

Short Communication

Test for the Presence of Autocorrelation in the Modified Gompertz model used in the 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*}

¹Department of Chemical Engineering and Process, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia.

²Snoc International Sdn Bhd, Lot 343, Jalan 7/16 Kawasan Perindustrian Nilai 7, Inland Port, 71800, Negeri Sembilan, Malaysia.

³Biotechnology Research Centre, MARDI, P. O. Box 12301, 50774 Kuala Lumpur, Malaysia

⁴Faculty of Medicine and Health Sciences, Universiti Sains Islam Malaysia, 13th Floor, Menara B, Persiaran MPAJ, Jalan Pandan Utama, Pandan Indah, 55100 Kuala Lumpur, Malaysia.

⁵Department 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: 17th of May 2015
Received in revised form: 14th of June 2015
Accepted: 26th of June 2015

KEYWORDS

Polyethylene Glycol
modified Gompertz
sludge microbes
ordinary least squares method
autocorrelation

ABSTRACT

Polyethylene glycols (PEGs), are nephrotoxic, and are employed in numerous industrial sectors. Their biodegradation by microbes could be a potential tool for bioremediation. A lot of bacterial growth reports overlook primary modelling despite the fact that modelling exercises can expose important parameters. Earlier, we have employed several growth models to model the growth of sludge microbes on PEG 600. We found out that the modified Gompertz model via nonlinear regression utilizing the least square method was the most effective model to describe the growth curve. Nonlinear regression using the least square method generally utilizes the assumption that data points do not depend on each other or the value of a data point is not dependent on the value of preceding or proceeding data points or do not exhibit autocorrelation. In this work, the Durbin-Watson statistic to check for the presence of autocorrelation in the growth model was carried out.

INTRODUCTION

Polyethylene glycols or PEGs have the common structural formula of $\text{HO}(\text{CH}_2\text{CH}_2\text{O})_n\text{CH}_2\text{CH}_2\text{OH}$ and are water-soluble polymers but the difference is in their molecular weights. PEGs are utilized in numerous industrial sectors for example cosmetics, lubricants, pharmaceuticals, and antifreeze for automobile radiators as well as in the production of non-ionic surfactants. PEGs are nephrotoxic. Injured rabbit subjected topically to polyethylene glycol-based antimicrobial cream model demonstrated proof of nephrotoxicity with signs and symptoms of failure. Some of the animals examined died within just 1 week of treatment [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]. 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]. The growth on this toxic substrate displays a substantial lag phase because of the needs of the cell to endure and trigger detoxification and degradation of enzymes upon contact with the substrate before assimilation can occur. The growth profile displays a number of phases in which the specific growth rate begins at the value of zero accompanied by a stagnation of the rate linked to 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]. Previously, we have utilized several growth models to model the

growth of sludge microbes on PEG 600. The data was obtained from the literature. We discovered that the modified Gompertz model via nonlinear regression utilizing the least square method was the best model to describe the growth curve. Nonlinear regression using the least square method normally uses the assumption that data points do not depend on each other or the value of a data point is not dependent on the value of preceding or proceeding data points. Autocorrelation between data can happen as a result of events for example temperature drift during time measurements or an overused tungsten lamp in a spectrophotometer. If one would count the quantity of animals each year in a given area the data will be extremely autocorrelated and nonindependence as the quantity of animals within an existing year will be highly influenced by the quantity of animals in the last year [7]. This is very similar to growth of microorganisms where the increase in cellular number in a given time frame can be exponentially fast and any event in time that effect the current or past number of cells would be seen in an amplified manner in future times.

In this work, the Durbin–Watson statistic for the presence of autocorrelation in the growth of the sludge microbes as modelled using the modified Gompertz model would be used. The method calculates the level of significance according to the method outlined by Draper and Smith [8].

MATERIALS AND METHODS

Acquisition of Data

In order to process the data, the graphs were scanned and electronically processed using WebPlotDigitizer 2.5 [9] which helps to digitize scanned plots into table of data with good enough precision [4]. Data were acquired from the works of Huang et al. [10], from Figure 1 which show the effect of different concentrations of the substrate PEG 600 on the growth of sludge microbes measured over several days, replotted, and then assessed using several growth models where the modified Gompertz model was found to be the best (Fig. 1, with permission) (Halimi, M.I.E., Shukor, M.S., Shamaan, N.A. and Shukor, M.Y. 2015. Evaluation of several mathematical models for fitting the growth of sludge microbes on PEG 600. Manuscript in preparation).

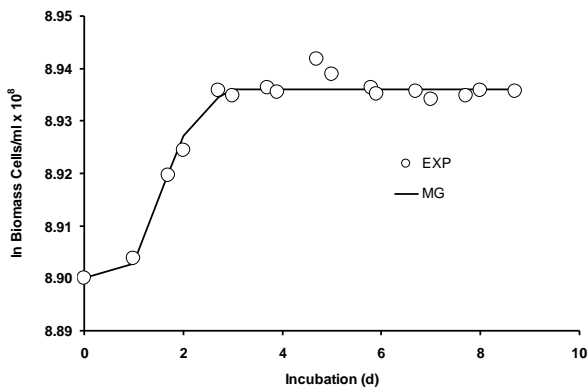


Fig. 1. Growth curves of sludge microbes on PEG 600 fitted by the modified Gompertz growth model.

Durbin-Watson test

The Durbin–Watson statistic calculates the level of significance according to the method outlined by Draper and Smith [8].

$$d = \frac{\sum_{i=2}^T (\hat{e}_i - \hat{e}_{i-1})^2}{\sum_{i=1}^T \hat{e}_i^2} \tag{1}$$

The hypothesis $H_0: \rho = 0$ versus the alternative $H_1: \rho > 0$ is tested. The statistic is about equal to $2(1 - \rho)$. The Durbin-Watson test statistic equals 2 when the ρ value is zero while a ρ value of one equals a Durbin-Watson test statistic of 0. Non-autocorrelation is specified by a d value near 2 while a value towards 0 indicates positive autocorrelation. Negative autocorrelation is indicated by d values nearing 4 (Eqn. 1).

The null hypothesis should be rejected for a low value of the Durbin-Watson test statistic indicating significant autocorrelation. Unlike the t- or z-statistics, the distribution of the Durbin-Watson test statistic is not available for ρ -value associated with d and tables must be used in the hypothesis testing.

The decision rule for the Durbin-Watson bounds test is

- if $d >$ upper bound, fail to reject the null hypothesis of no serial correlation,
- if $d <$ lower bound, reject the null hypothesis and conclude that positive autocorrelation is present,
- if lower bound $< d <$ upper bound, the test is inconclusive.

RESULTS AND DISCUSSION

The runs test has also been utilized as a technique to test for autocorrelation in time-series regression models. However, simulation studies using Monte Carlo have shown that the runs test produces distinctly asymmetrical error rates in the two tails [11]. The investigation is carried out to analyse the empirical properties of the runs test utilizing (a) sample sizes of between 12 and 100 (b) using non-intervention and intervention regression models, (c) utilizing directional and nondirectional tests (d) with three levels of α , and (e) with 19 levels of autocorrelation among the errors. In addition, both directional and nondirectional tests produce no satisfactory results with respect to Type I error. The increase of the ratio of degrees of freedom to sample size to as high as .98 could also not remedy the situation. Hence, the Durbin-Watson method would be the method of choice to assess autocorrelation.

The Durbin–Watson statistic (DW) can calculate for the presence of serial correlation of residuals. Autocorrelation, also known as serial correlation, is the cross-correlation of a signal with itself. The DW is used to test whether a model has been successful in describing the underlying trend. Informally, it is the similarity between observations as a function of the time lag between them. It is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise. [7,8,12].

The value of the Durbin-Watson statistics was $d = 0.000101/0.000279 = 0.363$. As usual the hypothesis $H_0: \rho = 0$ versus the alternative $H_1: \rho > 0$ is tested. The statistic is approximately equal to $2(1 - \rho)$. The Durbin-Watson test statistic equals 2 when the ρ value is zero while a ρ value of one equals a Durbin-Watson test statistic of 0. Non-autocorrelation is indicated by a d value near 2 while a value towards 0 indicates positive autocorrelation. Negative autocorrelation is indicated by d values nearing 4. The null hypothesis should be rejected for a low value of the Durbin-Watson test statistic

indicating significant autocorrelation. Unlike the *t*- or *z*-statistics, the distribution of the Durbin-Watson test statistic is not available for ρ -value associated with *d* and tables must be used in the hypothesis testing. For a three-parameter model like modified Gompertz, the upper critical value d_U was 1.432 while the lower critical value d_L was 0.672. Since *d* was lower than the lower critical value then the null hypothesis that no positive autocorrelation exist is rejected i.e. there appears to be evidence of positive autocorrelation and the modified Gompertz model used for fitting the growth curve might not be adequate. Other models need to be tested in the future to select the model that shows no positive or negative autocorrelations. Autocorrelation occurs when covariances of errors are not zero, a problem often seen in time series data such as microbial growth curves. A consequent of the presence of autocorrelation is that estimators for the models used even though are still considered linear and unbiased, but they there not efficient and not the best.

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):1875–81.
- [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):505–12.
- [7] McDonald JH, Dunn KW. Statistical tests for measures of colocalization in biological microscopy. *J Microsc.* 2013;252(3):295–302.
- [8] Draper NR, Smith H. *Applied Regression Analysis.* Wiley, New York; 1981.
- [9] Rohatgi, A. WebPlotDigitizer. <http://arohatgi.info/WebPlotDigitizer/app/> Accessed June 2 2014.;
- [10] 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):207–11.
- [11] Huitema BE, McKean JW, Zhao J. The runs test for autocorrelated errors: unacceptable properties. *J Educ Behav Stat.* 1996;21(4):390–404.
- [12] 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.