HIBISCUS PUBLISHER

JOURNAL OF ENVIRONMENTAL MICROBIOLOGY AND TOXICOLOGY

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

Short Communication

Test for the Presence of Autocorrelation in the Buchanan model used in the Fitting of the Growth of the Catechol-degrading *Candida parapsilopsis*

Shukor, M.S.[1] and Shukor, M.Y.*[1,2]

[1]Snoc International Sdn Bhd, Lot 343, Jalan 7/16 Kawasan Perindustrian Nilai 7, Inland Port, 71800, Negeri Sembilan, Malaysia.
[2]Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, UPM 43400 Serdang, Selangor, Malaysia.

***Corresponding author:**yunus.upm@gmail.com
Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, UPM 43400 Serdang, Selangor, Malaysia
Tel No: +603-89466722

HISTORY
Received: 28st October 2014
Received in revised form: 21st of December 2014
Accepted: 29th of December 2014

KEYWORDS
Buchanan three-phase model
catechol
Candida parapsilopsis
ordinary least squares method
autocorrelation

ABSTRACT

Catechol is a metabolic byproduct of phenol degradation by microbes. Its toxicity to human, mammals, insects and fishes has been long studied and its presence in the environment at toxic concentrations has been demonstrated. Fortunately there are microbes that could degrade catechol and can be used in bioremediation works. The growth of these microbes usually exhibit sigmoidal pattern due to the toxicity of the substrate. Previously, using the least square method in nonlinear regression, we report that the Buchanan three-phase model is the best model in fitting the growth of the yeast *Candida parapsilopsis* on this substrate. The ordinary least squares method relies heavily on several important assumptions such as residuals conformation to normal distribution, does not have outliers, is truly random, of equal variance (homoscedastic) and does not show autocorrelation. If all of these assumptions are satisfied, the test is said to be robust. In this work we perform statistical diagnosis test to test for the presence of autocorrelation as the growth model is time-dependent and many time-dependent curves shows evidence of autocorrelation.

INTRODUCTION

Catechol is toxic since it could react with sulphydryl groups of proteins and glutathione leading to cross-linking of protein and formation of glutathione dimer, which can cause cessation of enzyme and metabolic enzymatic activity [1]. Like many xenobiotics, the growth on this toxic substrate exhibits a significant lag phase due to the needs of the cell to tolerate and initiate detoxification and degradation of enzymes upon exposure to catechol before assimilation can take place. The growth profile exhibits lag time (λ), acceleration to a maximal value (μ_m) and a final phase where the rate decreases and eventually reaches zero or an asymptote (A) [2]. The maximum growth rate (μ_m) is important for the development of secondary models such as growth kinetics [2-5]. The sigmoidal curve can be fitted by different mathematical functions, such as Logistic (2,6), Gompertz (2,7), Richards (8), Schnute [2], Baranyi-Roberts [3] and Von Bertalanffy (9,10), Buchanan three-phase (11) and more recently Huang models (12). 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 amongst data can occur due to events such as temperature drift during time measurements or an overused tungsten lamp in a spectrophotometer. If one were to count the number of animals per year in a given area the data would be highly autocorrelated and nonindependence as the number of animals in a current year would be highly dependent upon the number of animals in the previous year [13]. 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 this yeast would be used. The method calculates the level of significance according to the method outlined by Draper and Smith [14].

METHODS

Acquisition of Data

In order to process the data, the cellular growth of the yeast data were obtained from the works of Rigo et al.[15], from Figures 2 and 3 which shows the effect of different concentration of the substrate catechol on the growth of *Candida parapsilopsis* measured over several hours. The graphs were scanned and electronically processed using WebPlotDigitizer 2.5 [16], and then assessed using several growth model where the Buchanan model was found to be the best (**Fig. 1**) and the results published elsewhere.

Durbin-Watson test

The Durbin–Watson statistic calculates the level of significance according to the method outlined by Draper and Smith [14].

$$d=\frac{\sum_{t=2}^{T}\left(\hat{e}_t-\hat{e}_{t-1}\right)^2}{\sum_{t=1}^{T}\hat{e}_t^2} \qquad (1)$$

As usual the hypothesis H_0: ρ= 0 versus the alternative $H1$: ρ > 0 is tested. The statistic is approximately equal to 2(1− p). 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 (**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 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. This is because most regression problems involving time series data exhibit positive autocorrelation. Autocorrelation amongst data can occur due to events such as temperature drift during time measurements or an overused tungsten lamp in a spectrophotometer. If one were to count the number of animals per year in a given area the data would be highly autocorrelated and nonindependence as the number of animals in a current year would be highly dependent upon the number of animals in the previous year [13,14,17].

The value of the Durbin-Watson statistics was d = 0.000213/0.000065=2.893. As usual the hypothesis H_0: ρ= 0 versus the alternative $H1$: ρ > 0 is tested. The statistic is approximately equal to 2(1− p). 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. The upper critical value d_U is 2.102 while the lower critical value d_L is 0.229. Since d was larger than the upper critical value then the null hypothesis is not rejected i.e. there appears to be no evidence of autocorrelation and the Buchanan three-phase model used for fitting the growth curve is adequate.

REFERENCES

[1]. Stoyanovsky DA, Goldman R, Claycamp HG, Kagan VE. Phenoxyl radical-induced thiol-dependent generation of reactive oxygen species: Implications for benzene toxicity. Arch Biochem Biophys. 1995;317(2):315–23.

[2]. Zwietering MH, Jongenburger I, Rombouts FM, Van't Riet K. Modeling of the bacterial growth curve. Appl Environ Microbiol. 1990;56(6):1875–81.

[3]. Baranyi J, Roberts TA. A dynamic approach to predicting bacterial growth in food. Int J Food Microbiol. 1994;23(3-4):277–94.

[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]. Halmi MIE, Shukor MS, Johari WLW, Shukor MY. Modeling the growth kinetics of *Chlorella vulgaris* cultivated in microfluidic devices. Asian J Plant Biol. 2014;2(1):7–10.

[6]. Ricker, F.J. 11 Growth Rates and Models. In: W.S. Hoar DJR and JRB, editor. Fish Physiology [Internet]. Academic Press; 1979 [cited 2014 Jun 27]. p. 677–743. Available from http://www.sciencedirect.com/science/article/pii/S1546509808600345

[7]. 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;115:513–85.

[8]. Richards, F.J. A flexible growth function for empirical use. J Exp Bot. 1959;10:290–300.

[9]. Babák L, Šupinová P, Burdychová R. Growth models of *Thermus aquaticus* and *Thermus scotoductus*. Acta Univ Agric Silvic Mendel Brun. 2012;60(5):19–26.

[10]. López S, Prieto M, Dijkstra J, Dhanoa MS, France J. Statistical evaluation of mathematical models for microbial growth. Int J Food Microbiol. 2004;96(3):289–300.

[11]. 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.

[12]. Huang L. Optimization of a new mathematical model for bacterial growth. Food Control. 2013;32(1):283–8.

[13]. Mcdonald JH, Dunn KW. Statistical tests for measures of colocalization in biological microscopy. J Microsc. 2013;252(3):295–302.

[14]. Draper NR, Smith H. Applied Regression Analysis. Wiley, New York; 1981.

[15]. Rigo M, Alegre RM, Vidal Bezerra JRM, Coelho N, Bastos RG. Catechol biodegradation kinetics using *Candida parapsilopsis*. Braz Arch Biol Technol. 2010;53(2):481–6.

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

[17]. 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):365–74.