - 0.7726^4)} \right\} \right\}$$

$$A = 0.366$$

By Eq. 30:

$$B = \ln^{-1} \left[\frac{(-26.1959) - (-17.8481) * (1 - 0.7726)}{(1 - 0.7726^4)^2} \right]$$

$$B = 0.0104$$

These values substituted into Eq. 27 yield the general equation for the data set:

$$Yi = 0.242 + 0.0366 * 0.0102^{0.7726^{(Xi/100 - 1)}}$$

This equation produces the calculated values found in column 4. Column 5 is the individual percent deviation of the calculated Cp from the actual Cp. The overall deviation is a very good 0.13%. Figure 7 is a graph of these results.

Reciprocal:

For this last group of correlations, one or both of the input variables (Xi,Yi) are simply inverted and the resulting variables related by the basic linear equation:

$$Yi = A + B*Xi \qquad \text{(Eq. 1)}$$

As there are no new concepts introduced in this section, an abbreviated outline of the porcedures involved should suffice without illustrations.

Reciprocal Y:

1) Let $Yi' = 1/Yi$
2) Then $Yi' = A + B*Xi$ (Eq. 31)
3) Solve for A & B by MOA or least squares method
4) General equation: $Yi = 1/(A + B*Xi)$ (Eq. 32)

Reciprocal X:

1) Let $Xi' = 1/Xi$
2) Then $Yi = A + B*Xi'$ (Eq. 33)
3) Solve for A & B by MOA or least squares method
4) General equation: $Yi = A + B/Xi$ (Eq. 34)

Reciprocal X,Y:

1) Let $Xi' = 1/Xi$; $Yi' = 1/Yi$
2) Then $Yi' = A + B*Xi'$ (Eq. 35)
3) Solve for A & B by MOA or least squares method
4) General equation: $Yi = 1/(A + B/Xi)$ (Eq. 36)

Figure 8 displays charateristic curves for each type of relationship.

References

1. Carlile, R.E., and B.E. Gillett, "FORTRAN and Computer Mathematics for the Engineer and Scientist," Library of Congress #72-95446, ISBN# 0-87814-016-6, The Petroleum Publishing Co., 211 So. Cheyanne, Tulsa, Oklahoma 74101.

2. Srini-Vasan, S., and C. Gatlin, "The Effect of Temperature on the Flow Properties of Clay-Water Drilling Mudgs," Journal of Petroleum Technology, December 1985, p.5^.

3. Gatlin, C., "Petroleum Engineering Drilling and Well Completions," Prentice Hall, Inc., Englewood Cliffs, N.J., p.74.

4. Melton, L.L., and R.S. Saunders, "Rheological Measurements of Non-Newtonian Fluids," Trans. AIME, 210,196 (1957).

5. Turk, E.A., and R.S. Metcalfe, L. Yarborough, R.L. Robinson Jr., "Phase Equilibria in Carbon Dioxide," SPE 9231, September 1980.

6. Black, W.M., "A Review of Drill-Stem Testing Techniques and Analysis," Journal of Petroleum Technology, June 1957, p.21.

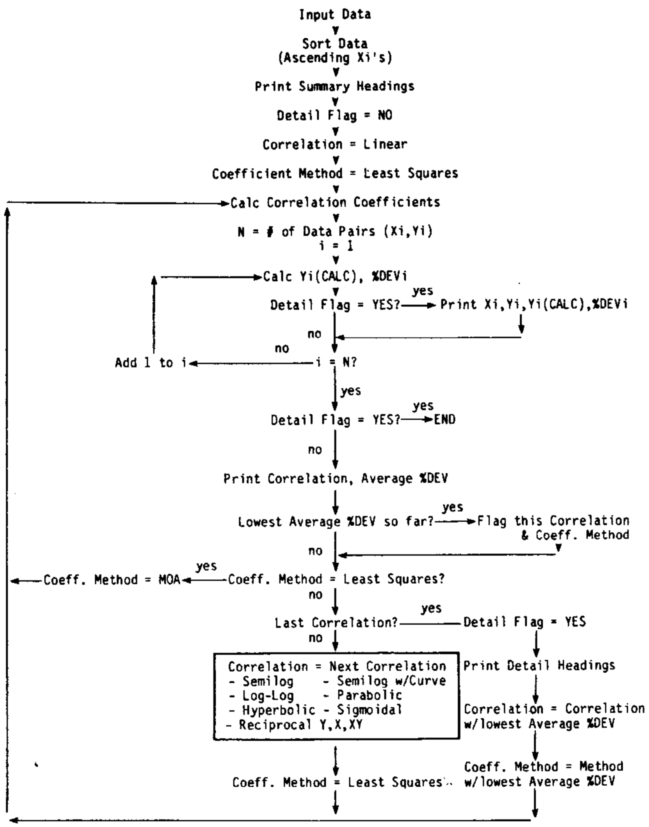
7. Keenan and Keyes, "Thermodynamic Properties of Steam," John Wiley and Sons (1936).

Bibliography

Carlile, R.E., and B.E. Gillett, "FORTRAN and Computer Mathematics for the Engineer and Scientist," Library of Congress #72-95446, ISBN# 0-87814-016-6, The Petroleum Publishing Co., 211 So. Cheyanne, Tulsa, Oklahoma 74101.

Poole, L. and M. McNiff, S. Cook, "APPLE II User's Guide," ISBN# 0-931988-46-2, OSBORNE/McGraw-Hill, Berkeley, California.

General Flow Chart



PROGRAM LISTING

```
10  REM  PROGRAM: LC  (LINEAR CORRELATION)
15  REM  AUTHOR: CLIFF ADAMS
20 D$ =  CHR$ (4)
30 FF$ =  CHR$ (12): REM  FORMFEED FOR EPSON
40  ONERR  GOTO 20000: REM  ERROR HANDLING ROUTINE
50  INPUT "DATA FILE: ";F$
60  GOSUB 17000: REM  LOAD TITLES
70  GOSUB 18000: REM  SCREEN HEADING
100  REM  ************************
110  REM  *      INPUT DATA      *
120  REM  ************************
130  PRINT D$;"OPEN ";F$
140  PRINT D$;"READ ";F$
150  INPUT N
160  DIM X(N),Y(N),XX(N),YY(N),X1(N),Y1(N)
170  FOR I = 1 TO N
180  INPUT X(I),Y(I)
190 Q = X(I): GOSUB 12000: REM  # OF SIGNIFICANT DIGITS
200  IF D > XD THEN XD = D: IF D > SD THEN SD = D
210 Q = Y(I): GOSUB 12000: REM  # OF SIGNIFICANT DIGITS
220  IF D > YD THEN YD = D: IF D > SD THEN SD = D
230  NEXT I
235  IF SD < 4 THEN SD = 4
240  PRINT D$;"CLOSE ";F$
300  REM  ******************************
310  REM  * BUBBLE SORT (ASCENDING X) *
320  REM  ******************************
330  FOR I = 1 TO N - 1
340  FOR J = N TO I + 1 STEP  - 1
350  IF X(J) > X(J - 1) THEN  GOTO 380
360 X = X(J):X(J) = X(J - 1):X(J - 1) = X
370 Y = Y(J):Y(J) = Y(J - 1):Y(J - 1) = Y
380  NEXT J
390  NEXT I
400 M = 1
410  ON M GOTO 1000,2000,3000,4000,5000,6000,7000,8000,8000,8000
1000  REM  **********************
1010  REM  *  LINEAR FUNCTION   *
1020  REM  **********************
1030 M = 1
1040  FOR I = 1 TO N
1050 X1(I) = X(I):Y1(I) = Y(I)
1060 XX(I) = X(I):YY(I) = Y(I)
1070  NEXT I
1080 NN = N:Z = 1: IF ZZ = 1 THEN Z = ZM
1090  ON Z GOSUB 9000,11000: REM  CALC A & B COEFFS.
1100 DT = 0
1110  FOR I = 1 TO NN
1120 RX = A + B * XX(I):D = YD: GOSUB 10000: REM  ROUND-OFF
1130 YC = RX
1140  GOSUB 21000: REM  DEVIATION CALC
1170  NEXT I
1180 RX = DT / NN:D = 2: GOSUB 10000: REM  ROUND-OFF
```

PROGRAM LISTING (Cont'd)

```
1190  IF Z = 1 THEN DM = RX:MM = 1:ZM = 1
1200  GOSUB 15000: REM  CHECK FOR MIN. DEV.
1210  IF ZZ = 0 THEN  GOSUB 19000: REM  SCREEN DISPLAY
1220  IF Z = 1 THEN Z = 2: GOTO 1090
2000  REM  **********************
2010  REM  *  SEMILOGARITHMIC    *
2020  REM  **********************
2040 M = 2
2050 NN = 0:Z = 1: IF ZZ = 1 THEN Z = ZM
2060  FOR I = 1 TO N
2070  IF Y(I) = 0 OR Y(I) < 0 THEN  GOTO 2110: REM  AVOID 0,NEG LOG
2080 NN = NN + 1
2090 X1(NN) = X(I):Y1(NN) = Y(I)
2100 XX(NN) = X(I):YY(NN) =  LOG (Y(I))
2110  NEXT I
2120  ON Z GOSUB 9000,11000: REM  CALC A & B COEFFS.
2130 DT = 0
2140  FOR I = 1 TO NN
2150 RX =  EXP (A + B * XX(I)):D = YD: GOSUB 10000: REM  ROUND-OFF
2160 YC = RX
2170  GOSUB 21000: REM   DEVIATION CALC
2200  NEXT I
2210 D = 2:RX = DT / NN: GOSUB 10000: REM  ROUND-OFF
2220  GOSUB 15000: REM  CHECK OF MIN. DEV.
2230  IF ZZ = 0 THEN  GOSUB 19000: REM  SCREEN DISPLAY
2240  IF Z = 1 THEN Z = 2: GOTO 2120
3000  REM  **********************
3010  REM  * SEMILOG W/CURVE     *
3020  REM  **********************
3030 M = 3
3040 NN = 0:Z = 1: IF ZZ = 1 THEN Z = ZM
3050  FOR I = 1 TO N
3060  IF Y(I) < 0 OR Y(I) = 0 THEN  GOTO 3090: REM  AVOID 0,NEG LOG
3070 NN = NN + 1
3080 X1(NN) = X(I):Y1(NN) = Y(I)
3090  NEXT I
3100  REM  **********************
3110  REM  * SEMILOG SIGMA CALC  *
3120  REM  **********************
3130 S = 0:XM = (X1(1) + X1(NN)) / 2
3140  FOR I = 1 TO NN
3150  IF X1(I) < XM THEN X1 = X1(I):Y1 =  LOG (Y1(I))
3160  IF X1(I) = XM THEN S = 1:YM = Y1(I):I = NN: GOTO 3180
3170  IF X1(I) > XM THEN X3 = X1(I):Y3 =  LOG (Y1(I)):I = NN
3180  NEXT I
3190  IF S = 1 THEN  GOTO 3210: REM   EXACT MATCH
3200 YM =  EXP (Y1 + (XM - X1) * (Y3 - Y1) / (X3 - X1))
3210 S = Y1(NN) + Y1(1) - 2 * YM
3220  IF S = 0 THEN  GOTO 3240
3230 S = (Y1(1) * Y1(NN) - YM ^ 2) / S
3232 RX = S:D = SD: GOSUB 10000
3234 S = RX
3240 J = 0
3250  FOR I = 1 TO NN
3260 Y = Y1(I) - S: IF Y <  = 0 THEN  GOTO 3290
3270 J = J + 1
3280 XX(J) = X1(J):YY(J) =  LOG (Y)
3290  NEXT I
3300 NN = J: IF NN = 0 THEN  GOTO 20000
3310  ON Z GOSUB 9000,11000: REM  CALC A & B COEFF.
3320 DT = 0
3330  FOR I = 1 TO NN
3340 RX =  EXP (A + B * XX(I)) + S:D = YD: GOSUB 10000: REM  ROUND-OFF
3350 YC = RX
3360  GOSUB 21000: REM   DEVIATION CALC
3390  NEXT I
3400 D = 2:RX = DT / NN: GOSUB 10000: REM  ROUND-OFF
3410  GOSUB 15000: REM  CHECK OF MIN. DEV.
3420  IF ZZ = 0 THEN  GOSUB 19000: REM  SCREEN DISPLAY
3430  IF Z = 1 THEN Z = 2: GOTO 3310
4000  REM  *********************
4010  REM  *     LOG-LOG        *
4020  REM  *********************
4030 M = 4
4040 NN = 0:Z = 1: IF ZZ = 1 THEN Z = ZM
4050  FOR I = 1 TO N
4060  IF X(I) < 0 OR X(I) = 0 THEN  GOTO 4110: REM  AVOID DIV. BY ZERO
4070  IF Y(I) < 0 OR Y(I) = 0 THEN  GOTO 4110: REM  AVOID DIV. BY ZERO
4080 NN = NN + 1
4090 X1(I) = X(I):Y1(I) = Y(I)
4100 XX(I) =  LOG (X(I)):YY(I) =  LOG (Y(I))
4110  NEXT I
4120  ON Z GOSUB 9000,11000: REM  CALC A & B COEFFS.
4130 DT = 0
4140 A =  EXP (A)
4142 RX = A:D = SD: GOSUB 10000
4144 A = RX
4150  FOR I = 1 TO NN
4160 RX = A * (X1(I) ^ B):D = YD: GOSUB 10000: REM  ROUND-OFF
4170 YC = RX
4180  GOSUB 21000: REM   DEVIATION CALC
4210  NEXT I
4220 RX = DT / NN:D = 2: GOSUB 10000: REM  ROUND-OFF
4230  GOSUB 15000: REM  CHECK FOR MIN. DEV.
4240  IF ZZ = 0 THEN  GOSUB 19000: REM  SCREEN DISPLAY
4250  IF Z = 1 THEN Z = 2: GOTO 4120
5000  REM  *********************
5010  REM  *     PARABOLIC      *
5020  REM  *********************
5030 M = 5
5040 NN = 0:Z = 1: IF ZZ = 1 THEN Z = ZM
5050  FOR I = 2 TO N
5060  IF X(I) - X(1) = 0 THEN  GOTO 5110: REM  AVOID DIVISION BY ZERO
5070 NN = NN + 1
5080 X1(NN) = X(I):Y1(NN) = Y(I)
5090 XX(NN) = X(I) + X(1)
```

PROGRAM LISTING (Cont'd)

```
5100 YY(NN) = (Y(I) - Y(1)) / (X(I) - X(1))
5110  NEXT I
5130  ON Z GOSUB 9000,11000: REM  CALC A & B COEFFS.
5140  FOR I = NN TO 1 STEP  - 1
5150 X1(I + 1) = X1(I):Y1(I + 1) = Y1(I)
5160  NEXT I
5170 X1(1) = X(1):Y1(1) = Y(1):NN = NN + 1
5180 DT = 0
5190  FOR I = 1 TO NN
5192 C = Y1(1) - A * X1(1) - B * X1(1) ^ 2
5200 RX = C + A * X1(I) + B * X1(I) ^ 2
5210 D = YD: GOSUB 10000: REM  ROUND-OFF
5220 YC = RX
5230  GOSUB 21000: REM   DEVIATION CALC
5270  NEXT I
5280 RX = DT / NN:D = 2: GOSUB 10000: REM  ROUND-OFF
5290  GOSUB 15000: REM  CHECK FOR MIN. DEV.
5300  IF ZZ = 0 THEN  GOSUB 19000: REM  SCREEN DISPLAY
5310  IF Z = 1 THEN Z = 2:NN = 0: GOTO 5050
6000  REM  ********************
6010  REM  *    HYPERBOLIC     *
6020  REM  ********************
6030 M = 6
6040 NN = 0:Z = 1: IF ZZ = 1 THEN Z = ZM
6050  FOR I = 2 TO N
6060  IF Y(I) - Y(1) = 0 THEN  GOTO 6110: REM  AVOID DIVISION BY ZERO
6070 NN = NN + 1
6080 X1(NN) = X(I):Y1(NN) = Y(I)
6090 XX(NN) = X(I)
6100 YY(NN) = (X(I) - X(1)) / (Y(I) - Y(1))
6110  NEXT I
6120  ON Z GOSUB 9000,11000: REM  CALC A & B COEFFS.
6130  FOR I = NN TO 1 STEP  - 1
6140 X1(I + 1) = X1(I):Y1(I + 1) = Y1(I)
6150  NEXT I
6160 X1(1) = X(1):Y1(1) = Y(1):NN = NN + 1
6170 DT = 0
6180  FOR I = 1 TO NN
6190  IF I = 1 THEN RX = Y1(1): GOTO 6210
6200 RX = Y1(1) + (X1(I) - X1(1)) / (A + B * X1(I))
6210 D = YD: GOSUB 10000: REM  ROUND-OFF
6220 YC = RX
6230  GOSUB 21000: REM   DEVIATION CALC
6270  NEXT I
6280 RX = DT / NN:D = 2: GOSUB 10000: REM  ROUND-OFF
6290  GOSUB 15000: REM  CHECK FOR MIN. DEV.
6300  IF ZZ = 0 THEN  GOSUB 19000: REM  SCREEN DISPLAY
6310  IF Z = 1 THEN Z = 2:NN = 0: GOTO 6050
7000  REM  ********************
7010  REM  *   SIGMOIDAL      *
7020  REM  ********************
7030 M = 7:Z = 1
7040 S = 0
```

PROGRAM LISTING (Cont'd)

```
7050 NN = 0
7060  FOR I = 1 TO N
7070  IF X(I) < 0 OR Y(I) <  = 0 THEN  GOTO 7100
7080 NN = NN + 1
7090 X1(NN) = X(I):Y1(NN) = Y(I)
7100  NEXT I
7112 IN = X1(2) - X1(1):S = Y1(1)
7114  FOR I = 1 TO NN
7115 X1(I) = X1(I) - X1(1)
7116  NEXT I
7120 ND = (NN - 1) - 3 *  INT ((NN - 1) / 3)
7121 NN = NN - ND
7210  REM  SORT Y'S ASCENDING
7211  GOTO 7290
7220  FOR I = 1 TO NN - 1
7230  FOR J = NN TO I + 1 STEP  - 1
7240  IF Y1(J) > Y1(J - 1) THEN  GOTO 7270
7250 X = X1(J):X1(J) = X1(J - 1):X1(J - 1) = X
7260 Y = Y1(J):Y1(J) = Y1(J - 1):Y1(J - 1) = Y
7270  NEXT J
7280  NEXT I
7290 S1 = 0:S2 = 0:S3 = 0
7300  FOR I = 2 TO NN
7310 Y =  LOG (Y1(I) - S)
7320  IF I - 1 <  = (NN - 1) / 3 THEN S1 = S1 + Y: GOTO 7350
7330  IF I - 1 <  = 2 * (NN - 1) / 3 THEN S2 = S2 + Y: GOTO 7350
7340 S3 = S3 + Y
7350  NEXT I
7360 NS = (NN - 1) / 3
7370  IF S1 - S2 = 0 THEN  GOTO 20000: REM  ERROR HANDLING
7380 C = ((S2 - S3) / (S1 - S2)) ^ (1 / NS)
7382 RX = C:D = SD: GOSUB 10000
7384 C = RX
7390  IF C = 1 THEN  GOTO 20000: REM  ERROR HANDLING
7400 A =  EXP ((S1 - (S1 - S2) / (1 - C ^ NS)) / NS)
7410 B =  EXP ((S1 - S2) * (1 - C) / (1 - C ^ NS) ^ 2)
7412 RX = A:D = SD: GOSUB 10000
7414 A = RX
7416 RX = B:D = SD: GOSUB 10000
7418 B = RX
7420 DT = 0
7440  FOR I = 1 TO NN
7450 RX = S + A * (B ^ (C ^ (X1(I) / IN - 1))):D = YD: GOSUB 10000: REM
     ROUND-OFF
7460 YC = RX
7470  GOSUB 21000: REM   DEVIATION CALC
7500  NEXT I
7510 RX = DT / NN:D = 2: GOSUB 10000: REM  ROUND-OFF
7520  GOSUB 15000: REM  CHECK FOR MIN. DEV.
7530  IF ZZ = 0 THEN  GOSUB 19000: REM  SCREEN DISPLAY
```

PROGRAM LISTING (Cont'd)

```
8000  REM  *********************
8010  REM  *  RECIPORCAL X,Y    *
8020  REM  *********************
8030 M = 8:Z = 1
8040 NN = 0
8060  IF ZZ = 1 THEN M = MM:Z = ZM
8070  FOR I = 1 TO N
8080  IF (M = 8 OR M = 10) AND Y(I) = 0 THEN  GOTO 8160: REM  NO /0
8090  IF (M = 9 OR M = 10) AND X(I) = 0 THEN  GOTO 8160: REM  NO 0
8100 NN = NN + 1
8110 X1(NN) = X(I):Y1(NN) = Y(I)
8120 XX(NN) = X(I):YY(NN) = Y(I)
8130  IF M = 8 THEN YY(NN) = 1 / Y(I)
8140  IF M = 9 THEN XX(NN) = 1 / X(I)
8150  IF M = 10 THEN XX(NN) = 1 / X(I):YY(NN) = 1 / Y(I)
8160  NEXT I
8170  ON Z GOSUB 9000,11000: REM  CALC A & B COEFFS.
8180 DT = 0
8190  FOR I = 1 TO NN
8200 RX = A + B * XX(I)
8210  IF (M = 8 OR M = 10) THEN RX = 1 / RX
8220 D = YD: GOSUB 10000: REM  ROUND-OFF
8230 YC = RX
8240  GOSUB 21000: REM   DEVIATION CALC
8270  NEXT I
8280 RX = DT / NN:D = 2: GOSUB 10000: REM  ROUND-OFF
8290  GOSUB 15000: REM  CHECK FOR MIN. DEV.
8300  IF ZZ = 0 THEN  GOSUB 19000: REM  SCREEN DISPLAY
8310  IF Z = 1 THEN Z = 2: GOTO 8040
8320  IF M = 8 THEN M = 9:Z = 1: GOTO 8040
8330  IF M = 9 THEN M = 10:Z = 1: GOTO 8040
8340  PRINT : INPUT "HIT RETURN KEY TO CONTINUE...";RT$
8350  HOME
8360 ZZ = 1:M = MM: GOTO 410
9000  REM  *********************
9010  REM  *  LEAST SQUARES     *
9020  REM  *********************
9030 XY = 0:SX = 0:SY = 0:X2 = 0
9040  FOR I = 1 TO NN
9050 XY = XY + XX(I) * YY(I)
9060 SX = SX + XX(I)
9070 SY = SY + YY(I)
9080 X2 = X2 + XX(I) * XX(I)
9090  NEXT I
9100 B = (NN * XY - SX * SY) / (NN * X2 - SX * SX)
9102 RX = B:D = SD: GOSUB 10000
9104 B = RX
9110 A = (SY - B * SX) / NN
9112 RX = A:D = SD: GOSUB 10000
9114 A = RX
9120  RETURN
```

PROGRAM LISTING (Cont'd)

```
10000  REM  ********************
10010  REM  * ROUND-OFF ROUTINE *
10020  REM  ********************
10030 RX =  INT (RX * 10 ^ D + .5) /  INT (10 ^ D + .5)
10040 RX =  INT (RX * 10 ^ D) / 10 ^ D
10050  RETURN
11000  REM  *********************
11010  REM  * METHOD OF AVERAGES *
11020  REM  *********************
11030 N1 =  INT (NN / 2):N2 = NN - N1
11040 N3 = N2: IF N2 = N1 THEN N3 = N3 + 1
11050 X1 = 0:X2 = 0:Y1 = 0:Y2 = 0
11060  FOR I = 1 TO N1
11070 X1 = X1 + XX(I):Y1 = Y1 + YY(I)
11080  NEXT I
11090  FOR I = N3 TO NN
11100 X2 = X2 + XX(I):Y2 = Y2 + YY(I)
11110  NEXT I
11120 B = (Y1 * N2 - Y2 * N1) / (X1 * N2 - X2 * N1)
11122 RX = B:D = SD: GOSUB 10000
11124 B = RX
11130 A = (Y1 - X1 * B) / N1
11132 RX = A:D = SD: GOSUB 10000
11134 A = RX
11140  RETURN
12000  REM  ********************
12010  REM  *  # SIGNIF. DIGITS *
12020  REM  ********************
12030 D = 0
12040 Q$ =  STR$ (Q):Q1$ =  STR$ ( INT (Q))
12050 D =  LEN (Q$) -  LEN (Q1$)
12060  IF Q > 1 AND D > 1 THEN D = D - 1
12070  RETURN
13000  REM  ********************
13010  REM  * PRINT-OUT ROUTINE *
13020  REM  ********************
13030 P$ = "":Z$ = "0000000000"
13040 BL$ = "          ": REM  10 BLANKS
13050 Q = X1(I):DD = XD: GOSUB 14000: REM  PAD ROUTINE
13060 Q = Y1(I):DD = YD: GOSUB 14000: REM  PAD ROUTINE
13070 Q = YC:DD = YD: GOSUB 14000: REM  PAD ROUTINE
13080 Q = DV:DD = 2: GOSUB 14000: REM  PAD ROUTINE
13090  IF I = 1 THEN  GOSUB 16000: REM  PRINT HEADING
13100  PRINT P$
13110  IF I < NN THEN  RETURN
13120 RX = DT / NN:D = 2: GOSUB 10000: REM  ROUND-OFF
13130  PRINT : PRINT "AVER.DEV : ";RX;"%"
13140  PRINT FF$: REM  FORMFEED
13150  END
```

PROGRAM LISTING (Cont'd)

```
14000  REM  *********************
14010  REM  *  PAD ROUTINE       *
14020  REM  *********************
14030  GOSUB 12000: REM  # OF SIGNIFICANT DIGITS
14040  IF DD - D < = 0 THEN  GOTO 14070
14050  IF D = 0 THEN Q$ = Q$ + "."
14060 Q$ = Q$ +  LEFT$ (Z$,(DD - D))
14070  IF  LEN (Q$) > =  LEN (BL$) THEN  GOTO 14090
14080 Q$ =  LEFT$ (BL$,( LEN (BL$) -  LEN (Q$))) + Q$: REM  RIGHT JUSTIF
     Y
14090 P$ = P$ + Q$
14100  RETURN
15000  REM  *********************
15010  REM  * LOWEST DEV CHECK  *
15020  REM  *********************
15030  IF RX < DM THEN DM = RX:MM = M:ZM = Z
15040  RETURN
16000  REM  *********************
16010  REM  *   HEADING ROUTINE *
16020  REM  *********************
16030  PRINT "DATA FILE: ";F$: PRINT
16040 A$ = "ACTUAL EQ: "
16050  ON MM GOTO 16060,16110,16170,16240,16290,16340,16390,16470,16510,
     16550
16060  PRINT "LINEAR CORRELATION: ";
16070  PRINT " Y = A + B*X"
16080  GOSUB 16660
16090  PRINT : PRINT A$;"Y = ";A;" + ";B;"*X"
16100  GOTO 16590
16110  PRINT "SEMILOGARITHMIC CORRELATION: ";
16120  RRINT " LOG Y = A + B*X"
16130  GOSUB 16660
16140  IF D = 2 THEN DT$ =  STR$ (RX)
16150  PRINT : PRINT A$;"LOG Y = ";A;" + ";B;"*X"
16160  GOTO 16590
16170  PRINT "SEMILOGARITHMIC W/CURVE CORRELATION: ";
16180  PRINT " LOG (Y-S) = A + B*X"
16190  GOSUB 16660
16200  PRINT : PRINT "S: ";S
16210  PRINT : PRINT A$;"LOG(Y - ";S;") = ";A;" + ";B;"*X"
16220  PRINT : PRINT "S = (YMAX-YMIN-YMID^2) / (YMAX+YMIN-2*YMID)"
16230  GOTO 16590
16240  PRINT "LOG-LOG CORRELATION: ";
16250  PRINT " LOG Y = LOG A + B*LOG X"
16260  GOSUB 16660
16262 LA =  LOG (A):RX = LA:D = SD: GOSUB 10000
16264 LA = RX
16270  PRINT : PRINT A$;"LOG Y = ";LA;" + ";B;"*LOGX"
16280  GOTO 16590
16290  PRINT "PARABOLIC CORRELATION: ";
16300  PRINT " Y = C + A*X + B*(X^2)"
16310  GOSUB 16660
16312  PRINT : PRINT "C: ";C
```

PROGRAM LISTING (Cont'd)

```
16320  PRINT : PRINT A$;"Y = ";C;" + (";A;"*X) + ";B;"*(X^2)"
16330  GOTO 16590
16340  PRINT "HYPERBOLIC CORRELATION: ";
16350  PRINT " Y = YMIN + (X-XMIN)/(A+B*X)"
16360  GOSUB 16660
16370  PRINT : PRINT A$;"Y = ";Y1(1);" + (X- ";X1(1);")/(";A;" + ";B;"*X
     )"
16380  GOTO 16590
16390  PRINT "SIGMOIDAL CORRELATION: ";
16400  PRINT " Y = S + A*(B^(C^(X/IN - 1)))"
16410  GOSUB 16660
16420  PRINT : PRINT "S: ";S
16430  PRINT : PRINT "C: ";C
16440  PRINT : PRINT "IN: ";IN
16450  PRINT : PRINT A$;"Y = ";S;" + ";A;"*(";B;"^(";C;"^(X/";IN;" - 1))
     )"
16460  GOTO 16630
16470  PRINT "RECIPORCAL Y CORRELATION: 1/Y = A + B*X"
16480  GOSUB 16660
16490  PRINT : PRINT A$;"1/Y = ";A;" + ";B;"*X"
16500  GOTO 16590
16510  PRINT "RECIPORCAL X CORRELATION: Y = A + B/X"
16520  GOSUB 16660
16530  PRINT : PRINT A$;"Y = ";A;" + ";B;"/X"
16540  GOTO 16590
16550  PRINT "RECIPORCAL X,Y CORRELATION: 1/Y = A + B/X"
16560  GOSUB 16660
16570  PRINT : PRINT A$;"1/Y = ";A;" + ";B;"/X"
16580  GOTO 16590
16590  PRINT : PRINT "A & B DETERMINATION BY ";
16600  ON ZM GOTO 16610,16620
16610  PRINT "LEAST SQUARES METHOD": GOTO 16630
16620  PRINT "METHOD OF AVERAGES": PRINT
16630  PRINT "      X          Y         YCALC       %DEV"
16640  PRINT "   -------    -------    -------    -------"
16650  RETURN
16660  PRINT : PRINT "A: ";A
16670  PRINT : PRINT "B: ";B: RETURN
17000  REM  *********************
17010  REM  *  LOADING TITLES   *
17020  REM  *********************
17030  FOR I = 1 TO 10: READ M$(I): NEXT I: RETURN
17040  DATA  "LINEAR","SEMILOG","SEMILOG W/CURVE","LOG-LOG","PARABOLIC"
17050  DATA   "HYPERBOLIC","SIGMOIDAL","RECIPORCAL Y","RECIPORCAL X","RE
     CIPORCAL X,Y"
```

Line No.	Description
10 - 15	General remark statements
20	All DOS commands in APPLESOFT must be preceded by a CHR$(4) which for this program will be D$.
30	FF$ will advance the paper to top of the next form. For an Epson-type printer, this is a CHR$(12).
40	All detected errors will branch to Line 20000.
50	Data file name is entered from keyboard and stored in F$. The data file had been previously created by the program FILEWRITER.
60 - 70	Loads arrays and strings that are used to print headings & titles.
130	Opens the serial data file (F$).
140	Informs the program that all further input will come from F$.
150	Reads the first record from F$ and stores the value in N. N is the number of X,Y data pairs that are to be read from F$.
160	Dimensions all the arrays used in the program.
170 : 230	Routine which reads all the data pairs from the input file and stores the values in the X(I) and Y(I) arrays respectively. As each value is read, the program branches to Line 12000 where the number of significant digits is determined. The maximum number of digits for each value is kept in the XD and YD variables for output formatting purposes. (APPLESOFT has no formatting functions.) The highest number of significant digits of either variable is stored in SD to be used as the number of digits of the derived coefficients.
235	The minimum number of significant digits for the coefficients is 4.
240	All data input is complete and the data file is closed.
330 : 390	This is a bubble-sort routine which sorts the data pairs by the X value in ascending order.
400	The variable M contains the CURRENT CORRELATION value which allows the subroutines to know which function is being investigated. In this case, M = 1 signals that the LINEAR function is in use.
410	This line branches the program to the CURRENT CORRELATION routine. The program returns to this line until all correlations have been tested at which time the program branches to the best-fit correlation for the detail print-out pass.
1030	CURRENT CORRELATION is LINEAR (M = 1).
1040 : 1070	The original input data pairs X(),Y() are stored into "working" arrays X1(),Y1() and XX(),YY(). Throughout this program, the original data pairs are edited at the beginning of each correlation to remove any points which would "blow-up" the routine (ie. divide by 0) The allowable data points are stored in X1(),Y1() and XX(),YY(). X1(),Y1() values are left untouched for output purposes while XX(),YY() values are "massaged" in whatever manner is required by the particular correlation. NN is the number of allowable data points reset in each correlation.
1080	NN is the number of allowable data points in this correlation. Z is the COEFFICIENT FLAG which is either 1 or 2. 1 indicates that the least squares method (LS) is to be used while 2 signals for the Method of Averages (MOA). ZZ is the DETAIL FLAG (See General Flow Chart). ZM is the variable which holds the lowest COEFFICIENT FLAG to be used for the detail print-out pass.
1090	Determine the coefficients (A & B) by either least squares method (Z=1) or MOA (Z=2).
1100	DT holds the DEVIATION TOTAL value from which the average deviation is determined. Here it is cleared to 0.
1110 : 1170	This routine determines a calculated Y (YC) for each allowable X based on Eq. 1 using the current coefficients A & B. The YC value is then rounded to YD number of significant digits. The deviation of each calculated Y from the original Y is then determined (Line 1140). by a sub-routine (21000). It is in this sub-routine that a detail print-out will occur if the DETAIL FLAG is set (ZZ=1).
1180	The average deviation is calculated and rounded off to 2 significant digits.
1190	DM holds the minimum average deviation value. MM holds the current lowest correlation. ZM holds the lowest coefficient method value. These are initially set to the linear correlation by the LS method.
1200	Branches to Line 15000 to check for lowest deviation.
1210	If DETAIL FLAG is not set (ZZ=0) then goto SCREEN DISPLAY routine.
1220	If the current coefficient determination is LS (Z=1) then set the current coefficient method to MOA (Z=2) and rerun correlation.
2040	CURRENT CORRELATION is SEMILOG (M=2)
2050	Same concept as 1080. NN is determined in Lines 2060 - 2110.
2060 : 2110	Same concept as Lines 1040 - 1070. Line 2070 avoids negative or 0 log values. Line 2100 sets YY() = Yi' and XX() = Xi as used in Eq. 13.
2120 : 2130	Same as 1090 - 1100

2140 : 2200	Same concept as 1110 - 1170. Lines 2150 - 2160 calculates and rounds YC by Eq. 12.
2210 : 2240	Same as 1180 - 1200
3030	CURRENT CORRELATION is SEMILOG W/CURVE (M=3).
3040	Same concept as 2050.
3050 : 3090	Same as 2160 - 2110
3130 : : : 3230	This section determines the SHIFT FACTOR (s) required in Eq. 16. Line 3130 determines Xm directly (XM). If none of the data points X() is exactly equal to Xm then an interpolation is made to find Ym. This is done in Lines 3140 - 3200. S is calculated in Line 3230 by Eq. 17.
3232 : 3234	S is rounded off the SD number of significant digits. (Line 235)
3240 : 3300	The left half of Eq. 16 is determined (ln(Yi-s)) and any invalid points are ommitted from consideration. The varible J keeps count of the number of allowable points which is given to NN.
3310 : 3430	Same concept as Lines 1080 - 1220. Line 3340 corresponds to Eq. 16.
4030	CURRENT CORRELATION is LOG-LOG.
4040	Same as 1080
4050 : 4110	Same concept as 1040 - 1070.
4120	Same concept as 1090
4130	Same concept as 1100
4140	A corresponds to A' used in Eq. 20.
4142 : 4144	A is rounded to SD significant digits.
4150 : 4210	Same concept as 1110 - 1170. YC is calculated by Eq. 20, rounded to YD significant digits and the deviation is determined.
4220	Same as 1180.
4230	Same as 1200.
4240	Same as 1210.
4250	Same as 1220.
5030	CURRENT CORRELATION is PARABOLIC.
5040	Same as 1180.
5050 : 5110	Same concept as 1040. XX() and YY() are set equal to Xi' and Yi' respectively as in Eq. 23.
5130	Same concept as 1090.
5140 : 5170	X1() and Y1() arrays are shifted back to their original positions.
5180	Same as 1100.
5190 : 5270	Same as 1110 - 1170 using Eq. 22. C is determined in Line 5192 using the first pair of allowable X1(),Y1() and the current coefficients.
5280	Same as 1180.
5290	Same as 1200.
5300	Same as 1210.
5310	Same as 1220.
6030	CURRENT CORRELATION is HYPERBOLIC.
6040	Same as 1080.
6050 : 6110	Same concept as 1040 - 1070. XX() and YY() are set equal to Xi and Yi' respectively by Eq. 26.
6120	Same as 1090.
6130 : 6160	Same as 5140 - 5170.
6170	Same as 1100.

6180 : 6270	Same concept as 1100 - 1170 using Eq. 25. The first point, X1(),Y1() is passed through so as to avoid a division by zero.
6280	Same as 1180.
6290	Same as 1200.
6300	Same as 1210.
6310	Same as 1220.
7030	CURRENT CORRELATION is SIGMOIDAL. For this correlation, the coefficients are determined by unique processes; the LS and MOA methods do not apply.
7040	S corresponds to the S variable used in Eq. 27.
7050 : 7100	Same concept as 1040 - 1070. Negative or zero values for Y() or negative values for X() would create an error and are avoided.
7112	IN is the increment amount between each X() value.
7114 : 7116	X1() values are shifted so as to have a zero value for X1(1).
7120 : 7121	ND is the number of points that will be dropped to allow for an even division of the points into thirds.
7210 : 7280	This routine is not currently in use and may be omitted.
7290 : 7350	S1, S2 and S3 are calculated in this section. Refer to the discussion in the text following Eq. 27.
7360	NS corresponds to the N used to calculate C in Eq. 27a.
7370 : 7390	C is calculated by Eq. 27a and rounded to SD significant digits. C is then checked for being equal to 1 as this would cause an error in Lines 7400 - 7410.
7400	A is determined by Eq. 29.
7410	B is determined by Eq. 30.
7412 : 7418	A and B are rounded off to SD significant digits.
7420	Same as 1100.
7440 : 7500	Same as 1040 - 1070 using Eq. 27.
7510	Same as 1180.
7520	Same as 1200.
7530	Same as 1210.
8030 : 8060	CURRENT CORRELATION is RECIPROCAL Y; COEFFICIENT FLAG is LS. Same as 1080. The M = MM is redundant and may be omitted.
8070 : 8120	This section is used by the RECIPROCAL Y, RECIPROCAL X and the RECIPROCAL XY correlations to avoid division by zero problems.
8130	Calculates Yi' used in Eq. 31 for the RECIPROCAL Y correlation.
8140	Calculates Xi' used in Eq. 33 for the RECIPROCAL X correlation.
8150	Calculates Xi' and Yi' used in Eq. 35 for the RECIPROCAL XY correlation.
8170	Same as 1090.
8180	Same as 1100.
8190 : 8270	Same as 1110 - 1170 using Eq. 1. Line 8210 inverts the results if the CURRENT CORRELATION is RECIPROCAL Y (M=8) or RECIPROCAL XY (M=10).
8280	Same as 1180.
8290	Same as 1200.
8300	Same as 1210.
8310	Same as 1220.
8320	If CURRENT CORRELATION is RECIPROCAL Y (M=8) then set CURRENT CORRELATION equal to RECIPROCAL X (M=9) and rerun correlation.
8330	If CURRENT CORRELATION is RECIPROCAL X (M=9) then set CURRENT CORRELATION equal to RECIPROCAL XY (M=10) and rerun correlation.
8340	Holds screen display until RETURN key is hit. RT$ is a dummy variable.
8350	Clears screen and moves cursor to upper left-hand corner of screen.

8360 All correlations have been checked at this point so the DETAIL FLAG is set (ZZ=1); the CURRENT CORRELATION is set equal to the best-fit correlation (M=MM) and the program branches back to Line 410 to rerun the desired correlation, this time displaying the point-by-point calculations and deviations.

9030 Zeros out the summation variables used in the least squares coefficient determination routine.

9040 Routine which sums the various input variables:
: XY = sum of (Xi*Yi) used in Eq. 11.
: SX = sum of Xi used in Eq. 10.
: SY = sum of Yi used in Eq. 10.
9090 X2 = sum of Xi*Xi used in Eq. 11.

9100 B is calculated here by solving Eqs. 10 and 11 simultaneously for B.

9102 B is rounded off to SD significant digits.
:
9104

9110 A is determined by plugging B back into Eq. 10 and solving for A.

9112 A is rounded off to SD significant digits.
:
9114

10030 Round-off Routine: Values are assigned to RX before branching here.
: RX is rounded to D significant digits and returned.
10050

11000 Determines coefficients A and B by the Method of Averages.
: N1 = number of points in 1st half of allowable points.
: N2 = number of points in 2nd half of allowable points.
: N3 = starting point for 2nd half of allowable points.
: X1 = sum of X values in 1st half of allowable points.
: X2 = sum of X values in 2nd half of allowable points.
: Y1 = sum of Y values in 1st half of allowable points.
: Y2 = sum of Y values in 2nd half of allowable points.
: B is determined in Line 11120 by solving Eqs. 3 and 4 simultaneously
: for B. B is then rounded to SD significant digits (Lines 11122 -
: 11124). A is calculated in Line 11130 by Eq. 3 and then rounded to
11140 SD significant digits (Lines 11132 - 11134).

12000 Determines the number of significant digits of the variable Q.
: Values are assigned to Q before branching here. The number of
12070 significant digits is assinged to D before returning.

13000 Print-out routine for the detailed display of the individual points
: X1(), Y1() and the corresponding calculated Y (YC) with its devia-
: tion (DV). P$ is initially blanked out. Z$ and BL$ are used to
: pad values to DD significant digits. Q is used to pass values to
: the PAD routine and back. If I=1 (first time thru) then it branches
: to print HEADINGS and returns. P$ is built up by the PAD routine
: before printing. If I NN (not the last point) then the program
: branches back to the calling line. Otherwise I=NN (last point) so
: Lines 13120 - 13130 calculates the AVERAGE DEVIATION, rounds it off
: to 2 significant digits and prints the line. Line 13140 advances
: the paper in the printer to the top of the next form (if in use)
13150 and in Line 13150 the program ends.

14000 PAD routine: Lines 14030 - 14050 determine if a decimal point is
: present. Q$ is the individual values being padded (either X1(),
: Y1(), YC or DV). Once the appropriate 0's or blanks have been
: added to Q$ it is appended onto the end of P$. The program then
: returns to the PRINT-OUT routine until all values have been added
14100 to P$ at which point P$ is printed.

15000 Checks if the CURRENT CORRELATION's average deviation (RX) is less
: than the lowest deviation thus far (DM). If so, DM is set equal to
: RX; the best-fit correlation flag (MM) is set to the CURRENT CORREL-
: ATION value (M) and the best-coefficient method flag (ZM) is set
: equal to the current COEFFICIENT FLAG (Z). Program returns to the
15040 calling line.

16000 Heading routines for the detailed display. MM branches execution
: to the correct correlation titles and equations before returning
16670 to the PRINT-OUT routine.

17000 Loads titles into string arrays used in the display routines.
:
17050

18000 Prints screen headings that appear at the start of the program.
:
18130

19000 This section prints each correlation and the average deviations
: calculated by using the coefficients determined by both the LS and
: MOA methods (DT$). If there was a problem with any part of a
19170 correlation, a "N/A" is printed in place of the deviation.

20000 This section basically aborts the current correlation/coefficient
: method if any problem arises, prints a "N/A" for that deviation
: and returns processing to the next correlation/coefficient combina-
20050 tion.

21000 Calculates the individual point deviation (DV) by Eq. 5. The sum
: of DV is stored in DT. Lines 21030 - 21040 avoid a division by
: zero error if Y1(I) happens to be 0. Program returns to the calling
21080 line.

INPUT PROGRAM LISTING

```
100  REM  ***********************************
102  REM  * PROGRAM: FILEWRITER             *
104  REM  * THIS PROGRAM IS INITIALLY       *
106  REM  * RUN WHEN THE DISK IS BOOTED.    *
108  REM  * IT WILL SHOW A MENU WHICH       *
110  REM  * ALLOWS THE USER TO EITHER       *
112  REM  * CREATE A DATA FILE OR RUN       *
114  REM  * THE PROGRAM "LC"                *
116  REM  ***********************************
200  HOME :D$ =  CHR$ (4):S$ = ""
445 S$ = ""
446 X$ = "          "
448  PRINT X$;
450  PRINT "**********************": PRINT X$;
500  PRINT "*                    *": PRINT X$;
600  PRINT "* LINEAR CORRELATION *": PRINT X$;
700  PRINT "*                    *": PRINT X$;
800  PRINT "*         BY         *": PRINT X$;
820  PRINT "*                    *": PRINT X$;
900  PRINT "*    CLIFF ADAMS     *": PRINT X$;
1000  PRINT "*                    *": PRINT X$;
1100  PRINT "**********************"
1102  PRINT : PRINT
1104  PRINT "WHICH WOULD YOU LIKE?"
1106  PRINT
1108  PRINT "   1. CREATE A DATA FILE
1110  PRINT
1112  PRINT "   2. RUN LINEAR CORRELATION
1114  PRINT : PRINT
1116  INPUT "ENTER YOUR SELECTION: ";S$
1118  IF S$ = "1" THEN  GOTO 1298
1120  IF S$ = "2" THEN  PRINT D$;"RUN LC"
1122  GOTO 100
1298  HOME
1300  INPUT "OUTPUT FILE NAME: ";F$
1400  INPUT "NUMBER OF ENTRIES: ";N
1500  DIM X(N),Y(N)
1600  FOR I = 1 TO N
1700  PRINT "X(";I;"),Y(";I;"): ";
1800  INPUT X(I),Y(I)
1900  NEXT I
2000  PRINT D$;"OPEN ";F$
2100  PRINT D$;"WRITE ";F$
2200  PRINT N
2300  FOR I = 1 TO N
2400  PRINT X(I)
2500  PRINT Y(I)
2600  NEXT I
2700  PRINT D$;"CLOSE ";F$
2800  HOME
2900  PRINT "VERIFYING FILE..."
3000  PRINT D$;"OPEN ";F$
3100  PRINT D$;"READ ";F$
3200  INPUT NN
3300  PRINT "N WRITTEN: ";N;"  N READ: ";NN
3400  PRINT
3500  FOR I = 1 TO NN
3600  INPUT X: INPUT Y
3700  IF X <  > X(I) OR Y <  > Y(I) THEN  PRINT "ERROR"
3800  NEXT I
3900  PRINT D$;"CLOSE ";F$
4000  GOTO 100
```

PRINT OUT EXAMPLE

Linear Correlation

File: Air

Correlation	% Deviation Least Squares	% Deviation Method of Averages
Linear	.46	.43
Semilog	.41	.39
Semilog w/curve	.34	.34
Log-Log	20.06	14.07
Parabolic	.51	.50
Hyperbolic	13.18	147.63
Sigmoidal	.13	.13
Reciprocal Y	.37	.34
Reciprocal X	2.79	3.34
Reciprocal X,Y	2.71	2.95

Hit return key to continue...

Data File: Air

Sigmoidal Correlation: Y = S + A*(B^(C^(X/IN - 1)))

A: .0366
B: .0104
S: .242
C: .7726
IN: 100

Actual EQ: Y = .242 + .0366*(.0104^(.7726^(X/100 - 1)))

X	Y	YCALC	% Dev
0	.2420	.2420	0.00
100	.2422	.2424	.08
200	.2434	.2431	.12
300	.2450	.2444	.24
400	.2470	.2464	.24
500	.2495	.2491	.15
600	.2523	.2524	.04
700	.2555	.2558	.13
800	.2590	.2591	.04
900	.2620	.2624	.16
1000	.2650	.2652	.08
1100	.2680	.2679	.03
1200	.2710	.2700	.37

Aver. Dev: .13%

Fig. 1
TEMPERATURE OF THE EARTH Vs. DEPTH

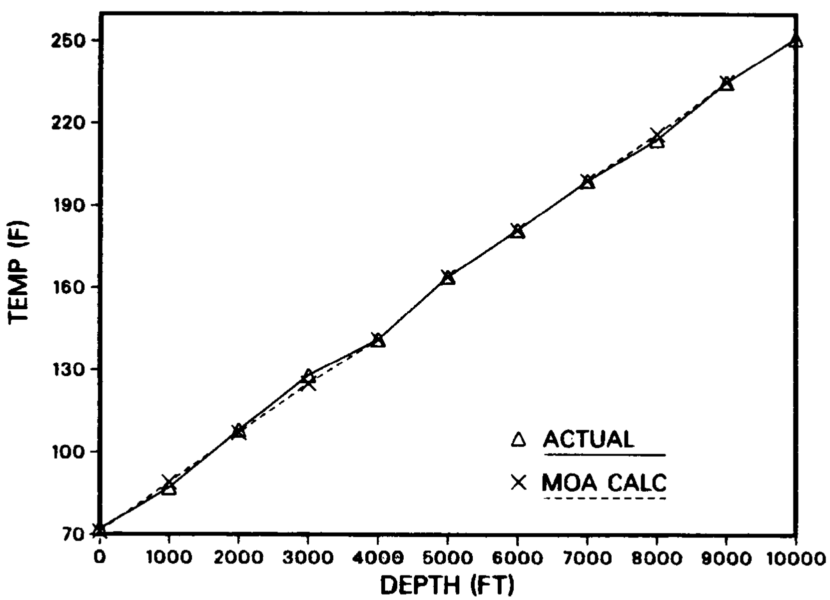


Table 1
Temperature versus Depth

(1) Depth (ft)	(2) Temp (deg F)	(3a) Calc Temp (deg F)	(4a) % Dev	(3b) Calc Temp (deg F)	(4b) % Dev
0	72	71	1.39	71	1.39
1000	87	89	2.30	89	2.30
2000	108	107	0.92	108	0.00
3000	128	125	2.33	126	1.56
4000	141	144	2.13	144	2.13
5000	164	162	1.22	162	1.22
6000	181	180	0.55	180	0.55
7000	199	198	0.50	198	0.50
8000	214	216	0.93	216	0.93
9000	235	235	0.00	234	0.43
10000	251	253	1.19	252	0.41
Average % Deviation:			1.19		1.04

Fig. 2

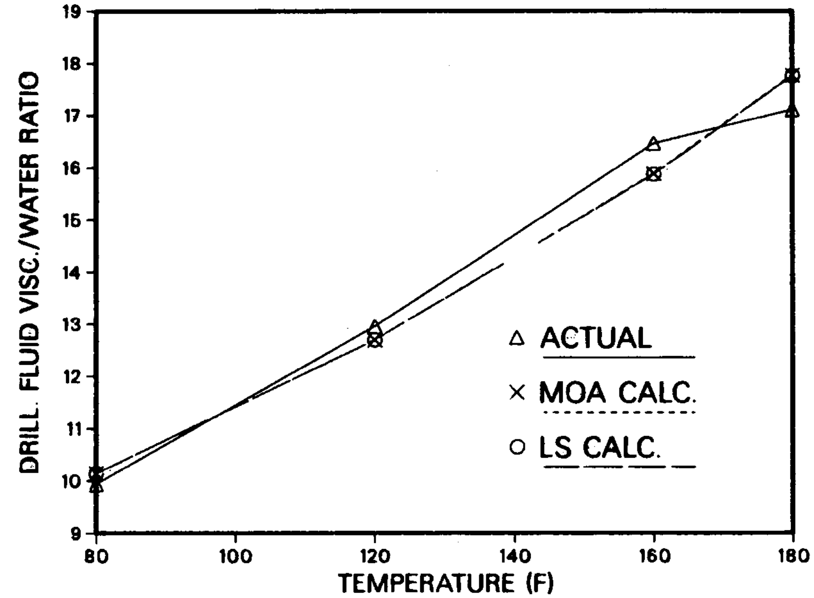


Table 2a

(1) Temp (deg F)	(2) FV (cp)	(3) WV (cp)	(4) FV/WV (actual)	(5) ln(FV/WV)	(6) FV/WV (calc)	(7) % Dev
80	8.75	0.88	9.9432	2.2969	10.1442	2.02
120	7.25	0.56	12.9464	2.5608	12.6911	1.97
160	6.75	0.41	16.4634	2.8011	15.8775	3.55
180	6.50	0.38	17.1053	2.8394	17.7590	3.83

Average Deviation: 2.84%

Table 2b

(1) Temp (deg F)	(2) FV (cp)	(3) WV (cp)	(4) FV/WV (actual)	(5) ln(FV/WV)	(6) FV/WV (calc)	(7) % Dev
80	8.75	0.88	9.9432	2.2969	10.1411	2.00
120	7.25	0.56	12.9464	2.5608	12.6873	2.00
160	6.75	0.41	16.4634	2.8011	15.8726	3.59
180	6.50	0.38	17.1053	2.8394	17.7537	3.80

Average Deviation: 2.84%

Fig. 3
VISCOSITY OF WATER AT VARIOUS TEMP.

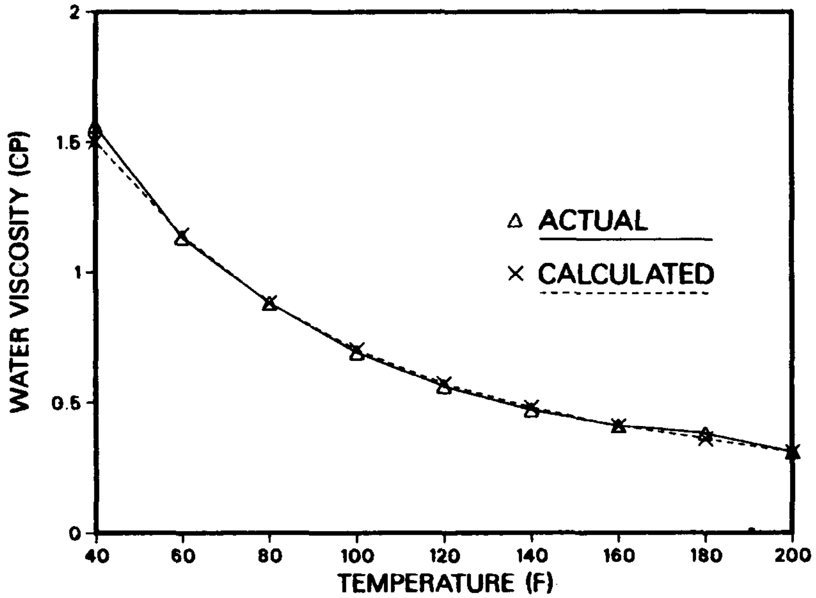


Table 3

(1) Temp (deg F)	(2) Actual Viscosity (cp)	(3) Calc Viscosity (cp)	(4) % Dev
40	1.56	1.50	3.84
60	1.13	1.14	0.88
80	0.88	0.88	0.00
100	0.69	0.70	1.46
120	0.56	0.57	1.80
140	0.47	0.48	2.13
160	0.41	0.41	0.00
180	0.38	0.36	5.26
200	0.31	0.31	0.00

Average Deviation: 1.71%

Fig. 4
PRESSURE DROP AT VARIOUS FLOW RATES
IN A LAMINAR FLOW OF WATER SUSPENSION

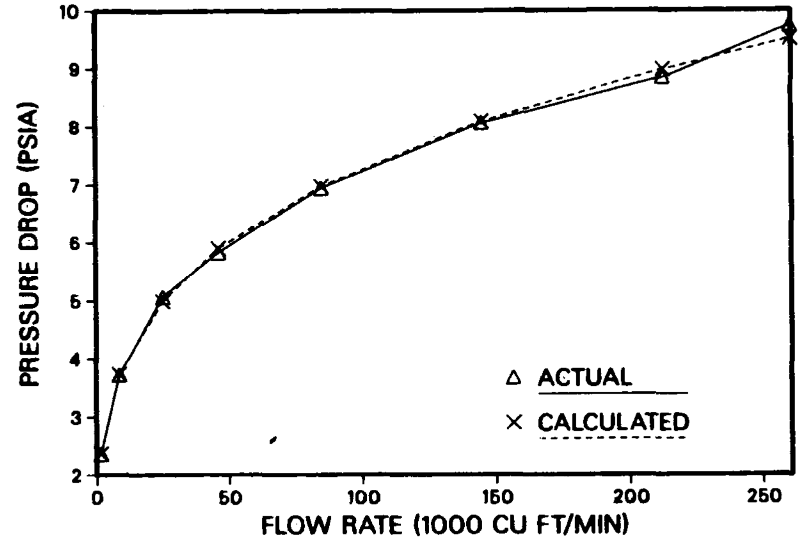


Table 4

(1) Flow Rate q x 1000 (cu ft/min)	(2) Pressure Drop (psia)	(3) Pressure Drop (calc)	(4) % Dev
1.674	2.37	2.37	0.00
8.789	3.73	3.74	0.28
25.110	5.07	4.99	1.58
46.040	5.83	5.90	1.21
84.540	6.94	6.97	0.44
144.800	8.06	8.09	0.38
212.600	8.84	8.98	1.59
260.700	9.74	9.51	2.36

Average Deviation: 0.98%

Fig. 5
SATURATION OF A SYNTHETIC OIL AT 322 DEG K
WITH VARIOUS CONCENTRATIONS OF CO2

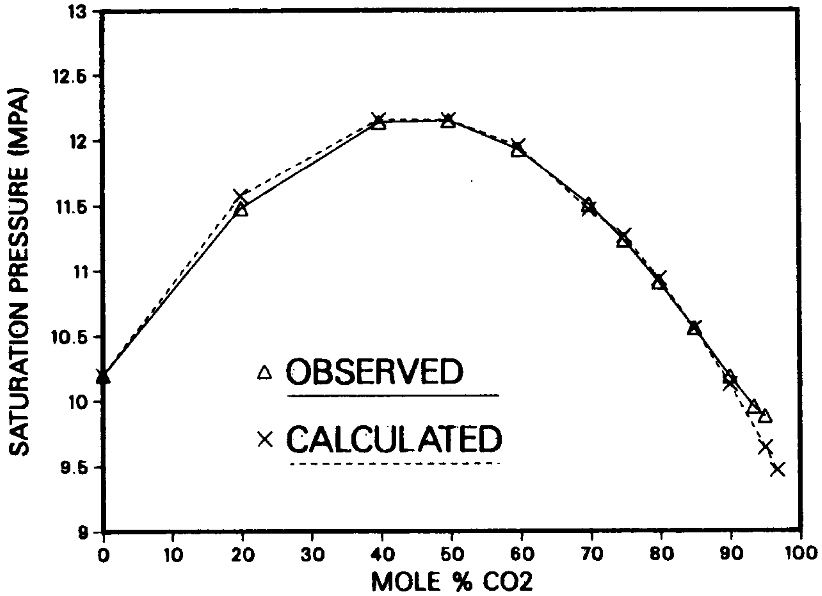


Table 5

(1) Mole Percent CO_2 in Mixture	(2) Observed Saturation Pressure, MPa	(3) Calculated Saturation Pressure, Mpa	(4) % Dev
0.00	10.197	10.197	0.00
19.87	11.480	11.573	0.82
39.73	12.135	12.158	0.20
49.69	12.149	12.155	0.06
59.70	11.928	11.952	0.21
69.77	11.507	11.456	0.34
74.82	11.225	11.265	0.37
79.85	10.908	10.936	0.27
84.87	10.556	10.556	0.00
89.93	10.1[illegible]4	10.122	0.61
93.38	9.953	9.797	1.56
94.98	9.880	9.639	2.44
96.66	9.729	9.466	2.70

Average Deviation: 0.74%

Fig. 6
API GRAVITY OF STOCK TANK OIL AS A FUNCTION
OF RESERVE OIL VISCOSITY

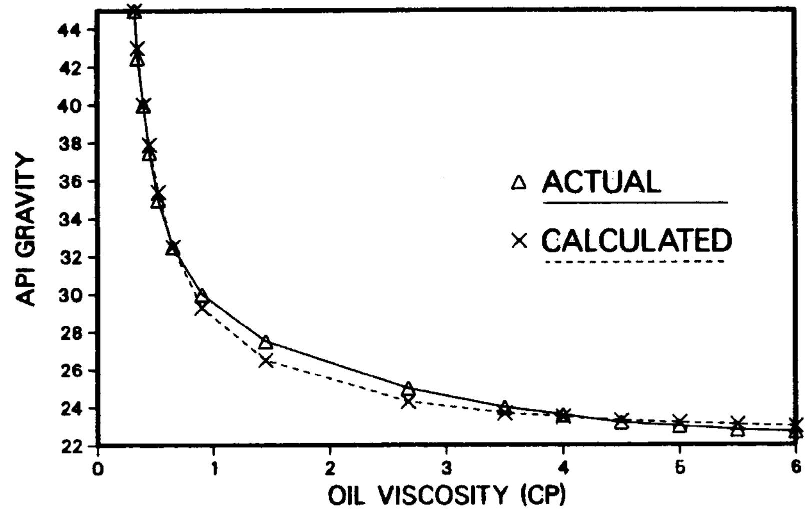


Table 6

(1) Oil Visc Xi	(2) API Yi	(3) Yi'	(4) Yi Calc	(5) % Dev
0.325	45.0	0.000	45.0	0.00
0.350	42.5	-0.010	43.0	1.19
0.400	40.0	-0.015	40.0	0.00
0.450	37.5	-0.017	37.9	1.07
0.525	35.0	-0.020	35.4	1.15
0.650	32.5	-0.026	32.5	0.00
0.900	30.0	-0.038	29.3	2.33
1.450	27.5	-0.064	26.5	3.63
2.675	25.0	-0.118	24.3	2.79
3.500	24.0	-0.151	23.7	1.25
4.000	23.6	-0.172	23.5	0.42
4.500	23.2	-0.192	23.3	0.44
5.000	23.0	0.213	23.2	0.88
5.500	22.8	-0.233	23.1	1.32
6.000	22.7	-0.254	23.0	1.32

Average Deviation: 1.19%

Fig. 7
SPECIFIC HEAT OF AIR AT 0-1 ATMOSPHERES
AT VARIOUS TEMPERATURES

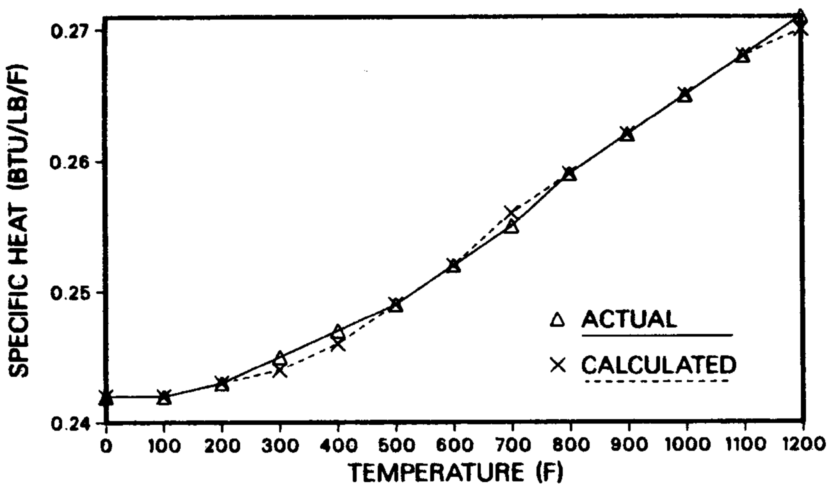


Table 7

(1) Temp (deg F)	(2) Cp (Btu/lb/deg F)		(3) ln(Yi-0.242)	(4) Cp Y Calc	(5) % Dev
0	0.2420		--	0.2420	0.00
100	0.2422		-8.5172	0.2424	0.08
200	0.2434		-6.5713	0.2431	0.12
300	0.2450		-5.8091	0.2444	0.24
400	0.2470		-5.2983	0.2464	0.24
		S1 =	-26.1959		
500	0.2495		-4.8929	0.2491	0.15
600	0.2523		-4.5756	0.2524	0.04
700	0.2555		-4.3051	0.2558	0.13
800	0.2590		-4.0745	0.2591	0.04
		S2 =	-17.8481		
900	0.2620		-3.9120	0.2624	0.16
1000	0.2650		-3.7723	0.2652	0.08
1100	0.2680		-3.6497	0.2679	0.03
1200	0.2710		-3.5405	0.2700	0.37
		S3 =	-14.8744		

Average Deviation: 0.13%

Fig. 8
RECIPROCAL CURVE EXAMPLES