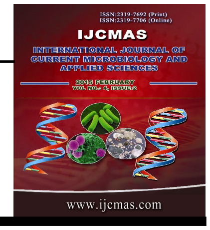


International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 4 Number 2 (2015) pp. 679-686**
http://www.ijcmas.com

Original Research Article

Lactobacillus helveticus growth modelling: behavior analysis of estimators of the parameters using measures of nonlinearity

Marcelo Teixeira Leite[1]*, Marcos Antonio de Souza Barrozo[2] and Eloizio Júlio Ribeiro[2]

[1]Center of Technology and Regional Development, Federal University of Paraíba, 58051-900, João Pessoa, PB, Brazil
[2]Chemical Engineering College, Federal University of Uberlândia, 38408-144, Uberlândia, MG, Brazil
**Corresponding author*

Keywords

Logistic model, Bias measure, Curvature Measures, *Lactobacillus helveticus*, Whey

A B S T R A C T

The logistic growth model was fitted to the experimental growth data of *Lactobacillus helveticus* cultivated on whey. Since it is a nonlinear model, the validity of the least squares estimators of the parameters was assessed using the bias measure of Box and the curvature measures of Bates and Watts. These measures demonstrated that there is statistical reliability for the least squares estimators of the parameters.The logistic model were able to accurately simulate the growth of *Lactobacillus helveticus* in the experimental range. The quadratic correlation coefficient shows that about 98.8% of the variability in the experimental results was explained by the model. The largest deviations occurred at the lag phase and in stationary phase, where the condition of instantaneous growth rate equal to zero is not predicted by the model.

Introduction

Lactic acid from whey

Lactic acid is a versatile chemical that finds applications in the food, cosmetics and pharmaceutical industries (Datta et al., 1995).Its isomers $L(+)$ and $D(-)$ can be polymerized to obtain compounds with different properties, depending on their intended application. In the medical area, lactic acid has been used in the production of biodegradable polymers, employed as scaffold in tissue transplants (Gao et al., 2009; Liu et al., 2005).

Lactic acid polymers are also used in the production of biodegradable packages (Bustos et al., 2004).

Currently, all lactic acid manufacture is based on carbohydrate fermentation technology (Datta and Henry, 2006). The microorganisms most frequently used are *Lactobacillus delbrueckii*, *Lactococcus lactis*, *Lactobacillus casei* and *Lactobacillus helveticus*. The carbon source commonly used can be either sugar in pure form such as glucose, sucrose, lactose etc., or sugar-containing materials such as whey, molasses and starch from crops or wastes (John et al., 2007). Whey is the most common substrate for production of lactic acid from renewable sources (Hofvendahl and Hahn-Hagerdal, 2000).

In our previous work, we used the response surface methodology to find the most suitable conditions to produce lactic acid from whey using a strain of *Lactobacillus helveticus* (Leite et al., 2012). In the present work a logistic model has been used to represent the experimental growth data of *Lactobacillus helveticus* cultivated in whey. As this model is nonlinear care should be taken when estimating its parameters from experimental data. In some situations, the estimators (especially, confidence intervals) may not be appropriate. Thus some procedures are available in the literature to validate the statistical properties of the least squares (LS) estimators of nonlinear models. Box (1971) presented a useful formula for estimating the bias in the LS estimators; Bates and Watts (1980) developed new measures of nonlinearity based on the geometric concept of curvature.

The present study is focused on use of nonlinearity measures to analyze the parameters of the logistic model applied to the experimental growth data of *Lactobacillus helveticus* cultivated in whey.

Nonlinear models and measures of nonlinear behavior

A question that has occupied researchers is how well some specified model fits the data. Extensive methodology has been developed to investigate whether a proposed model provides a good description of the data and if there is statistical reliability for the estimated parameters.

For a linear regression model, the least squares estimators of the parameters are linear combinations of the random variable. Consequently, the estimators are unbiased, normally distributed, and have the minimum possible variance. These properties are generally accepted to be the most desirable properties that estimators may have (Searle, 1971). However, for nonlinear models the parameter estimators are not linear combinations of the random variable and do not possess these qualities. The estimators achieve these properties only asymptotically, that is, as the sample sizes approach infinity. In small samples these properties are usually unknown (Jennrich, 1969).When theleast squares estimators of an onlinear model present small bias, near-normal distribution and true variances close to those calculated by the asymptotic variance-covariance matrix, it can be stated that the estimators exhibit behavior *close tolinear* (Seber and Wild, 1989). The closer is the linear behavior of an onlinear model, the more accurate the asymptotic results and consequently the more reliable inferences (Arnosti, et al., 1999).The extent of the bias, the deviation from normal distribution and the excess variance differ greatly from model to model. The extent of nonlinear behavior can be evaluated through nonlinearity measures. Thus, measures of nonlinearity are expressions used to evaluate the adequacy of the linear approximation of an onlinear model and their effects on he inferences (Ratkowsky, 1983).

One of the first relevant attempts to quantify the nonlinearity of a nonlinear regression was presented by Beale (1960), who proposed four measures. However, these measures should not be used in practice, since they tend to underestimate the true nonlinearity (Seber and Wild, 1989). Box (1971) presented a formula for estimating the bias in the LS estimators, and Gillis and

Ratkowsky (1978) concluded that this formula not only predicted bias to the correct order of magnitude but also gave a good indication of the extent of nonlinear behavior of the model, using simulation studies. Bates and Watts (1980) developed measures of nonlinearity based on the geometric concept of curvature. They demonstrated the relationship between their measures and those of Beale (1960), explained why measures of Beale generally tend to underestimate the true nonlinearity and also showed that the bias measure of Box is closely related to the their measure of parameter-effects nonlinearity (PE).

The bias measure of Box and the curvature measures of Bates and Watts are described summarized below.

Bias measure of box

A nonlinear regression model can be represented by:

$y_i= f(x_i,\boldsymbol{\theta}) + \varepsilon_i,\ i=1, ..., n$ (1)

where y_i is the *i*th response, x_i is a vector of known variables, θ is a p x 1 vector of unknown parameters, the response function f is a known, scalar-valued function that is twice continuously differentiable in θ, and the errors ε_i are independent and identically distributed normal random variables with mean zero and variance σ^2.

Box (1971) presented an equation to estimate the bias of least squares estimators of a nonlinear regression model:

$$\text{Bias}\left(\hat{\theta}\right) = -\frac{\sigma^2}{2}\left(\sum_{i=1}^{n}\mathbf{F}(\boldsymbol{\theta})\mathbf{F}(\boldsymbol{\theta})^t\right)^{-1}\sum_{i=1}^{n}\mathbf{F}(\boldsymbol{\theta})\text{tr}\left[\left(\sum_{i=1}^{n}\mathbf{F}(\boldsymbol{\theta})\mathbf{F}(\boldsymbol{\theta})^t\right)^{-1}\mathbf{H}(\boldsymbol{\theta})\right] \quad (2)$$

Where F (θ)is the vector of first order derivatives of f(xi, θ) called velocity vector and H(θ)is a matrix of second order derivatives with respect to each element ofθ.

In practice, $\hat{\theta}$ and $\hat{\sigma}^2$ are usually used in place of the unknown quantities.

It is common to express the Box's bias as a percentage value:

$$\%\text{Bias}\left(\hat{\theta}\right) = \frac{100.\text{Bias}(\hat{\theta})}{\hat{\theta}} \quad (3)$$

The bias expressed as a percentage of the least squares estimate is a useful quantity, as an absolute value in excess of 1% appears to be a good rule of thumb for the identification of which parameter, or parameters, are responsible for the nonlinear behavior (Souza et al., 2012).

Curvature measures of Bates and Watts

Bates and Watts (1980) divide the concept of nonlinearity into two parts: intrinsic nonlinearity (IN) and parameter effects nonlinearity (PE). The intrinsic nonlinearity (IN) measures the curvature of the solution locus in sample space, where the locus represents all possible solutions to the estimation problem. The least squares solution is the point on the solution locus closest to the observed response vector **y**. For a linear regression model, IN is zero since the solution locus is straight (a line, plane or hyperplane). For a nonlinear regression model, the solution locus is curved, with IN measuring the extent of that curvature.

The parameter-effect nonlinearity (PE) is a measure of the lack of parallelism and the inequality of spacing of parameter lines on the solution locus at the least-squares solution. The PE value is a scalar quantity, representing the maximum value of the effect of the parametrization, obtained from a three-dimensional array called *acceleration array*. This array is a scaled version of the matrix of second order derivatives with respect to the parameters. For a linear regression model, the parameters appear linearly, so that the second order derivatives are all zero, resulting in a PE value of zero. For nonlinear regression models having a given value of IN, PE increases the more the model departs from a linear model. In models with behavior far from linear, in which IN is small, the nonlinearity is mainly due to the effect of the parameters PE. In such cases, it is necessary to reparametrize the model (Ratkowsky, 1990).

The statistical significance of these measures can be evaluated by comparing the IN and PE values with $1/2\sqrt{F}$, where $F = F(\alpha, n\text{-}p, p)$ is the inverse of Fisher´s probability distribution obtained at significance level α, p is the number of parameters and n is the number of data points. The value $1/2\sqrt{F}$ may be regarded as the radius of the curvature of the $100(1-\alpha)\%$ confidence region. Hence, the solution locus may be considered to be sufficiently linear within an approximately 95% confidence region if $IN < 1/2\sqrt{F}$ ($\alpha = 0.05$). Similarly, if $PE < 1/2\sqrt{F}$, the projected parameter lines may be regarded as being sufficiently parallel and uniformly space, i.e., the least squares estimates of the parameter do not depend on the user being able to supply a good initial prediction and the tests of parameter invariance will be adequate (Babetto et al., 2011).

Details about the development, procedure and equations for determining IN and PE are found in Bates and Watts (1980).

The logistic model

The Belgian mathematician Verhulst (1838) presented a sigmoid function called *logistic growth model* to describe the self-limiting growth of a biological population:

$$\frac{dX}{dt} = \mu_{max}.X\left(1 - \frac{X}{K}\right) \quad (4)$$

Where

X is the cell mass concentration (g/L);

dX/dt is the growth rate (g/L.h);

μ_{max} is the maximum specific growth rate (h^{-1});

K is the saturation density or carrying capacity, and represents the theoretical maximum cell mass concentration(g/L) that can be achieved in a given environment.

The integrated form of Eq.(4)is given by:

$$X = \frac{X_0.K.\exp(\mu_{max} \cdot t)}{K + X_0(\exp(\mu_{max} \cdot t) + 1))} \quad (5)$$

Where X_0 is the cell mass concentration at t=0.

The logistic model is one of unstructured models commonly used to describe microbial growth. However, this model is based on certain assumptions which may not be satisfied in some populations:(i) there is an initial stable age distribution among individuals, (ii) the maximum specific growth is achieved under existing conditions, (iii) the relationship between the specific rates of birth and death and the population density is linear. The discrepancies between these assumptions and the actual characteristics of the population may become inadequate adoption of this model.

Materials and Methods

Inoculum, medium, fermentation and analyses

Lactobacillus helveticus ATCC 15009 was supplied by the André Tosello Foundation (Campinas, SP, Brazil). Stock cultures were maintained on MRS broth and deep frozen at −18◦C. As required, these cultures were thawed and reactivated by two transfers in MRS broth (24 h, 37 ºC, 120 rpm). The fermentation medium was made up of deproteinized reconstituted cheese whey, supplemented with yeast extract at three initial concentrations of lactose: 52 g/L, 82 g/L e 112 g/L. Batch fermentations were carried out at 150 rpm for 32 hours under anaerobic conditions in a 3.0 L fermentor NBS Bioflo 110 (New Brunswick Scientific, USA), in which 1800mL of sterile culture medium was inoculated with 200mL of seed culture at 1.0 × 107 cells/mL. In each fermentation, seventeen samples were withdrawn at regular time intervals (one sample every two hours from t =0 tot= 32hours)and the biomass concentration was estimated by optical density readings at 650 nm. More details on the experimental procedures are available in our previous work (Leite et al., 2012).

Model fitting to the experimental data

The logistic model (Eq.5) was fitted to the experimental growth data of *Lactobacillus helveticus* cultivated in whey. The parameters K and μ_{max} were estimated by nonlinear estimation, using the Statistica software. The cell mass concentration at t=0, X_0, was determined experimentally.

Nonlinearity measures of the parameters K and μ_{max} were calculated using a computer program in FORTRAN language.

Result and Discussion

Parameters estimation and nonlinearity measures

Table 1 presents the results obtained by the least squares parameters estimation for the logistic model (Eq. 5). These results include the estimated parameter values, as well as the respective values of the quadratic regression coefficient (R^2), the intrinsic curvature measure (IN), the parameter-effects (PE) measure and the Box's bias percentage.

As can be seen, there was obtained high values of $R^2$for the three initial concentrations of lactose. However, as mentioned previously a high R^2valueis not sufficient to ensure the statistical validity of the parameters obtained in a nonlinear regression. Other properties are desirable for nonlinear models such as the least squares (LS) estimators, whose parameters are almost unbiased, normally distributed, and whose variances are close to the minimum variance (Lira et al., 2010).Therefore, the adequacy of the linear approximation and its effect on the inferences was evaluated by the bias measure of Box and the curvature measures of Bates and Watts.

The significance levels of IN and PE were evaluated by comparing the estimates with the critical value $1/2\sqrt{F}$ in which F = F (p,n-p,α) is according to the F distribution with p and n-p degrees of freedom at significance level α.

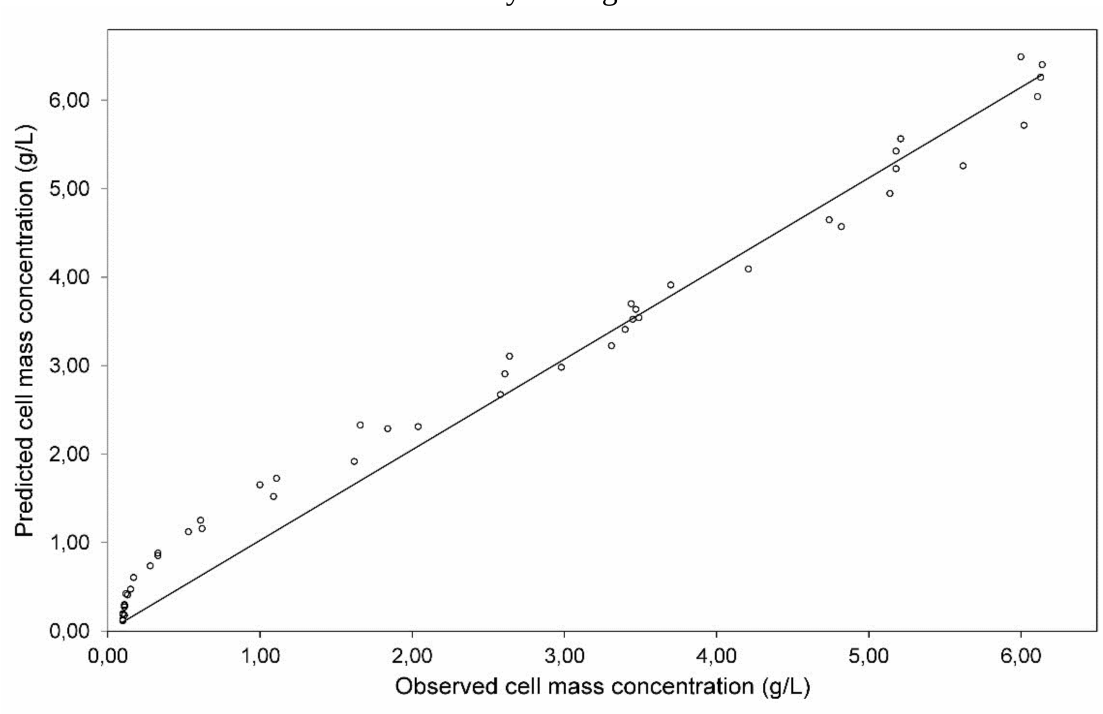
Table.1 Results of least square estimation and nonlinearity measures

Run	S_0 (g/L)	μ_{max} (h^{-1})	K(g/L)	R^2	IN	PE	% Box's Bias μ_{max}	% Box's Bias K
1	52	0.209	3.83	0.9881	0.094	0.34	0.09	0.24
2	82	0.244	6.64	0.9878	0.090	0.24	0.07	0.14
3	112	0.215	5.89	0.9872	0.092	0.34	0.07	0.25

Table.2 Theoretical (K) and experimental (X_{max}) maximum dried cell mass concentrations

	S_0 (g/L)		
	52	82	112
X_{max} (g/L)	3.49	6.14	5.21
K (g/L)	3.83	6.64	5.89
Error (%)	9.74	8.14	13.05

Figure.1 Experimental versus predict values of cell mass concentration predict by the logistic model



Thus, for α=0.05, n=17 (number of data points) and p=2 (number of parameters), F= 3.68 and therefore $1/2\sqrt{F}$ =0.26.It can be observed in Table 1 that all calculated values of IN are not significant (IN < $1/2\sqrt{F}$). Hence, the solution locus may be considered to be sufficiently linear within an approximately 95% confidence region.

At the initial lactose concentration equal to 82 g/L, the calculated value of PE was lower than the critical value $1/2\sqrt{F}$. On the other hand, the PE values exceeded the critical value at the initial lactose concentrations equal to 52 g/L and 112 g/L. However, these values were only slightly higher than the critical value and therefore this information may be insufficient for an accurate evaluation of the validity of the linear approximation. In this case, an analysis of the percentage of Box´s bias becomes very useful because, as stated previously, this measure can be used to identify parameters that are responsible for the nonlinear behavior. Parameters whose Box's bias percentage exceeds 1% are considered biased and responsible for the nonlinear behavior of the model. As can be seen in Table 1, all calculated values of the Box's bias percentage are less than1%. Thus, it can be considered that the estimates of μ_{max} and K are valid, the model has close to linear behavior and a reparametrization in this case is not necessary.

Model evaluation

Figure 1 shows the cell concentrations obtained experimentally and those predicted by logistic model, for the initial lactose concentrations of 52, 82 e 112 g/L. The quadratic correlation coefficient R^2 (Table 1) shows that about98.8% of the variability in the experimental results was explained by the model. The largest deviation occurred at the beginning of fermentation and in the stationary phase, where the concentrations remain fairly constant and therefore dX/dt is zero. The sigmoidal curve described by the logistic equation shows that this situation is only expected in the asymptotic condition, i.e., when $t \rightarrow \infty$.

The theoretical maximum cell concentrations, represented by Kin Eq.(5), were very close to their respective experimental maxima values X_{max}, as can be seen in Table2. This confirms the good suitability of the model to experimental results.

The bias measure of Box and the curvature measures of Bates and Watts showed that there is statistical reliability for the least squares estimators of the parameters of the logistic model fitted to the growth data of *Lactobacillus helveticus*, under the conditions used in this study.

The logistic model represented with good accuracy the growth of *Lactobacillus helveticus*. About98.8% of the variability in the experimental results was explained by the model. The largest deviations occurred at the beginning of fermentation and in the stationary phase, where the condition of growth rate equal to zero is not predicted by the model.

References

Arnosti Junior, S., Freire, J. T., Sartori, D. J. M., Barrozo, M. A. S. 1999. Equilibrium moisture contente of *Brachiaria brizantha*. *Seed Sci. Technol.*, 27, 273-282.

Babetto, A.C., Freire, F.B., Barrozo, M.A.S.,Freire, J.T.2011. Drying of garlic slices: Kinetics and nonlinearity measures for selecting the best equilibrium moisture content equation. *J. Food Eng.*, 107, 347-352. DOI: 10.1016/j.jfoodeng.2011.07.007

Bates, D. M. and Watts, D. G. 1980. Relative curvature measures of nonlinearity. *J. R. Statist. Soc.*, 42, 1-25.

Beale, E.M.L. 1960. Confidence regions in nonlinear estimation. *J. R. Statist. Soc.*, B, London, 22, 41-76.

Box, M. J. 1971. Bias in nonlinear estimation. *J. R. Statist. Soc.*, 33, 171-201.

Bustos,G., Moldes, A. B., Alonso, J. L., Vázquez, M. 2004.Optimization of D-lactic acid production by *Lactobacillus coryniformis* using response surface methodology, *Food Microbiol.*, 21(2), 143–148. DOI: 10.1016/S0740-0020(03)00061-3

Datta,R., Tsai, S. P., Bonsignore, P., Moon, S. H., Frank, J. R. 1995. Technological and economic potential of poly(lactic acid) and lactic acid derivatives, *FEMS Microbiol. Rev.*, 16**(**2-3), 221–231. DOI: 10.1016/0168-6445(94)00055-4

Datta, R., Henry, M. 2006. Lactic acid: recent advances in products, processes and technologies — a review. *J. Chem. Technol. Biot.*, 81**(**7), 1119–1129. DOI: 10.1002/jctb.1486

Gao, M. T., Shimamura, T., Ishida, N., Takahashi, H. 2009. Fermentative lactic acid production with a metabolically engineered yeast immobilized in photo-crosslinkable resins, *Biochem. Eng. J.*, 47**(**1–3), 66–70. DOI: 10.1016/j.bej.2009.07.004

Gillis, P.R.; Ratkowsky, D.A. 1978. The behavior of estimators of the parameters of various yield-density relationships. *Biometrics*, Washington, D.C., 34,191-198.

Hofvendahl, K., Hahn-Hagerdal, B. 2000. Factors affecting the fermentative lactic acid production from renewable resources, *Enzyme Microb. Tech.*, 26**(**2–4), 87–107.

Jennrich, R.I. 1969. Asymptotic properties of nonlinear least squares estimators. *Ann. Math. Stat.*, Beochwood, 40, 633-643.

John, R. P., Nampoothiri, K. M., Pandey, A. 2007. Fermentative production of lactic acid from biomass: an overview on process developments and future perspectives. *Appl. Microbiol. Biot.*, 74, 524-534. DOI 10.1007/s00253-006-0779-6

Leite, M. T., Barrozo, M. A. S., Ribeiro, E. J. 2012. Canonical Analysis Technique as an Approach to Determine Optimal Conditions for Lactic Acid Production by *Lactobacillus helveticus* ATCC 15009, *Int. J. Chem. Eng.*Article ID 303874, 9 pages.DOI:10.1155/2012/303874.

Lira, T. S., Santos, K. G., Murata, V. V., Gianesella, M., Barrozo, M. A. S. 2010. The use of nonlinearity measures in the estimation of kinetic parameters of sugarcane bagasse pyrolysis. *Chem. Eng. Technol.*, 33,1699-1705. DOI: 10.1002/ceat.201000137

Liu, X., Smith, L., Wei, G., Won, Y., Ma, P. X. 2005. Surface engineering of nano-fibrous poly(L-lactic acid) scaffolds via self-assembly technique for bone tissue engineering, *J. Biomed. Nanotech.*, 1**(**1), 54–60.DOI:10.1166/jbn.2005.013

Ratkowsky, D. A. 1983. *Nonlinear Regression Analysis*. Marcel Dekker Inc., New York.

Ratkowsky, D.A. 1990. *Handbook of Nonlinear Regression Models*. New York: Marcel Dekker, Inc.

Searle, S.R. 1971. *Linear Models*. New York: John Wiley and Sons.

Seber, G.A.F.; Wild, C. J. 1989.*Nonlinear Regression*. New York: John Wiley and Sons.

Souza, G. F. M. V., Miranda, R. F., Arruda, E. B., Mendoza, O. S. H., Barrozo, M. A. S.2012. Drying kinetics of silica gel: statistical discrimination using nonlinearity measures. *Chem. Eng. Technol.*, 35, 797-802. DOI: 10.1002/ceat.201100342

Verhulst, Pierre-François. 1838. Notice sur la loi que la population poursuit dans son accroissement. *Correspondance mathématique et physique*, 10, 113–121.