

## Critical points in logistic growth curves and treatment comparisons

José Raimundo de Souza Passos, Sheila Zambello de Pinho\*, Lídia Raquel de Carvalho, Martha Maria Mischan

UNESP/IBB – Depto. de Bioestatística,  
C.P. 510 – 18618-970 – Botucatu, SP – Brasil.  
\*Corresponding author <Sheila@ibb.unesp.br>

Edited by: Thomas Kumke

**ABSTRACT:** Several biological phenomena have a behavior over time mathematically characterized by a strong increasing function in the early stages of development, then by a less pronounced growth, sometimes showing stability. The separation between these phases is very important to the researcher, since the maintenance of a less productive phase results in uneconomical activity. In this report we present methods of determining critical points in logistic functions that separate the early stages of growth from the asymptotic phase, with the aim of establishing a stopping critical point in the growth and on this basis determine differences in treatments. The logistic growth model is fitted to experimental data of imbibition of araribá seeds (*Centrolobium tomentosum*). To determine stopping critical points the following methods were used: i) accelerating growth function, ii) tangent at the inflection point, iii) segmented regression; iv) modified segmented regression; v) non-significant difference; and vi) non-significant difference by simulation. The analysis of variance of the abscissas and ordinates of the breakpoints was performed with the objective of comparing treatments and methods used to determine the critical points. The methods of segmented regression and of the tangent at the inflection point lead to early stopping points, in comparison with other methods, with proportions ordinate/asymptote lower than 0.90. The non-significant difference method by simulation had higher values of abscissas for stopping point, with an average proportion ordinate/asymptote equal to 0.986. An intermediate proportion of 0.908 was observed for the acceleration function method.

Keywords: nonlinear regression, asymptotic regression, stopping critical level, seeds imbibition

Received August 01, 2011

Accepted March 13, 2012

### Introduction

Biological phenomena can have a mathematically characterized behavior as a function of time with a strong increasing function in early development stages, followed by a less pronounced growth, sometimes showing stability. The separation between these phases is of great importance since, in many processes, the maintenance of a less productive phase results in an uneconomical activity. Sometimes the growth phases have different biological meanings and it is important to separate them.

Several methods have the aim to determine critical points that separate growth phases, as in Cate Jr. and Nelson (1965) to determine critical level of nutrients in plants and Portz et al. (2000) that use the segmented regression in a fish study. The segmented regression method is also used in Cate Jr. and Nelson (1971) and discussed in Rayment (2005) which named it "Linear Response and Plateau Model".

Empirical methods separate growth stages with basis in yield percentages, generally using the levels of 90 to 95 % of maximum yield as an upper limit, as in Korndörfer et al. (2001) working with rice (*Oryza sativa*), Evans et al. (2008) with an ornamental bush (*Euonymus fortunei*), and Santos et al. (2004) with alfalfa (*Medicago sativa*). Other statistical and mathematical methods are employed with the same aim to determine critical points in logistic curves (Carvalho and Pinho, 1996 and Mischan et al., 2011) fitted to data of seed imbibition and weight of cattle, respectively.

The aim of this paper is to determine new methods to obtain critical points in logistic growth curves and

compare them with some already available in the literature, in order to offer alternatives for decisions on the best time to stop the process. We also indicate the possibility of comparing treatments through the analysis of variance of the critical points.

### Materials and Methods

The statistical model of growth used is the logistic one:

$$y_j = \alpha [1 + \exp(-\beta \cdot \gamma x_j)]^{-1} + e_j, \quad (1)$$

with parameters  $\alpha$ ,  $\beta$  and  $\gamma$ ,  $\alpha > 0$  and  $\gamma > 0$ , where  $y_j$  is the observed measure at time  $x_j$ , and  $e_j$  is the random error, with normal distribution  $(0, \sigma^2)$ . The estimated function is represented by

$$y = a [1 + \exp(-b \cdot cx)]^{-1}, \quad (2)$$

with  $a$ ,  $b$  and  $c$  the estimates of the parameters  $\alpha$ ,  $\beta$  and  $\gamma$ , respectively. To determine the stopping critical points in the logistic model six methods are used as described below.

#### $M_1$ - Acceleration function method with point $P_1(x_1; y_1)$

The method, described in Mischan et al. (2011), works with the acceleration function of the logistic model of growth, represented by its second derivative

$$y'' = \alpha \gamma^2 \exp(-\beta \cdot \gamma x) [\exp(-\beta \cdot \gamma x) - 1] [1 + \exp(-\beta \cdot \gamma x)]^{-3}, \quad (3)$$

which has two extreme points, a maximum and a minimum, and three inflection points. After the last point of inflection, called asymptotic deceleration point ( $P_1$ ), the deceleration of growth is very slow and approaches to zero when  $x$  tends to infinity. The point coordinates are obtained by equating the derivative of order 4 to zero:  $P_1 [-(\ln(5-2\sqrt{6}) + b) / c; a(3 + \sqrt{6}) / 6]$  or, approximately,  $P_1 [(2.29 - b)/c; 0.908a]$ .

**$M_2$  - Tangent at the inflection point method with point  $P_2 (x_2; y_2)$**

In the asymptotic growth function estimated, with an inflection point  $P_i (x_i; y_i)$ , we call  $y'_i$  be the value of the first derivative of the function at its inflection point. If we consider a tangent to the curve at this point as a linear approximation to the growth phase before the asymptote, its intersection with the asymptote of the function  $y = a$  can be interpreted as an indication of maximum growth and therefore a stopping point. The equation of this tangent line is

$$Y - Y_i = Y'_i (x - x_i), \tag{4}$$

and in its intersection with the asymptote ( $y = a$ ) is

$$a - Y_i = Y'_i (x_2 - x_i), \tag{5}$$

where  $x_2$  is the abscissa of the  $P_2$  point. Therefore,

$$x_2 = (a - Y_i)/Y'_i + x_i \tag{6}$$

The logistic function has  $P_i (-b/c; a/2)$  and  $y'_i = ac/4$ , hence  $P_2 [(2-b)/c; a/(1 + \exp(-2))]$  or, approximately,  $P_2 [(2-b)/c; 0.881a]$ . See Figure 1.

**$M_3$  - Segmented regression method with point  $P_3 (x_3; y_3)$**

For  $n$  pairs  $(x_j; y_j)$  from the observed data the segmented regression model can be represented by two lines, one parallel to the  $x$ -axis and one inclined,

$$y_j = \alpha_1 + \beta_1(\rho - x_j) + e_j, \quad j = 1, 2, \dots, n_1, n_1 + 1, \dots, n, \tag{7}$$

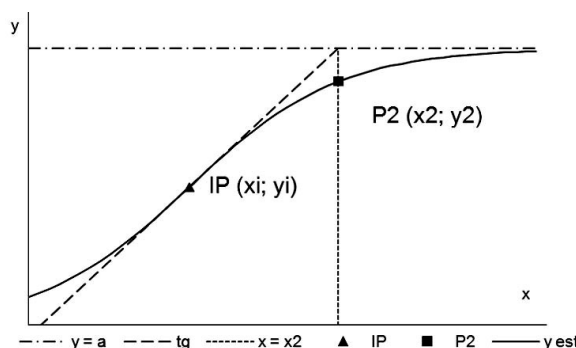


Figure 1 – Representation of the adjusted logistic function, the asymptote  $y = a$ , the tangent to the curve at the inflection point IP and the stopping point  $P_2$ .

with the restriction  $(\rho - x_j) = 0$  for  $j \geq n_1 + 1$ , where  $n_1 =$  number of observations before the intersection point of straight lines,  $\alpha_1 =$  intersection of line with the vertical axis  $y$ ,  $\beta_1 =$  slope,  $\rho =$  abscissa of the intersection point and  $e_j =$  experimental error. The estimates of parameters  $\alpha_1, \beta_1$  and  $\rho$  are represented by  $a_1, b_1$  and  $r$ , respectively. The estimate  $r$  is the abscissa of the sought critical point,  $r = x_3$ .

**$M_4$  - Modified segmented regression method with point  $P_4 (x_4; y_4)$**

When determining the intersection point in the segmented regression method, this point is the beginning of the asymptotic phase of growth and is therefore a point of early occurrence. Naming the horizontal line determined in the former method  $y = a_2$ , we see that the fitted curve does not reach the asymptote  $y = a$ , but intersects the line  $y = a_2$ ,  $a_2 < a$  at some point  $P_{4'}$ , occurring later in time as compared to  $P_3$ . Based on this information, the estimate  $a_2 =$  intersection of horizontal line with the  $y$ -axis, obtained at the segmented regression method  $M_{3'}$ , is used as the ordinate of the sought critical point. Observe that  $y = a_2$  is the least squares straight line fitted to the data that are distributed in a nearly horizontal manner, parallel to the  $x$ -axis. These data are placed after a point  $P_4 (x_4, y_4)$ , where  $y_4 = a_2$  and  $x_4$  is determined by the function fitted to the data in the logistic curve,

$$a_2 = a [1 + \exp(-b \cdot c x_4)]^{-1}, \tag{8}$$

where

$$x_4 = \frac{1}{c} [\ln(a_2/(a - a_1)) - b]. \tag{9}$$

See Figure 2.

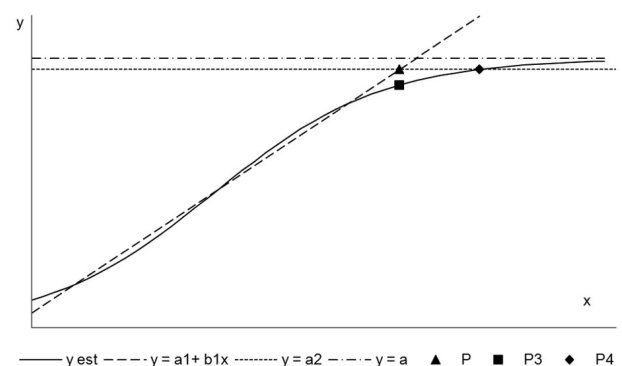


Figure 2 – Representation of the adjusted logistic function, its asymptote  $y = a$ , the straight lines fitted by the method of segmented regression,  $y = a_1 + b_1x$  and  $y = a_2$ , and the points  $P_3$  and  $P_4$  determined by the methods of segmented regression and modified segmented regression, respectively.

### $M_5$ - Non-significant difference method with point $P_5(x_5; y_5)$

The method for determining the point of non-significant difference is described in Carvalho and Pinho (1996). We consider the difference  $y^* = a - y$ , between the estimated function and its asymptote, and we verify from which point we can consider it non-significant by a Student t-test. The abscissa  $x_5$  of the critical point determined by this method is the solution of the equation

$$T = y^* x_5 / \{[\hat{V}(y^*)] x_5\}^{0.5}, \quad (10)$$

for a value of  $T = t_{\alpha, f}$ , where  $t_{\alpha, f}$  is the unilateral t-test at significance level  $\alpha$ , with  $f$  = number of degrees of freedom associated with the estimate of error variance. The resolution is made with the parameter estimates, their variances and covariances, and the estimated variance of  $y^*$  for  $x = x_5$ . The values of the abscissas of the points are generally beyond the range of observation data, which suggests that this method is quite strict in order to determine a stopping critical point to the observations of the experiment. This rigor can be mitigated considering a difference  $y^*$  not between  $y$  and its asymptote, but between  $y$  and a percentage,  $p$ , of this. In this case we have:  $y^* = pa - y$ . In this paper we adopted  $p = 0.90$ , which is considered in several articles in the literature on critical points, for example in Korndörfer et al. (2001), Evans et al. (2008) and Santos et al. (2004). This  $p$  value is very close to that determined in method  $M_1$ , 0.908.

### $M_6$ - Non-significant difference method by simulation with point $P_6(x_6; y_6)$

The proposed method for determining the critical point of stopping by simulation is based on the technique of Monte Carlo simulation (Shapiro and Gross, 1981). They also appear in literature applications of this technique in studies of sampling distribution of estimators in nonlinear regression models (Bates and Watts, 1988).

In this method, after adjusting the asymptotic growth function to data, new observations of the response variable are generated and repeated 1,000 times, based on the values of the dependent variable, the estimated parameters and sample variance. Then the nonlinear regression model is fitted again to the replicates obtained by simulation. From the inflection point of each of the simulated curves, the difference  $D$  between the ordinate of the considered point and the estimated asymptote is determined. This process is repeated at a step  $k$  to a maximum allotted time for the independent variable. The empirical sampling distribution of the random variable  $D$  and the percentiles  $p_{2.5}$  and  $p_{97.5}$  are determined for each empirical distributions obtained previously. The stopping point  $P_6(x_6; y_6)$  for the value of the dependent variable is the first value of  $p_{2.5} < 0$ .

The model is fitted to data of accumulated weight gain of fruits of araribá (*Centrolobium tomentosum*), in percentage, during soaking for 84 h in distilled water, subjected to treatment with sulfuric acid during the times {0, 10, 15, 20 min.} represented by T0, T10, T15 and T20 as described in Carvalho and Pinho (1996). The experiment had 25 repetitions. The parameter estimates, their standard errors, confidence intervals and coefficients of asymmetry of Hougaard (1985) were determined.

In the analysis of variance of the abscissas and ordinates were considered the factors 'treatments' and 'methods of determining breakpoints' and the interaction between them. The data used in these tests is obtained by fitting the logistic function and posterior determination of breakpoints for each of the 25 replicates for four treatments.

## Results and Discussion

Table 1 and Figure 3 present the results of the logistic function fit to the data of weight gain of fruits araribá for each of the four treatments, using all the replicates in

Table 1 – Parameter estimates of the logistic function, their approximate standard errors, approximate 95 % confidence limits and asymmetry coefficients of Hougaard, for four treatments followed, in brackets, by the residual variance ( $s^2$ ) and the degrees of freedom (df).

Parameters estimates	Approximate standard errors		Approximate confidence limits of 95 %		Asymmetry coefficients
<b>Treatment 0</b> ( $s^2 = 14.4710$ ; $df = 319$ )					
a	74.2706	0.3907	73.5019	75.0392	0.0525
b	-1.2765	0.0446	-1.3641	-1.1888	-0.0488
c	0.0836	0.00248	0.0787	0.0885	0.071
<b>Treatment 10</b> ( $s^2 = 12.9719$ ; $df = 333$ )					
a	57.803	0.8553	56.1204	59.4856	0.231
b	-1.2648	0.0398	-1.3431	-1.1864	-0.0713
c	0.0496	0.00206	0.0455	0.0536	0.0542
<b>Treatment 15</b> ( $s^2 = 12.0849$ ; $df = 333$ )					
a	59.8092	0.6489	58.5328	61.0857	0.157
b	-1.4526	0.0411	-1.5335	-1.3717	-0.0679
c	0.0573	0.00196	0.0534	0.0611	0.0555
<b>Treatment 20</b> ( $s^2 = 9.5192$ ; $df = 333$ )					
a	57.5378	0.5195	56.5159	58.5596	0.1252
b	-1.5147	0.0395	-1.5925	-1.437	-0.0615
c	0.0609	0.00184	0.0573	0.0646	0.0535

each time. Table 1 shows the parameter estimates, their standard errors and approximate confidence intervals limits at 95 % and the coefficients of asymmetry of Hougaard. All parameter estimates present values of  $t = \text{estimate} / \text{standard error of estimate}$  with  $p$ -values  $< 0.0001$ . The parameters estimates of  $\beta$  and  $\gamma$  present coefficients of asymmetry less than 0.1 which, by the classification of Ratkowsky (1989), suggest how the estimators are quite close to linear. For the parameter  $\alpha$ , the greater asymmetry coefficient was 0.231 in the treatment T10, which makes the estimator reasonably close to linear (coefficient between 0.1 and 0.25). The Durbin-Watson test to check for autocorrelation and the Breusch-Pagan for homogeneity of variances showed that the model of independent errors and constant variance can be employed.

Table 2 and Figure 4 show the stopping critical points  $(x_i, y_i)$  determined by methods  $M_1$  to  $M_6$  for each treatment, as well as the inflection point.

The analysis of variance of estimated abscissas showed significant interaction ( $p < 0.001$ ) among the factors 'treatment' and 'methods for determining the stopping points'. The  $M_2$  (tangent at the inflection point method) and  $M_3$  (the segmented regression method) did not differ in any

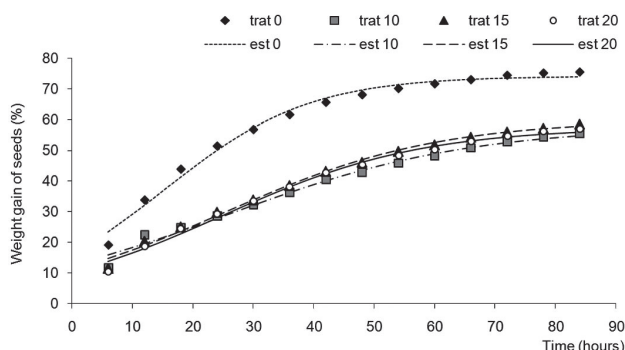


Figure 3 – Logistic function adjusted to data on weight gain of fruits araribá, subject to four treatments. Observed values: T0, T10, T15 and T20 (average of 25 replicates) and estimated est 0, est 10, est 15 and est 20.

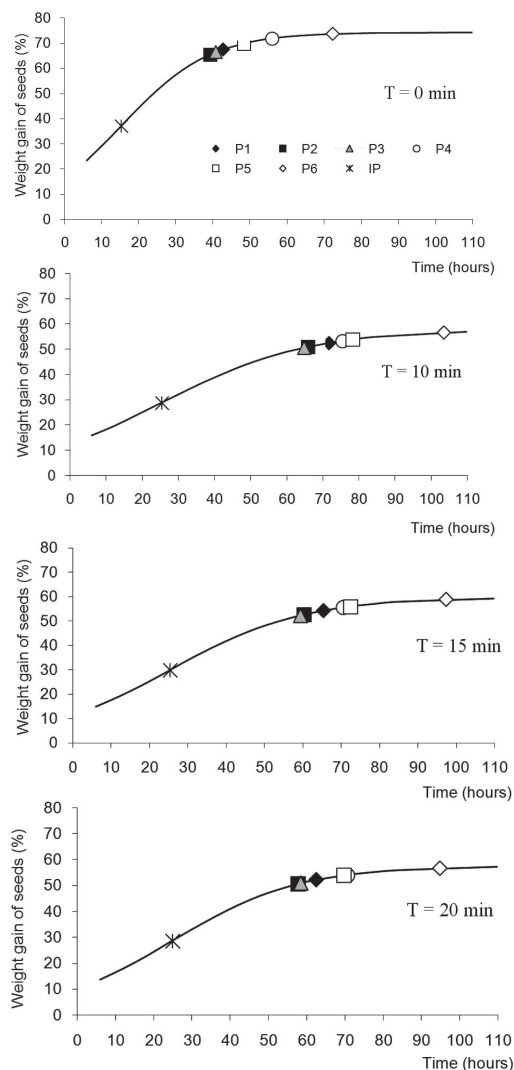


Figure 4 – Logistic function fitted to the four treatments and location of inflection (IP) and stopping points: asymptotic decelerating ( $P_1$ ), tangent to the curve at the inflection ( $P_2$ ), segmented regression ( $P_3$ ), modified segmented regression ( $P_4$ ), non-significant difference ( $P_5$ ) and non-significant difference by simulation ( $P_6$ ).

Table 2 – Critical points  $(x_i, y_i)$  and proportions  $(y_i/a)$  of the ordinate at the point relative to the asymptote,  $a$ , according to the treatments.

Treatments	IP	Methods						
		M1	M2	M3	M4	M5	M6	
T0	$x_i$	15.3	42.7	39.2	40.8	55.9	48.3	72.3
	$y_i$	37.135	67.456	65.417	66.396	71.872	69.846	73.643
	$y_i/a$	0.500	0.908	0.881	0.894	0.968	0.940	0.992
T10	$x_i$	25.5	71.7	65.8	64.9	75.4	78.4	103.5
	$y_i$	28.902	52.499	50.913	50.624	53.309	53.892	56.619
	$y_i/a$	0.500	0.908	0.881	0.876	0.922	0.932	0.980
T15	$x_i$	25.4	65.4	60.3	59.4	70.6	72.3	97.4
	$y_i$	29.905	54.322	52.680	52.356	55.630	56.010	58.857
	$y_i/a$	0.500	0.908	0.881	0.875	0.930	0.936	0.984
T20	$x_i$	24.9	62.5	57.7	58.4	70.5	69.6	94.9
	$y_i$	28.769	52.259	50.679	50.931	54.186	53.998	56.741
	$y_i/a$	0.500	0.908	0.881	0.885	0.942	0.938	0.986

treatment and had the lowest values of abscissas (Tukey,  $p < 0.05$ ). Both methods assume a straight line representing the initial growth that intersects with a horizontal line representing the asymptotic growth - the asymptote ( $a$ ) in the  $M_2$  method and the least squares line of  $y = a_2$  fitted to final growth in method  $M_3$ . The method  $M_6$  (non-significant difference between the function and its asymptote, determined by simulation) shows breakpoints with the highest values of abscissas,  $x_i$ , differing from other methods in all treatments. Abscissa values determined by the method  $M_4$  (modified segmented regression) are larger than those determined by the method  $M_3$ . So, the modified segmented regression method is an efficient alternative to the method of segmented regression  $M_3$ , when we want a stop point occurring later. The  $M_1$  method (accelerating growth function) leads the points with abscissas intermediate in comparison with other methods.

The comparison between treatments showed that T10 differs from other treatments having the highest abscissa of breakpoints. T15 and T20 did not differ and T0 leads to lower values. All the treated seeds, therefore, showed a reduced growth rate compared with the control T0, what is already evident in the parameter estimates presented in Table 1, where  $c$  (T0) = 0.071, a value higher than the others. A high value for  $c$  shows that the estimated growth rate is high, which implies a smaller value for the abscissa of the breakpoint.

Analyzing the ordinate values of the breakpoints, significant main effects were found for the factors 'treatment' and 'methods' and no significant interactions. In general the differences between methods are similar to those seen in the analysis of abscissas: the smallest ordinates are obtained by methods  $M_2$  and  $M_3$ , not showing difference among them. The method  $M_6$  has the highest ordinate value, differing from other methods;  $M_1$  is in the middle. If the ordinate criterion is used to compare treatments, T0 also differs from other treatments presenting, unlike the comparisons of abscissas, the highest average value; this is probably due to the high growth rate of T0 compared to the treated seeds.

The determination of a stopping point in the adjusted logistic growth curve to data from soaked seeds of araribá can be made by different methods, mathematical or statistical methods, which can be considered different regarding the values of their coordinates. The segmented regression and the tangent at the inflection point methods lead to the decision of a significant slowdown in growth from lower values of abscissa; they lead to early stopping points, in comparison with other methods, with proportions ordinate/asymptote a little lower than 0.90. The non-significant difference method by simulation determines higher values of abscissas for stopping points, with an average proportion ordinate/asymptote 0.986, a value that shows an ordinate excessively near to the asymptote. An intermediate proportion of 0.908 is obtained by the acceleration function method. The segmented regression method (M3), the acceleration function method (M1) and the modified segmented regression method (M4) are quite

simple in application and cover a wide range of variation in the amounts ordinate/asymptote, on average 0.883, 0.908 and 0.941, respectively, which is very useful in practice. The method M1 depends only on a mathematical formula that uses the parameters of the fitted model, M3 is a method already widely used in literature and M4 is a simple modification of M3. On the other hand, the method of the tangent at the inflection point (M2) shows very similar results to those obtained by the method M3, segmented regression, but the points are difficult to obtain; the non-significant difference methods (M5 and M6) are not simple to apply and M6 leads to excessively high values of abscissas with ordinates very close to the asymptote. The breakpoints can also be used as a criterion for comparing treatments in experiments where the growth functions are adjusted.

## References

- Bates, M.D.; Watts, D.G. 1988. Nonlinear Regression Analysis and Its Applications. John Wiley, New York, NY, USA.
- Carvalho, L.R.; Pinho, S.Z. 1996. Method for comparison of logistic growth curve with the presence of residual autocorrelation. *Energia na Agricultura* 11: 38–55 (in Portuguese, with abstract in English).
- Cate Jr., R.B.; Nelson, L.A. 1965. A Rapid Method for Correlation of Soil Tests Analyses with Plant Response Data. International Soil Testing, Raleigh, NC, USA. (Technical Bulletin. North Carolina Agricultural Experiment Station, 1).
- Cate Jr., R.B.; Nelson, L.A. 1971. A simple statistical procedure for partitioning soil test correlation data into two classes. *Soil Science Society of America Proceedings* 35: 658–659.
- Evans, R.Y.; Smith, S.J.; Paul, J.L. 2008. Nitrogen critical level determination in the woody ornamental shrub *Euonymus fortunei*. *Journal of Plant Nutrition* 31: 2075–2088.
- Hougaard, P. 1985. The appropriateness of the asymptotic distribution in a nonlinear regression model in relation to curvature. *Journal of the Royal Statistical Society, Serie B* 47: 103–114.
- Korndörfer, G.H.; Snyder, G.H.; Ulloa, M.; Powell, G.; Datnoff, L.E. 2001. Calibration of soil and plant silicon analysis for rice production. *Journal of Plant Nutrition*. 24: 1071–1084.
- Mischan, M.M.; Pinho, S.Z.; Carvalho, L.R. 2011. Determination of a point sufficiently close to the asymptote in nonlinear growth functions. *Scientia Agricola* 68: 109–114.
- Portz, L.; Dias, C.T.S.; Cyrino, J.E.P. 2000. A broken-line model to fit fish nutrition requirements. *Scientia Agricola* 57: 601–607.
- Ratkowsky, D.A. 1989. Handbook of Nonlinear Regression Models. Marcel Dekker, New York, NY, USA.
- Rayment, G.E. 2005. Statistical aspects of soil and plant test measurement and calibration in Australasia. *Communications in Soil Science and Plant Analysis* 36: 107–120.
- Santos, A.R.; Mattos, W.T.; Almeida, A.A.S.; Monteiro, F.A.; Corrêa, B.D.; Gupta, U.C. 2004. Boron nutrition and yield of alfalfa cultivar crioula in relation to boron supply. *Scientia Agricola* 61: 496–500.
- Shapiro, S.S.; Gross, A.J. 1981. Statistical Modeling Techniques: Statistical, Textbooks and Monographs. Marcel Dekker, New York, NY, USA.