Indian J. Fish., 60(2) : 67-71, 2013

Forecasting of fish production from ponds – a nonlinear model approach

N. OKENDRO SINGH, PREM KUMAR*, L. M. BHAR, K. N. SINGH AND PAL SINGH

Indian Agricultural Statistics Research Institute, New Delhi-110 012, India
**Directorate of Coldwater Fisheries Research, Bhimtal - 263 136, Nainital, Uttarakhand, India*
e-mail: okendro@iasri.res.in

ABSTRACT

Growth of fish is affected by many factors which are usually time dependent. The present study attempts to develop a forecasting model of fish production based on growth models. The concept of partial reparameterisation by expected-value parameters to mitigate nonlinear behaviour of parameters is highlighted. A reparameterised version of the Gompertz model is proposed for forecasting fish production, which is illustrated with the average growth data of grass carp rearing in polyponds of upland region. Two months ahead forecasting of grass carp production in the 11th and 12th months of rearing period is best predicted by the proposed model.

Keywords: Forecasting, Growth models, Nonlinear behaviour, Reparameterisation, Polyponds

Introduction

Timely and accurate forecasting of fish production from culture ponds will be of immense help for the farmers to plan for marketing their produce profitably. In the uplands of India, there is generally lack of market facility in the nearby locality and poor infrastructure is available for transportation of the fish produce from the farmers' pond. Hence most of the time, fish farmers compromise the net profit at the end. Till today, this important task has not been seriously taken into consideration. Walia *et al*. (1997) developed nonlinear models for forecasting fish production from cemented ponds based on classical approach. Nonlinear models are appropriate for developing forecasting models since fish weight data are usually collected over time. The above work was further extended by Bhar *et al*. (2002) in the natural ponds considering heteroscedastic error structure. However, the nonlinear models fitted to any type of growth data usually results in highly nonlinear behaviour. Many authors highlighted the importance of reparameterisation in nonlinear model fitting (Sarada and Prajneshu, 2005; Prajneshu, 2008; El-Shehawy, 2010; Ross *et al*., 2010). It apperas that no one so far highlighted the importance of close-to-linear behaviour of parameters in nonlinear growth models while they developed such type of forecasting models, although we come across such situations frequently. In fact, a little attention is given to the various reparameterisations and consequently, the parameter estimates hardly satisfy any of the optimum properties of good estimators. In view of the above, the present study aimed to develop the most appropriate forecasting model for fish production from polyponds of upland region. The methodology is illustrated with an example considering the growth data of grass carp rearing in polyponds of upland region during the period March 2009 to February 2010.

Materials and methods

Some of the most commonly used nonlinear growth models, where W_t represents average fish weight at time t, are given below:

Logistic model: $$W_t = \frac{\beta_1}{1+\beta_2 \exp(-\beta_3 t)} \quad (1)$$

Gompertz model: $$W_t = \beta_1 \exp[-\beta_2 \exp(-\beta_3 t)] \quad (2)$$

Von-Bertalanffy model: $$W_t = \frac{\beta_1}{[1-\beta_2 \exp(-\beta_3 t)]^3} \quad (3)$$

where β_1, β_2 and β_3 are the parameters to be estimated. The parameter β_1 represents the limiting growth value or asymptotic size, β_2 the scaling parameter and β_1, the rate of maturity.

If we assume that 'β_1' is an offensive parameter say, in terms of nonlinear behaviour in equation (2), then it can be partially reparameterised by expected-value parameter to get a possible solution. To obtain an expected-value parameter from above equation (2), we need to choose value t_1 of the regressor variable t, within the observed range t. Then, we get the expected value from equation (2) as follows:

$$W_1 = \beta_1 \exp[-\beta_2 \exp(-\beta_3 t_1)]$$

Solving this equation for the parameter 'β_1' we get,

$$\beta_1 = \frac{W_1}{\exp[-\beta_2 \exp(-\beta_3 t_1)]}$$

Substituting back into the original equation (2), we get,

$$W_t = W_1 \frac{\exp[-\beta_2 \exp(-\beta_3 t)]}{\exp[-\beta_2 \exp(-\beta_3 t_1)]} \quad (4)$$

The above model can be used to eliminate the nonlinear behaviour of parameters. Here, the likely offensive parameter 'β_1' is reparameterised by expected-value parameter while the other parameters are not changed. The above model given by equation (4) is referred to as 'Gompertz-II' in the subsequent discussions.

The model performance is generally evaluated based on summary statistics like root mean square error (RMSE) and mean absolute error (MAE):

$$RMSE = \left[\sum_{t=1}^{n} (W_t - \hat{W}_t)^2 / n\right]^{1/2};$$

$$MAE = \sum_{t=1}^{n} \left|(W_t - \hat{W}_t)\right| / n,$$

where,

$\hat{W}_t$ = Predicted fish weight of t^{th} observation;

$\overline{W}$ = Average fish weight;

n = Number of observations, $t = 1,2,...,n$.

The better model will have the least values of these statistics. It is, further, recommended for residual analysis to check the model assumptions such as independence or the randomness assumption of the residuals and the normality assumption. To test the independence assumption of residuals, run test procedure is available in the literature (Ratkowsky, 1990). Further, Shapiro-Wilk's test was applied to check the normality assumption.

Further, Hougaard's measure of skewness g_t, can be employed to assess whether a parameter is close to linear or whether it contains considerable nonlinearity. Hougaard's measure is computed as follows:

$$E\left[\hat{\beta}_t - E(\hat{\beta}_t)\right]^3 = -(MSE)^2 \sum_{jkl}^{p} L_{jk} L_{kl} L_{jl} (W_{jkl} + W_{kjl} + W_{lkj})$$

where the sum is a triple sum over the number of parameters,

$L = [J'J]^{-1}$,

$$W_{jkl} = \sum_{m=1}^{n} J_{mj} H_{mkl},$$

J is the Jacobian matrix, J_m is the Jacobian vector, H is the Hessian matrix, H_m is its component evaluated at observation m and β_τ is the t^{th} parameter. This third moment is normalised using the standard error to give Hougaard's measure of skewness as:

$$g_t = E\left[\hat{\beta}_t - E(\hat{\beta}_t)\right]^3 \Big/ (MSE * L_{tt})^{3/2}$$

According to Ratkowsky (1990), if $|g_2|<0.1$, the estimator $\hat{\beta}_t$ of parameter β_t is very close-to-linear in behaviour and, if $0.1<|g1_t|<0.25$, the estimator is reasonably close-to-linear. If $|g1_t|>0.25$, the skewness is very apparent. For $|g_t|>1$, the nonlinear behaviour is considerable.

Also, the bias of Box reveals which parameters are responsible for the nonlinear behaviour. The bias of Box is calculated in multivariate form as given by Cook *et al.* (1986):

$Bias = (D_2'D_2)^{-1} (D_2'H_2)$,

where D_2 is the nxp first derivative matrix; $H_2 = -1/2\sigma^2 trac\{(D_2'D_2)^{-1} F_{2i}\}$, the expected difference between the quadratic and linear components of the Taylor approximation and F_{2t}, t=1,2,…,p are pxp faces of the second derivative matrix, p is the number of parameters involved in the model.

$$\%Bias = \frac{Bias}{\hat{\beta}_t} \times 100$$

Here, $\hat{\beta}_t$ is the estimated value of β_t. Ratkowsky (1983) suggested using an absolute value of greater than 1% as an indicator of nonlinear behaviour.

Moreover, the curvature in a nonlinear model consists of two components: the intrinsic (IN) curvature and parameter effects (PE) curvature. Details of the root mean square (RMS) IN and PE measures of curvature and curvature critical value are given in Bates and Watts (1980, 1998). According to Ratkowsky (1983), the IN curvature is typically smaller than the PE curvature, which can be affected by altering the parameterisation of the model. Severe curvature effects are indicated by values of IN and PE exceeding the critical value *i.e.*, $1/\sqrt{F_{p,(n-p)}(0.05)}$. Usually, PE is computed when IN is within permissible limits and a lower value of PE suggests that the model exhibits close-to-linear behaviour (Ratkowsky, 1990).

Results and discussion

The average growth data of grass carp was considered for illustration and the basic statistics of the dataset is presented in Table 1. The dataset was generated from the NAIP (National Agricultural Innovative Project) Component-III entitled "Enhancement of livelihood security through sustainable farming systems and related farm enterprises in North-West Himalaya" conducted at Directorate of Coldwater Fisheries Research, Bhimtal, Uttarakhand. Polyponds were created for conducting experiments on integrated fish culture under NAIP by selecting three clusters of villages in Champawat District of Kumaun region, Uttarakhand. Polyponds are usually made by spreading a layer of low density polyethylene (LDPE) of 200 gsm (g m^{-2}) in the cemented tanks or earthen ponds. The average size and average volume of water for polyponds were 60 m^2 and 100 m^3 respectively. One polypond was selected from each cluster for this experiment. The three different species of exotic carps *viz*., silver carp grass carp and common carp were reared in polyculture systems and the species composition was in the pattern : silver carps - 30%, grass carp - 30% and common carp - 40%. The stocking density used was 3 fingerlings per cubic meter of water and thus on an average 300 fish in aggregate (90 numbers of grass carp) were reared in each polypond which was replicated in three polyponds. Further, a high rate of fish mortality ranging from 20 to 30% per pond was reported upto the end of the 10^{th} month during the rearing period. The growth data of a sample of size 30 per polypond comprising of 10 specimens from each fish group was randomly selected and data in terms of length and weight of fish was regularly observed for every month during March 2009 to February 2010. The average weight of grass carp obtained from 30 (10 individuals per pond) observations for each month was utilised for present study and thus there were 12 average data points. The first ten data points were used for developing the forecasting model and the rest two points were kept for model validation purposes. The SAS 9.3 version was used for various analyses.

The above nonlinear models were fitted to the available growth data of grass carp. von Bertalanffy model failed to give optimum solution however, the summary statistics for fitting of other models are presented in Table 2. Gompertz model performed well based on the criteria of having least values of RMSE and MAE. Also, randomness assumption follows since the run test $|Z|$ value of 1.01 is below the critical value of 1.96 at 5% level of significance. Shapiro-Wilk's test p-value of 0.10 indicated that the normality assumption follows.

Table 1. Basic statistics of the dataset used

Variable	n	Minimum	Maximum	Mean	Std. Deviation
Weight of fish (in g)	12	8.61	514.15	232.98	171.27

Table 2. Summary statistics of fitted models

	Logistic	Gompertz	Gompertz-II
Parameter estimation			
b_1 (or, W_1)	436.90 (26.38)	545.20 (32.61)	351.90 (3.47)
b_2	19.65 (3.51)	3.96 (0.18)	3.96 (0.18)
b_3	0.56 (0.05)	0.28 (0.02)	0.28 (0.02)
Hougaard's Skewness and Box's % Bias			
b_1 (or, W_1)	0.68 & 0.68	0.54 & 0.50	$0.0^{3}6$ & 0.01
b_2	0.92 & 3.49	0.46 & 0.44	0.46 & 0.44
b_3	0.22 & 0.54	0.10 & 0.20	0.10 & 0.20
Curvature effects of Bates and Watts			
RMS IN Curvature	0.06	0.04	0.04
RMS PE Curvature	0.66	0.85	0.22
Critical Value	0.48	0.48	0.48
Goodness of fit			
RMSE	12.42	6.40	6.40
MAE	9.78	4.94	4.94
Residual analysis			
Run test Z Value	0.91	1.01	1.01
Shapiro-Wilk's Test p-value	0.09	0.10	0.10

As Hougaard's skewness values are less than unity and Box's % bias are also less than 1%, we can say that parameters do not show any extreme nonlinear behaviour. However, RMS IN curvature (0.04) of Bates and Watts is less than the corresponding critical value 0.48. However, RMS PE curvature (0.85) is greater than the corresponding critical value which indicates that at least one parameter is showing nonlinear behaviour. To rectify the above problem, a partially reparameterised Gompertz model, given in equation (4) was used in which the parameter β_1 is considered to be an offensive parameter as it provides the maximum values of Hougaard's skewness and Box's% bias. Here, the parameter β_1 is replaced by W_1 in the process of reparameterisation as β_1 is considered to be the offensive parameter. A value of t_1=8 was chosen and the corresponding value of W_1=353.45 was taken as an initial value for computation of the final estimate of the parameter W_1, which gives the best result in terms of least magnitude of RMS PE curvature. The reparameterised model was refitted to the data and the results are again presented in Table 2. The RMS PE curvature value of 0.22 is now well below the corresponding critical value and it is acceptable. Further improvements in Hougaard's skewness and Box's% bias are also seen in this refitted model of Gompertz-II. The graph of fitted model along with observed fish weight is also depicted in Fig. 1, which shows the appropriateness of the proposed model. If there is no fish mortality during the rearing period, the grass carp production in the 11th and 12th months are best forecasted by the proposed Gompertz-II model as 38.12 kg and 40.51 kg respectively (Table 3a). Assuming 20% and 30% fish mortality in each pond upto the end of the 10th month, the forecasting of grass carp production for the 11th and 12th month are given in the Table 3b and 3c respectively. The modified version of the Gompertz model outperformed irrespective of the situations.

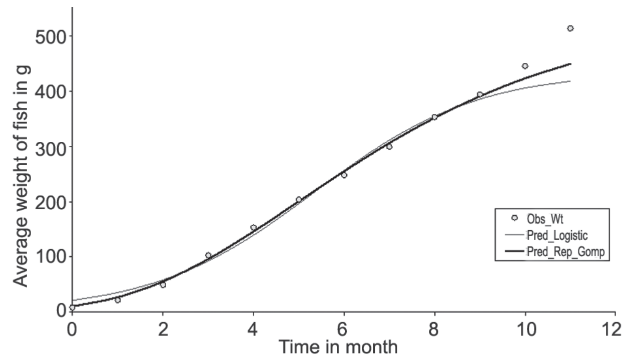

Fig. 1. Actual and predicted grass carp production obtained using different models

References

Bates, D. M. and Watts, D. G. 1980. Relative curvature measures of nonlinearity. *J. Roy. Stat. Soc.,* B*,* 42(1): 1-25.

Bates, D. M. and Watts, D. G. 1998. *Nonlinear regression analysis and its applications*. John Wiley and Sons, New York.

Table 3a. Actual and predicted production of grass carp (weight in kg) from polyponds (if there is no fish mortality upto the end of the 10th month)

Month	Actual	Logistic	Gompertz	Gompertz-II
11th	40.14	36.54	38.12	38.12
		(3.60)	(2.02)	(2.02)
12th	46.27	37.68	40.51	40.51
		(8.59)	(6.76)	(6.76)

Table 3b. Actual and predicted production of grass carp (weight in kg) from polyponds (if there is 20% fish mortality per pond upto the end of 10th month)

Month	Actual	Logistic	Gompertz	Gompertz-II
11th	32.11	29.23	30.49	30.49
		(2.88)	(1.62)	(1.62)
12th	37.02	30.14	32.40	32.40
		(6.88)	(4.62)	(4.62)

Table 3c. Actual and predicted production of grass carp (weight in kg) from polyponds (if there is 30% fish mortality per pond upto the end of 10th month)

Month	Actual	Logistic-I	Gompertz	Gompertz-II
11th	28.09	25.58	26.68	26.68
		(2.51)	(1.41)	(1.41)
12th	32.39	26.37	28.35	28.35
		(6.02)	(4.04)	(4.04)

The values in parentheses are the corresponding forecasting errors

Bhar, L. M., Walia, S. S. and Roy, A. K. 2002. Forecasting fish production from ponds. *Project Rep*ort, IASRI, New Delhi.

Cook, A. D., Tsai, C. L. and Wei, B. C. 1986. Bias in nonlinear regression. *Biometrika*, 3: 615-623.

El-Shehawy, S. A. 2010. On the selection of models in nonlinear regression. *Asian J. Mathemat. Stat.*, 3(4): 254-266.

Prajneshu 2008. Fitting of Cobb-Douglas production functions: revisited. *Agr. Econ. Res. Rev.*, 21: 289-292.

Ratkowsky, D. A. 1983. *Nonlinear regression modelling - A unified practical approach*. New York, Marcel Dekker.

Ratkowsky, D. A. 1990. *Handbook of nonlinear regression models*. New York, Marcel Dekker.

Ross, J. S. G., Prajneshu and Sarada, C. 2010. Reparameterisation of nonlinear statistical models: a case study. *J. Appl. Stat.*, 37(12): 2015-26.

Sarada, C. and Prajneshu 2005. On appropriate reparameterisation of a nonlinear statistical model. *J. Ind. Soc. Ag. Stat.,* 59(3): 237-242.

Walia, S. S., Jain, R. C. and Sehdev, R. S. 1997. Nonlinear statistical models for pre-harvest forecasting inland fish from ponds. *Project Report*, IASRI, New Delhi.

Date of Receipt : 02.01.2012
Date of Acceptance : 16.01.2013