HIBISCUS PUBLISHER

JOURNAL OF ENVIRONMENTAL MICROBIOLOGY AND TOXICOLOGY

Website: http://journal.hibiscuspublisher.com/index.php/JEMAT/index

Evaluation of Several Mathematical Models for Fitting the Growth of Sludge Microbes on PEG 600

Halmi M.I.E.[1], Shukor, M.S.[2], Masdor, N.A.[3], Shamaan, N.A.[4] and Shukor, M.Y.[2,5*]

[1]Department of Chemical Engineering and Process, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia.
[2]Snoc International Sdn Bhd, Lot 343, Jalan 7/16 Kawasan Perindustrian Nilai 7, Inland Port, 71800, Negeri Sembilan, Malaysia.
[3]Biotechnology Research Centre, MARDI, P. O. Box 12301, 50774 Kuala Lumpur, Malaysia
[4]Faculty of Medicine and Health Sciences, Universiti Sains Islam Malaysia, 13th Floor, Menara B, Persiaran MPAJ, Jalan Pandan Utama, Pandan Indah, 55100 Kuala Lumpur, Malaysia.
[5]Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, UPM 43400 Serdang, Selangor, Malaysia.

*Corresponding author:
Associate Prof. Dr. Mohd. Yunus Abd. Shukor
Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, UPM 43400 Serdang, Selangor, Malaysia.
Email: yunus.upm@gmail.com
Tel: +603-89466722
Fax: +603-89430913

HISTORY
Received: 11th May 2015
Received in revised form: 12th of June 2015
Accepted: 25th of June 2015

KEYWORDS
growth curve
mathematical model
sludge microbes
modified Gompertz
statistical analysis

ABSTRACT

Polyethylene glycols (PEGs) are employed in numerous sectors. PEGs are nephrotoxic and their biodegradation by microbes could be a potential tool for bioremediation. Numerous bacterial growth studies neglect primary modelling even though modelling exercises can reveal important parameters. In this work we modelled the growth of the sludge microbes on PEG 600 based on available published work in the literature using several growth models such as modified logistic, modified Gompertz, modified Richards, modified Schnute, Baranyi-Roberts, Von Bertalanffy, Huang and the Buchanan three-phase linear model. Statistical analysis results indicated that the modified Gompertz model was the best with highest adjusted R^2, lowest RMSE and AICc values and Bias and Accuracy Factor values closest to unity. The results from this work can be used in the further optimization works of this process in the future.

INTRODUCTION

Polyethylene glycols (PEGs) are widely-used in a great many industrial sectors for instance cosmetics, lubricants, pharmaceuticals, and antifreeze for automobile radiators and in the production of non-ionic surfactants. PEGs have been proved for being nephrotoxic. Wounded rabbit exposed topically to polyethylene glycol-based antimicrobial cream model showed evidence of nephrotoxicity with symptoms of failure. Several of the animals tested died within one week of therapy [1]. Several millions of tons of PEGs are manufactured globally. Effluents contaminated with PEGS usually reach conventional sewage treatment systems making them a significant pollutant [2]. PEGs have the common structural formula of $HO(CH_2CH_2O)nCH_2CH_2OH$ and are water-soluble polymers but the difference is in their molecular weights. From the last three decades, concern has been expressed about the fate of these polymers in the environment and several studies have been performed on their biodegradability. Biodegradation of PEG was first documented in 1965 [3] and further isolations of PEG-degrading microorganisms have been reported [2]. Like many xenobiotics, the growth on this toxic substrate exhibits a significant lag phase due to the needs of the cells to tolerate and initiate detoxification and degradation of enzymes upon exposure to this substrate before assimilation can take place. The growth profile exhibits several phases where the specific growth rate starts at the value of zero followed by a stagnation of the rate associated with the lag time (λ). This is followed by acceleration to a maximal value (μ_m) for a given period of time. Finally, the growth curves exhibit a final phase where the rate decreases and eventually reaches zero or an asymptote (A) [4]. A valuable parameter of the growth is the maximum growth rate (μ_m) [5]. This value is important for the development of secondary models such as growth kinetics [6]. In a large number of publications, this parameter is often estimated manually by deciding subjectively the part of the curve that is nearly linear and then the slope of this curve section is then determined usually by linear regression. A better method is to describe the entire set of data

with a nonlinear regression growth model and then estimate μmax, λ, and A from the model. In addition, many published works produced the growth curve but did not attempt any further fitting of the data to available models [5].

The sigmoidal curve can be fitted by various mathematical functions such as Logistic, Gompertz, Richards, Schnute [5], Baranyi-Roberts [7] and Von Bertalanffy [8,9], Buchanan three-phase [10] and more recently the Huang model [11] (**Table 1**). Apart from demonstrating predictive ability and internal consistency, which is a must, the usefulness of a model should also be judged by its mathematical simplicity, flexibility, the number of its adjustable parameters and, where appropriate, whether they have intuitive meaning.

Table 1. Growth models used in this study.

Model	n	Equation
Modified Logistic	3	$y = \frac{A}{\left\{1+\exp\left[\frac{4\mu_{max}}{A}(\lambda - t)+2\right]\right\}}$
Modified Gompertz	3	$y = A\exp\left\{-\exp\left[\frac{\mu_{max} e}{A}(\lambda - t)+1\right]\right\}$
Modified Richards	4	$y = A\left\{1+v\exp(1+v)\exp\left[\frac{\mu_{max}}{A}(1+v)\left(1+\frac{1}{v}\right)(\lambda - t)\right]\right\}^{\left(\frac{-1}{v}\right)}$
Modified Schnute	4	$y = \left(\mu_{max}\frac{(1-\beta)}{\alpha}\right)\left[\frac{1-\beta\exp(\alpha\lambda+1-\beta-\alpha t)}{1-\beta}\right]^{\frac{1}{\beta}}$
Baranyi-Roberts	4	$y = A+\mu_{max}x+\frac{1}{\mu_{max}}\ln\left(e^{-\mu_{max}x}+e^{-h_0}-e^{-\mu_{max}x-h_0}\right) - \ln\left(1+\frac{e^{\mu_{max}x+\frac{1}{\mu_{max}}\ln\left(e^{-\mu_{max}x}+e^{-h_0}-e^{-\mu_{max}x-h_0}\right)}-1}{e^{(y_{max}-A)}}\right)$
Von Bertalanffy	3	$y = K\left[1-\left[1-\left(\frac{A}{K}\right)^3\right]\exp^{-\left(rx/3K^{\frac{1}{3}}\right)}\right]^3$
Huang	4	$y = A+y_{max}-\ln\left(e^{A}+\left(e^{Y_{max}}-e^{A}\right)e^{-\mu_{max}B(x)}\right)$ $B(x) = x+\frac{1}{\alpha}\ln\frac{1+e^{-\alpha(x-\lambda)}}{1+e^{\alpha\lambda}}$
Buchanan Three-Phase Linear Model	3	Y = A, IF X < LAG Y=A + *K*(X−λ), IF λ ≤ X ≥ X_{MAX} Y = Y_{MAX}, IF X ≥ X_{MAX}

Note:
a= bacterial lower asymptote;
μ_{max}= maximum specific growth rate;
v= affects near which asymptote maximum growth occurs.
λ=lag time
y_{max}= bacterial upper asymptote;
e = exponent (2.718281828)
t = sampling time
α,β, k = curve fitting parameters
h_0 = a dimensionless parameter quantifying the initial physiological state of the cells. the lag time (day^{-1}) can be calculated as $h_0=\mu_{MAX}$

The objective of the first part of this work is to evaluate similarities and differences between the models using published available data from Huang et al. [12] that lacks the initial modelling and to deal with the question of which model(s) can be used, on the basis of statistical reasoning. This should give new data and results that could spur further information and improvement in the works already done by researchers.

MATERIALS AND METHODS

Acquisition of Data

In order to process the data, graphs were scanned and electronically processed using WebPlotDigitizer 2.5 [13]. The software helps to digitize scanned plots into table of data with good enough precision [4]. Data were acquired from the works of Huang et al. [12], from Figure 1 which shows the effect of different concentrations of the substrate PEG 600 on the growth of sludge microbes measured over several days and then replotted.

Fitting of the data

To decide whether there is a statistically substantial difference between models with different number of parameters, in terms of the quality of fit, data was statistically assessed through various methods such as the root-mean-square error (RMSE), adjusted coefficient of determination (R^2), bias factor (BF), accuracy factor (AF) and corrected AICc (Akaike Information Criterion) [14].

RMSE

The RMSE was calculated according to Eq. (1), where *Pdi* are the values predicted by the model and Ob_i are the experimental data, *n* is the number of experimental data, and *p* is the number of parameters of the assessed model. It is expected that the model with the smaller number of parameters will give a smaller RMSE values (Eqn. **1**).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Pd_i - Ob_i)^2}{n-p}} \quad (\mathbf{1})$$

In linear regression models, the coefficient of determination or R^2 is used to assess the quality of fit of a model. However, in nonlinear regression where difference in the number of parameters between one model to another is normal, the adoption of the method does not readily provides comparable analysis. Hence, an adjusted R^2 is used to calculate the quality of nonlinear models according to the formula where RMS is Residual Mean Square and s_y^2 is the total variance of the y-variable (Eqns. **2** and **3**).

$$Adjusted\left(R^2\right) = 1-\frac{RMS}{s_Y^2} \quad (\mathbf{2})$$

$$Adjusted\left(R^2\right) = 1-\frac{\left(1-R^2\right)(n-1)}{(n-p-1)} \quad (\mathbf{3})$$

Akaike information criterion with correction (AICc)

The Akaike information criterion (AIC) provides a means for model selection through measuring the relative quality of a given statistical model for a given set of experimental data [15]. AIC deals with the trade-off regarding the goodness of fit of the model along with the intricacy of the model. It is in reality founded on information theory. The procedure offers a comparative approximation of the information lost for every time a certain model is employed to signify the process that produces the information or data. For any output of a collection of predicted models, the most accepted model is the model demonstrating the minimum value for AIC. This value is often a negative value, with for example; an AICc value of -10 more preferred than the one with -1. The formula includes a number of parameters punishment, the greater the parameters, the less favoured the end result or the greater the AIC value. Therefore, AIC not simply returns goodness of fit, but additionally, doesn't really encourage utilizing more complex model (overfitting) for

fitting experimental data. Considering that the data within this work is smaller compared to the number of parameter employed a remedied version of AIC, the Akaike information criterion (AIC) with correction or AICc is employed in its place. The AICc is computed for each and every data set for each model based on the following equation (Eqn. **4**);

$$AICc = 2p + n\ln\left(\frac{RSS}{n}\right) + 2(p+1) + \frac{2(p+1)(p+2)}{n-p-2} \quad (4)$$

Where p is the number of parameters of the model and n is the number of data points. The procedure considers the alteration in goodness-of-fit and the improvement in number of parameters between two models. For each and every data set, the model having the smallest AICc value is extremely likely correct [15].

Accuracy Factor (AF) and Bias Factor (BF)

Accuracy Factor (AF) and Bias Factor (BF) to test for the goodness-of-fit of the models as recommended by Ross [16] were also employed. A Bias Factor equal to 1 indicates a perfect match between predicted and observed values. For microbial growth curves or degradation studies, a bias factor with values < 1 indicates a fail-dangerous model while a bias factor with values > 1indicates a fail-safe model. The Accuracy Factor is always ≥ 1, and higher AF values indicate less precise prediction (Eqns. **5** and **6**).

$$\text{Bias factor} = 10^{\left(\sum_{i=1}^{n}\log\frac{(Pd_i/Ob_i)}{n}\right)} \quad (5)$$

$$\text{Accuracy factor} = 10^{\left(\sum_{i=1}^{n}\log\frac{|(Pd_i/Ob_i)|}{n}\right)} \quad (6)$$

RESULTS AND DISCUSSION

Essentially, the most vital results from curve fitting in growth curve model is the capacity to utilize a growth model that have a good fundamental mechanistic function in accordance with good theoretical understanding of the system. Among the finest of such model is the Michaelis-Menten kinetics that models the effects substrate on the initial enzyme activity of an enzyme. To get the best model, eight various growth models were put to use for this study to suit the experimental data. The ensuing fitting illustrates visually sufficient fitting for the models of Huang, modified Gompertz, modified logistics, Von Bertalanffy, Baranyi-Roberts and Buchanan-3-models (**Figs. 1-6**). Other models gave poor fitting and were not shown. The statistical analysis results (**Table 2**) indicated that the modified Gompertz model was the best with highest adjusted R^2, lowest RMSE and AICc values, and Bias and Accuracy Factor values closest to unity. The modified Gompertz model was then used to fit the data and the resultant fitted values obtained (**Table 3**).

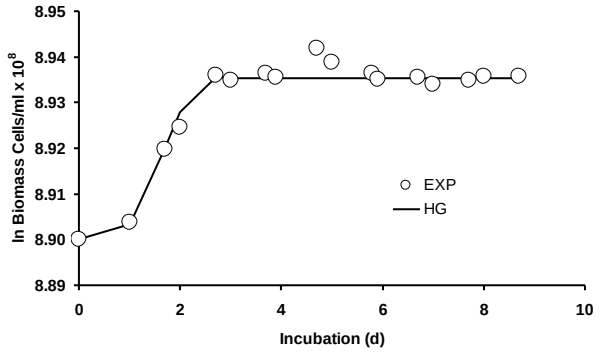

Fig. 1. Growth curves of sludge microbes on PEG 600 fitted by the Huang growth model.

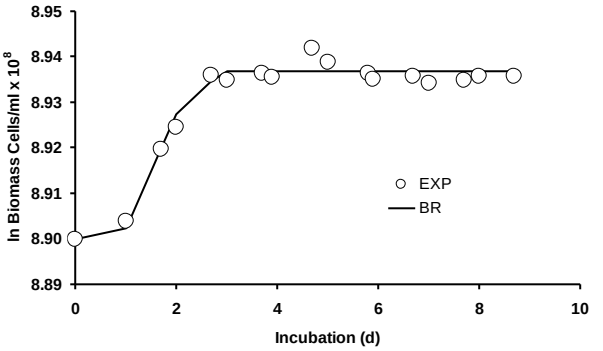

Fig. 2. Growth curves of sludge microbes on PEG 600 fitted by the Baranyi-Roberts growth model.

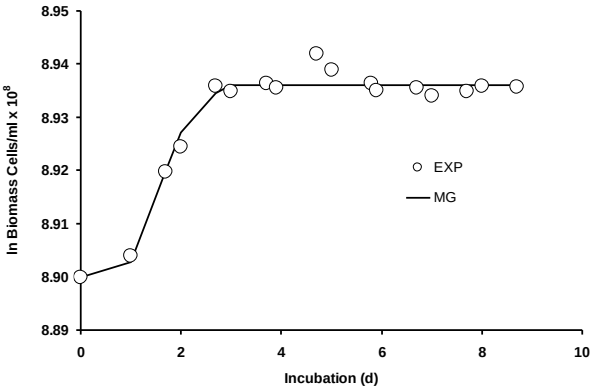

Fig. 3. Growth curves of sludge microbes on PEG 600 fitted by the modified Gompertz growth model.

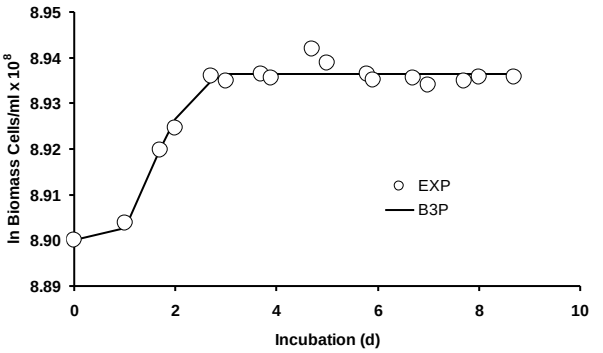

Fig. 4. Growth curves of sludge microbes on PEG 600 fitted by the Buchanan-3-phase growth model.

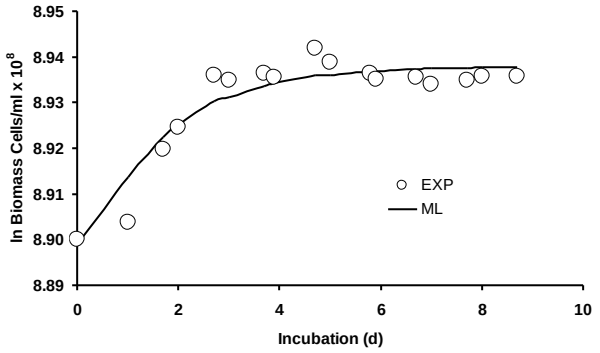

Fig. 5. Growth curves of sludge microbes on PEG 600 fitted by the modified logistics growth model.

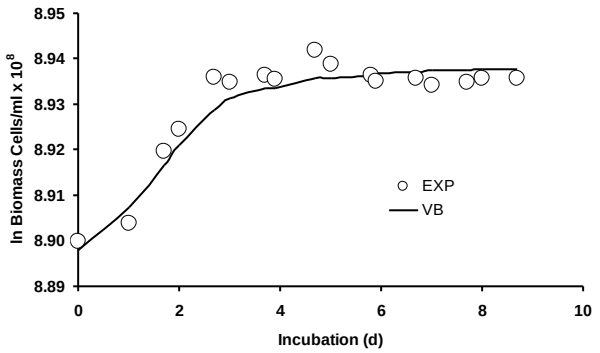

Fig. 6. Growth curves of sludge microbes on PEG 600 fitted by the von Bertalanffy growth model.

Table 2. Statistical analysis of the various fitting models.

Model	*n*	RMSE	R^2	adR^2	AF	BF	AICc
Huang	4	0.0047	0.90	0.87	1.00	1.00	-163.20
Baranyi-Roberts	4	0.0047	0.90	0.87	1.00	1.00	-163.06
Modified Gompertz	3	0.0004	1.00	1.00	1.00	1.00	-248.13
Buchanan-3-Phase	3	0.0045	0.90	0.88	1.00	1.00	-169.38
Modified Richards	4	na.	na.	na.	na.	na.	na.
Modified Schnute	4	na.	na.	na.	na.	na.	na.
Modified Logistics	3	0.0043	0.89	0.86	1.00	1.00	-171.30
Von Bertalanffy	3	0.0047	0.89	0.87	1.00	1.00	-168.55

Note:

SSE	Sums of Squared Errors
RMSE	Root Mean Squared Error
R^2	Coefficient of Determination
adR^2	Adjusted Coefficient of Determination
AICc	Corrected Akaike Information Criterion
BF	Bias Factor
AF	Accuracy Factor
N	No of parameter
n.a.	Not available

Table 3. Fitted growth parameters according to the modified Gompertz model.

Parameters	Value	Std-Error
Y_0	8.896	0.002
lag	0.762	0.149
y_{max}	8.937	0.001
μ_{max}	0.026	0.003

The modified Gompertz model is one of the classical growth models that include model such as the Verhulst [5,17]. The Gompertz function, named in 1844-1845 by Pierre François Verhulstis, is based on an exponential relationship between specific growth rate and population density. The initial stage of growth is approximately exponential; then, as saturation begins, the growth slows, and at maturity, growth stops. Gibson et al. [18] were the first ones to make use of the Gompertz equation to suit microbial growth curves plus the equation was used successfully to explain the exponential and stationary phases of the microbial growth curves which is sigmoidal. Nevertheless, the model had not been satisfactory to explain the lag phase. The model was altered by Gibson et al. to include the lag phase, and also have been used successfully in modelling numerous microbial growth curves so much that its popularity in mathematically modelling bacterial growth and product formation curves have been recognized [19,20].

CONCLUSION

In conclusion, the various models used to fit the growth of sludge microbes on PEG 600 as a substrate showed that the best model was modified Gompertz based on statistical analysis. The fitted data from this work can be used in the further optimization works of the microbe.

ACKNOWLEDGEMENT

This project was supported by a grant from Snoc International Sdn Bhd.

REFERENCES

[1] Herold DA, Rodeheaver GT, Bellamy WT, Fitton LA, Bruns DE, Edlich RF. Toxicity of topical polyethylene glycol. Toxicol Appl Pharmacol. 1982;65(2):329–335.

[2] Watanabe M, Kawai F. Study on biodegradation process of polyethylene glycol with exponential growth of microbial population. 2010;145-157.

[3] Payne WJ, Williams JP, Mayberry WR. Primary alcohol sulfatase in a Pseudomonas species. Appl Microbiol. 1965;13(5):698–701.

[4] Halmi MIE, Shukor MS, Johari WLW, Shukor MY. Evaluation of several mathematical models for fitting the growth of the algae Dunaliella tertiolecta. Asian J Plant Biol. 2014;2(1):1–6.

[5] Zwietering MH, Jongenburger I, Rombouts FM, Van't Riet K. Modeling of the bacterial growth curve. Appl Environ Microbiol. 1990;56(6):1875–1881.

[6] Ahmad SA, Ahamad KNEK, Johari WLW, Halmi MIE, Shukor MY, Yusof MT. Kinetics of diesel degradation by an acrylamide-degrading bacterium. Rendiconti Lincei. 2014;25(4):505–512.

[7] Baranyi J, Roberts TA. A dynamic approach to predicting bacterial growth in food. Int J Food Microbiol. 1994;23(3-4):277–294.

[8] Cloern JE, Nichols FH. A von Bertalanffy Growth Model with a Seasonally Varying Coefficient. J Fish Res Board Can. 1978;35(11):1479–1482.

[9] Darmani Kuhi H, Kebreab E, Lopez S, France J. A derivation and evaluation of the von Bertalanffy equation for describing growth in broilers over time. J Anim Feed Sci. 2002;11(1):109–125.

[10] Buchanan RL, Whiting RC, Damert WC. When is simple good enough: A comparison of the Gompertz, Baranyi, and three-phase linear models for fitting bacterial growth curves. Food Microbiol. 1997;14(4):313–26.

[11] Huang L. Growth kinetics of Escherichia coli O157:H7 in mechanically-tenderized beef. Int J Food Microbiol. 2010;140(1):40–48.

[12] Huang Y-L, Li Q-B, Deng X, Lu Y-H, Liao X-K, Hong M-Y, et al. Aerobic and anaerobic biodegradation of polyethylene glycols using sludge microbes. Process Biochem. 2005;40(1):207–211.

[13] Rohatgi, A. WebPlotDigitizer. http://arohatgi.info/WebPlotDigitizer/app/ Accessed June 2 2014.;

[14] Motulsky HJ, Ransnas LA. Fitting curves to data using nonlinear regression: a practical and nonmathematical review. FASEB J Off Publ Fed Am Soc Exp Biol. 1987;1(5):365–374.

[15] Akaike H. Statistical information processing system for prediction and control. AAPG Bull Am Assoc Pet Geol. 1981;237–241.

[16] Ross T. Indices for performance evaluation of predictive models in food microbiology. J Appl Bacteriol. 1996;81(5):501–508.

[17] Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philos Trans R Soc London. 1825;115:513–585.
[18] Gibson AM, Bratchell N, Roberts TA. The effect of sodium chloride and temperature on the rate and extent of growth of Clostridium botulinum type A in pasteurized pork slurry. J Appl Bacteriol. 1987;62(6):479–490.
[19] Çelekli A, Balci M, Bozkurt H. Modelling of Scenedesmus obliquus; function of nutrients with modified Gompertz model. Bioresour Technol. 2008;99(18):8742–8747.
[20] Johnsen AR, Binning PJ, Aamand J, Badawi N, Rosenbom AE. The Gompertz function can coherently describe microbial mineralization of growth-sustaining pesticides. Environ Sci Technol. 2013;47(15):8508–8514.