Separation Science and Technology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713708471

Curve Fitting in Science and Technology

A. S. Said[a]; R. S. Al-ameeri[a]
[a] DEPARTMENT OF CHEMICAL ENGINEERING, UNIVERSITY OF KUWAIT, KUWAIT

To cite this Article Said, A. S. and Al-ameeri, R. S.(1987) 'Curve Fitting in Science and Technology', Separation Science and Technology, 22: 1, 65 — 84
To link to this Article: DOI: 10.1080/01496398708056158
URL: http://dx.doi.org/10.1080/01496398708056158

SEPARATION SCIENCE AND TECHNOLOGY, **22**(1), pp. 65-84, 1987

Curve Fitting in Science and Technology

A. S. SAID and R. S. AL-AMEERI

DEPARTMENT OF CHEMICAL ENGINEERING
UNIVERSITY OF KUWAIT, KUWAIT

Abstract

The classical methods of curve fitting using polynomials are not necessarily the best methods in science and technology. In most cases other functions (the inverse linear, exponential and logarithmic functions) lead to a better fit of the original function or data. The best fitting function is the function which has the same asymmetry as the original function or data around the 45° line with reference to normalized coordinate axes. In the present paper this technique of curve fitting is applied to five different fitting functions. An extensive numerical example is given to illustrate the application of the derived formulas. It can be extended to more complex fitting functions, to the derivation of accurate interpolation and extrapolation formulas, and to the numerical solution of differential and partial differential equations of interest in chemistry, chemical engineering, separation science, and other branches of science and technology.

INTRODUCTION

In different branches of science and technology there are certain relationships which cannot be represented by easy analytic expressions, yet, on plotting these relationships, a simple curve is obtained which can be approximated quite accurately by a simple function. Examples of such relationships are the bubble point t_b and dew point t_d of a binary ideal liquid mixture and the relation between the relative volatility α of this mixture with temperature t or composition x, where x is the mole fraction of the more volatile component in the ideal liquid mixture.

On the other hand, the differential equation representing a certain phenomenon might not be of a standard form leading to an analytic solution, yet, on integrating it numerically and plotting the results, a simple curve is obtained which can be easily approximated by a simple mathematical expression.

The approximating function can be determined from the boundary conditions of the problem under consideration, particularly from the slopes at the beginning and end which can be easily determined even for complicated implicit functions and nonconventional differential equations. Another useful parameter is the area under the curve under consideration which can usually also be obtained from the boundary conditions. From these and other boundary conditions, the approximating function is deduced.

THE *Y-X* COORDINATES

For the purpose of generality we express all functions under consideration in terms of dimensionless parameters Y and X such that Y is the vertical dimensionless parameter and X is the horizontal dimensionless parameter. For example, if we desire a relation between vapor pressure P and temperature t in the temperature range from t_1 to t_2, the corresponding dimensionless parameters would be

$$Y = \frac{P - P_1}{P_2 - P_1} \quad \text{and} \quad X = \frac{t - t_1}{t_2 - t_1}$$

where P_2 and P_1 are the pressures at t_2 and t_1, respectively. Similarly, for a relation between relative volatility α and temperature, we define the dimensionless coordinates as

$$Y = \frac{\alpha - \alpha_0}{\alpha_1 - \alpha_0} \quad \text{and} \quad X = \frac{t_0 - t}{t_0 - t_1}$$

Subscript 0 corresponds to $x = 0$, where x is the mole fraction of the more volatile component, and subscript 1 corresponds to $x = 1$, therefore $t_0 \geqslant t \geqslant t_1$ and $\alpha_1 \geqslant \alpha \geqslant \alpha_0$.

Similarly, if an explicit relation for the bubble point of a binary ideal liquid mixture is required as a function of mole fraction, we define Y and X by

$$Y = \frac{t_0 - t}{t_0 - t_1} \quad \text{and} \quad X = x$$

t_0 and t_1 are the boiling points corresponding to $x = 0$ and 1, respectively, hence t_0 is the boiling point of the less volatile component and t_1 is the boiling point of the more volatile component.

In general, if $y = f(x)$, where y is any dependent variable and x is any independent variable not necessarily composition, and if it is desired to approximate this function in the region x_i to x_f where i and f refer to initial and final conditions, respectively, and if y_i and y_f are the corresponding values of the dependent variable, then

$$Y = \frac{y - y_i}{y_f - y_i} \quad (1)$$

$$X = \frac{x - x_i}{x_f - x_i} \quad (2)$$

$$\frac{dY}{dX} = \frac{dY}{dy}\frac{dy}{dx}\frac{dx}{dX}$$

or

$$\frac{dY}{dX} = \left(\frac{x_f - x_i}{y_f - y_i}\right)\frac{dy}{dx} = \theta\,\frac{dy}{dx} \quad (3)$$

After developing the relation between Y and X, the original parameters are substituted to get the relation between P and t, α and x, etc.

FITTING FUNCTIONS

One can think of a large number of functions where Y varies from 0 to 1 as X varies between 0 and 1. Only five important functions will be considered in this paper, namely,

I. The inverse linear function

$$Y = \frac{aX}{1 + (a - 1)X} \quad (4)$$

II. The exponential function

$$Y = \frac{g^X - 1}{g - 1} \quad (5)$$

III. The logarithmic function

$$Y = \frac{\ln [X(b - 1) + 1]}{\ln b} \tag{6}$$

IV. The quadratic function

$$Y = cX + (1 - c)X^2 \tag{7}$$

V. The Poisson-type function

$$Y = Xk^{1-X} \tag{8}$$

where *a, g, b, c,* and *k* are constants.

By differentiation and integration, m_0, m_1, and I were deduced for each function, where

$$m_0 = \text{slope at } (X = 0) = (dY/dX)_{X=0} \tag{9}$$

$$m_1 = \text{slope at } (X = 1) = (dY/dX)_{X=1} \tag{10}$$

and

$$I = \text{area under the } Y \text{ vs } X \text{ curve} = \int_0^1 Y dX \tag{11}$$

Values of I, m_0, m_1, and m_0m_1 are given in Table 1 for all five functions.

As will be shown later, m_0m_1 is a measure of the asymmetry of the function around the 45° line.

It should be noted that each of the above five functions contains one constant only and has finite slopes at $X = 0$ and $X = 1$. Other functions which have slopes equal to zero or infinity at one or both ends or functions containing more than one constant will be discussed only briefly later on. The emphasis in this paper is on the one-constant equations having finite initial and final slopes of which Eqs. (4) to (8) are important examples.

DETERMINATION OF THE CONSTANTS *a, g, b, c,* AND *k*

These constants can be evaluated from the coordinates *Y, X* of one point only on the curve and particularly from $Y_{1/2}$ or $X_{1/2}$ which leads to

TABLE 1

Function	I	m_0	m_1	$m_0 m_1$
Inverse linear (I.L.)	$\frac{a(a-1-\ln a)}{(a-1)^2}$	a	$\frac{1}{a}$	1
Exponential (Ex)	$\frac{1}{\ln g} - \frac{1}{g-1}$	$\frac{\ln g}{g-1}$	$\frac{g \ln g}{g-1}$	$\frac{g(\ln g)^2}{(g-1)^2}$
Logarithmic (Log)	$\frac{b}{b-1} - \frac{1}{\ln b}$	$\frac{b-1}{\ln b}$	$\frac{b-1}{b \ln b}$	$\frac{(b-1)^2}{b(\ln b)^2}$
Quadratic (Q)	$\frac{1}{3} + \frac{c}{6}$	c	$2-c$	$c(2-c)$
Poisson (P)	$\frac{k - \ln k - 1}{(\ln k)^2}$	k	$1 - \ln k$	$k(1 - \ln k)$

easier relations for the constants and also to a better fit. $Y_{1/2}$ is the value of Y at $X = 1/2$ and $X_{1/2}$ is the value of X at $Y = 1/2$. For convenience we denote $Y_{1/2}$ by the symbol Z and $X_{1/2}$ by the symbol $\bar{Z}$ The constants may be obtained also from m_0 or m_1 or even I. Relations between the constants and I are rather complex. They are deduced from m_0 or m_1 if the emphasis is on the first portion or last portion of the curve, respectively, or if the values of Z and $\bar{Z}$ are not easy to determine. In general, whenever possible, the constants are deduced from Z or $\bar{Z}$ which lie around the middle portion of the curve. This will, in general, lead to an overall better accuracy over the entire length of the curve.

The five values obtained from m_0, m_1, I, Z, and $\bar{Z}$ are identical only if the fit between the original and approximating function is perfect, otherwise they differ from one another. If the difference is significant, then the approximating function is not satisfactory.

The equations relating the different constants to (Y and X), Z, $\bar{Z}$, m_0, and m_1 can be deduced either from explicit functions directly or from implicit functions by trial and error.

In most practical cases, however, $Y_{1/2}$ is known or can be determined easily. The constants a, g, c, and k for the inverse linear, exponential, quadratic, and Poisson functions can be expressed easily in terms of $Z \equiv Y_{1/2}$, leading to

$$a = \frac{Z}{1 - Z} \tag{12}$$

$$g = \left(\frac{1 - Z}{Z}\right)^2 \tag{13}$$

$$c = 4Z - 1 \tag{14}$$

and

$$k = 4Z^2 \tag{15}$$

Substituting into Eqs. (4), (5), (7), and (8) gives the relations between Y and X in terms of Z as follows.

Inverse linear:

$$Y = \frac{ZX}{1 - Z + X(2Z - 1)} \tag{16}$$

Exponential:

$$Y = \frac{\left(\frac{1 - Z}{Z}\right)^{2X} - 1}{\left(\frac{1 - Z}{Z}\right)^2 - 1} \tag{17}$$

Quadratic:

$$Y = (4Z - 1)X + 2(1 - 2Z)X^2 \tag{18}$$

Poisson:

$$Y = X(2Z)^{2(1-X)} \tag{19}$$

An explicit relation for b in terms of Z is not possible. In this case, b is calculated by trial and error from Z and its values are substituted in Eq. (6) to get the fitting logarithmic function. An analytic expression is possible, however, in terms of $X_{1/2} \equiv \bar{Z}$.

ASYMMETRY OF THE *Y* vs *X* FUNCTION

With reference to Fig. 1, θ_0 and θ_1 are the two angles between the tangents to the Y-X curve and the 45° line at $X = 0$ and $X = 1$, respectively. A curve is symmetrical around the 45° line if $\theta_0 = \theta_1$. If $\theta_1 \neq \theta_0$, then the curve is asymetric around the diagonal. When $\theta_1 > \theta_0$, the curve has an asymmetry of the first kind, and when $\theta_1 < \theta_0$, the curve has an asymmetry of the second kind.

Since this paper is devoted to curves which have no maxima or minima or points of inflection between $X = 0$ and $X = 1$, then the curves of interest will lie totally above the diagonal, in which case both θ_0 and θ_1 have positive values, or lie totally below the diagonal, in which case both θ_0 and θ_1 have negative values.

θ_0 and θ_1 are related to m_0 and m_1, respectively, as follows:

$$m_0 = \tan(45 + \theta_0) \tag{20}$$

$$m_1 = \tan(45 - \theta_1) \tag{21}$$

Solving for m_0 and m_1 gives

$$m_0 = \frac{1 + \tan\theta_0}{1 - \tan\theta_0} \tag{22}$$

and

$$m_1 = \frac{1 - \tan\theta_1}{1 + \tan\theta_1} \tag{23}$$

and hence

$$\tan\theta_0 = \frac{m_0 - 1}{m_0 + 1} \tag{24}$$

and

$$\tan\theta_1 = \frac{1 - m_1}{1 + m_1} \tag{25}$$

or

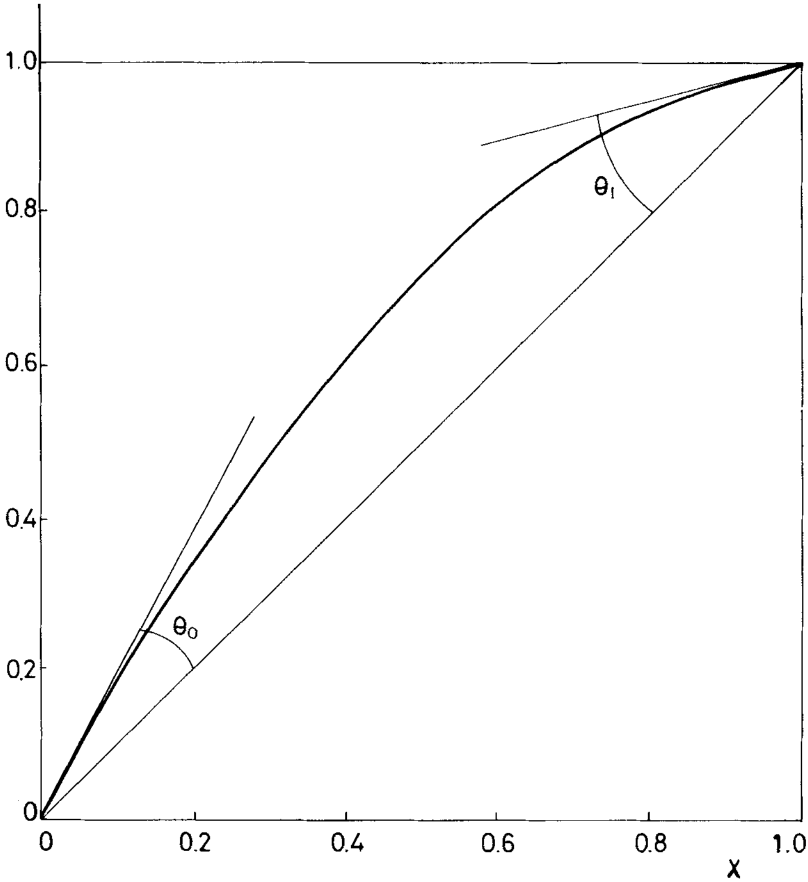

FIG. 1. Illustrative curve.

$$\theta_0 = \tan^{-1} \frac{m_0 - 1}{m_0 + 1} \tag{26}$$

and

$$\theta_1 = \tan^{-1} \frac{1 - m_1}{1 + m_1} \tag{27}$$

In the case of the inverse linear function (Eq. 4), $m_0 = a$ and $m_1 = 1/a$, and hence

$$m_0 = 1/m_1 \tag{28}$$

Substituting from Eq. (20) into Eq. (16) gives

$$\tan \theta_0 = \frac{1 - m_1}{1 + m_1} \tag{29}$$

Comparing with Eq. (17) leads to

$$\theta_0 = \theta_1$$

Therefore, the inverse linear function is symmetric around the 45° line.

One can also show that in the case of the Exponential (Eq. 5), Quadratic (Eq. 7), and Poisson-type (Eq. 8) functions, θ_1 is always greater than θ_0, and hence they possess asymmetries of the first kind. On the other hand, the logarithmic function (Eq. 6) has an asymmetry of the second kind where θ_1 is always less than θ_0. In Fig. 2 the Inverse linear ($\theta_0 = \theta_1$), the Exponential ($\theta_1 > \theta_0$), and the Logarithmic ($\theta_1 < \theta_0$) functions are plotted for $y_{1/2} = 0.2$ and 0.8.

As can be seen from Fig. 2, the difference between the three functions at these extreme values of $Y_{1/2}$ is significant. As the value of $Y_{1/2}$ or $X_{1/2}$ approaches 0.5, the difference between them becomes smaller and smaller, and for values of Z and $\bar{Z}$ between .45 and .55, the three functions lead to approximately identical results.

Furthermore, one can prove the following:

$$\theta_1 > \theta_0 \qquad \text{for } m_0 m_1 < 1 \tag{30}$$

$$\theta_1 < \theta_0 \qquad \text{for } m_0 m_1 > 1 \tag{31}$$

and

$$\theta_1 = \theta_0 \qquad \text{for } m_0 m_1 = 1 \tag{32}$$

This leads to a convenient asymmetry index As which is defined by the simple equation

$$As = m_0 m_1 \tag{33}$$

Substituting from Eqs. (12), (13), (14), and (15) into the corresponding equations for $m_0 m_1$ in Table 1 gives

Inverse linear:

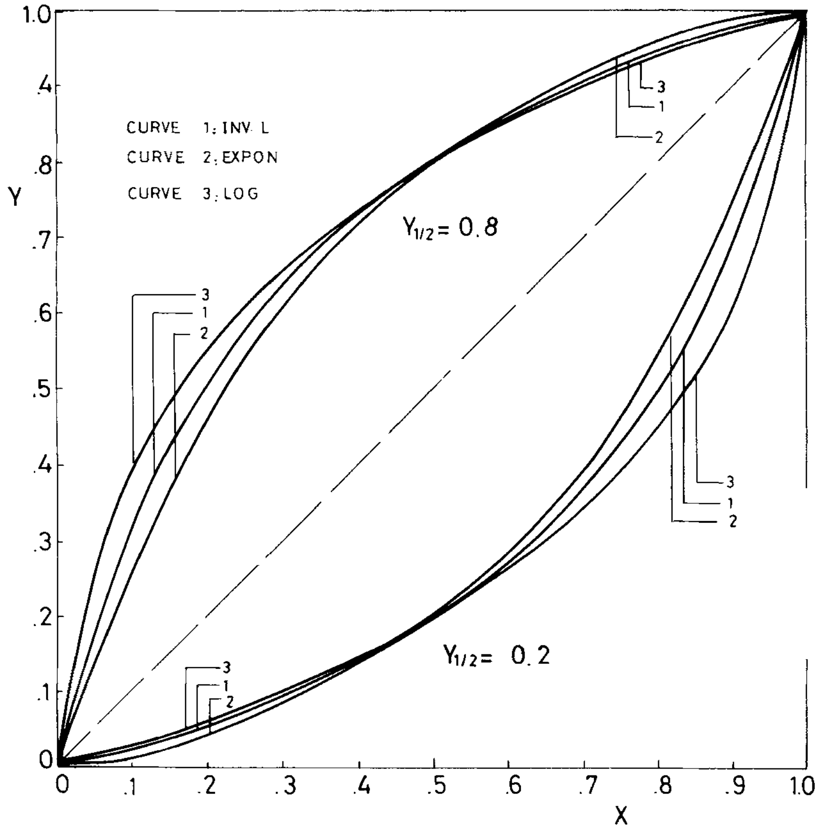

FIG. 2. Plots of inverse linear, exponential, and logarithmic functions.

$$m_0 m_1 = 1 \tag{34}$$

Exponential:

$$m_0 m_1 = \left[\frac{2Z(1 - Z)}{1 - 2Z} \ln \frac{1 - Z}{Z}\right]^2 \tag{35}$$

Poisson:

$$m_0 m_1 = 4Z^2(1 - 2 \ln 2Z) \tag{36}$$

Quadratic:

$$m_0 m_1 = (4Z - 1)(3 - 4Z) \tag{37}$$

m_0m_1 was calculated for different fitting functions using the above equations in the range from $Z = 0.2$ to $Z = 0.8$. The calculated values are given in Table 2 and are plotted in Fig. 3.

For the Inverse linear fit, m_0m_1 is equal to 1 regardless of the value of Z. Values of m_0m_1 vs Z for the logarithmic function were calculated from the corresponding equation in Table 1 by trial and error and are plotted in Fig. 3. They are not included in Table 2 because the logarithmic fit can always be accomplished through the exponential fit. If y vs x is fitted best by a logarithmic function, then x vs y is fitted best by an exponential function. As can be seen from Fig. 3, for the same value of Z, the asymmetry m_0m_1 is equal to $(m_0m_1)^{-1}$ for the exponential function.

If the original function is given in the form of an experimental curve, then m_0 and m_1 are obtained by measuring the slopes of the experimental curve at $X = 0$ and $X = 1$ after replotting it on dimensionless X–Y coordinates or by using Eq. (3).

The application of the above equations and methods are illustrated by the following example.

ILLUSTRATIVE EXAMPLE

The relation between the vapor pressure of benzene and temperature is given accurately by the Antoine equation:

$$\ln P = A - \frac{B}{t + C} \tag{38}$$

where P is the vapor pressure in mmHg and t is the temperature in °C. A, B, and C are Antoine constants. For benzene, they have the following values:

$$A = 15.871, \qquad B = 2771.23, \qquad C = 219.89$$

It is required to express the relation between P and t in terms of the five fitting functions given above in the range $t = 80°\text{C}$ to $t = 110°\text{C}$ and to compare the values given by all fitting functions with the values given by the original function.

TABLE 2

	m_0m_1				m_0m_1		
$Y_{1/2}$	Exponential	Poisson	Quadratic	$Y_{1/2}$	Exponential	Poisson	Quadratic
.20	.5466	.4532	−.4400	.50	1.0000	1.0000	1.0000
.21	.5745	.4825	−.3456	.51	.9995	.9992	.9984
.22	.6017	.5115	−.2544	.52	.9979	.9968	.9936
.23	.6282	.5402	−.1664	.53	.9952	.9927	.9856
.24	.6539	.5686	−.0816	.54	.9915	.9869	.9744
.25	.6789	.5966	.0000	.55	.9867	.9794	.9600
.26	.7031	.6240	.0784	.56	.9808	.9701	.9424
.27	.7265	.6510	.1536	.57	.9739	.9590	.9216
.28	.7490	.6773	.2256	.58	.9660	.9462	.8976
.29	.7707	.7029	.2944	.59	.9570	.9315	.8704
.30	.7915	.7278	.3600	.60	.9470	.9149	.8400
.31	.8114	.7519	.4224	.61	.9359	.8965	.8064
.32	.8303	.7752	.4816	.62	.9238	.8761	.7696
.33	.8483	.7976	.5376	.63	.9107	.8538	.7296
.34	.8654	.8191	.5904	.64	.8966	.8295	.6864
.35	.8815	.8395	.6400	.65	.8815	.8032	.6400
.36	.8966	.8590	.6864	.66	.8654	.7749	.5904
.37	.9107	.8774	.7296	.67	.8483	.7446	.5376
.38	.9238	.8946	.7696	.68	.8303	.7122	.4816
.39	.9359	.9107	.8064	.69	.8114	.6777	.4224
.40	.9470	.9256	.8400	.70	.7915	.6410	.3600
.41	.9570	.9393	.8704	.71	.7707	.6023	.2944
.42	.9660	.9516	.8976	.72	.7490	.5614	.2256
.43	.9739	.9627	.9216	.73	.7265	.5183	.1536
.44	.9808	.9724	.9424	.74	.7031	.4729	.0784
.45	.9867	.9807	.9600	.75	.6789	.4254	.0000
.46	.9915	.9875	.9744	.76	.6539	.3756	−.0816
.47	.9952	.9929	.9856	.77	.6282	.3236	−.1664
.48	.9979	.9968	.9936	.78	.6017	.2692	−.2544
.49	.9995	.9992	.9984	.79	.5745	.2126	−.3456
.50	1.0000	1.0000	1.0000	.80	.5467	.1536	−.4400

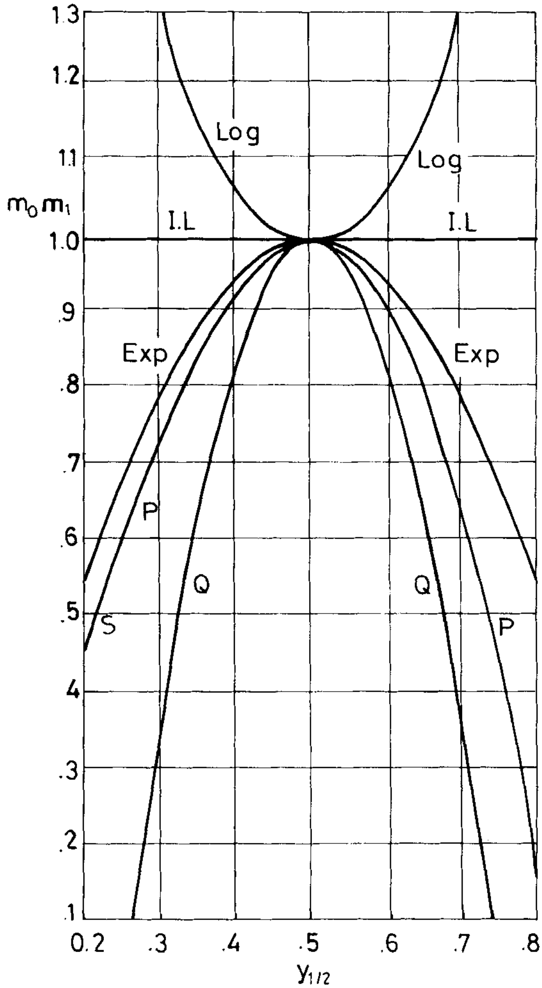

FIG. 3. Plots of $m_0 m_1$ versus $y_{1/2}$ for logarithmic, inverse linear, exponential, poisson, and quadratic functions.

Solution

The two dimensionless parameters are

$$Y = \frac{P - P_1}{P_2 - P_1} \quad \text{and} \quad X = \frac{t - t_1}{t_2 - t_1}$$

P_1 is the initial pressure corresponding to initial temperature t_1, and P_2 is the final pressure corresponding to final temperature t_2. Substituting in the Antoine equation gives $P_1 = 757.6$ mmHg and $P_2 = 1755.6$ mmHg. Therefore,

$$Y = \frac{P - 757.6}{998.0} \tag{39}$$

and

$$X = \frac{t - 80}{30} \tag{40}$$

At $t = 80$, $X = 0$ and $Y = 0$. At $t = 110$, $X = 1$ and $Y = 1$.

Dividing the temperature span into 10 equal intervals, the pressure corresponding to each intermediate temperature was calculated using Eq. (38), and the results are tabulated in Columns 1 and 2 of Table 3. The corresponding values of X and Y are also tabulated in Columns 3 and 4, respectively.

From the calculated values in Table 3, the value of Y at $X = ½$ is given by

$$Y_{1/2} = Z = 0.4198$$

Substituting this value in Eqs. (16)–(19) and substituting for Y and X from Eqs. (39) and (40), respectively, and rearranging:

For the inverse linear fit:

$$P = 757.6 + \frac{2611.4(t - 80)}{188.5 - t} \tag{41}$$

For the Exponential fit:

$$P = 757.6 + 1096.5(0.1775e^{0.0216t} - 1) \tag{42}$$

TABLE 3

				Inverse linear		Exponential		Logarithmic		Quadratic		Poisson	
t (°C)	P (mmHg)	X	Y	P_{calc}	E	P_{calc}	E	P_{calc}	E	P_{calc}	E	P_{calc}	E
80	757.6	0	0	757.6	0	757.6	0	757.6	0	757.6	0	757.6	0
83	830.1	.1	.0726	831.9	1.8	830.9	.8	832.7	2.6	828.6	−1.5	830.4	.3
86	908.2	.2	.1509	910.5	2.3	909.1	.9	911.6	3.4	906.0	−2.2	908.5	.3
89	991.7	.3	.2346	993.8	2.1	992.6	.9	994.9	3.2	989.8	−1.9	992.0	.3
92	1081.1	.4	.3241	1082.3	1.2	1081.6	.5	1083.0	1.9	1080.0	−1.1	1081.3	.2
95	1176.6	.5	.4198	1176.6	0	1176.6	0	1176.6	0	1176.6	0	1176.6	0
98	1278.5	.6	.5219	1277.0	−1.5	1277.9	−.6	1276.0	−2.5	1279.6	1.1	1278.2	−.3
101	1387.1	.7	.6308	1384.3	−2.8	1386.0	−1.1	1382.5	−4.6	1389.0	1.9	1386.6	−.5
104	1502.6	.8	.7465	1499.3	−3.3	1501.3	−1.3	1497.0	−5.6	1504.8	2.2	1502.1	−.5
107	1625.3	.9	.8694	1622.7	−2.6	1624.3	−1.0	1620.9	−4.4	1627.0	1.7	1624.9	−.4
110	1755.6	1.0	1	1755.6	0	1755.6	0	1755.6	0	1755.6	0	1755.6	0

For the Poisson fit:

$$P = 757.6 + 9.228(t - 80)e^{0.011657t} \tag{43}$$

For the Quadratic fit:

$$P = 1226.7 - 34.32t + 0.3557t^2 \tag{44}$$

In the case of the logarithmic fit, b is calculated by trial and error from the relation

$$.4198 = \frac{\ln [\frac{1}{2}(b + 1)]}{\ln b}$$

and we get $b = 0.5205$.

Substituting in Eq. (6) leads to

$$P = 7068 - 1526 \ln (142.5 - t) \tag{45}$$

P was calculated using the five equations (41 to 45) for different values of X between $X = 0$ and $X = 1$, and the results are listed in Table 3 together with the values obtained from the original function. The error E is also listed in Table 3 where $E = P_{\text{calc}} - P$.

As can be seen from the table, the Poisson function gives the best fit in this case. It leads to the minimum error among the five functions tested.

The best function could have been determined from the beginning by calculating m_0m_1 for the original function and comparing it with the asymmetry of the different fitting functions as given by Fig. 3 or Table 2.

To calculate m_0m_1 for the original function, we proceed as follows. Differentiating Eq. (38) gives

$$\frac{dP}{dt} = \frac{BP}{(t + c)^2} \tag{46}$$

From Eqs. (3) and (46), one gets

$$m_0m_1 = \left(\frac{dY}{dX}\right)_{x=0}\left(\frac{dY}{dX}\right)_{x=1} \tag{47}$$

or

$$m_0 m_1 = \left(\frac{t_2 - t_1}{P_2 - P_1}\right)^2 \frac{B^2 P_1 P_2}{(t_1 + c)^2 (t_2 + c)^2} \tag{48}$$

Substituting $t_2 = 110°C$, $t_1 = 80°C$, $P_2 = 1755.6$ mmHg, $P_1 = 757.6$ mmHg, $B = 2771.23$, and $C = 219.89$ gives $m_0 m_1 = 0.9430$.

Comparing this asymmetry with the asymmetries of the different functions at $Z = 0.4198$ as given in Fig. 3, one finds that the Poisson function has the nearest asymmetry and therefore gives the best fit. For a more accurate comparison, the asymmetries of the different functions were calculated from Eqs. (34)–(37) and from the corresponding formula given in Table 1 in the case of the Logarithmic function. The calculated values are listed in Table 4 together with the asymmetry deviation $|D|$ where $|D| = |(As)_f - (As)_o|$. The subscripts f and o refer to fitting and original functions, respectively.

Defining an average error E_{av} by the relation $E_{av} = \Sigma|E|/10$, where the figure 10 is the number of equal intervals at which the error E is measured, the average error was calculated for different fitting functions from the data tabulated in Table 3. These values are also listed in Table 4.

The last column in Table 4 lists values of the ratio $E_{av}/|D|$. The constancy of this ratio indicates that the average error is proportional to the difference in asymmetries between the fitting and original function.

Therefore, to choose the best fitting function, it is sufficient to calculate As for the original function and see where it stands on Fig. 3 and which curve is nearest to it. The nearest function gives the best fit.

TABLE 4

Function	$m_0 m_1$ at $Y_{1/2} = .4198$	$\|D\|$	E_{av} (mmHg)	$E_{av}/\|D\|$
Inverse linear	1	.0570	1.76	30.9
Exponential	0.9658	.0228	0.71	31.1
Logarithmic	1.0360	.0930	2.82	30.3
Quadratic	0.8971	.0459	1.36	29.6
Poisson	0.9514	.0084	0.28	33.3

SPECIAL AND MORE COMPLEX FITTING FUNCTIONS

Functions (4)–(8) may be used to fit relationships that arise in chemical engineering and particularly in distillation (bubble points and dew points of ideal mixtures, variation of the relative volatility α with temperature or composition, etc.). They are one-constant equations having finite slopes at $X = 0$ and $X = 1$.

For a more accurate fit, two constant equations may be utilized. From the many formulas which can be utilized, the following equation is probably the simplest and most useful.

$$Y = \frac{aX}{1 + (a - 1)X^n} \tag{49}$$

where a and n are constants.

Similar to Eqs. (4)–(8), Eq. (44) has finite slopes at $X = 0$ and $X = 1$. By differentiation one can show that

$$m_0 = a \tag{50}$$

$$m_1 = \frac{a + n - an}{a} \tag{51}$$

$$m_0 m_1 = \text{asymmetry}\, As = a + n - an \tag{52}$$

and

$$Y_{1/2} = \frac{\frac{1}{2}a}{1 + (a - 1)(\frac{1}{2})^n} \tag{53}$$

Any combination of two of the above four equations may be used to deduce the values of a and n. For better accuracy, the combination should contain $Y_{1/2}$. The values of $Y_{1/2}$, m_0, and/or m_1 are calculated for the original function and substituted in the chosen combination. The combination of $Y_{1/2}$ and m_0 leads directly to the values of a and n. The other two alternatives ($Y_{1/2}$ and m_1 or $Y_{1/2}$ and $m_0 m_1$) require trial and error for the determination of a and n. The combination of $Y_{1/2}$ and $m_0 m_1$ should give the best overall fit of the original function. The other two give a better fit at one or the other end of the original curve.

Original functions having slopes equal to 0 or ∞ at one of the two ends may be fitted by certain equations, the simplest among them being

$$Y = aX^n \tag{54}$$

There are also equations to fit curves having slopes equal to 0 and ∞ and $X = 0$ and $X = 1$ or vice versa. Curves having slopes equal to 0 and 0 or ∞ and ∞ have points of inflection between $X = 0$ and $X = 1$. Such curves also occur in chemical engineering as in the case of breakthrough curves. They can be fitted by more complex functions such as the normal distribution integral and the Poisson summation distribution.

Investigating these special and complex functions is important, but outside the scope of the present paper.

CONCLUSION

The above technique is not meant to replace least-square procedures abundantly available in modern software. It does not deal with finding the best fit for scattered experimental data. It is mainly concerned with representation of simple curves having no maximum or minimum or inflection points and for which simple or explicit relations are not available. Curves of this kind are quite abundant in science and technology. The variation of the relative volatility α of a binary ideal mixture with temperature or concentration is probably a good example.

In batch distillation calculations utilizing the Rayleigh equation, it has been common practice to utilize a constant α equal to the average of initial and final α values or equal to α at the average temperature T_{av}. A constant α in Rayleigh's integration leads to an analytic expression. By utilizing the present technique, one can show that the variation of α with mole fraction x of the more volatile component follows quite accurately the relation

$$\alpha = \alpha_0 + \frac{ax}{1 + bx} \tag{55}$$

which on substitution in the Rayleigh equation leads to integrals which are still simple and give analytic and more exact answers without the inconsistency in choosing α_{av}. The answers may not differ considerably in the case of binary distillations, but in multicomponent distillations where the temperature range is large, the errors resulting from a constant α could be appreciable. Comparing the results obtained from utilizing Eq. (55) with results obtained from utilizing such different average values of α

as log mean, geometric, and harmonic averages, it is possible to determine which average leads to the best results. The same equation for α can also be conveniently utilized in deriving a more accurate recursion formula for plate-to-plate calculations in the case of continuous fractionation of binary or multicomponent mixtures using a computer or a programmable calculator.

Received by editor January 29, 1985
Revised November 6, 1985