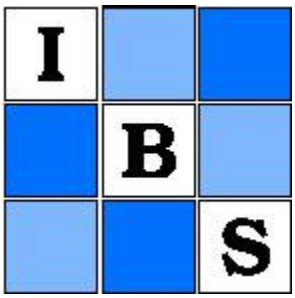

A COMPUTER PROGRAM FOR FITTING THE RICHARDS FUNCTION

D. R. CAUSTON[1]

Department of Botany, University College of Wales, Aberystwyth, U. K.

SUMMARY

A method is described for fitting Richards' generalisation of the logistic function by the method of least squares. The solution of the equations obtained requires the use of an iterative process from given starting values and these are obtained in the first part of the program by a procedure based on Hartley's [1948] method of 'internal least squares.'

1. INTRODUCTION

The Richards function is defined by the differential equation

$$\frac{dW}{dt} = \frac{kW}{nA^n}(A^n - W^n), \tag{1}$$

where W may be the weight of an organism at time t, and A, k, and n are constants (Richards [1959]). The function is a generalisation of the logistic which is given by (1) when $n = 1$. The integrated form of (1) is

$$W = A(1 \pm be^{-kt})^{-1/n}, \tag{2}$$

where b is associated with the constant of integration, while the upper sign within the brackets is applicable when n is positive and the lower sign when n lies in the range $-1 \leq n < 0$. The function is not defined for $n < -1$, nor for $n = 0$. Mathematical properties of the function are given by Richards [1959] and Causton [1967].

Although Richards' function is probably, as yet, the most realistic mathematical description of plant and animal growth, the problem of fitting (2) to experimental data is a formidable one. Richards describes an empirical method of fitting which is not only very laborious but, in the present author's experience, can produce misleading results. Nelder [1961] describes a method of fitting the curve by least squares based on a realistic statistical model, and also provides special tables to assist in the fitting of the curve by hand. The least squares equations obtained in this process have no explicit solution and have to be solved by iteration from given starting values, which themselves must be obtained by a rather time-consuming procedure.

The method of solution of the least squares equations is readily programmed for an electronic computer, but starting values have still to be provided. If

[1]Present address: Forestry Commission Research Branch, Government Buildings, Bankhead Ave., Edinburgh 11, Scotland.

many curves are to be fitted at any one time the provision of such values will be a long and tedious process. The object of this paper is to present a computer program for the fitting of Richards' function which also provides its own starting values. Thus only the actual data need be fed into the computer; no subjective estimation is required. First, however, we shall briefly consider the least squares theory relevant to the purpose in hand.

2. FITTING RICHARDS' FUNCTION BY LEAST SQUARES

We are required to fit a curve to a series of points (t_i , W_i): $i = 1, \ldots, N$. In describing biological growth it is usual to find that the variance of $\log W$ is approximately constant, and the t_i are presumed to be known exactly. Writing $w = \log W$ (all logarithms in this paper are to base e), $a = \log A$, $m = -1/n$, $K = -k$ so that this parameter may take its own sign implicitly which will be negative, and $B = |b|$, (2) becomes

$$w = a + m.\log (1 + Be^{Kt}) \tag{3}$$

and this is the form in which curves will be fitted to data. Hence the expression to be minimised is

$$S^2 = \Sigma\{w_i - a - m.\log (1 + Be^{Kt_i})\}^2. \tag{4}$$

Differentiating (4) successively with respect to a, B, K, and m, we have

$$\partial S^2/\partial \mathbf{a} = \mathbf{v}, \tag{5}$$

where $\mathbf{a}$ is a column vector of the parameters of (3) — a, B, K, and m; and the elements of $\mathbf{v}$ are given in the appendix to this paper. Solution of the equation

$$\mathbf{v} = \mathbf{0}, \tag{6}$$

where $\mathbf{0}$ is a column vector of four zeros, yields the least squares estimates of the four parameters. However, an explicit solution does not exist and (6) must be solved by iteration. The method used is the Newton–Raphson technique in four dimensions.

If we had a function of the form $y = f(x)$ and were required to find the roots of this function by the Newton–Raphson method, a guess, x', of the correct value would first be made and then a better estimate of the true value would be given by

$$x'' = x' - f(x')/f'(x'). \tag{7}$$

In our present four-dimensional case, the analogy with (7) is complete in matrix notation:

$$\mathbf{a}_1 = \mathbf{a}_0 - \mathbf{M}^{-1}\mathbf{v}, \tag{8}$$

where $\mathbf{a}_0$ and $\mathbf{a}_1$ are column vectors of the starting and adjusted values of the four parameters respectively, and $\mathbf{M}$ is the sample information matrix, the elements of which are given in the appendix.

3. OBTAINING THE STARTING VALUES

Richards' method

A convenient method of obtaining the starting values is the empirical method of fitting the curve given by Richards, which makes use of the linear form of the function:

$$\log |(A/W_i)^n - 1| = \log B - kt_i \,. \tag{9}$$

To calculate values of the left-hand side for each point, to be written as y_i, values of A and n must be first provided; the remaining two constants, B and k, may then be found by linear regression.

Now A is the asymptotic maximum value of W and so may be estimated from a graph of W_i against t_i. In the case of n, it can be shown that

$$W_I/A = (n + 1)^{-1/n}, \tag{10}$$

where W_I is the value of W at the point of inflexion of the curve (2). Hence from the graph of W_i against t_i it is possible to estimate the position of the point of inflexion, evaluate the ratio W_I/A, and obtain an estimate of n. From calculated values for A and n, a series of y_i's may be computed, one for each W_i, and these when plotted against t_i should give a straight line if the values of A and n are approximately correct, apart from random errors. Deviations from linearity may then be corrected by changing the values of A and/or n as described by Richards.

In practice it will be found that unless the data are very regular, considerable difficulty will be experienced in determining the correct adjustments of A and n to give a straight line; one never knows just how straight a line a particular set of data can give. If the Richards function is not an appropriate underlying model of the data in hand, or an approximation to it, a linear relationship between y_i and t_i will never be obtained by adjustment of A and n. Apart from these objections associated with the need to make estimations by eye, there are other disadvantages to this method. In order to be able to estimate A and n with any degree of confidence, values of W_i must extend over almost the whole range of W from near zero to near A. With less extensive data there is no hope of being able to make good estimates of the two parameters. Secondly, it may be found that one or more W_i's are greater than the estimated A. Such points will have to be eliminated for all further calculations involved in finding starting values since the corresponding y_i's will be indeterminate.

A new method

Hartley [1948] described a procedure whereby the parameters of certain non-linear functions could be estimated by fitting the constants of the linear 'generating law' of the function in question. Since Richards' function is derived from such a law (equation (1), not (9)), the principle of Hartley's method is applicable here although the actual procedure adopted differs in detail.

If we denote the relative growth rate $(dW/dt \cdot 1/W)$ as R, equation (1) may

be written as

$$R = \frac{k}{n} - \frac{kW^n}{nA^n} \tag{11}$$

which is of the general form

$$R = \alpha + \beta W^c. \tag{12}$$

Comparing coefficients of (11) and (12), we have

$$n = c, \qquad k = \alpha n, \qquad A = (-k/\beta n)^{1/n}. \tag{13)-(15)}$$

Thus if (12) is fitted, the constants of (11) may be determined and so provide starting values of three of the parameters. A starting value of B is then readily determined by use of (9).

The data from a growth experiment consist of values of W_i at times t_i ; values of R must be calculated from these observations. Fisher [1921] has shown that, whatever form a growth curve may take, the mean value of R over a period of time between harvests at t_1 and t_2 , when the weights are W_1 and W_2 respectively, is given by

$$\bar{R} = (\log W_2 - \log W_1)/(t_2 - t_1). \tag{16}$$

The present procedure is to use (16) to estimate mean relative growth rates between alternate harvests, i.e.

$$\bar{R}_{i,i+2} = (\log W_{i+2} - \log W_i)/(t_{i+2} - t_i), \tag{17}$$

and this, for the purpose of fitting equation (12) may be considered as corresponding to W_{i+1} . Thus a series of points $(W_{i+1}$, $\bar{R}_{i,i+2})$ are obtained corresponding to W_2 , ... , W_{N-1} , the two extreme values W_1 and W_N being, of necessity, omitted. From here on the co-ordinates of the points will be referred to as $(W_i$, $R_i)$. The variance of the R_i may be considered to be constant since the log W_i are independent of one another and of approximately constant variance. The variance of the W_i are roughly proportional to W^2. Thus we have a bivariate population and, strictly, the least squares model with only one of the variables subject to error is not applicable. However, since we require only approximate starting values, the simple regression model has been used.

The method for obtaining the least squares fit of (12) would also require the use of an iterative process from given starting values and thus defeat the object of this procedure. Hence the actual method employed is to fit a straight line to R_i against W_i^c . The parameter c is first set equal to -1 which is the lowest sensible value for Richards' function and represents the monomolecular curve. Values of W_i^c are next calculated, a linear regression is executed, and then the value of c is increased by 0.1 and the whole process repeated. The residual sum of squares obtained in the first regression is compared with that obtained in the second, and if the latter is lower, 0.1 is again added to c and a further regression computed. The process continues until the latest calculated residual sum of squares is greater than the one previous. At this stage the minimum residual sum of squares has been reached (over-shot slightly in fact) and so the value

of c at this stage gives almost the straightest line that the series of points (W_i, R_i) are capable of yielding. From this value of c together with the values of α and β given by the latest regression, the starting values of A, k, and n are given by equations (13) to (15).

In the calculation of a starting value for B from (9), two alternative procedures are possible. A value of this parameter can be calculated from the relationship

$$\log B' = (\Sigma y_i - k'\Sigma t_i)/N', \tag{18}$$

where N' is the number of points remaining when those in which $W_i \geq A$ have been eliminated, and the priming of the parameters denotes starting values. Alternatively, both $\log B'$ and a new value of k' may be simultaneously computed by linear regression. Now the variance of y_i is small when W_i is small, and approaches infinity as $W_i \to A'$. A least squares model could be formulated for this situation, but here the simpler course is adopted of using a model which assumes that the variance of y_i is proportional to t_i . It has been found in practice that merely calculating B' from (18) is not only adequate, but is better than estimating both B' and k' simultaneously unless the value of the former is small, which will be the case when n' is near -1. Thus when n' lies in the range (say) $-1 \leq n' \leq -0.7$, the second method wherein both B' and k' are simultaneously estimated is much superior.

4. THE COMPUTER PROGRAM AND ITS LIMITATIONS

The program was written in FORTRAN II and developed on an IBM 1620 machine. The whole program consists of two distinct parts—that for obtaining the starting values, and that for carrying out the least squares iterations. If the computer working store is not large enough to carry the complete program, a split into two linked programs may be conveniently made between these two parts.

The general structure of the first part of the program is shown in the flow diagram in Figure 1. The only aspect which requires comment is concerned with the sign of k'. When n' is positive, k' will always be positive since the straight line (12) will be situated as shown in Figure 2A and $k' = \alpha n'$. When n' is negative the line should occur as shown in Figure 2B with α negative. Now if the data consist of a few rather irregular points where the W_i are small compared with A', it may be that linear regression yields a line in which α is positive. As may be seen from examination of Figure 2B, A' will then be indeterminate and the program will be unable to proceed further. Hence the provision for testing the sign of k' (which will be negative under these circumstances) is provided for, before the computer attempts an impossible calculation.

The second part of the program, the formulation and solution of the least squares equations, is completely straightforward in structure and it has usually been found that the residual sum of squares converges to a minimum quite rapidly as the number of iterations increases. Often convergence is not smooth but consists of two or three iterations in which the residual sum of squares decreases, followed by a single iteration in which the residual increases somewhat.

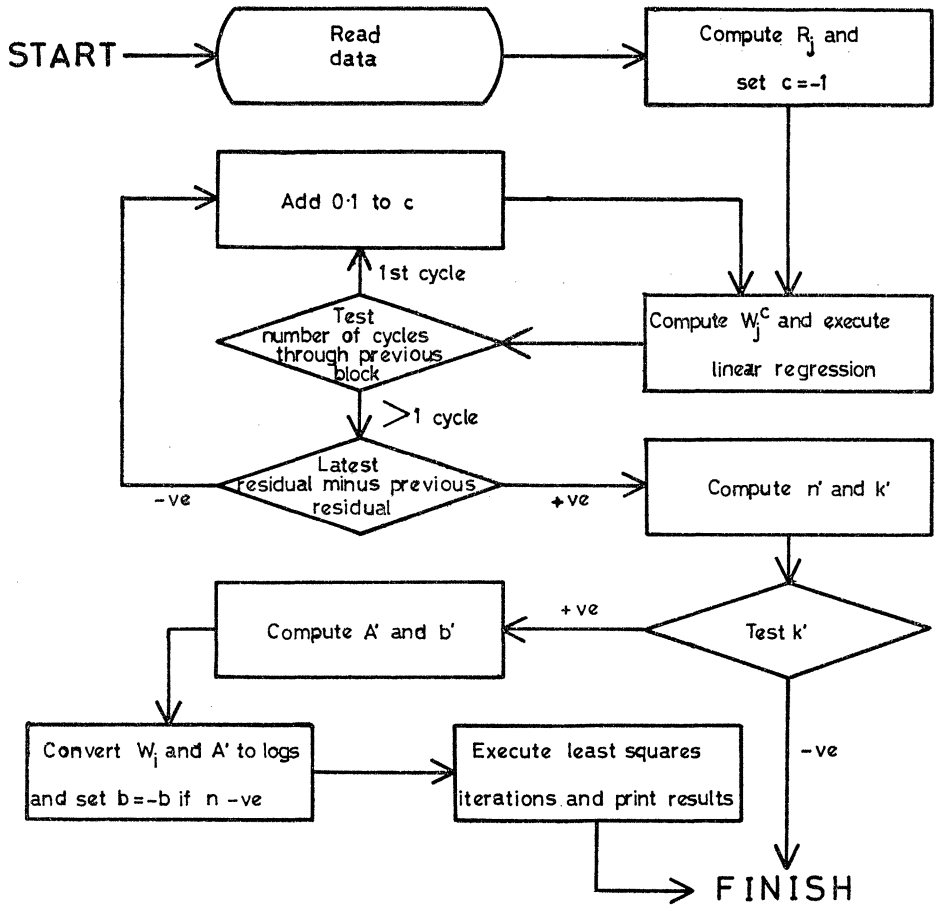

FIGURE 1

FLOW DIAGRAM OF THE FIRST PART OF THE PROGRAM FOR FINDING THE STARTING VALUES OF THE PARAMETERS FOR FITTING RICHARDS' FUNCTION

If a sufficient number of iterations is computed for a set of data a definite pattern of convergence may emerge. Occasionally, with very irregular data, complete instability is suddenly shown, the residual increasing by a hundred- or even a thousand-fold, and this may happen at any time, even at the very beginning; stability is then never recovered. An explanation for this behaviour is given by Marquardt [1963] who also gives an algorithm which may overcome the situation, but this algorithm has not been tried by the present author.

5. WORKED EXAMPLE

The data below show the growth of first year seedlings of sycamore (*Acer pseudoplatanus*). Harvests were taken at 2-weekly intervals commencing 5th May, 1965.

t (weeks)	W (gm.)	t (weeks)	W (gm.)
2	0.3898	14	5.610
4	0.8110	16	9.698
6	1.293	18	12.56
8	1.840	20	21.85
10	2.590	22	23.12
12	3.770	24	20.03

The value of n of the fitted Richards function was 25, showing that growth was exponential throughout the season with a sharp decline in the growth rate to zero after the 20th week. The goodness-of-fit of the Richards function should thus be substantially better than that of a logistic function fitted to the same data. The table below compares relevant quantities obtained by fitting the two types of curve.

Function	Richards'	Logistic
Starting value of a	3.1367	3.5267
B	5.7094×10^{45}	124.0186
K	−5.1795	−0.2412
n	24.5	1*
Residual sum of squares about curve defined by the above parameters	0.1248	0.2160
Converged value of a	3.0772	3.5435
B	7.7020×10^{45}	124.0136
K	−5.2460	−0.2387
n	25.0180	1*
Residual sum of squares	0.1159	0.2151

* By definition

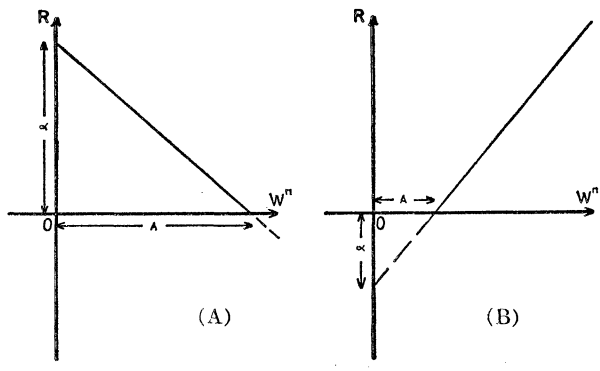

FIGURE 2

THE LINEAR RELATIONSHIP BETWEEN W^n AND R WHEN n IS POSITIVE (A), AND WHEN n LIES IN THE RANGE $-1 \leq n < 0(B)$

Since the logistic function is merely a special case of Richards' function and differing by one parameter, a comparison of the goodness-of-fit of the curves may be made as shown below.

	S.S.	D.F.	M.S.	F
Deviations from logistic curve	0.2151	9		
Deviations from Richards' curve	0.1159	8	0.0145	
Difference	0.0992	1	0.0992	6.84

The F-ratio is significant at $P(0.05)$.

ACKNOWLEDGEMENTS

I am grateful to Professor D. V. Lindley, and to Dr. S. C. Pearce for helpful discussions; I also wish to thank a referee for drawing my attention to Marquardt's work. The work was carried out during the tenure of a Forestry Commission Grant.

UN PROGRAMME SUR CALCULATEUR POUR AJUSTER LA FONCTION DE RICHARD

RESUME

On décrit une méthode pour ajuster la généralisation de Richard de la fonction logistique par la méthode des moindres carrés. La solution des équations obtenues implique l'utilisation d'un processus itératif à partir de valeurs initiales données, et celles-ci sont obtenues au début du programme grâce à une procédure établie à partir de la méthode de Hartley [1948] des 'moindres carrés internes.'

REFERENCES

Causton, D. R. [1967]. Some mathematical properties of growth curves and applications to plant growth analysis. Ph. D. thesis, University of Wales.

Fisher, R. A. [1921]. Some remarks on the methods formulated in a recent article on the quantitative analysis of plant growth. *Ann. Appl. Biol. 7*, 367–72.

Hartley, H. O. [1948]. The estimation of non-linear parameters by 'internal least squares.' *Biometrika 35*, 32–45.

Marquardt, D. W. [1963]. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Industr. Appl. Math. 11*, 431–41.

Nelder, J. A. [1961]. The fitting of a generalisation of the logistic curve. *Biometrics 17*, 89–110.

Richards, F. J. [1959]. A flexible growth function for empirical use. *J. Exp. Bot. 10*, 290–300.

APPENDIX

The details of the vectors of equation (5) are:

$$\begin{bmatrix} \frac{\partial S^2}{\partial a} \\ \frac{\partial S^2}{\partial B} \\ \frac{\partial S^2}{\partial K} \\ \frac{\partial S^2}{\partial m} \end{bmatrix} = \begin{bmatrix} -2\Sigma\{w_i - a - m\cdot\log(1 + Be^{Kt_i})\} \\ -2m\Sigma\{w_i - a - m\cdot\log(1 + Be^{Kt_i})\}\left\{\frac{e^{Kt_i}}{1 + Be^{Kt_i}}\right\} \\ -2Bm\Sigma\{w_i - a - m\cdot\log(1 + Be^{Kt_i})\}\left\{\frac{e^{Kt_i}t_i}{1 + Be^{Kt_i}}\right\} \\ -2\Sigma\{w_i - a - m\cdot\log(1 + Be^{Kt_i})\}\{\log(1 + Be^{Kt_i})\} \end{bmatrix}$$

The elements of **M** are:

$$\frac{\partial^2 S^2}{\partial a^2} = 2N \qquad \frac{\partial^2 S^2}{\partial a\,\partial B} = 2m\Sigma\left\{\frac{e^{Kt_i}}{1+Be^{Kt_i}}\right\} \qquad \frac{\partial^2 S^2}{\partial a\,\partial K} = 2Bm\Sigma\left\{\frac{e^{Kt_i}t_i}{1+Be^{Kt_i}}\right\}$$

$$\frac{\partial^2 S^2}{\partial a\,\partial m} = 2\Sigma \log\,(1+Be^{Kt_i}) \qquad \frac{\partial^2 S^2}{\partial m^2} = 2\Sigma \log^2\,(1+Be^{Kt_i})$$

$$\frac{\partial^2 S^2}{\partial B^2} = 2m\Sigma[w_i - a - m\{\log\,(1+Be^{Kt_i}) - 1\}]\left[\frac{e^{Kt_i}}{1+Be^{Kt_i}}\right]^2$$

$$\frac{\partial^2 S^2}{\partial B\,\partial K} = -2m\Sigma[w_i - a - m\{\log\,(1+Be^{Kt_i}) + Be^{Kt_i}\}]\left[\frac{e^{Kt_i}t_i}{(1+Be^{Kt_i})^2}\right]$$

$$\frac{\partial^2 S^2}{\partial B\,\partial m} = -2\Sigma\{w_i - a - 2m\cdot\log\,(1+Be^{Kt_i})\}\left\{\frac{e^{Kt_i}}{1+Be^{Kt_i}}\right\}$$

$$\frac{\partial^2 S^2}{\partial K^2} = -2Bm\Sigma[w_i - a - m\{\log\,(1+Be^{Kt_i}) + Be^{Kt_i}\}]\left[\frac{e^{Kt_i}t_i^2}{(1+Be^{Kt_i})^2}\right]$$

$$\frac{\partial^2 S^2}{\partial K\,\partial m} = -2B\Sigma\{w_i - a - 2m\cdot\log\,(1+Be^{Kt_i})\}\left\{\frac{e^{Kt_i}t_i}{1+Be^{Kt_i}}\right\}$$

Received February 1968, *Revised October* 1968