American Journal of Engineering Research (AJER)
E-ISSN: 2320-0847 p-ISSN: 2320-0936
Volume-8, Issue-1, pp-01-09
www.ajer.org
Research Paper Open Access

Sigmoidal Models Fitted to the Saccharomyces cerevisiae Growth Curve: Statistical Analysis Using Measures of Nonlinearity

Marcelo Teixeira Leite
Department of Sugar and Ethanol Technology, Federal University of Paraíba, Brazil, Zip code 58058-600
Corresponding Author:Marcelo Teixeira Leite

ABSTRACT: *Sigmoidal models were compared to describe the growth curve of Saccharomyces cerevisiae in batch ethanol fermentation of sugarcane molasses. Five classical functions and some of their reparameterizations were fitted to three experimental data sets: logistic, Gompertz, Chapman-Richards, Morgan-Mercer-Flodin and a Weibull-type model. Since these models are nonlinear, measures of nonlinearity were used to evaluate the statistical properties of the least squares estimators. The measures used were the intrinsic (IN) and parameter-effects (PE) curvatures of Bates and Watts, the bias measure of Box, and the Hougaard's measure of skewness. Among the models analyzed, only a reparameterization of the Weibull-type model presented close to linear behavior, ensuring the statistical validity of the parameters estimated by the method of least squares.*

KEYWORDS*: Saccharomyces cerevisiae, ethanol fermentation, sigmoidal models, measures of nonlinearity.*

Date of Submission: 20-12-2018 Date of acceptance: 04-01-2019

I. INTRODUCTION

Saccharomyces cerevisiae growth

Saccharomyces cerevisiae, commonly known as baker's yeast, is well-known and commercially significant among the yeasts. This microorganism has been used for a long time in the food industry, mainly in the baking and alcoholic beverages industries. It is also used to produce ethanol by using fermentation of sugarcane juice or molasses and, more recently, from lignocellulosic biomass, which further increases the importance of this microorganism in the biofuel industry.

Yeast cell grows in three main phases: lag, exponential and stationary. When a culture of yeast cells is inoculated in a fresh growth medium, it enters a lag phase or adaptation time, where the cells are biochemically active but not dividing. After this, the cell growth increases slowly initially, in a positive acceleration phase; then it increases rapidly, approaching an exponential growth rate; but then, it declines in a negative acceleration phase until a zero-growth rate, when the population stabilizes. This slowdown in the rate of growth results in an increase of environmental resistance, which becomes proportionately more important at higher cell population densities. This type of growth is termed "density-dependent", since the growth rate depends on the number of cells available in the population. The point of stabilization, or zero growth rate, is termed the "saturation value" or "carrying capacity" of the environment for that microorganism (Allaby, 2014). If the cell concentration is plotted against time, a typical sigmoidal or "S-shaped" growth curve is obtained.

Processes producing sigmoidal or "S-shaped" growth curves are widespread in biology, agriculture, engineering, and economics. Such curves start at some fixed point and increase their growth rate monotonically to reach an inflexion point; after this, the growth rate decreases to approach asymptotically some final value. Numerous mathematical functions have been proposed for modeling sigmoidal growth curves, many of which are claimed to have some underlying theoretical basis. Among these are the logistic equation and the Gompertz model (Ratkowsky, 1983).

In this work, sigmoidal models were compared to describe the growth curve of *Saccharomyces cerevisiae* in batch ethanol fermentation of sugarcane molasses. Five sigmoidal functions were used: logistic, Gompertz (1825), Chapman-Richards (Pienaar and Turnbull, 1973), Morgan-Mercer-Flodin (1975) and a model derived from the Weibull (1951) distribution, here designated as Weibull-type model. Since these models are nonlinear, measures of nonlinearity were used to validate the inference results based on asymptotic approximations assumed in the least squares nonlinear estimates. The measures utilized were the intrinsic (IN) and parameter-effects (PE) curvatures of Bates and Watts (1980), the bias measure of Box (1971), and the Hougaard's measure of skewness (1985). The paper also discusses the misuse of the determination coefficient R^2 as a measure of goodness of fit in nonlinear models.

Nonlinear regression models

Sigmoidal growth models are nonlinear regression models. The concept of nonlinear model and its consequences to statistical inference can be explained by using the following regression model:

$\boldsymbol{y} = f(\boldsymbol{x}, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}$ (1)

Where **y** is the vector of the response variables, **x** is the vector of the independent variables, **θ** is the vector of regression parameters, **ε** is the vector of the random errors and f (**x**, **θ**) is a function of the independent variables and the parameters, known as regression function. Generally, it is assumed that the errors ε_i are independently and normally distributed with mean zero and constant variance. When $\partial f(x_i, \boldsymbol{\theta})/\partial \theta_j$ is independent of **θ** or $\partial f^2(x_i, \boldsymbol{\theta})/\partial \theta_j^2 = 0$, the regression model is called *linear* with respect to the parameters. If at least one derivative of **y** with respect to a parameter is a function of that parameter, the regression model is called *nonlinear*(Seber and Wild, 1989). An important consequence of the fact that a regression model is nonlinear is that the least squares estimators of its parameters do not possess the desirable properties of their counterparts in linear regression models, that is, they are not unbiased, minimum variance, normally distributed estimators (Ratkowsky, 1983).

Nonlinear estimation

Assuming that the regression function in equation (1) is twice continuously differentiable in **θ**, the residual sum of squares is given by:

$S(\boldsymbol{\theta}) = \sum_{i=1}^{n}[y_i - f(x_i, \boldsymbol{\theta})]^2$ (2)

The least-squares estimators $\hat{\theta}$ of the parameters are the values of **θ** which minimize the sum of squares S(**θ**). For linear models, there is an analytical solution which leads to the minimum value of S(**θ**). For nonlinear models, the search for the minimum value of S(**θ**) is performed by iterative numerical methods, and the algorithms used is based on a linear approximation to the regression function $f(x_i, \boldsymbol{\theta})$ (Draper and Smith, 1998). Thus, in nonlinear estimation, the degree of nonlinearity must be sufficiently small so that the usual estimation techniques developed for linear regression can be used as a reliable approximation for the nonlinear model. Both the effectiveness of least squares algorithms and the validity of inferences made regarding the parameters of a nonlinear model will be affected by the closeness of the linear approximation to the model (Bates and Watts, 1980). The closer the linear behavior of a nonlinear model is, the more accurate the asymptotic results and consequently the more reliable inferences are. There are some nonlinear regression models whose estimators come close to being unbiased, normally distributed, minimum variance estimators. Such models have been termed *close to linear* models by Ratkowsky (1990).

Measures of nonlinear behavior

Since most asymptotic inferences for nonlinear regression models are based on analogy with linear models, and since these inferences are approximate, some measures of nonlinearity have been proposed as a guide for understanding how good linear approximations are likely to be (El-Shaarawi and Piegorsch, 2002). The most used measures of nonlinearity are the curvature measures of Bates and Watts, the bias measure of Box, and the Hougaard's measure of skewness. These measures are described summarized as follows. For further details, see the original works.

Curvature measures

Bates and Watts (1980) divide the concept of nonlinearity into two parts: intrinsic nonlinearity (IN) and parameter-effects nonlinearity (PE). Relative intrinsic and parameter-effects curvatures can be used to quantify the global nonlinearity of a nonlinear regression model. The intrinsic nonlinearity (IN) measures the curvature of the solution locus in sample space. 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 (Ratkowsky, 1990).

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. As a result, the parameter-effects curvature can be reduced by reparameterization of the model, whereas intrinsic curvature is an inherent property of the model that cannot be affected by reparameterization (Bates and Watts, 1980).

Bias and skewness

In practice, only a few of the parameters might dominate the global nonlinearity, in which case these parameters are the reparameterization parameters of interest. Unfortunately, the global nonlinearity measures of

Bates and Watts do not differentiate the parameters based on their contribution to the overall curvature. Manifestations of nonlinear behavior include significant bias and skewness. Hence, it is essential to estimate at least these basic statistical properties of the parameter estimates in order to identify the reparameterization parameters of interest (Gebremariam, 2014).

The bias and skewness of the parameter estimates of a nonlinear regression model can be estimated by using the bias measure of Box and the Hougaard measure of skewness. The Box's bias represents the discrepancy between the estimates of the parameters and the true values. Skewness is a measure of lack of symmetry. Hougaard's measure of skewness can be employed to assess whether a parameter is close to linear or whether it contains considerable nonlinearity, because of the close link between the extent of nonlinear behavior of an estimator and the extent of nonnormality in the sampling distribution of this estimator (Ratkowsky, 1990).

II. MATERIAL AND METHODS

Inoculum, medium, fermentation and analyses

This work utilized the experimental results available in the study published by Tosetto(2002). Below follows a brief description of the methodology used to obtain these results. For further details, see the original work.

The yeast strain used was *Saccharomyces cerevisiae* (Mauriferm Y904, Mauri, Brazil). The inoculum preparation included suspending 120 g of the dried yeast in 1200 g of water, maintained at 100 rpm and 34 °C for 30 minutes. Three different sugar mills supplied the molasses used as a fermentation medium and in this work are identified merely as molasses A, B, and C. Table I shows the physical and chemical properties of these molasses. Before fermentation, the molasses were diluted with deionized water to reduce the concentration of Total Reducing Sugars to 200 g/L. For each diluted molasses, two batch fermentations were carried out at 100 rpm and 34 ºC in a 6.0 L working volume fermenter, which included 1200 g of inoculum suspension to 4800 ml of sterile diluted molasses. Samples were withdrawn at regular time intervals (one sample at every one hour up to ten hours) and measurements of cell dry weight quantified the biomass concentration.

Table I - Chemical and physical properties of the molasses utilized in fermentation media.

Molasses	**pH**	**Sulphuric Acidity (g/L)**	**Density (g/cm^3)**	**Brix (%)**	**Total Reducing Sugars (g/L)**	**Apparent Purity (%)**
A	5.88	5.58	1.3505	81.6	718.14	65.16
B	6.08	4.98	1.3584	83.4	662.52	58.48
C	6.02	7.62	1.3558	82.8	667.36	59.45

Models

The sigmoidal models fitted to the experimental growth data of *Saccharomyces cerevisiae* are shown in Table II. Due to the high initial cell concentrations used in the fermentations (see Table III, in Results and Discussion), all the models used in this work allow a nonzero lower asymptote.

Table II – Sigmoidal models fitted to the *Saccharomyces cerevisiae* growth data.

Model	**Equation**	
Logistic	$y = \delta + \alpha/(1 + \exp(\beta - \gamma X))$	(3)
Gompertz	$y = \delta + \alpha\exp(-\exp(\beta - \gamma X)$	(4)
Chapman-Richards	$y = \alpha/(1 + \exp(\beta - \gamma X))^{1/\delta}$	(5)
Morgan-Mercer-Flodin	$y = (\beta\gamma + \alpha X^{\delta})/(\gamma + X^{\delta})$	(6)
Weibull	$y = \alpha - \beta\exp(-\gamma X^{\delta})$	(7)

Measures of nonlinearity

In this study, the parameter estimates and their respective measures of nonlinearity were obtained by using the NLIN procedure of the SAS software. Details about the development, procedure, and equations for determining the curvature measures of nonlinearity of Bates and Watts(1980), the bias measure of Box(1971), and the Hougaard (1985) measure of skewness are found in the original works.

Curvature measures

The statistical significance of the intrinsic nonlinearity (IN) and parameter-effects nonlinearity (PE) were evaluated by comparing these values with $1/\sqrt{F}$, where $F = F\,(\alpha, n - p, p)$ is the inverse of Fisher´s probability distribution obtained at significance level α = 0.05, p is the number of parameters and n is the number of observations. The value $1/\sqrt{F}$ 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 $<1/\sqrt{F}$ (α= 0.05). Similarly, if PE $<1/\sqrt{F}$, the projected parameter lines may be regarded as being sufficiently parallel and uniformly spaced (Ratkowsky, 1990).

Box's bias

The Box's bias in the least square estimates of the parameters in nonlinear regression can be expressed as a percentage of the least square estimate. This *percentage bias* is estimated by:

$$\% \text{ Bias}(\hat{\theta}) = \frac{100.\text{Bias}(\hat{\theta})}{\hat{\theta}} \quad (8)$$

According to Ratkowsky (1983), a percentage bias greater than 1% in absolute value is considered to be significantly nonlinear.

Hougaard's skewness

The degree to which a parameter estimator exhibits nonlinear behavior can be assessed with Hougaard's measure of skewness, g_{1i}. According to Ratkowsky (1990), if $|g_{1i}| < 0.1$, the estimator of the parameter is very close to linear and, if $0.1 < |g_{1i}| < 0.25$, the estimator is reasonably close to linear. For $|g_{1i}| > 0.25$, the skewness is very apparent, and $|g_{1i}| > 1$ indicates considerable nonlinear behavior.

III. RESULTS AND DISCUSSION

Table III shows the results of the growth of Saccharomyces cerevisiae in the three fermentation media.

Table III – Cell dry weight concentration (X) of *Saccharomyces cerevisiae* in the three-fermentation media A, B and C. Each data set corresponds to the arithmetical mean of two runs.

t (h)	X (g/L)		
	A	**B**	**C**
0	17.50	17.45	16.82
1	16.78	17.01	16.57
2	17.55	17.74	17.28
3	18.53	18.34	18.42
4	20.13	19.74	19.97
5	22.10	21.20	21.86
6	23.94	22.36	23.41
7	24.65	23.09	24.34
8	24.88	23.35	24.49
9	24.79	23.54	24.34
10	24.52	23.41	24.41

Measures of nonlinearity

Sigmoidal models in Table II were fitted to the experimental results shown in Table III. Tables IV, V and VI show the least squares parameter estimates with the respective values for the measures of nonlinearity to the three experimental data sets. For all models, IN $<1/\sqrt{F}$ (α=0.05) and therefore the solution locus may be considered to be sufficiently linear within an approximately 95% confidence interval. On the other hand, all models exhibited high parameter-effects curvature (PE $<1/\sqrt{F}$). This indicates that at least one parameter in each model is departing from linear behavior, and the Hougaard's skewness and Box's bias indicate which parameter or parameters are responsible. According to Ratkowsky[12], parameter estimates that present a percentage bias greater than 1% in absolute value are considered to be significantly nonlinear. Similarly, a value of the standardized Hougaard's skewness measure greater than 0.25 in absolute value indicates nonlinear behavior. As can be seen in Tables IV, V and VI, according to these guidelines it is possible to conclude that the parameter estimates responsible for the far from linear behavior of the models are, in most cases, skewed but unbiased.

For the logistic model, the parameter estimates for each data set are unbiased. The nonlinear behavior demonstrated in Table IV can be attributed to the high skewness of the estimates of β and γ. In Table V, the estimate of γ is skewed and can be considered responsible for the significant parameter-effects curvature. In Table VI, the skewness of all parameter estimates for this model can be considered reasonably close-to-linear. However, the parameter-effects curvature exceeded the critical value. In this case, the nonlinear behavior is probably due to the estimate of γ, that presented the highest skewness among the parameter estimates. As in the logistic model, all parameter estimates of the Gompertz model are unbiased. The high skewness of β and γ parameter estimates explains the far from the linear behavior of the model. For the Chapman-Richards model, the estimates of α, β and γ presented high skewness. The estimates of γ have very high biases and the estimates of β are moderately biased, except in Table V, where the Box' bias of β is less than 1%. The estimates of α, γ and δ for the MMF model are very skewed, and the estimates of γ have the most significant biases among all models. For the Weibull model, the high parameter-effect curvature is undoubtedly due to the estimates of γ.

Table IV: Statistical results of the least-squares estimation for the models in Table II. Fermentation medium A.

Model	IN	PE	P	Parameter	Estimate	Std. Error	Skewness	% Bias
Logistic	0.192	0.627	0.493	α	7.627	0.246	0.18	0.18
				β	5.298	0.575	0.34	0.90
				γ	1.201	0.125	0.35	0.90
				δ	17.196	0.175	-0.17	-0.05
Gompertz	0.247	0.786	0.493	α	-7.753	0.263	-0.23	0.20
				β	-3.860	0.417	-0.28	0.79
				γ	-0.790	0.083	-0.32	0.83
				δ	24.708	0.118	0.04	0.01
Chapman-Richards	0.335	6.856	0.470	α	7.693	0.383	0.26	0.44
				β	0.789	0.123	0.39	1.63
				γ	21.209	11.134	1.95	19.0
				δ	17.291	0.257	-0.14	-0.09
MMF	0.278	23.688	0.470	α	25.037	0.257	0.43	0.09
				β	17.268	0.243	-0.12	-0.06
				γ	1857.7	2094.3	3.87	76.8
				δ	5.084	0.758	0.46	1.71
Weibull	0.192	8.118	0.470	α	24.774	0.096	0.07	0.01
				β	7.680	0.185	0.10	0.08
				γ	0.004	0.002	0.94	5.39
				δ	3.470	0.248	0.25	0.45

ρ = critical curvature value ($\rho = 1/\sqrt{F}$), F (α, n-p, p) is the inverse of Fisher´s probability distribution obtained at significance level α = 0.05, p is the number of parameters and n is the number of observations.

Table V: Statistical results of the least-squares estimation for the models in Table II. Fermentation medium B.

Model	IN	PE	P	Parameter	Estimate	Std. Error	Skewness	% Bias
Logistic	0.152	0.660	0.493	α	6.388	0.207	0.235	0.19
				β	4.397	0.423	0.241	0.60
				γ	0.995	0.090	0.261	0.60
				δ	17.129	0.145	-0.219	-0.04
Gompertz	0.263	0.708	0.493	α	6.373	0.216	0.205	0.20
				β	2.629	0.268	0.317	0.77
				γ	0.673	0.065	0.294	0.72
				δ	17.325	0.123	-0.089	-0.03
Chapman-Richards	0.288	4.379	0.470	α	6.476	0.281	0.292	0.34
				β	0.633	0.078	0.250	0.87
				γ	11.056	3.756	1.257	7.98
				δ	17.260	0.173	-0.153	-0.05
MMF	0.197	13.717	0.470	α	23.845	0.207	0.476	0.07
				β	17.233	0.159	-0.13	-0.04
				γ	460.6	311.2	2.346	28.1
				δ	4.098	0.459	0.282	0.82
Weibull	0.141	5.411	0.470	α	23.491	0.078	0.127	0.01
				β	6.421	0.148	0.136	0.09
				γ	0.010	0.003	0.657	2.59
				δ	2.871	0.171	0.177	0.27

ρ = critical curvature value ($\rho = 1/\sqrt{F}$), F (α, n-p, p) is the inverse of Fisher´s probability distribution obtained at significance level α = 0.05, p is the number of parameters and n is the number of observations.

Table VI: Statistical results of the least-squares estimation for the models in Table II. Fermentation medium C.

Model	IN	PE	P	Parameter	Estimate	Std. Error	Skewness	% Bias
Logistic	0.136	0.568	0.493	α	7.978	0.222	-0.197	0.14
				β	4.419	0.377	-0.218	0.48
				γ	1.034	0.083	-0.239	0.48
				δ	16.570	0.160	-0.189	-0.04
Gompertz	0.199	0.738	0.493	α	-8.212	0.243	-0.171	0.40
				β	-3.162	0.266	-0.245	0.15
				γ	-0.661	0.052	-0.213	0.43
				δ	24.407	0.092	0.037	0.01
Chapman-Richards	0.356	5.333	0.470	α	7.915	0.389	0.297	0.43
				β	0.685	0.099	0.326	1.28
				γ	12.675	5.3147	1.5632	12.2
				δ	16.833	0.255	-0.172	-0.09
MMF	0.245	17.264	0.470	α	24.853	0.278	0.510	0.11

				β	16.818	0.242	-0.147	-0.07
				γ	575.9	502.7	3.029	46.9
				δ	4.340	0.597	0.374	1.32
Weibull	0.156	5.752	0.470	α	24.482	0.088	0.086	0.01
				β	7.840	0.175	0.111	0.08
				γ	0.008	0.002	0.697	2.95
				δ	3.055	0.186	0.201	0.30

ρ = critical curvature value ($\rho = 1/\sqrt{F}$), F (α, n-p, p) is the inverse of Fisher´s probability distribution obtained at significance level α = 0.05, p is the number of parameters and n is the number of observations.

The parameter-effects curvature can be reduced by reparameterization of the model. Therefore, some reparameterizations of the logistic, Gompertz, Chapman-Richards, MMF and Weibull models, proposed by Ratkowsky[12] were tested. The goal was to find ones that have smaller parameter-effects nonlinearity, whose behavior may thereby more closely approach that of a linear model. Reparameterization techniques are beyond the scope of this work and will not be discussed.

Tables VII, VIII and IX show the new parameterizations of the models and the statistical results. The values of the intrinsic curvature were omitted, because reparameterization does not alter the position of the solution locus. The reparameterizations of the logistic model (Eqs. 19 and 20) show an increase in parameter-effects curvatures in comparison to the original model in Tables IV, V and VI. For the first reparameterization of this model (Eq. 19) the high parameter-effects curvature is due to the estimates of α, β and γ. For the second reparameterization (Eq. 20) the nonlinear behavior is due to the estimates of γ and δ. For the first reparameterization of the Gompertz model (Eq. 21), the significant parameter-effect curvature can be attributed to the estimate of γ (Table VII), α (Table IX) or both (Table VIII). For the second reparameterization (Eq. 22), the nonlinear behavior is due only to the estimates of *ν*. The reparameterization of the Chapman-Richards model (Eq. 23) shows a considerable decrease in parameter-effect curvatures compared to the original model in Tables IV, V and VI. However, the values of PE are even larger than the critical values. Tables VII and IX show that the estimates of α are skewed and the estimates of β and γ are skewed and biased. In Table VIII, the only responsible for the nonlinear behavior of this model is the estimate of α. For the reparameterization of the Morgan-Mercer-Flodin model (Eq.24), Tables VII, VIII and IX show that the estimates of α, γ and δ are responsible for the far from linear behavior.

Unlike the original model (Eq. 7 in Table II), for the reparameterization of the Weibull model (Eq. 27 in Tables VII, VIII and IX) the parameter-effects curvatures are less than the critical values. Thus, this reparameterized model can be considered close to linear, that is, the parameter estimates are almost unbiased, normally distributed, and have close-to-minimal variance. The fit of this model to each data set is shown in Figure 1. The other models presented far from linear behavior due to skewed or biased parameter estimates, which implies invalid inference results based on asymptotic approximations for the least squares estimators. Therefore, it does not make sense to display graphs representing the fit of these models to the experimental data.

Table VII – Reparameterizations and statistical results of the least-squares estimation. Fermentation medium A.

Parameterization		PE	ρ	Parameter	Estimate	Std. Error	Skewness	% Bias
$y = \delta + (\frac{\alpha - \delta}{1 + \exp(\beta - \gamma.\log(x))})$	(19)	1.146	0.470	α	25.037	0.257	0.43	0.09
				β	7.527	1.127	0.49	1.76
				γ	5.084	0.758	0.46	1.71
				δ	17.268	0.243	-0.123	-0.06
$y = \delta + (\frac{\alpha - \delta}{1 + \exp(\gamma.(\beta - \log(x)))})$	(20)	1.097	0.470	α	17.268	0.243	-0.123	-0.06
				β	1.481	0.031	0.010	0.07
				γ	-5.084	0.758	-0.46	1.71
				δ	25.037	0.257	0.43	0.09
$y = \delta + \alpha.\exp(-\exp(\gamma.(x - \varepsilon)))$	(21)	0.751	0.493	α	-7.754	0.263	-0.23	0.01
				γ	0.790	0.083	0.32	0.83
				δ	24.708	0.118	0.04	0.01
				ε	4.885	0.092	-0.01	-0.06
$y = \delta + \exp(\eta - \nu.\theta^{x})$	(22)	6.593	0.493	δ	17.359	0.191	-0.08	-0.05
				η	2.030	0.041	0.069	0.10
				N	24.510	10.91	1.791	14.7
				θ	0.443	0.048	-0.114	-0.58
$y = \alpha.(1 - \exp(-\beta.x))^{\gamma} + \delta$	(23)	0.960	0.470	α	7.693	0.383	0.26	0.44
				β	0.789	0.123	0.39	1.63
				γ	3.054	0.525	0.38	1.71
				δ	17.291	0.257	-0.137	-0.09
$y = \frac{(\beta.\exp(\gamma) + \alpha.x^{\delta})}{(\exp(\gamma) + x^{\delta})}$	(24)	1.146	0.470	α	25.037	0.257	0.43	0.09
				β	17.268	0.243	-0.12	-0.06
				γ	7.527	1.127	0.47	1.76

				δ	5.084	0.758	0.46	1.71
$y = \alpha - \beta.\exp(-\exp(-\gamma).x^{\delta})$	(25)	0.398	0.470	α	24.774	0.096	0.07	0.01
				β	7.680	0.185	0.10	0.08
				γ	5.480	0.396	0.25	0.45
				δ	3.470	0.248	0.25	0.45

ρ = critical curvature value ($\rho = 1/\sqrt{F}$), F (α, n-p, p) is the inverse of Fisher´s probability distribution obtained at significance level α = 0.05, p is the number of parameters and n is the number of data observations.

Table VIII – Reparameterizations and statistical results of the least-squares estimation. Fermentation medium B.

Parameterization		PE	ρ	Parameter	Estimate	Std. Error	Skewness	% Bias
$y = \delta + \left(\frac{\alpha - \delta}{1 + \exp(\beta - \gamma.\log(x))}\right)$	(19)	1.144	0.470	α	23.845	0.207	0.476	0.07
				β	6.133	0.676	0.319	0.86
				γ	4.098	0.459	0.282	0.82
				δ	17.233	0.159	-0.13	-0.04
$y = \delta + \left(\frac{\alpha - \delta}{1 + \exp(\gamma.(\beta - \log(x)))}\right)$	(20)	1.051	0.470	α	17.233	0.159	-0.13	-0.04
				β	1.497	0.028	0.09	0.08
				γ	-4.098	0.459	-0.28	0.82
				δ	23.845	0.207	0.48	0.07
$y = \delta + \alpha.\exp(-\exp(\gamma.(x - \varepsilon)))$	(21)	1.098	0.493	α	-6.554	0.312	-0.39	0.39
				γ	0.639	0.080	0.34	1.08
				δ	23.386	0.121	0.07	0.02
				ε	4.935	0.129	-0.08	-0.13
$y = \delta + \exp(\eta - \nu.\theta^{x})$	(22)	3.735	0.493	δ	17.325	0.123	-0.09	-0.03
				η	1.852	0.034	0.10	0.08
				ν	13.857	3.716	1.12	5.61
				θ	0.510	0.033	-0.10	-0.27
$y = \alpha.(1 - \exp(-\beta.x))^{\gamma} + \delta$	(23)	0.960	0.470	α	6.476	0.281	0.29	0.34
				β	0.633	0.078	0.25	0.87
				γ	2.403	0.340	0.24	0.92
				δ	17.260	0.173	-0.15	-0.05
$y = \frac{(\beta.\exp(\gamma) + \alpha.x^{\delta})}{(\exp(\gamma) + x^{\delta})}$	(24)	1.144	0.470	α	23.845	0.207	0.48	0.07
				β	17.233	0.159	-0.13	-0.04
				γ	6.133	0.676	0.32	0.86
				δ	4.098	0.459	0.28	0.82
$y = \alpha - \beta.\exp(-\exp(-\gamma).x^{\delta})$	(25)	0.406	0.470	α	23.491	0.078	0.13	0.01
				β	6.421	0.148	0.14	0.09
				γ	4.593	0.278	0.18	0.28
				δ	2.871	0.171	0.18	0.27

ρ = critical curvature value ($\rho = 1/\sqrt{F}$), F (α, n-p, p) is the inverse of Fisher´s probability distribution obtained at significance level α = 0.05, p is the number of parameters and n is the number of data observations.

Table IX – Reparameterizations and statistical results of the least-squares estimation. Fermentation medium C.

Parameterization		PE	ρ	Parameter	Estimate	Std. Error	Skewness	% Bias
$y = \delta + \left(\frac{\alpha - \delta}{1 + \exp(\beta - \gamma.\log(x))}\right)$	(19)	1.252	0.470	α	24.853	0.278	0.51	0.11
				β	6.356	0.873	0.41	1.38
				γ	4.340	0.597	0.37	1.32
				δ	16.818	0.242	-0.15	-0.07
$y = \delta + \left(\frac{\alpha - \delta}{1 + \exp(\gamma.(\beta - \log(x)))}\right)$	(20)	1.178	0.470	α	16.818	0.242	-0.15	-0.07
				β	1.465	0.033	0.05	0.09
				γ	-4.340	0.597	-0.37	1.32
				δ	24.853	0.278	0.51	0.11
$y = \delta + \alpha.\exp(-\exp(\gamma.(x - \varepsilon)))$	(21)	0.683	0.493	α	-8.212	0.243	-0.26	0.15
				γ	0.661	0.052	0.21	0.43
				δ	24.407	0.092	0.04	0.01
				ε	4.781	0.078	-0.06	-0.05
$y = \delta + \exp(\eta - \nu.\theta^{x})$	(22)	4.198	0.493	δ	16.801	0.179	-0.01	-0.05
				η	2.072	0.037	0.10	-0.08
				ν	13.959	4.285	1.29	7.37
				θ	0.496	0.038	-0.12	-0.37
$y = \alpha.(1 - \exp(-\beta.x))^{\gamma} + \delta$	(23)	1.013	0.470	α	7.915	0.389	0.30	0.43
				β	0.685	0.099	0.33	1.28
				γ	2.540	0.419	0.31	1.35
				δ	16.833	0.255	-0.17	-0.09
$y = \frac{(\beta.\exp(\gamma) + \alpha.x^{\delta})}{(\exp(\gamma) + x^{\delta})}$	(24)	1.252	0.470	α	24.853	0.278	0.51	0.11
				β	16.818	0.242	-0.15	-0.07

				γ	6.356	0.873	0.41	1.38
				δ	4.340	0.597	0.37	1.32
$y = \alpha - \beta.\exp(-\exp(-\gamma).x^{\delta})$	(25)	0.377	0.470	α	24.482	0.088	0.09	0.01
				β	7.840	0.175	0.11	0.08
				γ	4.794	0.298	0.20	0.31
				δ	3.055	0.186	0.20	0.30

ρ = critical curvature value ($\rho = 1/\sqrt{F}$), F (α, n-p, p) is the inverse of Fisher´s probability distribution obtained at significance level α = 0.05, p is the number of parameters and n is the number of data observations.

The misuse of R^2

As readers may have noticed, the coefficient of determination R^2 was not used as a measure of goodness of fit of the models. A widespread error in nonlinear regression is to believe that R^2, the “proportion of explained variation”,is of use in deciding whether the nonlinear model provides a good fit to the experimental data. It is only when one has a linear model with a constant term that R^2 represents the proportion of variation explained by the model (Draper, 1984,; Healy, 1984, Helland, 1987). For a nonlinear regression model, R^2 does not have any obvious meaning (Ratkowsky, 1990).

Although it has been demonstrated for some time that R^2 is an inadequate measure for nonlinear regression, many scientists and also reviewers insist on it being supplied in papers dealing with nonlinear data analysis. There is still a high occurrence in the current literature where the goodness of fit of nonlinear models is based only on R^2 values. In addition, several commercially available statistical software packages calculate R^2 values for nonlinear fits, which contributes to the misuse of this criterion (Spiess and Neumeyer, 2010). A more in-depth discussion of the misuse of R^2 is beyond the scope of this paper. For further details, see the works of Kvalseth (1985), Willet and Singer (1988), and others.

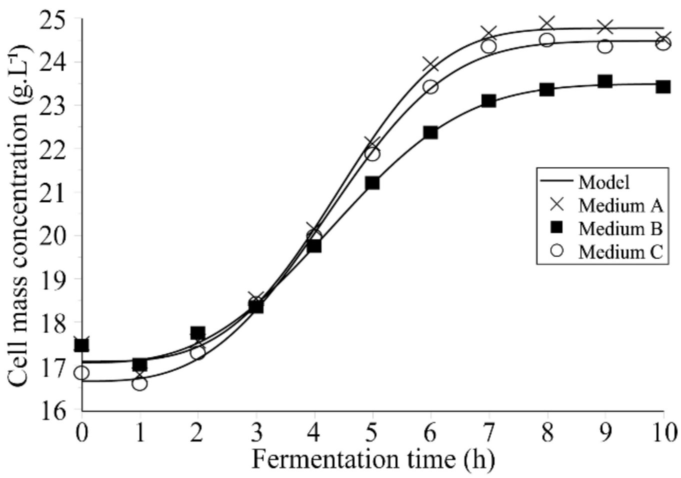

Figure 1 – Growth curve of *Saccharomyces cerevisiae* fitted with the reparameterized Weibull-type model(Eq. 25).

IV. CONCLUSIONS

In this work, sigmoidal models were compared to describe the growth curve of *Saccharomyces cerevisiae* in batch ethanol fermentation of sugarcane molasses. Among the twelve models fitted to the experimental growth data, only the reparameterized Weibull-type model presented close to linear behavior, ensuring the statistical validity of the parameters estimated by the method of least squares. For this model, the parameter estimates are almost unbiased, normally distributed and have close-to-minimal variance.

The other models presented far from linear behavior due to skewed or biased parameter estimates, which implies invalid inference results based on asymptotic approximations for the least squares estimators.

REFERENCES

[1]. Allaby, M. A DictionaryofZoology. 4thedition. New York: Oxford University Press; 2014.
[2]. Bates, DM, Watts, DG. 1980. Relativecurvaturemeasuresofnonlinearity. J R StatSoc Series B StatMethodol. 1980; 42(1): 1-25.
[3]. Box, MJ. Bias in nonlinearestimation. *J R StatSoc Series B StatMethodol*. 1971; 33: 171-201.
[4]. Draper, NR. The Box–Wetzcriterion versus R^2. *J R StatSoc Series A StatMethodol*. 1984; 147(1):100-103.
[5]. Draper, NR, Smith, H. Appliedregressionanalysis. 3rd ed. New York: John Wiley& Sons; 1998.
[6]. El-Shaarawi, AH, Piegorsch, WW. Encyclopediaofenvironmetrics. New York: Wiley; 2002.
[7]. Gebremariam, B. Isnonlinearregressionthrowingyou a curve? New diagnosticandinference tools in the NLIN procedure. SAS InstituteInc; 2014 Paper SAS384-2014.

[8]. Gompertz, B. Onthenatureofthefunctionexpressiveofthelawofhumanmortality, andon a new modeofdeterminingthevalueoflife contingencies. *PhilosTrans R SocLond* [Internet]. 1825 Jun; [cited 2017 Oct. 12]; 115: [71 p.]. Availablefrom:https://www.jstor.org/stable/107756?seq=1#metadata_info_tab_contents

[9]. Healy, MJR. 1984. The use ofR^2 as a measureofgoodnessoffit. *J R StatSoc Series B StatMethodol*. 1984; 147(4): 608-609.

[10]. Helland, IS. 1987. Ontheinterpretationand use of R^2 in regressionanalysis. *Biometrics.*1987; 43: 61-69.

[11]. Hougaard, P. The Appropriatenessoftheasymptoticdistribution in a regressionmodel in relationtocurvature. *J R StatSoc Series B StatMethodol*. 1985; 47: 103-114.

[12]. Kvalseth, TO. Cautionary note about R^2. *Am Stat.*1985; 39: 279–285.

[13]. Morgan, PH, Mercer, LP, Flodin, NW. 1975. General model for nutricional responses ofhigherorganisms. *ProcNatlAcadSci USA*. 1975; 72(11): 4327-4331.

[14]. Pienaar, LV, Turnbull, KJ. 1973. The Chapman-Richards generalizationof Von Bertalanffy'sgrowthmodel for basal área growthandyield in even-aged stands. *J For Sci*. 1973; 19(1): 2–22.

[15]. Ratkowsky, DA. Nonlinearregressionmodeling. New York: Marcel Dekker; 1983.

[16]. Ratkowsky, D. Handbookofnonlinearregressionmodels. New York: Marcel Dekker; 1990.

[17]. Seber, GAF, Wild, CJ. NonlinearRegression. New York: John Wileyand Sons; 1989.

[18]. Spiess,NA, Neumeyer N. Anevaluationof R^2 as aninadequatemeasure for nonlinearmodels in pharmacologicalandbiochemicalresearch: a Monte Carlo approach. *BMC PharmacolToxicol* [Internet]. 2010 Jun; [cited 2018 Apr. 08];10(6):[16 p.]. Availablefrom: https://doi.org/10.1186/1471-2210-10-6

[19]. Tosetto, GM. [Influenceofraw material onkineticbehaviorofyeast in etanol production] [dissertation]. Campinas (SP): Universidade Estadual de Campinas; 2002. Brazil.

[20]. Weibull, W. A statisticaldistributionfunctionofwideapplicability. *J ApplMech*. 1951; 18: 293–296.

[21]. Willet, JB, Singer, JD. Anothercautionary note about R-square: Its use in weightedleast-squaresregressionanalysis. *AmStat.* 1988; 42: 236-238.

Marcelo Teixeira Leite. "Sigmoidal Models Fitted to the Saccharomyces cerevisiae Growth Curve: Statistical Analysis Using Measures of Nonlinearity."American Journal of Engineering Research (AJER), vol.8, no.01, 2018,pp.01-09