A USEFUL SHORT CUT APPROACH TO PHYSICAL AND CHEMICAL CURVE FITTINGS

M.S.E. Abdo*

Chemical Engineering Department, Alexandria University, Egypt.

K.T. Ewida

Dept. Of Eng. Phys. and Math., Faculty of Engineering, Zagazig Univ., Egypt. A.S. Said

Dept. of Chemical Engineering, Cairo University, Giza, Egypt.

M.S.H. Bader

School Of Chemical Engineering, Oklahoma State University, Stillwater, USA.

ABSTRACT

By proper choice of a simple two-constants fitting function, it is now possible to get better accuracies than those achieved from complex fitting functions having three or more constants. The recommended new technique is highly dependent on comparing the asymmetry of the original function or data to be fitted with those of different simple fitting functions. The one that has the nearest asymmetry gives the best fit. The proposed technique has been successfully tested in predicting the variation of some properties frequently used in chemical engineering against an independent variable. The technique has also proved to be useful in the reverse prediction if needed and is less time consuming.

INTRODUCTION

In various branches of science and technology and particularly in chemical engineering, there are properties which vary with different parameters like temperature or composition. When these properties are plotted versus the independent variable, a simple curve is obtained which has no maximum or minimum or a point of inflection in the region of interest. However, a simple analytic relation cannot be deduced theoretically for these simple curves. Examples of such properties are numerous. The variation of relative volatility, bubble point and dew point with composition for binary ideal mixtures are well known examples. Other examples include the variation of the humidity ratio of saturated air with temperature and the variation of the viscosity of water with temperature.

This work presents a systematic technique for finding the best functions which fit these simple curves. The last two examples were chosen specifically to compare with another rapid curve fitting technique [1].

THE PREVIOUS TECHNIQUE

In this technique, the author utilized the following fitting function to represent the variation of a property (F) with temperature (T):

$$F = F_i + a (T-T_i) + b (T-T_i)^n$$
 (1)

Here: F = the property of interest which may be humidity ratio (H), or viscosity (μ) at temperature (T);

F_i = the value of the property at initial temperature (T_i);

a, b, n = constants.

The author utilized tables of the humidity ratio and viscosity at equal temperature intervals. Such tables could be based on experimental data or on values calculated from complex or difficult implicit relations. From a plot of F versus T, the author obtained the slope (dF/dT) at T_i and equated it to a. Knowing F at the final temperature T_f and the average temperature $T_{ave} = (T_i + T_f)/2$, he

^{*} Author to whom correspondence should be addressed.

deduced the values of b and n. Finally by trial and error, he adjusted the values of the constants to give the best possible fit. Of course, a better approach, but probably more time consuming is to get the values of a, b, and n from three simultaneous equations based on the values of F at T_f and two other temperatures at equal intervals (or nearly so) in the temperature range T_i and T_f .

Table (1) gives the values of H at different values of T between 70°F and 140 °F [2]. From this table, one finds:

H at
$$70^{\circ}$$
F = 0.01582
H at 140° F = 0.1534
H at T_{ave} (105°F) = 0.05070
slope at 70° F = 5.575E-4

Following the procedure outlined by Totman [1] and substituting in Eq. (1) we get:

$$H = 0.01582 + 5.575E - 4(T - 70) + 1.1142E - 6(T - 70)^{2.681}$$
 (2)

The equation obtained from solving three simultaneous equations based on the values of H at 94, 118 and 140°F is given by:

$$H = 0.01582 + 6.421E-4(T-70) + 4.1735E-7(T-70)^{2.8977}$$
 (3)

The equation given by Totman after adjusting the constants of Eq. (2) by trial and error is the following:

$$H = 0.01582 + 6.3114E-4(T-70) + 4.7625E-7(T-70)^{2.8684}$$
 (4)

Eqs. (3) and (4) are quite close to one another which shows that the adjustment of Eq. (2) was nothing but an effort to reach Eq. (3) without the trouble of solving three simultaneous equation in three unknowns.

Table 1. Saturated Humidity versus Temperature [2]

Temp. (°F)	Sat. Humidity	Temp.	Sat. Humidity	Temp.	Sat. Humidity
70	0.01582	94	0.03556	118	0.07652
72	0.01697	96	0.03795	120	0.08149
74	0.01819	98	0.04049	122	0.08678
76	0.01948	100	0.04319	124	0.09242
78	0.02086	102	0.04606	126	0.09841
80	0.02233	104	0.04911	128	0.10480
82	0.02389	106	0.05234	130	0.11160
84	0.02555	108	0.05578	132	0.11890
86	0.02731	110	0.05944	134	0.12670
88	0.02919	112	0.06333	136	0.13500
90	0.03118	114	0.06746	138	0.14390
92	0.03330	116	0.07185	140	0.15340

The maximum errors obtained from utilizing Eqs. (2), (3), and (4) in the range 70°F to 140°F are 2.9%, 1.9%, and 2.0% respectively.

In the same paper, Totman used Eq. (1) to fit the relation between the viscosity of water in the region 0°C to 30°C and temperature. He got the following relation:

$$\mu = 1.7921 - 0.06267 \text{ T} + 0.0035 \text{ T}^{1.627}$$
 (5)

with a maximum error equal to 0.9%.

It may be pointed out also that both Eqs. (4) and (5) can be utilized to calculate H or μ at a given value of T and not vice versa unless trial and error is used. No explicit relation for T in terms of H or μ can be derived.

With the present technique, it will be shown that, following a systematic approach, a two constant equation is obtained which is simpler and more accurate than the three constant equations (4) and (5) at the same time leads to a simpler explicit relation for T in terms of H or μ .

THE PRESENT TECHNIQUE

In the present technique, the fitting equation is a twoconstant chosen from five fitting functions depending on which one of them has the nearest asymmetry to that of the original function or data to be fitted.

In the dimensionless forms, the five functions are the following:

Inverse linear function (I.L.F.):

$$Y = [aX] / [1 + (a-1)X]$$
 (6)

Exponential function (E.F.):

$$Y = [g^{x} - 1] / [g - 1]$$
 (7)

Logarithmic function (L.F.):

$$Y = \ln [X (b-1) + 1] / \ln [b]$$
 (8)

Poisson function (P.F.):

$$Y = X k^{1-x}$$
 (9)

Quadratic function (Q.F):

$$Y = cX + (1-c)X^2$$
 (10)

Here: Y = dimensionless dependent variable;

ABDO, SAID, EWIDA and BADER: A Useful Short Cut Approach To Physical And Chemical Curve Fittings

$$a, g, b, c, and k = constants.$$

For all the above functions

$$Y = 0$$
 at $X = 0$

$$Y = 1$$
 at $X = 1$

For a function Y = f(X):

$$Y = [y-y_i] / [y_f y_i]$$
 (11)

and

$$X = [x-x_i] / [x_f x_i]$$
 (12)

The asymmetry of any of the above functions is given by mo, m1.

Here:

$$m_0 = \text{slope at } (X=0) = [dY/dX]_{x=0}$$
 (13)

$$m_1 = \text{slope at } (X=1) = [dY/dX]_{x=1}$$
 (14)

and therefore

$$m_0 m_1 = m_i m_f [(x_f - x_i) / (y_f y_i)^2 = m_1 m_f \theta^2 (15)$$

Here:

 θ = the scale factor

$$m_i = initial slope = [dY/dx]_{xi};$$
 (16)

$$m_f = \text{final slope} = [dY/dx]_{xf};$$
 (17)

The asymmetry of each function depends on the value of the corresponding constant a, g, b, k, or c which can be calculated from one value only of Y at any value of X between X_i and X_f .

For easy estimation and also for better accuracy, X is taken equal to $X_{ave} = [(X_i + X_f)/2]$ in which case X = [1/2] and we call the corresponding value $Y_{1/2}$ or Z and hence

$$Z = Y_{1/2} = Y \text{ at } X = 1/2$$

One can show that

$$a = [Z/(1-Z)]$$
 (18)

$$g = [(1-Z)/Z]^2$$
 (19)

$$K = 4 Z^2$$
 (20)

$$c = 4Z - 1$$
 (21)

No analytical expression can be derived for the logarithmic function. However, there is an analytic one in terms of Z which is the value of X at Y = 1/2. The logarithmic function can be accomplished through the exponential function and therefore will not be discussed in this paper.

The asymmetries of the other fitting functions were derived in terms of Z by differentiating Eqs. (6), (7), and (10) and substituting from Eqs. (18-21). They are given by the following relations:

I.L.F.
$$m_0 m_1 = 1$$
 (22)

E.F.
$$m_0 m_1 = \{ [2Z(1-Z)/(1-2Z)] \ln (1-Z)/Z \}^2$$
 (23)

P.F.
$$m_0 m_1 = 4Z^2 [1-2 \ln 2Z]$$
 (24)

Q.F.
$$m_0 m_1 = [4Z-1) (3-4Z)$$
 (25)

Values of m_0 m_1 as given by Eqs. (22-25) were calculated for different values of Z between 0.2 and 0.8 and the results are given in Table (2) and plotted in Figure (1) together with m_0 m_1 for logarithmic function which was calculated by trial and error.

If the original function is given in the form of difficult formula, $m_0 m_1$ and Z are obtained by differentiation and substitution respectively.

In the case of two simultaneous equations namely

$$Y = f_1$$
 (v) and $v = f_2$ (X), then

$$m = [dy/dx] = [dy/dv] [dv/dx]$$
 (26)

and in case of an implicit function, one utilizes the following relation:

$$[dy/dx] = [1/dx/dy)]$$
 (27)

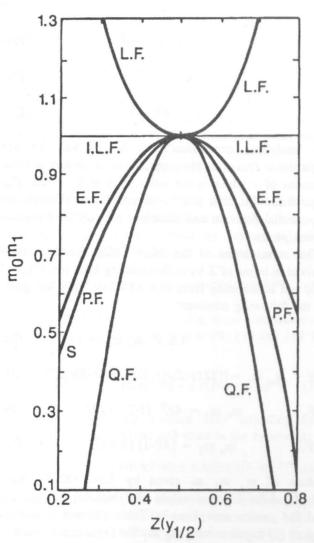


Figure 1. The asymmetries values for different values of Z.

If the original function is given in the form of a table where Y is given at equal intervals of X as in Table (1), the slopes are calculated from formulas given in the next section. In Table (1), the value of Y at X_{ave} . (H at T_{ave}) is missing and has to be calculated by interpolation using a convenient through accurate interpolation formula. This will be shown also in the following sections.

1. Calculation of Slopes from Tabulated Values

With reference to Figure (2), slope m_o at point 0 can be expressed in terms of

Hence:
$$m_1 = [y_1/h]$$
; $m_2 [y_2/2h]$: $m_3 = [y_3/3h$... etc.

and m-1 =
$$[y-1/h]$$
; m-2 = $|[y-2/2h]$; m-3

$$= |[y_3/3h]|]...etc$$

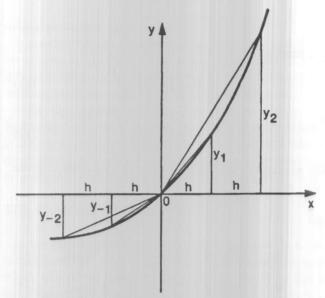


Figure 2. Illustrative curve.

when point o is an initial point,. then fitting a quadratic polynomial gives

$$m_0 = 2m_1 - m_2$$
 (28)

By fitting a polynomial of the 4th degree, we get:

$$m_0 = 3m_1 - 3m_2 + m_3 \tag{29}$$

and by fitting by a polynomial of the 4th degree, we get:

$$m_0 = 4m_1 - 6m_2 + 4m_3 - m_4$$
 (30)

etc.

When point 0 is a final point, the corresponding movalues are the following:

$$m_0 = 2m-1 - m-2$$
 (31)

$$m_0 = 3m-1 - 3m-2 + m-3$$
 (32)

$$m_0 = 4m-1 - 6m-2 + 4m-3 - m-4$$
 (33)

etc.

Usually Eqs. (28) and (31) are accurate enough. For better accuracy, Eqs. (29) and (32) are utilized and very seldom higher order equations are needed in this work.

If the point 0 is not an end point, then m_o is more accurately determined from equations obtained from fitting polynomials to points after and points before the origin.

Fitting a quadratic polynomial to one point after and one point before the origin gives

$$m_0 = m_1 - 1$$
 (34)

Here:
$$m_1-1 = [(y_1-y_{-1})/2h]$$
 (35)

Fitting a polynomial of the 4th degree to two points after and two points before the origin leads to

$$m_0 = [(4m_1-1 - m_2-2)/3]$$
 (36)

Here:
$$m_2-2 = [(y_2-y-2)/4h]$$
 (37)

and hence

$$m_0 = [8(y_1-y_{-1}) - (y_2-y_{-2})]/12h$$
 (38)

2. Interpolation in Tables

If the original function is given in the form of a table at equal intervals of the independent variable then the value of Y corresponding to any value of X between X_n and X_{n+1} can be obtained using any of the fitting Eqs. (6-10). When h is small enough as in the case of Table (1), Z will not differ much from 0.5 in which case, all fitting functions have asymmetries close to 1 and are practically identical as can be seen from Table (2) and Figure (1).

The fitting function should, however, be in a form suitable for interpolation. For example, in the case of the inverse linear function, on combining Eqs. (6) and (18) and then substituting for Y, X, and Z in terms of Y_n , Δ_1 , Δ_2 , and f one gets

$$Y = Y_n + \{ [\Delta_1(\Delta_1 + \Delta_2)f] / [2\Delta_2 - (\Delta_2 - \Delta_1)f] \}$$
 (39)

with reference to Figure (3),

$$\Delta_1 = Y_{n+1} - Y_n \tag{40}$$

$$\Delta_2 = Y_{n+2} - Y_{n+1} \tag{41}$$

$$Z = [\Delta_1/(\Delta_1 + \Delta_2)] \tag{42}$$

$$Y = [(y-y_n)/(y_{n+2}-y_n)] = [(y-y_n)/(\Delta_1 + \Delta_2)]$$
 (43)

$$X = [(x-x_n)/(x_{n+2}-x_n)] = [(x-x_n)/2h] = f/2)]$$
(44)

Here:

$$f = [(x-x_n)/h]. (45)$$

Table 2. The asymmetries values (m_o m₁) for different values of Z.

Z	mo m1			Z			
	E.F.	P.F.	Q.F.		E.F.	P.F.	Q.F.
0.20	0.5466	0.4532	-0.4400	0.51	0.9995	0.9992	0.9984
0.21	0.5745	0.4825	-0.3456	0.52	0.9979	0.9968	0.9936
0.22	0.6017	0.5115	-0.2544	0.53	0.9952	0.9927	0.9856
0.23	0.6282	0.5402	-0.1664	0.54	0.9915	0.9869	0.9744
0.24	0.6539	0.5686	-0.0816	0.55	0.9867	0.9794	0.9600
0.25	0.6789	0.5966	0.0000	0.56	0.9808	0.9701	0.9424
0.26	0.7031	0.6240	0.0784	0.57	0.9739	0.9590	0.9216
0.27	0.7265	0.6510	0.1536	0.58	0.9660	0.9462	0.8976
0.28	0.7490	0.6773	0.2256	0.59	0.9570	0.9315	0.8704
0.29	0.7707	0.7029	0.2944	0.60	0.9470	0.9149	0.8400
0.30	0.7915	0.7278	0.3600	0.61	0.9359	0.8965	0.8064
0.31	0.8114	0.7519	0.4224	0.62	0.9238	0.8761	0.7696
0.32	0.8303	0.7752	0.4816	0.63	0.9107	0.8538	0.7296
0.33	0.8483	0.7976	0.5376	0.64	0.8966	0.8295	0.6864
0.34	0.8654	0.8191	0.5904	0.65	0.8815	0.8032	0.6400
0.35	0.8815	0.8395	0.6400	0.66	0.8654	0.7749	0.5904
0.36	0.8966	0.8590	0.6864	0.67	0.8483	0.7446	0.5376
0.37	0.9107	0.8774	0.7296	0.68	0.8303	0.7122	0.4816
0.38	0.9238	0.8946	0.7696	0.69	0.8114	0.6777	0.4224
0.39	0.9359	0.9107	0.8064	0.70	0.7915	0.6410	0.3600
0.40	0.9470	0.9256	0.8400	0.71	0.7707	0.6023	0.2944
0.41	0.9570	0.9393	0.8704	0.72	0.7490	0.5614	0.2256
0.42	0.9660	0.9516	0.8976	0.73	0.7265	0.5183	0.1536
0.43	0.9739	0.9627	0.9216	0.74	0.7031	0.4729	0.0784
0.44	0.9808	0.9724	0.9424	0.75	0.6789	0.4254	0.0000
0.45	0.9867	0.9807	0.9600	0.76	0.6539	0.3756	-0.0818
0.46	0.9915	0.9875	0.9744	0.77	0.6282	0.3236	-0.1664
0.47	0.9952	0.9929	0.9856	0.78	0.6017	0.2692	-0.2544
0.48	0.9979	0.9968	0.9936	0.79	0.5745	0.2126	-0.3456
0.49	0.9995	0.9992	0.9984	0.80	0.5467	0.1536	-0.4400
0.50	0.10000	1.0000	1.0000				

E.F. = Exponential function;

P.F. = Poisson function;

Q.F. = Quadratic function.

Similar interpolation formulas can be deduced from the exponential, Poisson, and quadratic functions.

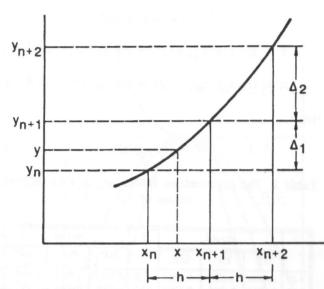


Figure 3. Illustrative curve.

Inverse interpolation is easily achieved by solving for f in Eq. (39) and then substituting from Eq. (45) leading to:

$$X = x_0 + [(2h\Delta_2)(Y - Y_n)]/[(Y - Y_n)(\Delta_2 - \Delta_1) + \Delta_1(\Delta_1 + \Delta_2)]$$
 (46)

Inverse interpolation formulas can be obtained for the exponential and quadratic fitting functions but not for the poisson function where no analytical expression can be deduced.

APPLICATIONS

To illustrate the application of the present method and compare with results achieved by a previous technique [1], the best fitting Z-constant function will be deduced for the humidity ratio of saturated air versus temperature and for viscosity of water between 0°C and 30°C.

1. The Humidity Ratio of Saturated Air Versus Temperature Between 70°F and 140°F.

From Table (1).

$$m_1 = [(0.01697 - 0.01582)/2]$$

$$m_2 = [(0.01819 - 0.01582)/4]$$

$$m_3 = [(0.01948 - 0.01582)/6]$$

Substituting in Eq. (29) gives

$$m_i = 3m_1 - 3m_2 + m_3 = 5.575E-4$$

similarly

$$m-1 = [(0.1534 - 0.1439)/2]$$

$$m-2 = [(0.1534 - 0.1350)/4]$$

$$m-3 = [(0.1534 - 0.1267)/6]$$

Substituting in Eq. (32) gives

$$m_f = 3m-1 - 3m-2 + m-3 = 4.90E-3$$

$$m_0 m_1 = m_i m_f \theta^2 = [5.575E-4] [4.90E-3] [(140-70)$$

$$/(.1534-.01582)]^2$$

$$m_0 m_1 = 0.707$$

$$T_{ave} = [(70+140)/2] = 105 \, ^{\circ}F$$

Interpolation between T=104 °F and T=106 °F in Table (1) using Eq. (39) one finds

$$Y_n = h_{104} = 0.04911$$

$$\Delta_1 = 0.05234 - 0.04911 = 0.00323$$

$$\Delta_2 = 0.05578 - 0.05234 = 0.00344$$

$$f = [(x - x_n)/h] = [(105 - 104)/2] = 0.5$$

Substituting these values in Eq. (39) gives

$$H_{105} = 0.05070$$

All other interpolating formulas resulting from the exponential, Poisson or quadratic fitting functions lead to the same value for H₁₀₅

$$Z = Y_{1/2} = [(0.0507-0.01582)/(0.1534-0.01582)] = 0.2535$$

Substituting in Eqs. (22-25) one finds

I.L.F.:

$$m_0 m_1 = 1$$

E.F.:

$$m_0 m_1 = 0.6875$$

P.F.:
$$m_0 m_1 = 0.6062$$

Q.F.:
$$m_0 m_1 = 0.0278$$

Comparing with the actual m_0 m_1 value which is equal to 0.707, one finds that the exponential function gives the best fit since 0.6875 is the nearest value to 0.707.

Formulas for different fitting functions in terms of Z are easily deduced by substituting for the constants a,g,k, and c from Eqs. (18-21). Substituting Z = 0.2535, one gets equations for all fitting functions, namely

I.L.F.

$$H=0.01582+[6.675E-4(T-70)]/[1-9.43E-3(T-70)]$$
 (47)

E.F.

$$H = 0.01582 + \{0.01794 [(2.9443) [(t-70)/35] -1]\}(48)$$

P.F.

$$H=0.01582+\{5.0531E-4(T-70)(1.9722)[(T-70)/35\}(49)$$

Q.F.

$$H=0.01582+[2.76E-5(T-70)]+[2.77E-5(T-70)^2]$$
 (50)

The value of H was calculated using all five fitting functions including Totman's {Eq. (4)] at different values of T. The percentage error was calculated for all five functions and is tabulated in Table (3).

Table (3) shows that different functions lead to varying errors. The Totman equation leads to a maximum error equal to 2.0% while the maximum error is less than 0.5% in the case of the exponential fitting function and is equal to 1.7% for the poisson function, 4% for the inverse linear and 16% in the case of the quadratic fitting function.

The exponential function [Eq. (48)] is therefore the best two-constant fitting function and leads to an error which is only one quarter the error of Totman's three-constant fitting function. Furthermore, an explicit relation for T in terms of H can be deduced easily from the exponential fitting function.

Table 3. The percentage Error of all functions.

Temp (°F)	Exact	Totman		E.F.		P.F.	LLF.	Q.F.
	Н	Н	E%	Н	E%	E%	E%	E%
70	0.01582	0.01582	0.00	0.01582	0.00	0.0	0.0	0.0
75	0.01882	0.01902	1.08	0.01881	-0.04	-1.2	2.7	-11.5
80	0.02233	0.02248	0.67	0.02230	-0.12	-1.7	3.9	-15.5
85	0.02642	0.02641	-0.03	0.02638	-0.16	-1.7	4.0	-15.0
90	0.03118	0.03101	-0.58	0.03113	-0.16	-1.5	3.5	-13.6
95	0.03674	0.03647	-0.73	0.03668	-0.16	-1.1	2.3	-8.0
100	0.04319	0.04297	-0.51	0.04315	-0.09	-0.6	1.3	-3.8
105	0.05070	0.05070	0.00	0.05070	0.00	0.0	0.0	0.0
110	0.05944	0.05983	0.66	0.05951	0.12	0.5	-1.2	2.3
115	0.06962	0.07052	1.29	0.06979	0.25	0.9	-2.3	5.0
120	0.08149	0.08296	1.80	0.08179	0.36	1.2	-3.1	6.0
125	0.09537	0.09729	2.20	0.09578	0.43	1.3	-3.4	6.0
130	0.11160	0.11370	1.89	0.11210	0.46	1.2	-3.1	5.0
135	0.13080	0.13240	1.19	0.13120	0.28	0.7	-3.1	5.0
140	0.15340	0.15340	0.00	0.15340	0.00	0.0	0.0	0.0

H = Saturated Humidity;

E% = Error percentage.

 Viscosity of Water V_s. Temperature between 0°C and 30°C

Applying the present technique to data obtained from Handbook of chemistry and Physics [3], one gets $(d\mu/dT)_i = -0.0607 \text{ cp/}^{\circ}\text{C}$

$$(d\mu/dT)_f = -0.0170 \text{ cp/}^{\circ}C$$

$$\theta = [30/0.9895] = 30.32$$

$$m_0 m_1 = (-0.0607) (-0.0170) (30.32)^2 = 0.949$$

$$Z = [(1.139-0.7975)/(1.787-0.7975)] = 0.345$$

The corresponding asymmetries were calculated from Eq. (22-25) leading to

$$I.L.F. m_o m_1 = 1$$

P.F.
$$m_0 m_1 = 0.83$$

Q.F.
$$m_0 m_1 = 0.62$$

The functions having the nearest asymmetry to that of the original function is the inverse linear function. Substituting ABDO, SAID, EWIDA and BADER: A Useful Short Cut Approach To Physical And Chemical Curve Fittings

$$Z = 0.345$$

$$Y = [(\mu - 0.7975) / (1.787-0.7975)]$$

and

in Eqs. (16) and (18), one gets the following simple relation for μ as a function of temperature.

$$\mu = 1.787 - [0.0627T/(1+0.03) T)]$$
 (51)

Eq. (51) leads to a maximum error equal to 0.13% or 0.26% which is approximately one quarter the maximum error of 0.9% or 0.6% resulting from using Totmn's three-constant equation. Again, as can be seen from Eq. (51), an explicit relation in T can be easily achieved.

REFERENCES

- [1] Ralph S. Totman. Rapid Curve-Fitting Technique. Chemical Engineering Dec. 24: pp. 63-65 1984.
- [2] Chemical Engineering Handbook, 5th, edition: New York, McGraw Hill 1973: p. 12-7
- [3] Handbook of Chemistry and Physics, 56th edition: p. F-49.