International Journal of Ecology and Environmental Sciences 40 (1): 41-47, 2014
ISSN: 0377-015X (Print); 2320-5199 (Online)

Nonlinear Statistical Model for Characterizing Culm Growth of *Bambusa cacharensis*

GITASREE DAS *
Department of Statistics, North-Eastern Hill University, Shillong, Meghalaya, 793022, India

ARUN JYOTI NATH # AND ASHESH KUMAR DAS +
Department of Ecology and Environmental Science, Assam University, Silchar, Assam, 788011, India.

* Corresponding author; *Email: gitasree22@gmail.com; # arunjyotinath@gmail.com, + asheshdas684@hotmail.com*

ABSTRACT

Bambusa cacharensis R.Majumder, locally known as betua, is an important village bamboo species in the Barak valley of Assam, India. Villagers prefer to grow this species in their home gardens because of its multiple uses and having straight culm. The study on growth pattern of bamboos is important for proper scientific management to obtain optimum yield. In this article nonlinear statistical models are used to characterise the culm height of *B. cacharensis* growing in clumps of different age groups (2-40 years). The adequacy of the models is judged by testing the validity of the assumptions of randomness and normality of residuals. Gompertz model with additive and AR (1) error structure is found to be the most suitable model for characterizing the culm growth of *B. cacharensis*. The physical interpretation of the parameters is also discussed. Each growth curve may be summarized by its parameter estimates as a single low-dimensional multivariate observation as (k, q, r, a). These observations may then be subjected to an analysis of variance. Wilk's Lambda test confirms significant difference between the growth parameters of bamboo culms for clumps of different age groups. Finally, Hotelling T^2 statistic and one-way analysis of variance of k-values was used to find out the best clump age group on the basis of growth parameters.

Key Words: Bamboo Culm; Bamboo Clump; Growth Model; Nonlinear; Gompertz; Wilk's Lambda; Hotelling t^2.

INTRODUCTION

Nonlinear statistical models have been used to describe growth behaviour, as it varies in time. The type of model needed in a specific area and specific situation depends on the type of growth that occurs. A nonlinear model is one in which at least one of the parameters appears nonlinearly. Generally growth models are mechanistic, arising as a result of making assumptions about the type of growth, writing down differential or difference equations that represent these assumptions and then solving these equations to obtain a growth model. Most of the mechanistic modeling has been done in a biological context. This type of work will be of interest mainly to biologists who examine growth and its underlying mechanism. For example, it is important from management point of view to know how large the plants will grow, how fast they grow and how they respond to environmental conditions or treatments. Unlike empirical models, here the parameters have meaningful biological interpretation. If the growth curve changes nonlinearly, it may be possible to find a simple mathematically defined curve which describes the change and which can be estimated from the data. The estimated parameters can then be treated as a summary of the growth pattern. These values may be used for comparison across species or varieties or same species growing under different environmental conditions.

Richards (1959) discussed application of growth functions viz. monomolecular, logistic, gompertz, von Bertalanffy's extended form etc. for plant data. Venugopal and Prajneshu (1997) studied generalized allometric growth model for length-weight relationship of Indian pearl oyster, observed for a period of 36 months. They also tested the validity of various assumptions and explored the possibility of autocorrelation in error term. The use of sigmoidal growth models viz. logistic, gompertz etc. has been discussed by Gore and Paranjpe (2000) for describing the growth of single species population.

While single growth curves may be of interest in their own right, very often growth curve data are collected to determine how growth responds to various treatments or other covariate information. Therefore it is essential to reduce each individual growth curve to a small number of parameters so that changing patterns of growth can be understood in terms of changes in these parameters. Each growth curve may be characterised by its parameter estimates (k, q, r, a) as a single low-dimensional multivariate observation. These observations may then be subjected to an analysis of variance or to a regression analysis (Seber and Wild 1989). Das et al. (2006) have studied nonlinear statistical model for culm growth of *Melocanna baccifera*. The model parameters k, q, r and a were used to describe the growth of *M. baccifera.* In this article an attempt is being made to obtain the most suitable model for describing the culm growth of *B. cacharensis* corresponding to clumps of different age groups. We further claim, with appropriate justification, that the parameters can even be used to characterize the growth of *B. cacharensis* corresponding to clumps of different age groups. We are yet to come across such literature where growth model parameters are used to characterize growth of different bamboo species, for making comparison between growth of different bamboo species or comparing growth of a bamboo species in different age groups. If the growth models are developed appropriately, taking care of validity of the assumptions etc., this may prove to be a handy tool for the biologists for making comparative studies on growth.

MATERIALS AND METHODS

Study Site and Culm Selection

The study was conducted in a bamboo grove located in Cachar district of Barak Valley of North East India at longitude 92° 45' East and latitude 24° 41' North. The climate of the area is subtropical, warm and humid. Culm growth was observed during July to November. Because bamboo is a quick growing plant and its height is stabilized over a comparatively short span of time, longitudinal study was preferred over random samples at different times (Hills 1974). Five clumps from each age group, viz. 2, 5, 10, 15 and 40 years, were selected randomly. Ten newly sprouted culms, two culms from each selected clump were selected and tagged with numbered aluminium foil. For 2-year old clumps only nine culms were available. Height of the elongating culm was measured every alternate day for the first 80 days. Thereafter, the height of culm was measured every week until growth of the culm ceased.

Sigmoidal Growth Model

The shape of the growth curve is described by the rate of change of growth at different time t. For certain types of growth data, the growth rate does not steadily decline, but rather increases to a maximum before declining to zero. This criterion, as discussed in detail by Das et al. (2006), Nandy et al. (2004) and Nath et al. (2004), is observed in the growth-rate curve for *Bambusa cacharensis* (Figure 1(b), Nath et al. (2004)). Sigmoidal growth models can best represent the growth curves in such cases. For such models the position of the point of inflection being the time when the growth rate is maximum. Sigmoidal behaviour can be achieved by modeling the current growth rate as proportional to the product of functions of the current size and remaining growth, namely

$$\frac{df}{dt} \propto g(f)[h(k) - h(f)] \quad (1)$$

where g and h are increasing functions with $g(0) = h(0) = 0$, f is the size at time t and k is the maximum possible growth size.

Logistic Model

The simplest form of (1) is with $g(f) = h(f) = f$, i.e. the growth rate is proportional to the product of the present size and the future amount of growth. So that

$$\frac{df}{dt} = \frac{r}{k} f(k - f) \quad (2)$$

where $r > 0$ and $0 < f < k$ and constant of proportionality being used is r / k. One may observe from (2) that the relative growth rate $f^{-1} df / dt$ decreases linearly in f as f approaches k. The equation (2) has a general solution as follows:

$$f(t) = \frac{k}{1 + e^{-r(t-q)}} \quad (3)$$

known as logistic or autocatalytic model. From (2), one can obtain the second derivative. It is easily seen that growth rate is maximum when $f = k/2$ and from (3), $f = k/2$ occurs when $t = q$. The maximum growth rate is $rk/4$ and the growth rate is symmetrical about $t = q$, giving a symmetrical sigmoidal growth curve. The curve increases to an upper limit k when t is large. The constant k is known as the 'carrying capacity' of the environment.

Logistic model requires growth rate be symmetrical about the point of inflection. However for *Bambusa cacharensis* (Figure 1(b), Nath et al. (2004)), growth rate does not seem to be symmetrical. Therefore a more generalized form of sigmoidal model known as Gompertz model is used.

Gompertz Model

The gompertz model is often used where the growth rate is not symmetrical about the point of inflection. The growth rate is

$$\frac{df}{dt} = rf(\log k - \log f) \quad r > 0, k > 0 \quad (4)$$

where $r > 0$ and $0 < f < k$ and the relative growth rate $f^{-1} df / dt$ declines non-linearly in f, or more precisely linearly with $\log f$. The equation (4) has a general solution as follows:

$$f(t) = k * \exp(-\exp(-r(t-q))) \quad (5)$$

known as gompertz model. It is easily seen that growth rate is maximum when $f = k/e$, and from (5), $f = k/e$ occurs when $t = q$, the point of inflection. The maximum growth rate is rk/e and the growth rate is not symmetrical about $t = q$. The curve increases to an upper limit k when t is large. Here also the constant k is known as the 'carrying capacity' of the environment.

Additive Error Structure

The above models have been presented deterministically, which is obviously unrealistic. We replace these nonlinear deterministic models with nonlinear statistical models by including additive error structure $\varepsilon(t)$ with appropriate assumptions, on the right hand side of the equation. Thus the logistic and gompertz models respectively appear as follows:

$$f(t) = \frac{k}{1 + e^{-r(t-q)}} + \varepsilon(t) \quad (6)$$

$$f(t) = k * \exp(-\exp(-r(t-q))) + \varepsilon(t)$$

Estimation of Parameters

In order to estimate the parameters $k, q,$ and r by 'method of least squares', ε (t) is assumed to be independently and identically distributed (i.i.d) following normal distribution. However, normal equations obtained by minimization of residual sum of squares are nonlinear in parameters. Since it is not possible to obtain explicit solutions for nonlinear equations, the next alternative is to obtain approximate solutions by employing iterative numerical procedures. In this work, the most commonly used method known as Levenberg-Marquardt non-linear iterative procedure (Seber and Wild 1989) was used for fitting the models to the data. To start the iterative procedure, initial estimates of the parameters of the models were required. Many sets of initial values were tried to ensure global convergence. The iterative procedure was stopped when the reduction between the successive residual sums of squares was found to be negligible.

Model Adequacy

It is important to remember that our confidence in statistical inference procedures is related to the validity of the assumptions about them. A mechanically made inference may be misleading if some model assumption is grossly violated. An examination of the residuals is an important part of regression analysis, because it helps to detect any inconsistency between the data and the postulated model. If no abnormalities are exposed in this process, then we can consider the model adequate and proceed with the relevant inferences. Otherwise we must search for a more appropriate model.

The adequacy of the models can be judged by testing the validity of the underlying assumptions of randomness and normality of residuals. Residuals were examined by using one-sample run test for inferring about randomness (Draper and Smith 1981) and Shapiro-Wilk test for testing normality.

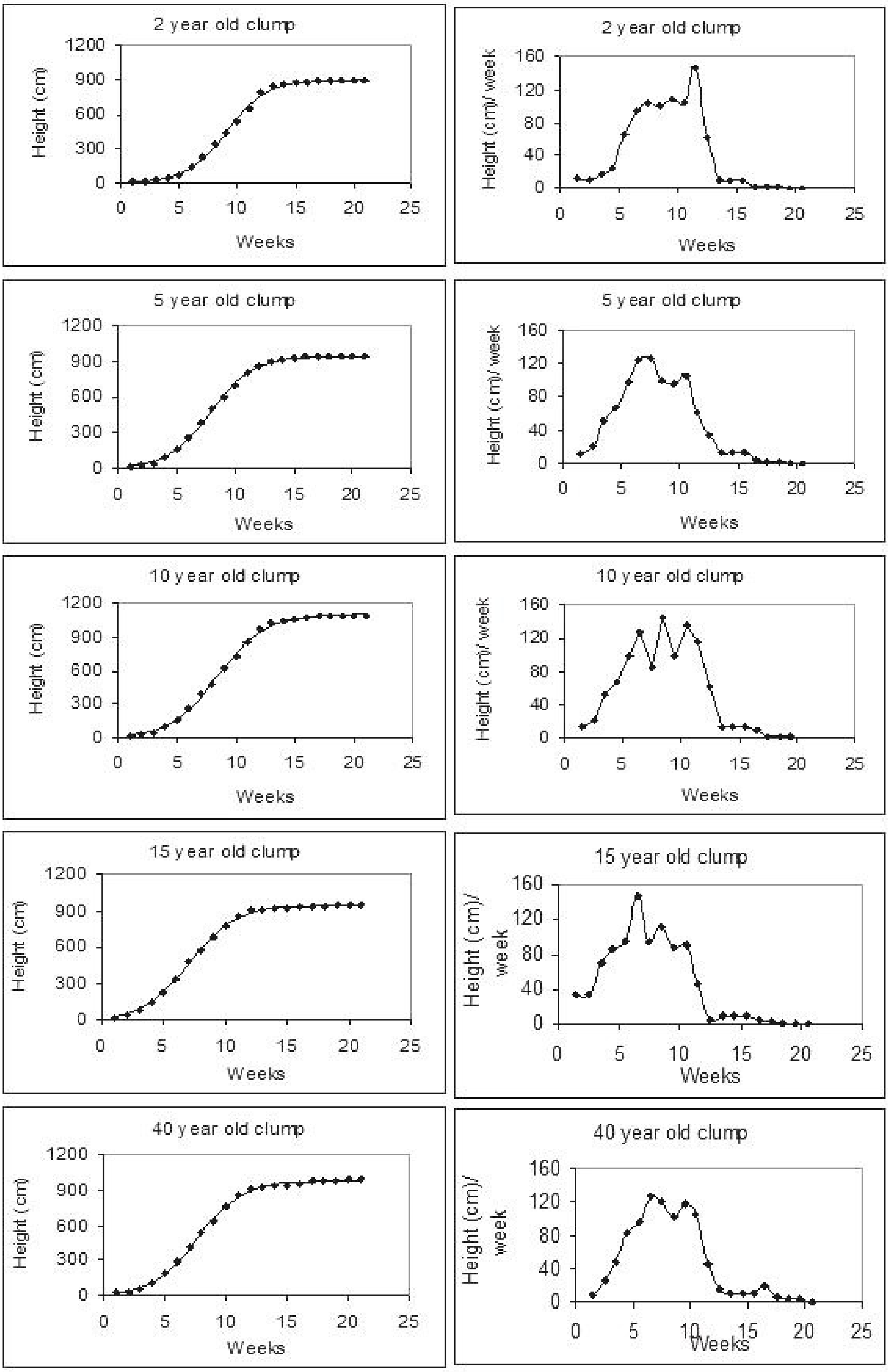

Figure 1. Culm height growth showing fitted gompertz model (----) and sample values (·) (1a) and culm height growth rate (1b) of *B. cacharensis.*

Growth data will often contain correlated errors if the data are collected as repeated measurements over time on the same experimental units. Growth models with uncorrelated errors become untenable because of long runs of residuals with the same sign (Seber and Wild 1989). When the models suggested are found to be inappropriate for a data set, using Durbin-Watson test one can examine the presence or absence of auto-correlation (Chatterjee and Price 1977). This test examines the presence or absence of first order auto-correlations, i.e. AR (1), among the residuals. If this test reveals presence of AR (1) error structure, the term ε (t) in equation (6) or (7) shall be replaced by

$$\varepsilon (t) = \rho\, \varepsilon (t-1) + u(t) \qquad (8)$$

where ρ is the autoregressive model parameter and $u(t)$ is assumed to be i.i.d. normally distributed with mean zero and variance σ^2. The unknown model parameters together with the AR parameter can be estimated by using two-stage estimation procedure.

The estimated parameters can further be used to characterize the growth pattern of culms for clumps of different age groups. Bamboo clumps of different age groups can further be compared on the basis of these parameters. Each growth curve may be summarized by its parameter estimates as a single low-dimensional multivariate observation. Gompertz model with additive and AR(1) error structure was fitted to individual culms from clumps of different age groups. Wilk's Λ-Test was used to test the equality of mean vectors (k, q, r, a) corresponding to clumps of different age groups. Hotelling T^2 statistic was computed to test equality of two mean vectors (k, q, r, a) corresponding to clumps of different age groups. Finally, one-way analysis of variance of k-values for clumps of different age groups was done to determine the clump age group with best culm growth. For fitting of nonlinear models software SPSS 10.0 was used. For simpler computations i.e. run test, Shapiro-Wilk test, Wilk's Λ-Test etc. Excel worksheet was used.

RESULTS AND DISCUSSION

At first the logistic model with additive error term given by eqn. (6) was fitted to the height growth data of 10 year old clump. However it was found that the assumption of randomness and normality of residuals are not satisfied. Thus this model is inappropriate for the present data set. The application of Durbin-Watson test indicates possibility of AR(1) error structure. Therefore eqn. (6) with errors at two consecutive time epochs having AR(1) structure given by eqn. (8) was fitted to the data. However Shapiro-Wilk's test and also the run test show that even this modification is not adequate. Thus, the logistic model with additive error and AR(1) structure is not found satisfactory for the data. And as such functional form of the model needs to be altered.

Next, another form of sigmoidal growth model, where the growth is not symmetrical about the point of inflection viz. gompertz model with additive error structure (eqn. (7)) was considered for height growth data corresponding to clumps of different age groups viz. 2, 5, 10, 15 and 40 years. The estimated parameters are shown in Table 1. Now, for some age group assumption of normality is violated while in others the assumption of randomness of residuals is violated. Nevertheless, in all the cases the low value of Durbin-Watson statistic indicates possibility of AR(1) error structure. Therefore,

Table 1. Parameter estimates of fitting Gompertz model to culm growth data along with their coefficients of variation (%) given in parentheses.

Clump Age (years)	Parameter Estimates k	q	r	R^2
2	917.832	7.937	0.379	0.9938
	(9.23)	(12.78)	(17.04)	
5	966.194	6.730	0.362	0.9981
	(8.95)	(12.44)	(20.48)	
10	1127.933	7.214	0.329	0.9955
	(4.53)	(12.84)	(9.11)	
15	961.521	5.977	0.365	0.9964
	(15.17)	(22.13)	(15.35)	
40	1001.339	6.519	0.356	0.9970
	(11.67)	(13.17)	(14.92)	

Gompertz model with additive error (eqn. (7)) and AR(1) structure (eqn.(8)) was tried and the results are shown in Table 2. The residual analysis confirms that with this modification the assumption of randomness and normality of residuals is not violated. Also the calculated value of Durbin-Watson statistic d in Table 2 indicates absence of autocorrelation ($d_u = 1.39$). Hence, gompertz model with additive and AR(1) error structure appears to be the most suitable model for the height growth data corresponding to clumps of all the age groups. The fitted values of culm height by gompertz model with additive and AR (1) error structure along with the observed values are shown in Figure 1(a) and height growth rate in Figure 1(b). Coefficients of variation (CV) of height growth at different time points were computed for clumps of all age groups. However variability in the data was not shown in the Figure so as not to make it overcrowded. It was found that the CV (%) was high only during the first four to five weeks of growth period ranging from 20-60%. However as the bamboo culms grow further in height, the CV's continue to reduce and during the last four to five weeks the CV's are as low as 4-14%.

Table 2. Parameter estimates of fitting Gompertz model with AR(1) error structure to culm growth data

Clump Age (years)	k	Q	r	a	R^2	d
2	908.347	7.971	0.405	0.798	0.9970	1.739
	(9.03)	(13.10)	(19.56)	(31.47)		
5	960.211	6.751	0.376	0.856	0.9992	1.760
	(9.01)	(12.22)	(22.71)	(54.61)		
10	1117.513	7.228	0.345	0.793	0.9977	2.015
	(4.52)	(11.79)	(8.32)	(26.35)		
15	955.758	6.009	0.382	0.782	0.9987	2.055
	(15.08)	(21.93)	(16.57)	(26.10)		
40	995.847	6.548	0.371	0.783	0.9987	1.630
	(11.56)	(12.95)	(17.03)	(24.29)		

The estimated parameters can further be used to characterize the growth pattern of culms for clumps of different age groups. Bamboo clumps of different age groups can further be compared on the basis of these parameters. Each growth curve may be summarized by its parameter estimates as a single low-dimensional multivariate observation. Gompertz model with additive and AR(1) error structure was fitted to individual culms from clumps of different age groups. Wilk's Λ-Test was used to test the equality of mean vectors (k, q, r, a) corresponding to clumps of different age groups. Since P ($F_{8, 82} > 6.9956$) = 5.351E-07, clumps of different age groups are significantly different with respect to their growth parameters.

Hotelling T^2 statistic was computed to test equality of two mean vectors (k, q, r, a) corresponding to clumps of different age groups and the p-values are shown in Table 3. It can be seen easily that 10-year old clumps are significantly different from the rest in terms of the parameters k, q, r, a. The constant *k* is known as the 'carrying capacity' of the environment. The growth curve increases to an upper limit *k* when *t* is large. Thus k is the most important parameter showing the maximum height that a bamboo culm can attain. One-way analysis of variance of k-values for clumps of different age groups indicates ($p < 0.001$) best culm growth for 10-year old clumps.

Table 3. p-values for testing equality of two mean vectors (k, q, r, a) using Hotelling T^2 corresponding to clumps of different age groups

Clump Age (years)	2	5	10	15	40
2	-	0.1188	0.0003	0.0142	0.0539
5			0.0003	0.2225	0.3564
10				0.0578	0.0589
15					0.3456
40					-

Thus, the parameters (k, q, r, a) given in Table 2 can be used to characterize the culm height growth of *B. cacharensis* growing in clumps of different age group. Das et al (2006) have shown that for *Melocanna baccifera* $k = 1698.855$, $q = 10.839$, $r = 0.182$ and $a = 0.833$.These are distinctly different from the growth parameters given in Table 2 corresponding to clumps of different age groups for *B. cacharensis*. Thus the parameters can also be used to characterize the growth of different species. The parameter *q* gives the position of the point of inflection i.e. the time when the growth rate is maximum. For *B. cacharensis* this period varies between 6 to 8 weeks, whereas for *M. baccifera* the same is attained at almost 11th week. For *M. baccifera* estimated maximum growth rate $rk / e = 113.75$ cm per week. Whereas for *B. cacharensis* it varies from 132.82 cm to 141.83 cm per week. Finally the growth attained by *M. baccifera* at this point is estimated as $f = k / e =$ 624.974 cm and the same for *B. cacharensis* varies between 334.16 cm and 411.11 cm. Such comparisons are important in developing proper scientific management systems for optimum yield.

Nath et al. (2006) studied the traditional harvest regimes of three bamboo species, *Bambusa cacharensis, B. vulgaris* and *B. balcooa* in Barak Valley, Assam, with respect to their population structure. After studying the selective and clearfelling harvest regimes within each species, they suggested the need for developing management strategies for enhancing bamboo productivity through restricting the clear-felling system. Nath et al. (2007) studied population structure of dolu bamboo (*Schizostachyum dullooa*), a prioritized forest bamboo species, in a forest patch of Cachar district, Barak Valley, Assam, North-East India and observed that the unscientific management strategies being practiced exert greater pressure on the sustainable productivity of the stand. Management strategies for the restoration of bamboo stands are proposed mainly on the basis of age of culms. However, the present study reveals that the age of clumps influences the commercially most important parameter viz. the culm height (Table 2). It is worth mentioning here that further investigation is needed to confirm this fact by considering clumps of some more age groups between 15 and 40 years. Thus, while developing management strategies, age of the culm as well as the clump needs to be considered. However, detailed discussion on the same is beyond the scope of this work.

REFERENCES

Chatterjee, S. and Price, B. 1977. Regression Analysis by Example. John Wiley and Sons, New York. 228 pages.

Das, G.; Das, A.K. and Nandy, S. 2006. Nonlinear Statistical Model for Culm Growth of Muli Bamboo- *Melocanna baccifera*. International Journal of Ecology and Environmental Sciences 32: 221-225.

Draper, N.R. and Smith, H. 1981. Applied Regression Analysis. John Wiley and Sons, New York. 709 pages.

Gore, A.P. and Paranjpe, S.A. 2000. A Course on Mathematical and Statistical Ecology. Kluwer Academic Publishers, Dordrecht. 288 pages.

Hills, M. 1974. Statistics for Comparative Studies. Chapman and Hall, London. 194 pages.

Nandy, S.; Das, A.K. and Das, G. 2004. Phenology and culm growth of *Melocanna baccifera* (Roxb.) Kurz in Barak Valley, North-East India. Journal of Bamboo and Rattan 3: 27 - 34.

Nath, A.J.; Das, G. and Das, A.K. 2004. Phenology and Culm Growth of *Bambusa cacharensis* R. Majumdar in Barak Valley, Assam, North-East India. Bamboo Science & Culture (The Journal of American Bamboo Society) 18: 19-23.

Nath, A.J.; Das, G. and Das, A.K. 2006. Population structure and culm production of bamboos under traditional harvest regimes in Assam, Northeast India. Journal of Bamboo and Rattan 5(1&2): 91-100.

Nath, A.J.; Das, G. and Das, A.K. 2007. Culm characteristics and population structure of dolu bamboo (*Schizostachyum dullooa* (Gamble) Majumder) in Barak Valley, Northeast India, the need for conservation and implications for management. Bamboo Science & Culture (The Journal of American Bamboo Society) 20(1): 15-20.

Richards, F.J. 1959. A flexible growth function for empirical use. Journal of Experimental Botany 10: 290-300.

Seber, G.A.F. and Wild, C.J. 1989. Nonlinear Regression. John Wiley, New York. 768 pages.

Venugopalan, R. and Prajneshu. 1997. A generalized allometric model for determining length-weight relationship. Biometrical Journal 39: 733-739.

Received 18 August 2013;
Accepted 5 April 2014