Jour. Ind. Soc. Ag. Statistics
56(1), 2003 : 19-24

Mixed-Influence Nonlinear Growth Model

Prajneshu and V.M. Kandala
Indian Agricultural Statistics Research Institute, New Delhi–110012
(Received : March, 2002)

SUMMARY

Nonlinear mixed-influence growth model, obtained by combining well-known logistic and monomolecular models, is thoroughly studied. As an illustration, proportion of area under high yielding varieties of wheat in the country during post-Green revolution era is modelled through this model. Employing nonlinear estimation procedures, it is found that some parameter estimates did not converge to biologically meaningful values. Subsequently, one parameter of the model is reparameterized by using the concept of 'Expected-value parameter'. Consequently, the resultant form is successfully fitted to the data. Another advantage is that, by computing Bates and Watts measures of nonlinearity, it is shown that the model-data set combination depicted close-to-linear behaviour. Finally, Wald's test is applied to test the hypothesis of equality of some parameters.

Key words : Monomolecular model, Logistic model, Nonlinear estimation, Measures of nonlinearity, Expected-value parameters.

1. Introduction

Nonlinear growth models are often employed to get insight into the dynamics of a population. Prajneshu and Sharma [4] provided a justification for nonlinear monomolecular growth model for describing path of adoption of high yielding varieties (HYVs) of wheat at State level. Also, wheat production data at State level in post-Green revolution era was modelled by Prajneshu and Das [3] employing nonlinear logistic and Gompertz growth models. These types of models are also gaining importance in "Marketing research" to explain the first-purchase sales growth of a new product over time and space. Monomolecular model is appropriate if diffusion occurs due to "Innovation" while logistic and Gompertz models are appropriate if it is due to "Imitation" (Mahajan and Wind [2]). However, in reality, both of these play important roles in the process of diffusion of a new product. Thus, it is highly desirable to include the aspects of innovation as well as imitation in a more realistic modelling exercise.

One main reason for wide applicability of above types of nonlinear models is that these contain fewer number of parameters, viz. three. Consequently, their estimation, assuming additive errors, can be carried out without much difficulty by using "Nonlinear estimation procedures". However, if a combination of nonlinear growth models, as envisaged above, is considered, number of parameters becomes at least four and their estimation poses challenging problems.

The objective of present paper is to study four-parameter mixed-influence growth model, that is a combination of monomolecular and logistic growth models. Subsequently, application of a reparameterization procedure, viz. "Expected-value parameters" is highlighted for estimation purposes. Finally, the model is fitted to some real data and hypothesis testing about equality of some parameters is carried out.

2. *Mixed-Influence Model*

The differential equation governing mixed-influence growth model for describing growth of a variable x is

$$dx / dt = a (c - x) + bx (1 - x / c) , x(0) = d \tag{1}$$

where a and b are respectively coefficients of "Innovation" and "Imitation" and c is carrying capacity. It may be noted that monomolecular and logistic models are particular cases of eq. (1) obtained respectively by putting b or a equal to zero. To solve eq. (1), separation of variables yield

$$\left(\frac{1}{x - c} - \frac{1}{x + ac / b} \right) dx = -(a + b)\, dt \tag{2}$$

Integrating and simplifying

$$x(t) = \frac{ac(d-c)\, e^{-(a+b)t} + c(ac + bd)}{ac + bd + b(c - d)\, e^{-(a+b)t}} \tag{3}$$

The point of inflexion for this model, obtained on solving $d^2x / dt^2 = 0$ for which $d^3x / dt^3 \neq 0$, is $x^* = c(-a + b) / (2b)$ if $b > a$ and occurs at time $t^* = (a + b)^{-1} \log_e (b / a)$. Thus this model is extremely flexible as point of inflexion can occur at any fraction of half the carrying capacity depending on the relative magnitude of coefficients of innovation and imitation.

3. *Estimation of Parameters*

Evidently, the four parameters, viz. a, b, c, and d appear nonlinearly in eq. (3). Fluctuations in the system are assumed to be described by adding an error term on the R.H.S. of this equation. Generally, Levenberg-Marquardt nonlinear estimation procedure, which is available in any standard software package, like SPSS, SAS, or SPLUS is employed for estimation of parameters. However, it is extremely rare that, for a data set, all four parameter estimates converge to meaningful values with low asymptotic standard errors.

To this end, sometimes reparameterization of one or more parameters makes it possible to fit the model to data. However, success of this procedure depends on the extent of nonlinearity of a model data-set combination. As discussed in Bates and Watts [1], this is determined by two measures of curvature, viz. Intrinsic curvature (IN) and Parameter-effect curvature (PE). The former is inherent to the model data-set combination and cannot be altered by any reparameterization, while latter is dependent on a particular parameterization. Evidently, lower the values of IN and PE, closer is the model-data set combination to linearity. To the best of our knowledge, only SPLUS software package contains programs for computing IN and PE.

The main method of reparameterization is to use the concept of "Expected-value parameters", which was introduced by Ross [6] to provide efficient estimates for nonlinear models and elaborated by Ratkowsky [5]. In this approach, initial values of one or more parameters are taken as expected-values of dependent variable for some data-values. Suppose, for some value of independent variable t, say t_1, the observed value of x(t) is x_1. In principle, any one of four parameters, viz. a, b, c, d appearing in eq. (3), may be replaced by an "Expected-value parameter". However, for the nonlinear model under consideration, it may be noted that it is not possible to express parameter a or parameter b explicitly in terms of parameters (b, c, d, x_1) or (a, c, d, x_1) respectively due to presence of exponential functions in eq. (3). Further, numerator of eq. (3) involves a complicated quadratic term in parameter c while it is only linear in parameter d. So, in short, as far as nonlinear model given by eq. (3) is concerned, best choice is to express parameter d in terms of "Expected-value parameter". To this end, using eq. (3), a straightforward algebra yields the expression for the fourth parameter d as

$$d = \frac{ac^2\,[1 - e^{-(a+b)t_1}] - cx_1\,[a + b\,e^{-(a+b)t_1}]}{bx_1\,[1 - e^{-(a+b)t_1}] - c\,[b + a\,e^{-(a+b)t_1}]} \tag{4}$$

In order to fit nonlinear model given by eq. (3) with additive errors to data, the fourth parameter d is replaced by the expected-value parameter x_1 , as given in eq. (4). As mentioned earlier, initial value of x_1 is taken as the corresponding data-value for some value t_1 of independent variable t. Final estimates of parameters (a, b, c, x_1) and hence of (a, b, c, d) can be obtained using "Nonlinear estimation procedures". Subsequently, PE-measure of nonlinearity may be computed to see whether or not there is a substantial reduction. If not, possibility of employing two or more expected-value parameters, on similar lines as above, may be explored.

4. An Illustration

All India proportion of area under high yielding varieties of wheat during the post-Green revolution period, i.e. from 1966-67 to 1996-97 is considered. Attempts are made to estimate parameters of the nonlinear model given by eq. (3), using Levenberg–Marquardt estimation procedure available in SAS package. However, it is found that global convergence to biologically meaningful values, particularly for parameter b, did not take place. Further, asymptotic standard errors of estimates of two parameters, viz. b and d are too high. Also, IN- and PE-values, using SPLUS software package are computed as

$$\text{IN} = 0.095 \qquad\qquad \text{PE} = 7.662$$

High value of PE also indicates the need for reparameterization. Subsequently, the reparameterization, as given in Eq. (4), is used. Different values of t_l are tried and it is found that $t_l = 3$ gives the best results in terms of least asymptotic standard errors and least PE-value. However, any other value of t_l would serve the purpose. Specifically, taking $t_l = 3$ and using SAS package

$$\hat{a} = \underset{(0.019)}{0.108} \qquad \hat{b} = \underset{(0.005)}{0.029} \qquad \hat{c} = \underset{(0.046)}{0.911} \qquad \hat{d} = \underset{(0.021)}{0.045}$$

where figures within brackets ( ) indicate corresponding asymptotic standard errors. It is worth noting that PE-value has now substantially decreased to only 0.812, which reflects that the model-data set combination exhibits closer-to-linear behaviour. This has several advantages in the sense that desirable properties of estimators, viz. unbiasedness, minimum variance, distribution being normal are more likely to be satisfied. Residual analysis reveals that the assumptions of independence and normality of error terms are not violated at 5% level of significance. The graph of residuals against time is depicted in Fig.1. Goodness of fit of the fitted model is examined by computing mean square error (MSE) which is as low as 0.0012. Thus, it is concluded that the model provides a good fit to the data. To get a visual idea, the fitted model along with data is displayed in Fig.2.

Some conclusions that may have policy implications can be drawn on the basis of results obtained. The estimate of carrying capacity c is 0.911; that means that only 91% area or so under wheat in the country is ultimately likely to be covered under the high yielding varieties. Further, the estimate of coefficient of innovation $(\hat{a})$, i.e. 0.108, is larger than that of coefficient of imitation $(\hat{b})$, which is only 0.029. However, to test H_0: $a - b = 0$ against H_1: $a - b > 0$, Wald's test, as given in Gallant [7], is carried out. The calculated value of W = 5.598, being more than the tabulated value $F_{2,27}(0.05) = 3.350$, implies that innovation played a significantly greater role vis-à-vis imitation as far as diffusion of high yielding varieties of wheat in the country is concerned.

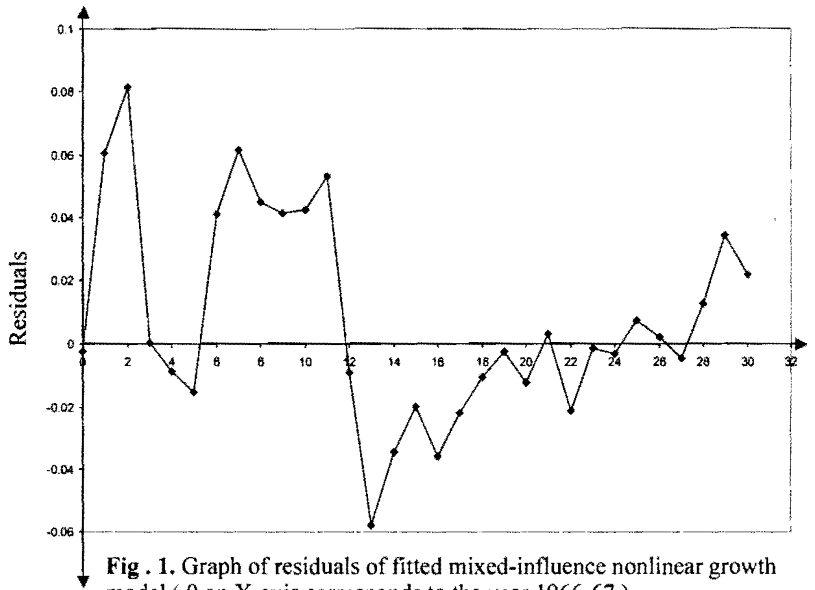

Fig . 1. Graph of residuals of fitted mixed-influence nonlinear growth model (0 on X-axis corresponds to the year 1966-67)

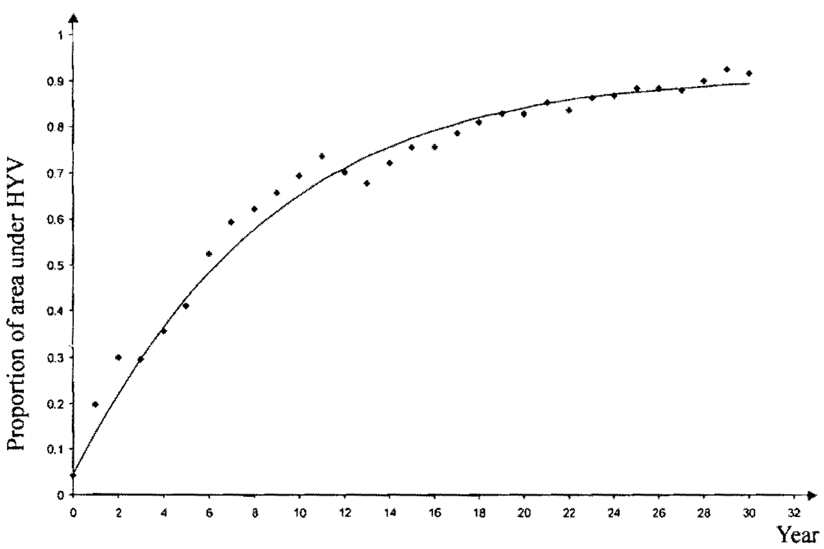

Fig . 2. Graph of fitted mixed-influence nonlinear growth model along with data points (0 on X-axis corresponds to the year 1966-67)

ACKNOWLEDGEMENT

Authors are grateful to the referee for valuable comments.

REFERENCES

[1] Bates, D. M. and Watts, D. G. (1988). *Nonlinear Regression Analysis and its Applications.* John Wiley and Sons, U.S.A.

[2] Mahajan, V. and Wind, Y. (1986). *Innovation Diffusion Models for New Product Acceptance.* Ballinger Publishing Co., U.S.A.

[3] Prajneshu and Das, P. K. (1998). Modelling wheat (*Triticum aestivum*) production at state level in post-green revolution era. *Ind. Jour. Agril. Sci.*, **68**, 23-26.

[4] Prajneshu and Sharma, V.K. (1992). A nonlinear statistical model for adoption of high yielding varieties. *Ind. J. Appl. Stat.*, **1**, 12-18.

[5] Ratkowsky, D. A. (1990). *Handbook of Nonlinear Regression Models.* Marcel Dekker, U.S.A.

[6] Ross, G. J. S. (1970). The efficient use of function minimization in nonlinear maximum likelihood estimation. *Appl. Stat.*, **19**, 205-221.

[7] Gallant, A. R. (1987). *Nonlinear Statistical Models.* John Wiley and Sons, U.S.A.