Proceedings of the Iowa Academy of Science

Volume 73 | Annual Issue Article 18

1966

Correlating Data Which Fit the Gompertz Equation of Growth

Robert G. Sheppard
University of Iowa

James O. Osburn
University of Iowa

were tested in some of the better solvents. These sheets are precoated plastic and are popular among experimenters because they are convenient and eliminate much of the labor needed to coat glass plates. Fair to good results were obtained in most cases, although the spots were less distinct, harder to visualize, faded quicker, and were subject to more tailing and spreading than their glass counterparts.

Literature Cited

Estes, J. Worth and Paul Dudley White, *Scientific American*, June 1965. p. 110-119.

———, "Lanoxin (brand) Digoxin", information from the Medical Department, Burroughs Wellcome & Company, Tuckahoe, N.Y. 1964. p. 1-6.

Stoll, Arthur, *Chemistry and Industry*, 1959. p. 1558-1567.

Randerath, Kurt, *Thin-Layer Chromatography*, Academic Press, New York, 1964. p. 111-125.

Oeth, Dennis, "Thin-Layer Chromatography of Selected Cardiac Glycosides", thesis, Loras College Library, Dubuque. 1966.

Correlating Data Which Fit the Gompertz Equation of Growth

Robert G. Sheppard and James O. Osburn[1]

Abstract. The Gompertz equation describes the decay of biological systems in which the rate of decrease of the concentration of an organism varies directly with its concentration and inversely with the concentration of an inhibitor. The inhibitor concentration falls exponentially with time. Four methods are compared for correlating data on the destruction of Ascites tumor cells which fit the Gompertz equation: analytical, graphical, analog computer, and digital computer.

The Gompertz equation (1) describes the rate of decay of a substance, where the decay is inhibited by a second substance which is also decaying. A common form of the equation as it is encountered in biology is

$$\frac{dy}{dt} = \frac{-k_1 y}{x} \cdot \qquad (1)$$

Here y is the concentration of the primary substance which is decaying, and x the concentration of the inhibitor. The concentration, x, can be expressed as a falling exponential

$$x = x_o e^{-k_2 t} \qquad (2)$$

[1] University of Iowa, Iowa City

where x_0 is the original concentration taken as 1.0, and k_2 the decay constant.

When equation 1 and 2 are combined and integrated, this equation results:

$$y = \exp(-A(\exp(Bt)-1)) \qquad (3)$$

In this form, $A = k_1/(k_2x_0)$, $B = k_2$, and $y = 1$ at $t = 0$.

Usually, one measures y at several different times, and wants to find the values of k_1 and k_2 which give the best fit of the equation to the data. Moressi (3), for example, measured the rate of destruction of Ascites Tumor cells by microwave irradiation, and this is the data that will be employed to demonstrate the use of four methods of solving for the constants in the Gompertz equation. These methods are: analytical, graphical, analog computer and digital computer.

Discussion

Analytical Method (1)

The analytical method for determining the constants of the Gompertz equation requires that the independent variable be equally spaced. In this case, values of y must be available for equal intervals of time, t. If determining y for equal time intervals is not convenient in the experimental procedure then the y values are plotted and a curve is drawn through the data; see Figure (1). Then y values are chosen at equal time spacing. This requires fitting the data "by eye" and may result in some error.

The Gompertz equation of interest, written in the following form, is

$$y = e^{-A(e^{Bt}-1)} = e^{A}(e^{-A})^{(e^{Bt})} \qquad (4)$$

or the more general form:

$$y = ab^{c^x}$$

where $a = e^{A}$

$b = e^{-A}$

$$c = e^{B\Delta t}$$

$x = 1, 2, 3, \ldots, 3n$

Δt = interval of time between data points.

The constants of this equation can be determined by putting the data into sequential order with respect to the independent variable, then dividing the data into three groups of n entries each. Let S_1, S_2, and S_3 be the sums of the three groups of values of the logarithm of y values which gives

Solving these equations yields

$$c = \left(\frac{S_2 - S_3}{S_1 - S_2}\right)^{\frac{1}{n}}; \qquad (5)$$

$$S_1 = \sum_{y_1}^{y_n} \log y = n\log a + \left(\sum_{x=1}^{n} c^x\right) \log b;$$

$$S_2 = \sum_{y_{n+1}}^{y_{2n}} \log y = n\log a + \left(\sum_{x=n+1}^{2n} c^x\right) \log b; \qquad (6)$$

$$S_3 = \sum_{y_{2n+1}}^{y_{3n}} \log y = n\log a + \left(\sum_{x=2n+1}^{3n} c^x\right) \log b.$$

$$\log b = \frac{(S_1 - S_2)(1 - c)}{(1 - c^n)^2} \qquad (7)$$

$$\log a = \frac{1}{n}\left(S_1 - \frac{S_1 - S_2}{1 - c^n}\right). \qquad (8)$$

Illustration 1; Analytical Method: Study 11x. Rate of destruction of Ascites Tumor cells by microwave irradiation = y; t = time.

Data points taken from "eye fit' of actual data.

i	TIME	y_i	$\ln y_i$		
1	0	1.000	0.0		
2	10	0.895	-0.11093	Σ =	S_1 = -0.46191
3	20	0.704	-0.35098		
4	30	0.443	-0.81419		
5	40	0.288	-1.24479	Σ =	S_2 = -3.81924
6	50	0.172	-1,76026		
7	60	0.090	-2.40795		
8	70	0.039	-3.24419	Σ =	S_3 = -10.25731
9	80	0.010	-4.60517		

$c = 1.242$; $\ln b = -.9650$

$A = \ln b = 0.9650$ sample variance = 0.0036

$$B = \frac{\ln}{\Delta t} c = \frac{\ln 1.242}{10} = 0.0217$$

The curve calculated from the Gompertz equation using these values for the constants is shown in Figure (1).

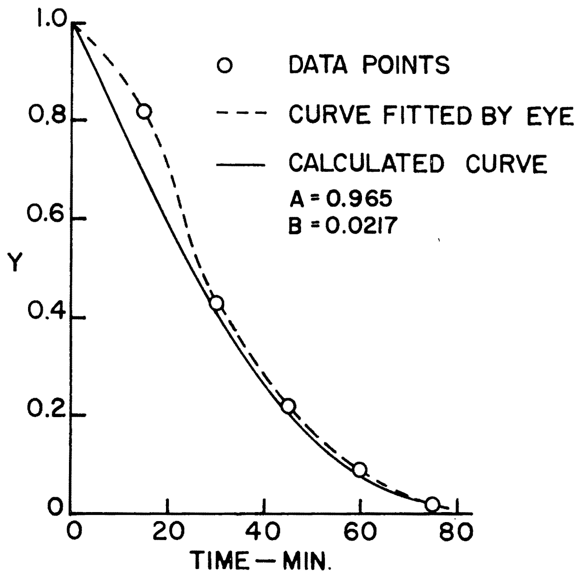

Fig. 1. Analytically Fitted Curve.

Graphical Method

The graphical determination of constants k_1 and k_2 of the Gompertz equation by numerical differentiation.
Gompertz equation,

$$y = e^{(-k_2/k_1)(e^{k_2 t}-1)}. \tag{9}$$

Linearized Gompertz equation,

$$\ln [d(-\ln y)/dt] = \ln k_1 + k_2 t. \tag{10}$$

Since the changes of ln y and t are not infinitesimally small, equation (10) may not be used in this form for analyzing the data; however, for purpose of approximating k_1 and k_2 we may express equation (10) in finite difference form as:

$$\ln [\triangle(-\ln y)/\triangle t] = \ln k_1 + k_2 t$$

where $\triangle$ refers to observed changes in ln y and t, and t_m refers to the median time between such changes.

If values of $\ln(\triangle(-\ln y)/\triangle t)$ are plotted against t_m, the slope of the resulting straight line is k_2, and the intercept is ln k_1, as seen in Figure (2).

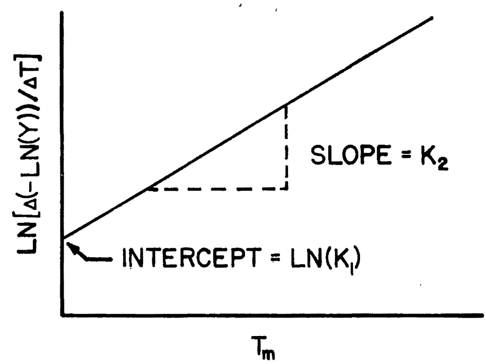

Fig. 2. Plot of Equation 10A.

Illustration 2; Graphical Method: Study 11x. Rate of destruction of Ascites Tumor cells by microwave irradication = y; t = time.

y	t	t_m	$-\Delta \ln y$	Δt	$-\Delta \ln y/\Delta t$	$\ln(-\Delta \ln y/\Delta t)$
1.00	0					
		> 7.5	0.1984	15	1.323×10^{-2}	-4.32
.82	15					
		>22.5	0.6455	15	4.303×10^{-2}	-3.14
.43	30					
		>37.5	0.6702	15	4.468×10^{-2}	-3.11
.22	45					
		>52.5	0.8938	15	5.959×10^{-2}	-2.82
.09	60					
		>67.5	1.5041	15	10.027×10^{-2}	-2.30
.02	75					

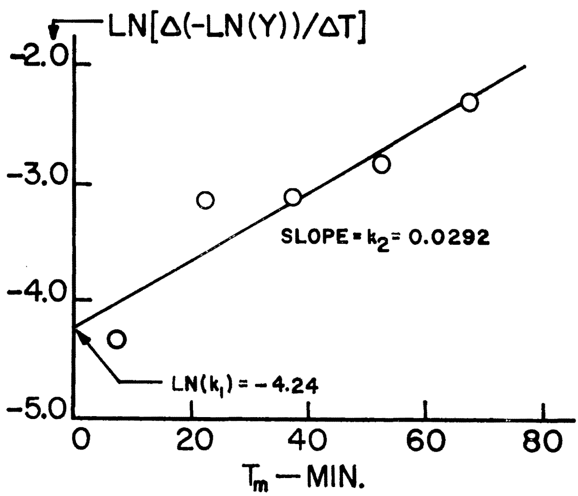

Fig. 3. Determining K_1 and K_2.

A straight line least squares fit of ln (△ (-lny)/ △t) versus t_m in Figure (3) yields:

Intercept $=$ ln $k_1 = -4.24$ $k_1 = 0.0145$
Slope $= k_2 = 0.0292$
$A = k_1/k_2 = 0.0145/0.0292 = 0.497$
$B = k_2 = 0.0292$
Sample Variance $= 0.00191$

It can be seen from Figure (3) that unless the data are quite good, the constants obtained by this method may be of questionable accuracy.

The calculated results and the actual data are compared in Figure (4).

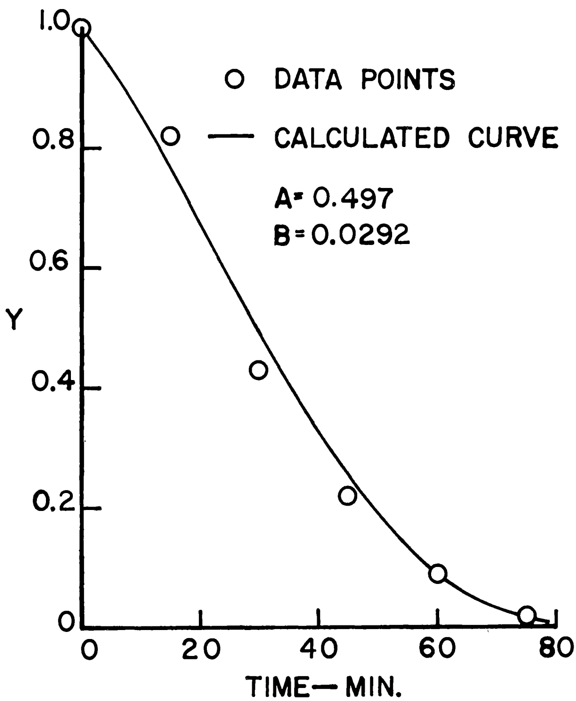

Fig. 4. Graphically Fitted Curve.

Analog Computer Simulation

Finding the constants k_1 and k_2 of the Gompertz equation with the analog computer is accomplished by programming the proper equations on the computer and by allowing k_1 and k_2 to be represented by two potentiometers. These potentiometers are varied by a trial and error procedure so as to produce the best fit of the experimental data. The experimental data points (y versus t) are plotted on an oscilliscope screen and the computer results are also displayed on the screen. The best fit of the experimental data is achieved by comparing the computed curve with the experimental data. This method requires the exercise of judgment in deciding where to draw the line through the given points.

If y is the concentration of the primary substance which is decaying and x is the concentration of the inhibitor, then y decays according to equation (1),

$$\frac{dy}{dt} = \frac{-k_1 y}{x}$$

where

$$x = x_o e^{-k_2 t}$$

or in differential form

$$\frac{dx}{dt} = -k_2 x \qquad (11)$$

These two simultaneous differential equations can be solved for y as a function of time by programming them on the analog computer.

The analog computer circuit which integrates equation (11), shown in Figure (5), requires only one integrating amplifier.

Equation (1) is integrated for y using one integrating amplifier, a function divider and a sign changing amplifier in Figure (6).

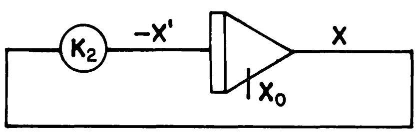

Fig. 5. Circuit for $X' = -K_2X$.

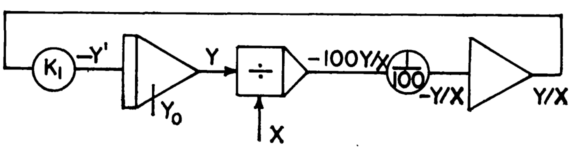

Fig. 6. Circuit for $Y' = -K_1Y/X$.

The combination of the circuits in Figure (5) and Figure (6), which give the complete solution of the two equations is given in Figure (7).

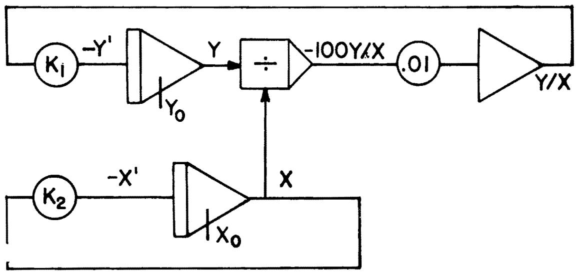

Fig. 7. Circuit for Gompertz Equation.

Illustration 3. A comparison of actual data and the fitted curve for Study 11x by Moressi (3) is shown in Figure (8) with the constants determined to be:

$k_1 = 0.0137$
$k_2 = 0.04$

or

$A = k_1/k_2 = 0.342$
$B = k_2 = 0.04$
Sample variance $= 0.0019$.

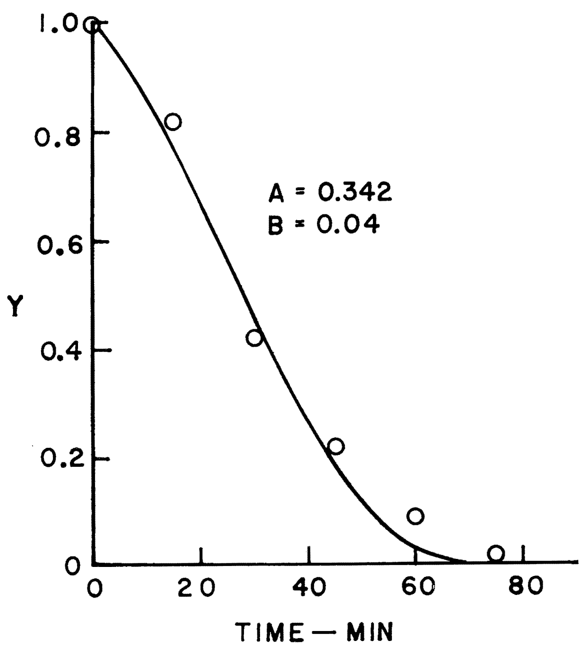

Fig. 8. Analog Computer Fit.

Digital Computer Method

This method utilized a non-linear least squares fit of data to determine the constants, A and B, of the Gompertz equation. The least squares equations are solved by iteration on a digital computer. This allows the constants to be determined for a large number of studies accurately and very rapidly.

Derivation: Assume that n values of y versus t are to be used, where y is the fraction of the initial value; values of y decrease with time t. These points are to be fitted by the Gompertz equation in the form of equation (3),

$$y = \exp(-A(\exp(Bt)-1)).$$

If this equation is transformed into the logarithmic form,

$$\ln y = -A\,(e^{Bt}-1), \qquad (12)$$

values of A and B can be found which minimize the function

$$\sum_{i=1}^{n} [(\ln y_i)_{CALCULATED} - (\ln y_i)_{OBSERVED}]^2 . \qquad (13)$$

However, because the difference of logarithms is calculated, data points at low values of y are given undue weight. A deviation of .01 unit at $y = .1$ is ten times as important as the same deviation at $y = 1.0$. A least squares fit on this basis fits low values very well, but the high values are almost neglected. To correct this, let us look at the least squares criterion which requires that we minimize the function.

$$\sum_{i=1}^{n} (y_i^* - y_i)^2 \qquad (13A)$$

but the function we are faced with from equation (13) is

$$\Sigma(\ln y_i^* - \ln y_i)^2$$

where y_i^* is a calculated value of y and y_i is an observed value. Thus, we would like to find some function f(y) which will satisfy the following relationship.

$$f(y)\,(\ln y^* - \ln y) = y^* - y \qquad (14)$$

Now letting $y^* = y + \Delta y$ (this says that the calculated y differs from the observed value of y by some small increment Δy) and solving equation (14) for $1/f(y)$ we obtain:

$$\frac{1}{f(y)} = \frac{\ln(y + \Delta y) - \ln y}{\Delta y} \qquad (15)$$

and if we take the limit of both sides as Δy approaches zero

$$\lim_{\Delta \to 0} \frac{1}{f(y)} = \lim_{\Delta \to 0} \frac{\ln(y + \Delta y) - \ln y}{\Delta y} \qquad (16)$$

we can see that the right side of this equation is the definition of the derivative of ln y. Thus, if $\triangle$y is small

$$\frac{1}{f(y)} \cong \frac{d(\ln y)}{dy} = \frac{1}{y} \qquad (17)$$

and

$$f(y) = y.$$

By squaring both sides of equation (14) and summing over all y_i's, we obtain a minimizing equation which is approximately equal to that of the least squares criterion:

$$\sum_{i=1}^{n} (y_i^* - y_i)^2 \cong \sum_{i=1}^{n} (f(y))^2 (\ln y_i^* - \ln y_i)^2$$
$$\cong \sum_{i=1}^{n} y_i^2 (\ln y_i^* - \ln y_i)^2. \qquad (18)$$

Replacing y_i^* by its calculated value, a function, γ, is defined, weighted as described:

$$\gamma = \sum_{i=1}^{n} y_i^2 [\ln y_i + A(e^{Bti} - 1)]^2. \qquad (19)$$

For the optimum fit, the derivatives of γ with respect to A and to B should both be zero:

$$\partial\gamma/\partial A = 2\sum_{i=1}^{n} y_i^2 [(\ln y_i + A(e^{Bt_i} - 1)(e^{Bt_i} - 1)] = o \qquad (20)$$

$$\partial\gamma/\partial B = 2\sum_{i=1}^{n} y_i^2 [(\ln y_i + A(e^{Bt_i} - 1)(At_i e^{Bt_i})] = o \qquad (21)$$

The solution of equations 20 and 21 require an interation procedure. This starts with an assumed value of B of .001. With this estimate of B, A can be calculated by manipulation of equation (20):

$$\sum_{i=1}^{n} y_i^2 \ln y_i (e^{Bt_i} - 1) = -\sum_{i=1}^{n} (y_i^2)(A)(e^{Bt_i} - 1)^2. \qquad (22)$$

Since A is constant, it can be taken out of the summation, giving

$$A = -\frac{\sum_{i=1}^{n} y_i^2 \ln y_i^2 (e^{Bt_i} - 1)}{\sum_{i=1}^{n} y_i^2 (e^{Bt_i} - 1)^2} \qquad (23)$$

With this value of A, equation (21) is used to calculate $\partial\gamma/\partial B$. This will, in general, not be equal to zero, so B is increased and the process repeated until $\partial\gamma/\partial B$ differs from zero by less than an arbitrary amount. A difference of .001 has given satisfactory results.

Testing: A computer program to perform the calculations described above was written in *Fortran IV*, and the IBM 7044 computer at the University of Iowa computer center was used to evaluate the constants for some data of Moressi (3) on the destruction of Ascites Tumor cells by microwaves.

Illustration 4; Digital Computer Fit: Study 11x. y = Rate of destruction of Ascites Tumor cells by microwave irradiation; t = time.

A = 0.424
B = 0.033
Sample Variance = 0.0014

The calculated curve and data points are shown in Figure 9.

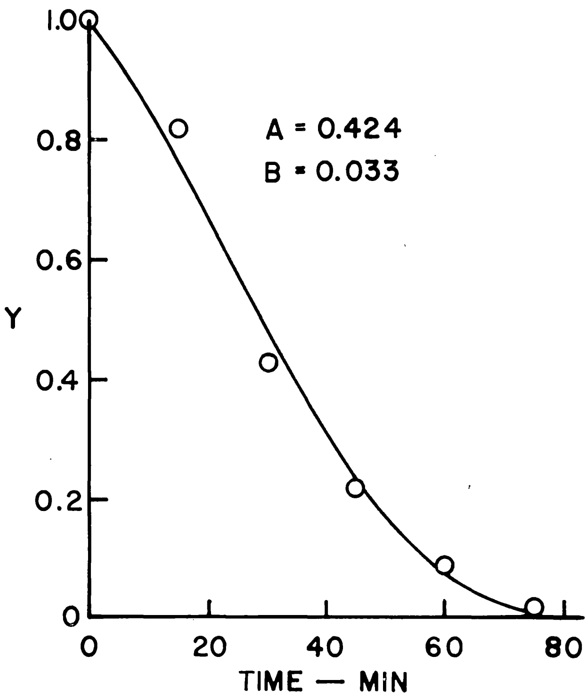

Fig. 9. Digital Computer Fit.

A comparison of the constants determined by the four methods is given in Table 1.

Table 1. Comparison of the Coefficients Determined for Study 11x

	Methods			
	Analytical	Graphical	Analog	Digital
A	0.965	0.497	0.342	0.424
B	0.022	0.029	0.040	0.033
Sample Variance	0.0036	0.0019	0.0019	0.0014

A comparison of the constants found by Moressi with those given by the least squares program is shown in Table 2. Points calculated from the constants were compared with the data; the variance given by the digital fit is lower in all cases than that for the analog fit.

Table 2. Coefficients in Moressi studies.

Study	Analog Computer Fit by Moressi			Digital Computer Fit		
	A	B	Sample Variance	A	B	Sample Variance
14x	0.354	0.04	0.0082	0.524	0.035	0.0005
11x	0.342	0.04	0.0019	0.424	0.033	0.0014
44c	2.29	0.023	0.00033	3.00	0.022	0.00015
35x	2.09	0.018	0.0016	5.76	0.008	0.0012

Figure (9) is a typical comparison of the fitted curves obtained by the analog computer with those obtained by the digital computer program.

Convergence: Convergence is usually obtained in 10 trials or less, if the data fit the assumed model. If the inhibitor concentration remains constant, however, $k_2 = 0$ and the program gives a rapid divergence. In this case, the data are fitted to the simpler model, for which

$$\frac{dy}{dt} = \frac{-k_1 y}{x_0} \qquad (24)$$

One integration, this gives

$$y = e^{-k_1 y t/x_0}. \qquad (25)$$

Methods are available for fitting data to this equation.

Literature Cited

1. Davis, D. S., (1955). *Nomography and Empirical Equations,* McGraw-Hill, New York, 74-79.
2. Gompertz, B., (1825). *Phil. Trans. Roy. Soc.* London *115,* 513.
3. Moressi, W. J., (1964). *Expt'l Cell Research* 33, 240.