Indian J. Fish., 65(3): 95-98, 2018
DOI: 10.21077/ijf.2018.65.3.80950-12

Note

Gompertz Stochastic Differential Equation growth model with discrete delay: application in forecasting all India marine and inland fish production

BISHAL GURUNG AND PRAJNESHU
ICAR-Indian Agricultural Statistics Research Institute, New Delhi - 110 012, India
e-mail: prajneshug@gmail.com

ABSTRACT

The present study considered the well-known Gompertz growth model. The effect of random environmental fluctuations in the parameters is modelled employing Stochastic Differential Equation (SDE) approach. Importance of incorporating discrete delay in this model is highlighted. The methodology for fitting the resultant model in terms of Stochastic Delay Differential Equation (SDDE) is described. Relevant computer program for fitting Gompertz SDDE (GSDDE) model is written in SAS package. Finally, as an illustration, superiority of GSDDE model over Gompertz SDE (GSDE) model for fitting and forecasting purposes is shown for all India marine and inland catch data sets.

Keywords: All India marine and inland fish catch data, Gompertz growth model, SAS software, Stochastic Delay Differential Equation

Nonlinear growth models such as Gompertz, Logistic and von Bertalanffy models are quite often employed for modelling and forecasting in fisheries (Prajneshu, 1991; Venugopalan and Prajneshu, 1996; Menon *et al.*, 2015). A heartening aspect of these models is that they are mechanistic in nature and so the underlying parameters have specific biological interpretations. For example, Gompertz model has three parameters, *viz.,* intrinsic growth rate, carrying capacity and initial population/biomass size. Growth models are generally expressed in terms of nonlinear differential equations. An attractive feature of these models is that they can be converted to linear forms by means of some transformations, like logarithmic and reciprocal transormations. Consequently, exact solutions of the underlying differential equations can be obtained, which are nonlinear in parameters. Usual practice for applying them to data is to add an additive term, with suitable assumptions, on the right hand side of the deterministic solution and applying nonlinear approximation methods, such as Levenberg-Marquardt procedure for estimation of the parameters. Almost all the standard statistical software packages, such as SAS or SPSS contain programs for fitting nonlinear models by employing nonlinear estimation procedures.

Although the above methodology has served many purposes in the past, it suffers from two main limitations. The first one is that it is applicable only when the data are equidistant. However, collection of growth data over time involves constraints of time, personnel and budget, that do not always satisfy this requirement. Undoubtedly, the data that do exist in studies with missing data or data at unequal time-intervals are potentially informative, and precluding such data from analysis could affect conclusions adversely (Dennis and Ponciano, 2014). The other limitation is that, simply by adding an error term, a nonlinear statistical model is not capable of describing underlying fluctuations of the system satisfactorily, particularly for longitudinal data. Both the above issues can, however, be tackled by employing the more general approach of Stochastic Differential Equations (SDE). These are generally obtained by adding a stochastic term to the differential equation form of the deterministic model. Further, in a physical situation, random environmental fluctuations due to variations in the parameters generally occur with great rapidity as compared to the time-scale of growth. Therefore, the stochastic term is generally assumed to be a Gaussian white noise stochastic process. To this end, two types of stochastic calculi due respectively to Stratonovich and Ito have been developed in the literature. However, for the present study, both these calculi yield identical results as we are concerned here only with the case of additive noise, which is independent of state variable. A good description of the above methodology is available, for example, in Venugopalan and Prajneshu (1999), Prajneshu and Ghosh (2017) and Prajneshu *et al.* (2017).

In our future discussions, we shall confine our attention to only the Gompertz growth model. The aim of the present investigation was to study a generalisation of Gompertz Stochastic Differential Equation (GSDE) growth model by incorporating the aspect of discrete delay.

The differential equation form of Gompertz growth model (Seber and Wild, 2003) is given by:

$$dy(t)/dt = \alpha y(t)-\beta y(t) \text{ In } y(t) \quad (1)$$

where, y(t) is the variable of interest at time t and α, β and y(0) are the parameters of the model. Incorporating the effect of random environmental fluctuations, Gompertz SDE model becomes:

$$dy(t) = \{\alpha y(t) - \beta y(t) \text{ In } y(t)\} \, dt + \sigma \, y(t) \ln y(t) \, dW(t), \quad (2)$$

where, $dW(t)$ is a Gaussian white noise stochastic process.

One drawback of the above model is that the rate of change of variable is assumed to depend only on its current value. However, in the real world, growth rate of a population often does not respond immediately to changes in its own population or that of an interacting population but rather would do so after some time lag. Therefore, for a more realistic modelling, growth rate should depend on the previous history. One reason for including a time lag in a model is that of age structure. The presence of a maturation period or a gestation period has the effect that the growth rate of a population to increased population densities is delayed. Some other factors for the delay effects can be replenishment or regeneration time for resources. Another mechanism which introduces a time lag is that the actual process has a spatial structure so that different stages in the process take place in different regions. At least one aspect of spatial structure, *i.e.* the finite time for an intermediate component to cross from region to region, can be modelled by introducing a time lag. A good description of time lags, for example, is available in Allen (2014). Incorporating this aspect of discrete lag in eq. (2), we get:

$$dy(t) = \{\alpha y(t)-\beta y(t-\tau) \text{ In } y(t)\} \, dt + \sigma \, y(t) \ln y(t) \, dW(t) \quad (3)$$

where, τ denotes the discrete lag. Note that eq. (3) is in terms of a Stochastic Delay Differential Equation. Further, eq. (3) reduces to eq. (2) when τ = 0.

As it is not possible to solve eq. (3) exactly, the next best alternative is to obtain its approximate analytic solution by employing iterative procedures. Four main methods of this kind are: Linearisation (or Taylor Series) method, Steepest Descent method, Gauss-Newton method and Levenberg-Marquardt (L-M) method. Among them, the L-M method is the most powerful and widely employed. Therefore, in this study, we have employed it to solve eq. (3) approximately.

As an illustration, all India marine and inland annual fish production data from 1980-81 to 2015-16, collected from Agricultural Statistics at a Glance (2016), Directorate of Economics and Statistics, Ministry of Agriculture and Farmers' Welfare, were considered for analysis. The data for the first 30 years, *i.e.,* from 1980-81 to 2009-10 were employed for fitting of models, while data from 2010-11 to 2015-16 were employed for validation of models.

In the first instance, Gompertz Stochastic Delay Differential Equation (GSDDE) model given by eq. (3) was fitted to the two data sets for τ = 0, 1, 2, ..., 5 and the results are reported in Table 1. A perusal of this table shows that, for both the data sets, GSDDE model with τ = 2 is the best as it has minimum Root Mean Square Error (RMSE) value. In the second instance, performance of fitted models for validation purpose was examined and the results are reported in Table 2. A perusal again shows that for both the data sets, GSDDE model with τ = 2 has performed the best as it has got minimum Root Mean

Table 1. Fitting of Gompertz SDE growth model with discrete delay

(a) All India marine fish production ('000 t)

Alpha	Beta	Tau	RMSE
0.002990	0.0000	0	131.80
-0.000040	-0.1275	1	99.07
-0.000020	-0.0637	2	93.19
0.000261	-0.0381	3	118.36
0.000085	-0.0330	4	118.42
-0.000200	-0.0283	5	120.65

(b) All India inland fish production ('000 t)

Alpha	Beta	Tau	RMSE
0.097200	0.0900	0	98.42
-0.000260	-0.1236	1	84.44
0.000339	-0.0555	2	82.26
0.000835	-0.0331	3	94.43
0.000496	-0.0249	4	97.93
0.001030	-0.0178	5	101.72

Table 2. Validation of fitted Gompertz SDE growth model with discrete delay

(a) All India marine fish production ('000 t)

Tau	RMSPE	RAMPE
0	86.30	8.73
1	87.33	9.34
2	71.06	8.27
3	119.90	10.70
4	85.95	8.66
5	115.57	10.12

(b) All India inland fish production ('000 t)

Tau	RMSPE	RAMPE
0	132.03	11.01
1	73.85	8.58
2	57.75	6.73
3	83.67	7.59
4	128.40	10.35
5	137.13	11.21

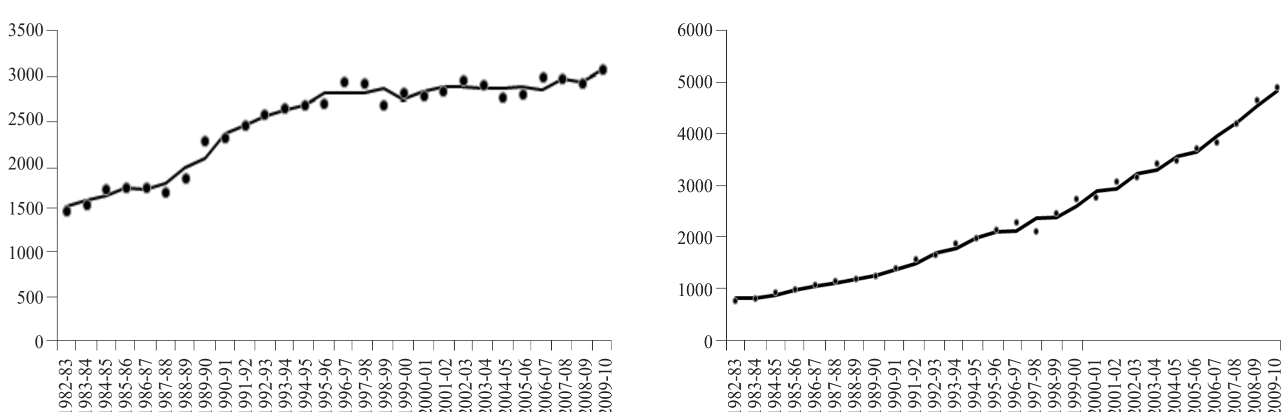
Square Prediction Error (RMSPE) and Root Absolute Mean Prediction Error (RAMPE) values. Subsequently, using estimates of parameters for this model (Table 1), fitted values for various years were computed and the same are reported in Table 3. A perusal of this table shows that, for both the data sets, fitted as well as predicted values throughout are quite close to actual values. In order to get a visual analogy, fitted and predicted values for both the data sets are exhibited in Fig. 1. Subsequently, forecasting for 6 years in respect of both the data sets for GSDDE model with $\tau = 2$ are carried out and the results are reported in Table 4. Evidently, for both the data sets, the forecast values are found to be quite close to actual values, implying thereby that the GSDDE model has performed quite well for validation purposes also. Finally, the out-of-sample forecasts (in '000 t) for 2016-17 and 2017-18, using the identified GSDDE model, for (a) all India marine fish production are respectively computed as 7780.6922 and 8209.9744 and (b) all India inland fish production are respectively computed as 3827.2671 and 4027.5944.

The purpose of this study was to incorporate the aspect of discrete time lag in Gompertz Stochastic Differential Equation (GSDE) model. As an illustration, all India annual marine fish and inland fish data sets were considered. Superiority of the model with discrete time lag over the one without discrete time lag was shown for the considered data sets. It may not be out of place to mention here that several other statistical modelling approaches have been employed by researchers in the past for forecasting such data sets. Sathianandan (2007) employed vector time-series modelling approach, while Ravichandran and Prajneshu (2001, 2005) used State space and Structural time-series modelling approaches, respectively. In this article, the potential of an alternative modelling technique, *viz.,* Gompertz Stochastic Differential Equation model

Table 3. Comparison of fitted values with actual values ('000 t) for the two data sets

Year	Marine fish production		Inland fish production	
	Actual	Fitted	Actual	Fitted
1982-83	1427.00	1479.00	940.00	996.62
1983-84	1519.00	1553.21	987.00	1004.66
1984-85	1698.00	1612.01	1103.00	1056.86
1985-86	1716.00	1704.87	1160.00	1155.10
1986-87	1713.00	1688.25	1229.00	1218.61
1987-88	1658.00	1761.86	1301.00	1274.33
1988-89	1817.00	1952.06	1335.00	1344.58
1989-90	2275.00	2054.24	1402.00	1419.34
1990-91	2299.65	2359.58	1536.25	1530.92
1991-92	2447.28	2436.72	1709.33	1644.96
1992-93	2576.25	2547.71	1789.05	1832.78
1993-94	2648.84	2633.85	1995.50	1923.91
1994-95	2691.81	2677.72	2096.76	2105.66
1995-96	2707.06	2829.81	2242.32	2224.47
1996-97	2966.81	2830.53	2381.43	2240.08
1997-98	2950.46	2828.81	2222.59	2464.97
1998-99	2696.46	2891.39	2563.24	2490.73
1999-00	2833.85	2754.01	2822.70	2695.41
2000-01	2811.00	2846.65	2845.83	2964.28
2001-02	2860.00	2901.28	3126.16	3016.89
2002-03	2990.00	2901.33	3209.86	3283.54
2003-04	2942.00	2882.23	3457.89	3362.27
2004-05	2779.00	2879.97	3525.88	3602.74
2005-06	2820.00	2900.50	3756.00	3682.04
2006-07	3020.00	2870.05	3844.89	3973.45
2007-08	2998.21	2998.21	4207.35	4223.92
2008-09	2954.71	2954.71	4637.90	4527.62
2009-10	3115.65	3115.66	4862.84	4813.60

with discrete lag was demonstrated. Research on various aspects of discrete lag may be incorporated in future for other growth models useful in fisheries, such as Logistic, Pella-Tomlinson and von Bertalanffy models. Finally,



(a) (b)

Fig. 1. Graph of fitted GSDDE model with $\tau = 2$ for (a) all India marine fish production, (b) all India inland fish production, along with actual values

Table 4. Forecasts for both the data sets using GSDDE model with $\tau = 2$

Year	Marine fish production		Inland fish production	
	Actual	Forecast	Actual	Forecast
2010-11	3249.00	3186.20	4981.25	5081.19
2011-12	3372.00	3285.72	5294.70	5347.86
2012-13	3321.00	3406.08	5719.56	5713.57
2013-14	3439.00	3478.62	6132.42	6154.56
2014-15	3626.00	3596.62	6581.12	6567.42
2015-16	3627.00	3783.62	7212.56	7016.12

effect of more general time lag, *viz.,* Distributed time delay may be investigated in various growth models.

Acknowledgements

The second author is grateful to Science and Engineering Research Board, New Delhi for providing financial assistance to carry out the above research work under Research Project No. SB/S4/MS/880/2014. Thanks are also due to the reviewers for valuable comments.

References

Allen, E. J. 2014. Derivation and computation of discrete-delay and continuous delay SDEs in mathematical biology. *Mathl. Biosci. Engng.*, 11: 403-25. DOI: 10.3934/mbe. 2014.11.403.

Dennis, B. and Ponciano, J. M. 2014. Density dependent state space model for population abundance data with unequal time intervals. *Ecology*, 95: 2069-76.

Menon, M., Maheswarudu, G., Rohit, P., Laxmilatha, P., Das, M. and Rao, K. N. 2015. Biology and stock assessment of the bigeye croaker *Pennahia anea* (Bloch, 1793) landed along Andhra Pradesh, north-east coast of India. *Indian J. Fish.*, 62(1): 46-51.

Prajneshu 1991. Cautionary note about non-linear models in fisheries. *Indian J. Fish.*, 38(4): 231-33.

Prajneshu and Ghosh, H. 2017. A new approach for fitting growth models in random environment. *Indian J. Ani. Sci.*, 87(12): 1531-35.

Prajneshu, Ghosh, H. and Pandey, N. N. 2017. Fitting of von Bertalanffy growth model: Stochastic differential equation approach. *Indian J. Fish.*, 64(3): 24-28. doi.org/10.21077/ijf.2017.64.3.71014-04.

Ravichandran, S. and Prajneshu 2001. State space modeling versus ARIMA time-series modeling. *J. Ind. Soc. Agric. Statist*., 54(1): 43-51.

Ravichandran, S. and Prajneshu 2005. On the modelling of seasonal fluctuations in fish landings using structural time-series approach. *Ind. J. Ani. Sci.*, 75(8): 982-84.

Sathianandan, T. V. 2007. Vector time series modeling of marine fish landings in Kerala. *J. Mar. Biol. Ass. India,* 49(2): 197-205.

Seber, G. A. F. and Wild, C. J. 2003. *Nonlinear regression.* John Wiley, New York.

Venugopalan, R. and Prajneshu 1996. Trend analysis in all-India marine products exports using statistical modelling techniques. *Indian J. Fish.*, 43(2): 107-13.

Venugopalan, R. and Prajneshu 1999. von Bertalanffy growth model in a random environment. *Can. J. Fish. Aquat. Sci.*, 56: 1026-30.

Date of Receipt : 25.06.2018
Date of Acceptance : 07.08.2018