Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 72(2) 2018 97–104

Gompertz Stochastic Differential Equation Growth Model with Exogenous Variables and Time-Dependent Diffusion

Himadri Ghosh and Prajneshu
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 27 February 2018; Revised 19 March 2018; Accepted 21 March 2018

SUMMARY

Gutierrez *et al.* (2005) studied Gompertz homogeneous diffusion process with exogenous variables that affect the trend. The purpose of present article is to modify this work in two directions. The first one is that the diffusion term is taken as time-dependent, which is more realistic. Another direction of modification is that the powerful technique of Wavelet analysis is employed to estimate the drift term, rather than the elementary piece-wise time-varying linear functions approach. The methodology, after incorporating the above two aspects is developed. Relevant computer programs for its application are written and the same are included as an Appendix. Finally, as an illustration, India's total foodgrain production time-series data dependent on rainfall, fertilizer, and pesticide time-series data as exogenous variables are considered and superiority of our proposed model is shown over the model proposed by Gutierrez *et al.* (2005) for given data.

Keywords: Gompertz nonlinear growth model, Stochastic differential equation, Time-dependent diffusion, Wavelet analysis.

1. INTRODUCTION

Gompertz nonlinear growth model is widely employed in various disciplines, such as agriculture, medicine, and industry. A heartening aspect of this model is that it is mechanistic in nature and so the underlying parameters have specific biological interpretations. This model is generally expressed in terms of a nonlinear differential equation, which can be converted to linear form by means of logarithmic transformation. Consequently, exact solution of the underlying differential equation can be obtained, which is nonlinear in parameters. Usual practice for applying the corresponding Gompertz nonlinear statistical model to data is to assume an additive term, with suitable assumptions, on the right hand side of the deterministic solution and to apply Nonlinear estimation procedures, such as Levenberg-Marquardt procedure (Seber and Wild 2003) for estimation of parameters.

Although the above methodology has served many useful purposes in the past, its main limitation is that, fitting a nonlinear deterministic model by simply adding an error term, is not capable of describing the underlying fluctuations of the system satisfactorily, particularly for longitudinal data (Seber and Wild 2003). Accordingly, Stochastic modelling approach based on 'Stochastic Differential Equation (SDE)' (See e.g. Cohen and Elliott 2015) has been employed. Several articles dealing with various aspects of Stochastic Gompertz model have so far been published (See e.g. Ferrante *et al.* 2000, Behera and O'Rourke 2008, Gutierrez *et al.* 2009, Skiadas 2010, and Ghosh and Prajneshu 2017). However, one drawback of all the above models is that they model the development of a process only over time. In a seminal paper, Gutierrez *et al.* (2005) developed the methodology for application of Stochastic Gompertz model by introducing the exogenous variables as functions of time that affect its trend. The purpose of the present article is to modify the work of Gutierrez *et al.* (2005) (Called Gutierrez model hereinafter) in two directions. In Gutierrez model, the diffusion term is time-independent, so for a more realistic modelling,

Corresponding author: Prajneshu
E-mail address: prajneshug@gmail.com

we relax this assumption and take the diffusion term as time-dependent. Another direction of modification is that, in Gutierrez model, the exogenous factors were used to approximate the drift term by piecewise time-varying linear functions, which is not realistic. So, the drift term would be estimated by applying the powerful technique of Wavelet analysis (Ghosh *et al.*, 2010). The resultant model hereinafter would be called 'Modified Gutierrez model'.

2. METHODOLOGY

Gutierrez *et al.* (2005) studied nonlinear Gompertz homogeneous diffusion process by introducing time functions that affect its trend. To this end, the stochastic system processes driven by almost everywhere continuous function is widely used. Among these processes, diffusions (strong Markovian process) have been widely considered, which is studied by characterizing drift and diffusion coefficients. The SDE as one of the characterizations of diffusion process is capable to describe a continuous stochastic process by appropriate limiting operation of changes in state and time of corresponding discrete parameter process. Let $\{X_G(t): t\epsilon[t_0,T]\}$ be the Gompertz one-dimensional diffusion process taking values on R^+ and with infinitesimal first order moment (drift) $a(x,t) = h(t)x - \beta x log x,$ and half of second order infinitesimal moment (diffusion) $b(x,t) = \sigma^2 x^2$, where $\sigma^2 > 0, \beta \epsilon R, h(t) = \alpha_0 + \sum_{i=1}^{q} \alpha_i g_i(t), \alpha_i \epsilon R$ and $g_i(t)$ are continuous functions in $[t_0, T]$. This process can be studied from the viewpoint of Kolmogorov forward and backward partial differential equations of transition probability density $f(y,\tau|x,t)$ satisfying

$$\frac{\partial f(y,\tau|x,t)}{\partial \tau} = -\frac{\partial a(y,\tau) f(y,\tau|x,t)}{\partial y} + \frac{\partial^2 b(y,\tau) f(y,\tau|x,t)}{\partial y^2}$$

$$\frac{\partial f(y,\tau|x,t)}{\partial t} + a(x,t)\frac{\partial f(y,\tau|x,t)}{\partial x} + b(x,t)\frac{\partial^2 f(y,\tau|x,t)}{\partial x^2} = 0.$$

Since, it is known that $a(x,t) = (\mu + \sigma^2)x$ and $b(x,t) = \sigma^2 x^2$ for geometric Brownian motion $X(t) = exp(W_t)$, where W_t is Brownian motion with drift μ and $b(x,t) = \sigma^2$ given by $dW_t = \mu dt + \sigma dB_t, B_t$ is zero-mean Brownian motion with scale parameter 2, therefore as $\beta \to 0$, the Gompertz homogeneous diffusion process leads to results of lognormal transition probability density $f(y,\tau|x,t)$ in the limit. Note that, non-homogeneous analogue of above type of lognormal diffusion process has been widely used in stochastic modelling in stochastic economics and environmental sciences (See *e.g.* Gutierrez *et al.* 2005). In this case, infinitesimal first order moment $a(x,t) = xh(t)$.

However, one limitation of the above methodology is that it is not capable to describe the data in certain time-intervals when there is presence of exponential trend. It may be noted that, in this case, the conditional mean of $X_G(t)$ is non-homogeneous due to time-varying exponential trend only in the Gaussian mean function of the logarithmic process, i.e. $X_G^{(\ln)}(t) = log X_G(t)$. Therefore, in this article, we propose to generalize the process $X_G(t)$ to the process $X_{G,d}(t)$, say, by taking into account the effect of time trend in the diffusion, i.e. $b(x,t)$ is now of the form $b(x,t) = g^2(t)x^2$, where $g(t) = \exp(K + \gamma t)$.

In an article that is apparently little known in the western world, Cherkasov has proved an interesting theorem characterizing the class of one-dimensional diffusion process (See e.g. Ricciardi 1976). The process can be derived from Wiener process provided it satisfies Kolmogorov partial differential equation. Following Cherkasov's notation, denote the following transformations for change of state and time by

$$x' = \psi(x,t) \text{ and } t' = \phi(t), \qquad (1)$$

so that it is capable to transform Kolmogorov backward partial differential equation into the differential equation of process with $b'(x,t) = 1$ and $a'(x,t) = 0$. In other words, transition probability density $f'(y',\tau'|x',t')$ of the transformed process $X'(t')$ satisfies the equation

$$\frac{\partial f'(y',\tau'|x',t')}{\partial t'} + \frac{\partial^2 f'(y',\tau'|x',t')}{\partial x'^2} = 0. \qquad (2)$$

To this end, transition probability density $f(y,\tau|x,t)$ of $X_G(t)$ is obtained by

$$f(y,\tau|x,t) = \frac{\partial \psi(y,\tau)}{\partial y} f'(y',\tau'|x',t'). \qquad (3)$$

It is important to get partial differential with respect to state and time for $f(y,\tau|x,t)$ in terms of $f'(y',\tau'|x',t')$ in above eq.(3) satisfying eq. (2), so that both transformations given by eq.(1) are related by the identity

$$\psi(x,t) = [\phi'(t)]^{1/2} \int_z^x \frac{dy}{[b(y,t)]^{1/2}} + \omega(t), \qquad (4)$$

where $\omega(t)$ is a function of the time variable alone and z belongs to state space. Eq. (4) can now be used to establish necessary and sufficient condition for Wiener process transformation (Ricciardi 1976), which is essentially the relation between $a(x,t)$ and $b(x,t)$ given by

$$a(x,t) = b'_x(x,t) + \frac{[b(t,x)]^{1/2}}{2}\Big\{c_1(t) + \int_z^x \left(\frac{c_2(t)b(y,t)+b_t'(y,t)}{[b(y,t)]^{3/2}}\right)dy\Big\}. \quad (5)$$

Using eqs. (3) and (5), Gutierrez *et al.* (2005) obtained $f(y,\tau|x,t)$, which is lognormal. However, it may be pointed out that there are typographical errors in the expressions for mean, variance, and $\psi(x,t)$ of the process $X_G^{(\ln)}(t)$. The correct formulae are given below:

$$E\left[X_G^{(\ln)}(\tau)|\mathcal{F}_t\right] = \Big\{exp(-\beta(\tau-t))\log x - \frac{\sigma^2}{\beta}\left(1-exp(-\beta(\tau-t))\right) + \int_t^\tau h(s)\,exp(-\beta(\tau-s))ds\Big\} \quad (6a)$$

$$V\left[X_G^{(\ln)}(\tau)|\mathcal{F}_t\right] = \frac{\sigma^2\left(1-exp(-2\beta(\tau-t))\right)}{\beta} \quad (6b)$$

$$\psi(x,t) = k_1^{1/2}\left(exp(\beta(t-t_0))\log(x/z) - \int_{t_2}^t h(\tau)exp(\beta(\tau-t_0))d\tau\right) + \frac{k_1^{1/2}}{\beta\sigma}(\sigma^2+\beta\log z)\left(exp(\beta(t-t_0)) - exp(\beta(t_2-t_0))\right) + k_3, \quad (6c)$$

where k_1, k_2 and k_3 are suitable constants for obtaining transition probability density function.

Further, for generalizing the nonlinear Gompertz SDE model with time-dependent diffusion $b(x,t) = \exp\{2(K+\gamma t)\}x^2$ for the process $X_{G,d}(t)$, following along similar lines as Gutierrez *et al.* (2005), it is found that $f(y,\tau|x,t)$ is lognormal with mean of the process $X_{G,d}^{(\ln)}(t)$ given by

$$E\left[X_{G,d}^{(\ln)}(\tau)|\mathcal{F}_t\right] = \Big\{exp(-\beta(\tau-t))\log x - \frac{exp(2K)}{(\beta+2\gamma)}\big(exp(2\gamma\tau) - exp(-\beta(\tau-t)+2\gamma t)\big) + \int_t^\tau h(s)\,exp(-\beta(\tau-s))ds\Big\},$$

$$V\left[X_{G,d}^{(\ln)}(\tau)|\mathcal{F}_t\right] = \frac{\left(1-exp(-2(\beta+\gamma)(\tau-t))\right)exp(2(\gamma\tau+K))}{(\beta+\gamma)} \quad (7)$$

Note that, the conditional variance of $X_{G,d}^{(\ln)}(\tau)$ is non-homogeneous, therefore it is capable to describe exponential trend by specifying both the time varying drift and diffusion coefficients. Using eq.(5), we obtain

$$c_1(t) = (2/g(t))(h(t) - g^2(t) - \beta\log z),$$
$$c_2(t) = -2\big(\beta + (g_t'/g_t)\big),$$

and using eq.(7), we get

$$\phi(t) = \frac{k_1}{2(\beta+\gamma)}\Big(exp\big(2(\beta+\gamma)(t-t_0)\big) - exp\big(2(\beta+\gamma)(t_1-t_0)\big)\Big) + k_3, \quad (8a)$$

$$\psi(x,t) = \frac{k_1^{1/2}}{exp(K)}\Big(\big(exp((\beta+\gamma)(t-t_0))\log(x/z)/g(t)\big) - exp(-\gamma t_0)\int_{t_2}^t h(\tau)exp(\beta(\tau-t_0))d\tau\Big) + k_1^{1/2}exp(-(\beta+\gamma)t_0)\Big(\big((\log z)(exp(\beta t) - exp(\beta t_2))/exp(K)\big) + \big(exp(K)\big(exp((\beta+2\gamma)t) - exp((\beta+2\gamma)t_2)\big)/(\beta+2\gamma)\big)\Big), \quad (8b)$$

$$\frac{\partial\psi(y,\tau)}{\partial y} = \frac{k_1^{1/2}exp(\beta t)}{y\,exp(K)}. \quad (8c)$$

We now discuss the method for fitting our proposed model. Let $x_1, x_2, \ldots, x_n$ be the observed values. Now we transform these values by means of $v_1 = x_1$ and $v_{i,\beta} = \log x_i - exp(-\beta)\log x_{i-1}$. Therefore, the likelihood function for the transformed sample is

$$L_{v_{2,\beta},v_{3,\beta},\ldots,v_{n,\beta}} = \left(\frac{\beta}{2\pi\sigma^2(1-exp(-2\beta))}\right)^{\left(\frac{n-1}{2}\right)} exp\left(-\frac{\beta\left(v_\beta-\gamma_\beta U_\beta'\alpha\right)'\left(v_\beta-\gamma_\beta U_\beta'\alpha\right)}{2\sigma^2(1-exp(-2\beta))}\right), \quad (9)$$

where $\boldsymbol{v}_\beta = (v_{2,\beta}, v_{3,\beta}, \ldots, v_{n,\beta})'$, $\boldsymbol{\alpha} = (\alpha_0-\sigma^2, \alpha_1, \ldots, \alpha_q)'$, $\gamma_\beta = ((1-exp(-\beta))/exp(-\beta))$, and $U_\beta = (\boldsymbol{u}_{2,\beta}, \boldsymbol{u}_{3,\beta}, \ldots, \boldsymbol{u}_{n,\beta})$ is the $(q+1)\times(n-1)$ matrix with

$$\boldsymbol{u}_{i,\beta} = \left(1, \frac{1}{\gamma_\beta}\int_{t_{i-1}}^{t_i} g_1(\tau)exp\,(-\beta(t_i-\tau))d\tau, \ldots, \frac{1}{\gamma_\beta}\int_{t_{i-1}}^{t_i} g_1(\tau)exp\,(-\beta(t_i-\tau))d\tau\right)'.$$

Following along similar lines and using eq. (7), the likelihood function of the process $X_{G,d}(t)$ is given by

$$L_{d;v_{2,\beta},v_{3,\beta},\ldots,v_{n,\beta}} = \left(\frac{(\beta+\gamma)}{2\pi\left(1-exp(-2(\beta+\gamma)(\tau-t))\right)exp(2(\gamma\tau+K))}\right)^{\left(\frac{n-1}{2}\right)} exp\left(-\frac{(\beta+\gamma)\left(v_\beta-U_\beta'\alpha\right)'\left(v_\beta-U_\beta'\alpha\right)}{2\left(1-exp(-2(\beta+\gamma)(\tau-t))\right)exp(2(\gamma\tau+K))}\right), \quad (10)$$

where

$\boldsymbol{v}_\beta = (v_{2,\beta}, v_{3,\beta}, \dots, v_{n,\beta})'$,

$$\boldsymbol{a} = \Big(\alpha_0\gamma_\beta - \frac{exp(2K)}{(\beta+2\gamma)}\big(exp(2\gamma\tau) - exp(-\beta(\tau - t) + 2\gamma t)\big), \alpha_1, \dots, \alpha_q\Big)',$$

$\gamma_\beta = ((1 - exp(-\beta))/exp(-\beta))$, and

$U_\beta = (\boldsymbol{u}_{2,\beta}, \boldsymbol{u}_{3,\beta}, \dots, \boldsymbol{u}_{n,\beta})$ is the $(q+1)\times(n-1)$ matrix with

$$\boldsymbol{u}_{i,\beta} = \Big(1, \int_{t_{i-1}}^{t_i} g_1(\tau) exp\,(-\beta(t_i - \tau))d\tau, \dots, \int_{t_{i-1}}^{t_i} g_1(\tau) exp\,(-\beta(t_i - \tau))d\tau\Big)'.$$

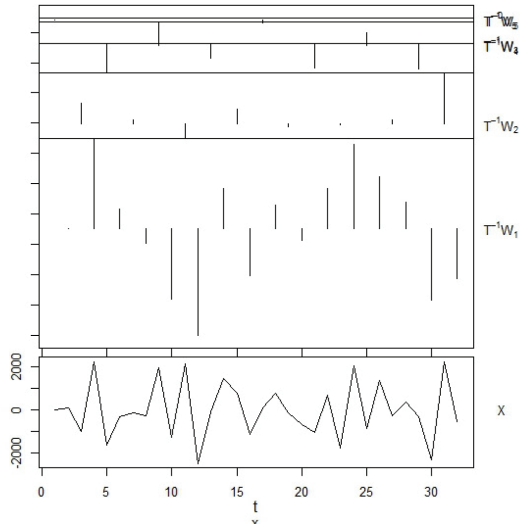
In line with Gutierrez *et al.* (2005), we also assume for simplicity, that the length of the time-intervals $[t_{i-1}, t_i]$ is equal to one. This assumption may not be too restrictive because quite often time-series data are available at equal intervals. Gutierrez *et al.* (2005) employed the polygonal function comprising $(q-1)$ piecewise linear functions $g_{ij}(t)$; $i = 2,3,\dots,n$; $j = 1,2,\dots,q$ passing through the coordinates (t_i, y_i), where y_i are the increments. However, this may not appropriately describe a realistic situation. Therefore we adopt the more general approach of Wavelet analysis (See e.g. Ghosh *et al.* 2010). Unlike Fourier analysis, wavelet analysis is used to decompose observed data into various components with respect to frequency or scale and time. Multiresolution analysis based on Pyramid algorithm is then employed to obtain wavelet transforms at various scales, which is used to explain smooth part of exogenous variables in the synthesis phase of wavelet analysis. Basis functions, like Haar and Daubechies filters are tried, which are capable to describe high-pass filter with significant gain function of respective spectral densities in the nominal pass-bands at various scales. The smoothed part is treated as mean function over time and is used to construct functions $g_j(t)$. Accordingly, a new approach of approximating polygonal function by using wavelet transform in place of observed values of exogenous variables is proposed. To this end, the sub-matrix of wavelet transformation $\mathcal{W}$ is multiplied with the vector of wavelet transforms of exogenous variables to smooth the effect of drift term in the process $\{X_G(t): t\epsilon[t_0, T]\}$. Relevant computer code in R and SAS software package for fitting the Modified Gompertz model is developed and the same is appended as an Appendix.

3. ILLUSTRATION

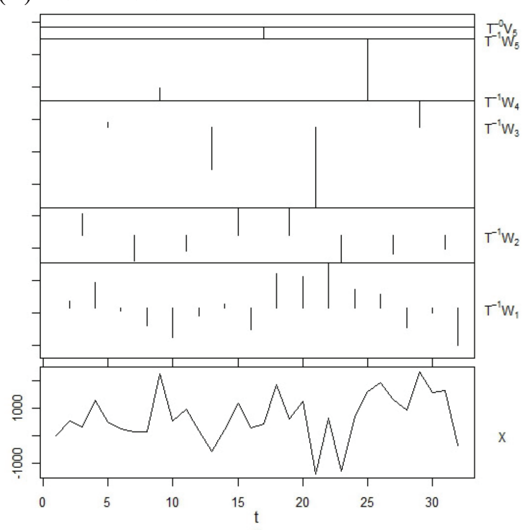
As an illustration, India's total foodgrain production time-series data for the period 1982-83 to 2015-16, available in 'Agricultural Statistics at a Glance 2015', are considered. The data for the period up to 2013-14 are used for fitting of the model while those for the remaining two years are used for validation purpose. As total foodgrain production depends mainly on rainfall, fertilizer consumption, and pesticide consumption in the country, therefore these three variables are taken as exogenous variables. Time-series data on rainfall are available at the website *www.tropmet.res.in* of the Indian Institute of Tropical Meteorology, while those for fertilizer and pesticide consumption are obtained from various issues of Fertilizer Statistics and from the website *www.Indiastat.com* respectively.

In the first instance, the data on three exogenous variables were smoothed using the R code given in the Appendix. Several basis functions, such as Haar and Daubechies filters were tried and it was found that Daubechies (db4) filter provided the best results and the same are exhibited in Fig.1. Subsequently, Gutierrez model was fitted to the data and the results are reported in the second column of Table 1. Using the SAS code given in the Appendix, Modified Gutierrez model was fitted to the data and the results are reported in the third column of Table 1. Using the estimates of parameters given in Table 1, year-wise India's foodgrain production (in million tonnes) for the two models are computed and the results are reported in Table 2. The last row of this table gives the Average Mean Square Error (MSE) computed for the two models. Evidently, the lower value, viz. 209.61 for the fitted Modified Gutierrez model vis-a-vis the value 224.82 for the fitted Gutierrez model, reflects superiority of the former over the latter for fitting purpose for given data. In order to compare the performance of the two models for validation purpose, forecasts were developed for two years and the results are reported in Table 3. A perusal of this table shows that the forecasts obtained by Modified Gutierrez model are much closer to actual values than the forecasts obtained by Gutierrez model. Thus, for forecasting also, Modified Gutierrez model has performed better than Gutierrez model for given data. To get a visual idea, the graph of fitted Modified Gutierrez model along with data is exhibited in Fig. 2.

(i) Rainfall:



(ii) Fertilizer:



(iii) Pesticide:

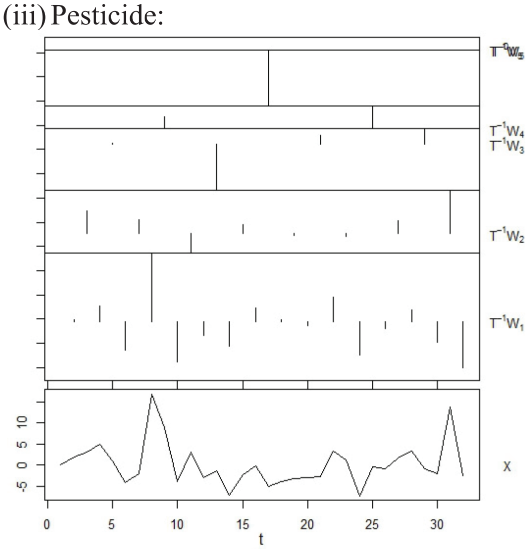


Fig. 1. Smoothing of exogenous variables by Daubechies (db4) filter

Table 1. Estimates of Parameters of Gutierrez and Modified Gutierrez models

Parameter	Estimate	
	Gutierrez model	Modified Gutierrez model
α_0	$4.700X10^{-2}$	$1.800X10^{-2}$
α_1	$3.000X10^{-5}$	$-1.607X10^{-5}$
α_2	$1.412X10^{-5}$	$-1.266X10^{-10}$
α_3	$-8.249X10^{-3}$	$-1.007X10^{-2}$
β	$4.172X10^{-3}$	$8.021X10^{-9}$
σ^2	$3.213X10^{-3}$	-
k	-	$2.749X10^{-3}$
γ	-	$7.401X10^{-3}$

Table 2. Year-wise fitting of India's Foodgrain production (in million tonnes) for various models

Year	Actual	Gutierrez model	Modified Gutierrez model
1983-84	152.37	133.31	132.32
1984-85	145.54	160.97	154.11
1985-86	150.44	151.86	157.27
1986-87	143.42	152.46	139.47
1987-88	140.35	146.56	155.79
1988-89	169.92	143.07	143.50
1989-90	171.04	174.58	173.87
1990-91	176.39	175.35	173.79
1991-92	168.38	180.95	179.67
1992-93	179.48	172.73	171.43
1993-94	184.26	184.07	182.74
1994-95	191.50	188.96	187.61
1995-96	180.42	196.35	194.98
1996-97	199.34	185.00	183.70
1997-98	192.26	204.53	202.87
1998-99	203.61	196.80	196.16
1999-00	209.80	204.69	206.17
2000-01	196.81	211.78	208.45
2001-02	212.85	198.12	206.09
2002-03	174.77	214.32	210.13
2003-04	213.19	179.42	178.42
2004-05	198.36	218.41	216.58
2005-06	208.60	203.08	202.24
2006-07	217.28	214.53	211.75
2007-08	230.78	222.69	222.83
2008-09	234.47	230.07	232.15
2009-10	218.11	234.28	230.80
2010-11	244.49	225.70	220.60
2011-12	259.29	261.27	254.30
2012-13	257.13	275.20	274.23
2013-14	265.04	266.69	265.50
Average MSE		224.82	209.61

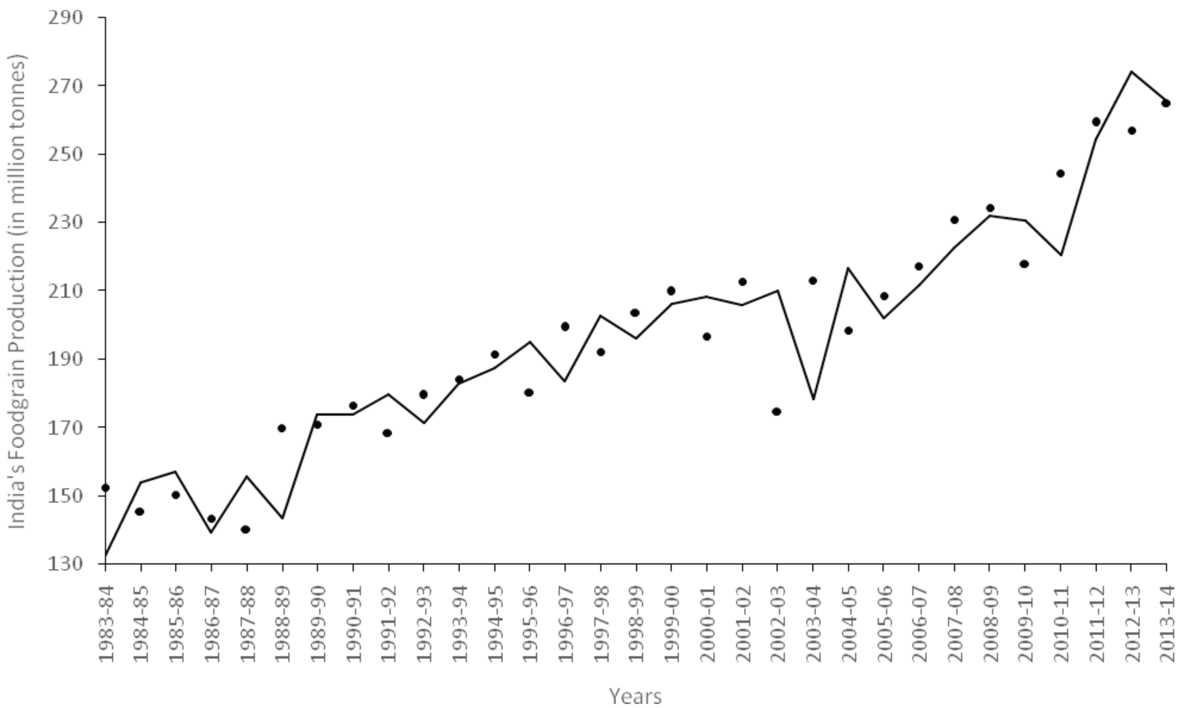


Fig. 2. Fitting of Modified Gutierrez model along with data

Table 3. Year-wise forecasting of India's Foodgrain production (in million tonnes) for various models

Year	Actual	Gutierrez model	Modified Gutierrez model
2014-15	252.02	274.43	263.18
2015-16	251.57	280.02	272.37

4. CONCLUDING REMARKS

Here, the methodology is developed for Gompertz nonlinear diffusion model with exogenous variables and time-dependent diffusion. However, this methodology is applicable only when time-series data on all the variables are available at equal intervals. It may be noted that collection of time-series data involves constraints of time, personnel, and budget, etc., which may not always be possible. Dennis and Ponciano (2014) emphasized that the data with missing observations or data at unequal time-intervals are potentially informative, and precluding such data from analysis could affect conclusions adversely. Thus, there is a need to extend the methodology developed in this article for the case when data are available only at unequal intervals. Work is in progress in this direction and shall be reported separately in due course of time. Further, as future work, similar type of work as done in this article may be carried out for other growth models, such as Richards and Logistic models.

ACKNOWLEDGEMENTS

The authors are grateful to Science and Engineering Research Board, New Delhi for providing financial assistance under Research Project No. SB/S4/MS/880/2014. Thanks are also due to an anonymous reviewer for valuable comments.

REFERENCES

Behera, A., and O'Rourke, S.F.C. (2008). The effect of correlated noise in a Gompertz growth model. *Brazilian J. Phys.*, **38**, 272-278.

Cohen, S. and Elliott, R.J. (2015). *Stochastic Calculus and Applications,* 2nd Ed. Birkhauser, Switzerland.

Dennis, B. and Ponciano, J.M. (2014). Density dependent state space model for population abundance data with unequal time intervals. *Ecology*, **95**, 2069-2076.

Ferrante, L., Bompadre, S., Possati, L. and Leone, L. (2000). Parameter estimation in a Gompertzian stochastic model for tumor growth. *Biometrics*, **56**, 1076-1081.

Ghosh, H., Paul, R.K. and Prajneshu (2010). Wavelet frequency domain approach for statistical modelling of Indian monsoon rainfall time-series data. *J. Statistl. Theor. Prac.*, **4**, 813-825.

Ghosh, H. and Prajneshu (2017). Gompertz growth model in random environment with time-dependent diffusion. *J. Statistl. Theor. Prac.*, **11**, 746-758.

Gutierrez, R., Gutierrez-Sanchez, R., Nafidi, A., Roman, P. and Torres, F. (2005). Inference in Gompertz-type nonhomogeneous stochastic systems by means of discrete sampling. *Cybern. Systs.*, **36**, 203-16.

Gutierrez, R., Gutierrez-Sanchez, R. and Nafidi, A. (2009). Modelling and forecasting vehicle stocks using the trends of stochastic Gompertz diffusion models: The case of Spain. *Appl. Stoch. Mods. Busi. Indus.*, **25**, 385-405.

Ricciardi, L.M. (1976). On the transformation of diffusion processes into the Wiener processes. *J. Mathl. Anal. Applns.*, **54**, 185-199.

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

Skiadas, C.H. (2010). Exact solutions of stochastic differential equations: Gompertz, Generalized logistic and Revised Exponential. *Method. Comput. Appl. Probab.*, **12**, 261-270.

APPENDIX

(i) R code for computation of DWT:

```
library (wavelets)
mydata<-read.csv ("LINK_TO_DATA_FILE", header
=TRUE)
attach (mydata)
mydata
y<-ts (mydata)
#for computing dwt
dwt<-dwt (y,filter="NAME_OF_FILTER", n.levels=
VALUE_OF_LEVEL)
dwt
plot(dwt)
```

(ii) SAS code for fitting Modified Gutierrez model:

(a) Estimation of parameters

```
proc optmodel;
number x{1..n,1..1} = [RESPONSE_VARIABLE_
DATA];
number xx{1..1,1..n} = [RESPONSE_VARIABLE_
DATA];
number r{1..n,1..1} = [EXPLANATORY_VARIA-
BLE1_DATA];
number rr{1..1,1..n} = [EXPLANATORY_VARIA-
BLE1_DATA];
number f{1..n,1..1} = [EXPLANATORY_VARIA-
BLE2_DATA];
number ff{1..1,1..n} = [EXPLANATORY_VARIA-
BLE2_DATA];
number p{1..n,1..1} = [EXPLANATORY_VARIA-
BLE3_DATA];
number pp{1..1,1..n} = [EXPLANATORY_VARIA-
BLE3_DATA];
set dim={2..n};/*n represents number of data points */
set len={1..n-1};
var zz1>=-0.0000000001 <=0.0000000001;
var zz2>=-0.0000000001 <=0.0000000001;
var zz3>=-0.0000000001 <=0.0000000001;
var zz4>=-0.0000000001 <=0.0000000001;
var zz5>=0 <=0.0001;
var zz6>=-0.00000001 <=0.00000001;
var zz7>=-0.00000001 <=0.00000001;
max func= (sum{t in len}((zz5+(2*zz7))/
(2*3.14*(1-exp(-2*(zz5+zz7)))*exp(2*(zz7*(t+1)+z
z6)))**0.5))*exp(((-1*(zz5+(2*zz7)))*(sum{tindim}
((((log(xx[1,t])-(exp(-zz5)*log(xx[1,t-1])))-
(((exp(2*zz7*(t+1))-exp(-zz5+2*zz7*t))*(-exp
(2*zz6)/(zz5+2*zz7)))+zz1*((1-exp(-zz5))/(zz5))+
((1-exp(-zz5))/zz5)**(1)*(rr[1,t-1]+((rr[1,t]-rr[1,
t-1])*((zz5-1+exp(-zz5))/(zz5*(1-exp(-zz5))))))*
zz2+((1-exp(-zz5))/zz5)**(1)*(ff[1,t-1]+((ff[1,t]-
ff[1,t-1])*((zz5-1+exp(-zz5))/(zz5*(1-exp(-zz5))))))*
zz3+((1-exp(-zz5))/zz5)**(1)*(pp[1,t-1]+((pp[1,t]-
pp[1,t-1])*((zz5-1+exp(-zz5))/(zz5*(1-exp
(-zz5))))))*zz4))*((1-exp(-2*(zz5+zz7)))*(exp(2*
(zz7*(t+1)+zz6))))**(-1))*((log(x[t,1])-(exp(-zz5)*
log(x[t-1,1])))-(((exp(2*zz7*(t+1))-exp(-zz5+
2*zz7*t))*(-exp(2*zz6)/(zz5+2*zz7)))+zz1*((1-exp
(-zz5))/(zz5))+((1-exp(-zz5))/zz5)**(1)*(r[t-1,1]+
((r[t,1]-r[t-1,1])*((zz5-1+exp(-zz5))/(zz5*(1-exp
(-zz5))))))*zz2+((1-exp(-zz5))/zz5)**(1)*(f[t-1,1]+
((f[t,1]-f[t-1,1])*((zz5-1+exp(-zz5))/(zz5*(1-exp
(-zz5))))))*zz3+((1-exp(-zz5))/zz5)**(1)*(p[t-1,1]+
((p[t,1]-p[t-1,1])*((zz5-1+exp(-zz5))/(zz5*(1-exp
(-zz5))))))*zz4)))))/2);
solve;
print func;
print zz1 zz2 zz3 zz4 zz5 zz6 zz7;
quit;
```

(b) Fitting and Forecasting

```
proc iml;
rr={EXPLANATORY_VARIABLE1_DATA};
ff={EXPLANATORY_VARIABLE2_DATA};
pp={EXPLANATORY_VARIABLE3_DATA};
xx={RESPONSE_VARIABLE_DATA};
alpha0=ESTIMATED_VALUE;
alpha1=ESTIMATED_VALUE;
alpha2=ESTIMATED_VALUE;
alpha3=ESTIMATED_VALUE;
```

```
beta=ESTIMATED_VALUE;
k=ESTIMATED_VALUE;
gamma=ESTIMATED_VALUE;
do i=2 to n;/*n represents number of data points */
b1=(1-exp(-beta))/beta;
b2=exp(-beta);
b3=(exp(2*k)/(beta+(2*gamma)))*(exp(2*gamma*
(i+1))-exp(-1*beta+2* gamma*i));
var=((1-exp(-2*(beta+gamma)))*exp(2*(gamma*
(i+1)+k)))/(beta+gamma);
bexpr=(beta-1+exp(-beta))/(beta*(1-exp(-beta)));
rexpr=alpha1*(rr[i-1]+((rr[i]-rr[i-1])*bexpr));
fexpr=alpha2*(ff[i-1]+((ff[i]-ff[i-1])*bexpr));
pexpr=alpha3*(pp[i-1]+((pp[i]-pp[i-1])*bexpr));
expr=alpha0*b1+rexpr +fexpr+pexpr;
final=b2*log(xx[i-1])-b3+expr;
value=exp(final+0.5*var);
value1=value1//value;
end;
xy=xx[,2:n];
v=t(value1);
diff=abs(xy-v);
print value1;
difff=diff##2;
summ=sum(difff);
av=summ/(n-1);
print av;
quit;
```