

Inflection and stability points of diphasic logistic analysis of growth

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

São Paulo State University/IBB – Dept. of Biostatistics, C.P. 510 – 18618-970 – Botucatu, SP – Brazil.

*Corresponding author <sheila@ibb.unesp.br>

Edited by: Marcin Kozak

Received July 11, 2014

Accepted November 29, 2014

ABSTRACT: Growth functions with inflection points following a diphasic model, can be adjusted by two approaches using segmented regression or the sum of two functions. In both cases, there are two functions, one for each phase, with inflection and stability points. However, when they are summed, the result is a new function and the points of inflection and stability are different from those obtained from using each function individually. A method to determine these points in a diphasic logistics sum of functions is suggested and the results obtained from fitting the models to eucalyptus growth data showed a better fit of the logistic diphasic sum as compared with segmented regression and monophasic logistic models.

Keywords: nonlinear models, segmented regression, sum of logistics

Introduction

Monophasic models of linear or non-linear growth are a description of the total growth cycle, from birth to adulthood of the individual, which may be greatly simplified. However, multiple growth cycles can be found in the literature (Seber and Wild, 1989) from the beginning of the 20th century (Brody and Ragsdale, 1921). Since then, curve adjustment by multiphasic models has been employed in various areas, such as human height growth (Bock et al., 1973), growth in weight of rats (Koops et al., 1987; Kurnianto et al., 1999), chickens (Grossman and Koops, 1988), rice crop biomass (Sheehy et al., 2004) and cows (Mendes et al., 2008; Mendes et al., 2009).

The logistic function has been the one most frequently used in multiphasic models due to its properties of symmetry in the growth velocity curve (Koops, 1986; Kurnianto et al., 1999; Mendes et al., 2008). This model may be fitted either by segmented regression (Portz et al., 2000; Robbins et al., 2006) or by the sum of functions – the most commonly employed methods (Koops, 1986; Koops and Grossman, 1991; Kurnianto et al., 1999; Özkan, 2004; Nešetřilová, 2005; Mendes et al., 2008; Mendes et al., 2009; Fenner et al., 2013).

In the nonlinear multiphasic model with inflection points, adjusted by the sum of functions, the points are determined in the functions that correspond to each growth phase (Koops, 1986; Koops et al., 1987; Koops and Grossman, 1991; Kurnianto et al., 1999; Mendes et al., 2009; Fenner et al., 2013), but the inflection points of the sum function are not determined.

In this study, a methodology for the determination of inflection and stability points of logistic diphasic models as adjusted by the sum of functions is presented, which are compared with the points obtained by segmented regression and the monophasic model.

Materials and Methods

Consider the following logistic models, where $\alpha > 0$ and $\gamma > 0$, fitted to a data set (x_i, y_i) from observed data,

$i = 1, 2, \dots, n$, where errors are considered independent, normal and homoscedastic, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

Model I - monophasic logistic model

$$y_i = F(x_i, \theta) = \frac{\alpha}{(1 + X_i)} + \varepsilon_i, X_i = \exp(-\beta - \gamma x_i), \theta = [\alpha \beta \gamma]' \quad (1)$$

Model II - diphasic logistic model segmented regression

$$y_i = F(x_i, \theta) = \frac{\alpha_1}{(1 + X_{1i})} + \varepsilon_i, X_{1i} = \exp(-\beta_1 - \gamma_1 x_i), \text{ if } x \leq \iota,$$

$$y_i = F(x_i, \theta) = \frac{\alpha_2}{(1 + X_{2i})} + \varepsilon_i, X_{2i} = \exp(-\beta_2 - \gamma_2 x_i), \text{ if } x \geq \iota, \quad (2)$$

$\theta = [\alpha_k \beta_k \gamma_k \iota]'$, $k = 1, 2$, and ι = abscissa of the intersection point of the two functions.

Model III - diphasic logistic sum of functions

$$y_i = F(x_i, \theta) = \sum_{k=1}^2 \frac{\alpha_k}{(1 + X_{ki})} + \varepsilon_i, X_{ki} = \exp(-\beta_k - \gamma_k x_i),$$

$$\theta = [\alpha_k \beta_k \gamma_k]', k = 1, 2. \quad (3)$$

Residual homoscedasticity is checked by the Breusch-Pagan test (Breusch and Pagan, 1979) and normality by the Shapiro-Wilk test. The residual autocorrelation is verified by the use of tables for testing randomness of grouping in a sequence of residual signs (Drapier and Smith, 1998).

Adjustments to the models can be checked by the usual criteria for comparing models: the residual mean square of the model, RMS; the square of the correlation coefficient between observed and fitted values by the model, $r_{\hat{y}, y}^2$ (Schinckel and Craig, 2002), and the F criterion

$$F = \frac{(SSR_1 - SSR_2)/(df_1 - df_2)}{SSR_2/df_2} \quad (4)$$

where: SSR_1 is the sum of the squares of residuals of the model with a smaller number of parameters, SSR_2 is the

sum of the squares of residuals of the model with a larger number of parameters, and df_1 and df_2 are degrees of freedom associated with SSR_1 and SSR_2 , respectively; the corrected Akaike Information Criterion:

$$AICc = AIC + 2p(p+1)/(n-p-1), AIC = n \log_e(SSR/n) + 2p \quad (5)$$

where: SSR is the sum of the squares of residuals of the model, and p is the number of model parameters, (Narinc et al., 2010). The Akaike information criterion evaluate whether the model adequately describes the studied population: the lower the value, the better the model.

To check the fit of the models in the early growth stage the residual square mean of m values, RSMm, was calculated as follows:

$$RSMm = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \quad (6)$$

where: y_i is the i -th observed value, \hat{y}_i is the i -th value estimated by the model and m is the number of observations corresponding to the first phase of the fitted logistic by segmented regression (MII).

The inflection points of the logistic function, determined by equating its second order derivative to zero, are obtained by the usual formula $(-\beta/\gamma; \alpha/2)$ for model I $(-\beta_k/\gamma_k; \alpha_k/2)$, $k = 1, 2$, for model II. In model III, however, the function (3) is a sum and its second order derivative equated to zero:

$$y'' = \sum_{k=1}^2 \frac{\alpha_k \gamma_k^2 X_k (X_k - 1)}{(1 + X_k)^3} = 0 \quad (7)$$

is an equation without explicit solution, which makes the problem more complex (Beyene and Ramakrishnan, 2013), though the solution can be approximated by iterative techniques. In this paper, the Newton-Raphson method is used, in the SAS *proc model* (Statistical Analysis System, version 9.2).

The abscissas of the inflection points of the diphasic logistic sum of functions, therefore, cannot be determined by common formulas for each plot of (7), since there is no fitted function for each phase, but a sum of two functions in the same phase. The solutions of equation (7), denoted by v , are contained in the interval $\tau_1 < x < \tau_2$, where $\tau_k = -\beta_k/\gamma_k$, $k = 1, 2$; this is shown below:

Consider the logistic functions defined for $\alpha_k > 0$, $\beta_k < 0$ and $\gamma_k > 0$,

$$Y_k = \frac{\alpha_k}{(1 + X_k)}, X_k = \exp(-\beta_k - \gamma_k x) = \exp[\gamma_k (\tau_k - x)], \quad \tau_k = -\beta_k/\gamma_k, k = 1, 2. \quad (8)$$

The functions are continuous, differentiable, positive and increasing in the interval $(-\infty, \infty)$ and have an inflection point where x equals $\tau_k = -\beta_k/\gamma_k$. The first order derivatives of y_k are continuous, positive and have maxima at $x = \tau_k$, whose values are the roots of the second order derivative functions:

$$y_k'' = \frac{\alpha_k \gamma_k^2 X_k (X_k - 1)}{(1 + X_k)^3}, \quad (9)$$

that are continuous, positive in the interval $-\infty < x < \tau_k$ and negative in $\tau_k < x < \infty$. Then the sum function

$$y = \sum_{k=1}^2 \frac{\alpha_k}{(1 + X_k)}, X_k = \exp[\gamma_k (\tau_k - x)], \quad (10)$$

is also continuous, differentiable, positive and increasing at $(-\infty, \infty)$. Assuming $\tau_1 < \tau_2$ we have in the interval $-\infty < x < \tau_1$, by (8), $X_1 > 1$ and $X_2 > 1$ and by (9), $y_1'' > 0$ and $y_2'' > 0$. Therefore $y'' = y_1'' + y_2'' > 0$ in this interval. At the point where $x = \tau_1$, $X_1 = 1$, $X_2 > 1$, $y_1'' = 0$, $y_2'' > 0$ and $y'' > 0$.

Then,

in $-\infty < x \leq \tau_1$, $y'' > 0$; there are no inflections in y in this interval. (11)

In the interval $\tau_2 < x < \infty$, by (8), $0 < X_1 < 1$ and $0 < X_2 < 1$ and by (9), $y_1'' < 0$ and $y_2'' < 0$. Therefore $y'' < 0$ in this interval. At the value where $x = \tau_2$, $X_1 < 1$, $X_2 = 1$, $y_1'' < 0$, $y_2'' = 0$ and $y'' < 0$. Then,

where $\tau_2 \leq x < \infty$, $y'' < 0$; there are no inflections in y in this interval. (12)

By (11) and (12), as y'' is continuous on $(-\infty, \infty)$ it follows that there is at least one value of x in the interval (τ_1, τ_2) where y'' is equal to zero. This value where $x = v$ is the abscissa of the inflection point of y in (10).

The stability points of the logistic function, mono or diphasic can be determined by various methods (Pazos et al., 2012). The method that equates the fourth-order derivative of the function y to zero is employed here (Mischan et al., 2011):

$$y^{(4)} = \frac{\alpha \gamma^4 X (X^3 - 11X^2 + 11X - 1)}{(1 + X)^5}, X = \exp(-\beta - \gamma x) \quad (13)$$

which gives the points

$$[-\beta - \log_e(5 + 2\sqrt{6})]/\gamma; \alpha/[2(3 + \sqrt{6})] \quad (14)$$

in model I and

$$\{-\beta_k - \log_e(5 + 2\sqrt{6})\}/\gamma_k; \alpha_k/[2(3 + \sqrt{6})\}], k = 1, 2 \quad (15)$$

in model II.

In model III, where there is a sum of functions, the equation

$$y^{(4)} = \sum_{k=1}^2 \frac{\alpha_k \gamma_k^4 X_k (X_k^3 - 11X_k^2 + 11X_k - 1)}{(1 + X_k)^5} = 0, \quad X_k = \exp(-\beta_k - \gamma_k x), k = 1, 2, \quad (16)$$

has no explicit solution. The solution method is the same used for the determination of inflection points.

The models were adjusted using the procedure *model*, *method=marquardt*, from SAS (Statistical Analysis System, version 9.2). The options *breusch 'pagan'* and *normal* in the *proc model* were employed to verify the homoscedasticity and normality of the residuals. All tests were verified at the significance level of $\alpha = 0.05$.

Observational growth data (volume-age) of the trunks of *Eucalyptus grandis* L. of a reforestation zone in Jacareí, in the state of São Paulo (23°22'27" S, 46°1'34" W) were used to illustrate the methodology. The reforestation zone was available for research and consisted of 150 plants, arranged in three rows of approximately 50 trees each, with a spacing of 3.0 × 2.5 m between plants. One row of 50 plants selected at random and 29 trees with measurements taken at all times were considered. Individual settings for each tree were marked by 11 observations made from 8 to 50 months with the following values of $x = \{8, 10, 12, 15, 19, 21, 25, 27, 30, 36, 50 \text{ months}\}$. The trunk volume data (m³) were calculated from the diameter at breast height (m) and tree total height (m).

Results and discussion

The estimates of parameters α , β , γ , in model I are denoted by a , b and c , respectively; in models II and III, the parameters α_k , β_k , γ_k by a_k , b_k , c_k , $k = 1, 2$, respectively; and in model II, the parameter ι is denoted by ri . Table 1 shows the parameter estimates of models I, II and III. The criteria used to check the fit to the data of eucalyptus are presented in Table 2.

The analysis of the residuals showed that the null hypothesis of homoscedasticity cannot be rejected, for the three models in all fits. The normality hypothesis was not rejected in 100 % of the cases for model I, 84 % for model II and 88 % for model III. The residuals can

be considered independent for 100 % of the adjustments of model I, 88 % for model II and 96 % for model III.

In model III all parameter estimates were in accordance with the constraints $\alpha_k > 0$, $\beta_k < 0$ and $\gamma_k > 0$, $k = 1, 2$, characterizing positive and increasing functions. The fitted values of asymptote in phase 2 of model II, with mean $a_2 = 287.8842$, and the values of the sum of the two asymptotes in the model III, where mean $a_1 + a_2 = 265.7759$, are near to the asymptote estimated in model I where mean $a = 278.9652$.

Compared with the monophasic model, the diphasic models fit better, not only throughout the measuring interval of trees, as shown by the values of residual mean square, RMS, as well as during the initial phase, as shown by the residual squares mean of m initial values, RSMm. Model III is more efficient, both in the initial period fit as for the whole. On average, the reduction in RMS values compared to model I were 75 % for model II and 82 % for model III. The AICc and $r_{y,y}^2$ criteria are similar in the three models, with slight improvements in the diphasic models.

Table 3 shows the values of the abscissas of inflection and stability points in adjusted models and Figure 1 illustrates the adjustments.

Considering the average values of the inflection points obtained to each plant, Table 3, it is seen that the monophasic model has an inflection point with abscissa $x = 33.46$ months. In model II, diphasic segmented regression, the averages are $x = 22.07$ in the first phase and $x = 34.47$ months in the second, the latter being comparable to the average of the monophasic model. For model III – the diphasic sum of two logistics – 20 plants had three inflection points and six plants only one; the means were $x = 21.26$ months which corresponds to the inflection point in the first phase, $x = 26.49$, which is the abscissa of the point that separates the two phases and $x = 36.60$ months, which corresponds to the inflection point

Table 1 – Mean (standard deviation) of parameters estimates of models I - monophasic, II - diphasic segmented and III - diphasic sum. (n = number of trees).

Models (n)	a_1	b_1	c_1	a_2	b_2	c_2	ri
Model I (29)	278.96 (60.374)	-5.17 (0.425)	0.1544 (0.0093)
Model II (26)	101.64 (54.832)	-7.67 (2.307)	0.3542 (0.1266)	287.88 (66.376)	-5.08 (0.833)	0.1478 (0.0260)	25.59 (3.749)
Model III (26)	65.49 (33.958)	-8.00 (3.359)	0.3822 (0.1791)	200.28 (61.590)	-8.32 (2.245)	0.2316 (0.0634)	..

Table 2 – Mean (standard deviation) of criteria for comparison of models, I - monophasic, II - diphasic segmented and III - diphasic sum. Residual mean square, RMS, number of parameters, p , corrected Akaike Information Criterion, AICc, square of correlation coefficient, $r_{y,y}^2$, and residual squares mean, RSMm. (n = number of trees).

Model (n)	RMS	p	AICc	$r_{y,y}^2$	RSMm
Model I (29)	46.05 (32.669)	3	45.52 (7.784)	0.9953 (0.0027)	35.29 (25.439)
Model II (26)	11.39 (14.374)	6	45.08 (12.822)	0.9993 (0.0006)	4.01 (4.620)
Model III (26)	8.38 (6.498)	6	44.98 (7.878)	0.9994 (0.0004)	3.30 (2.406)

m is the number of observations in interval $x = 0$ to $x = ri$, the interception value of the two logistics in model II.

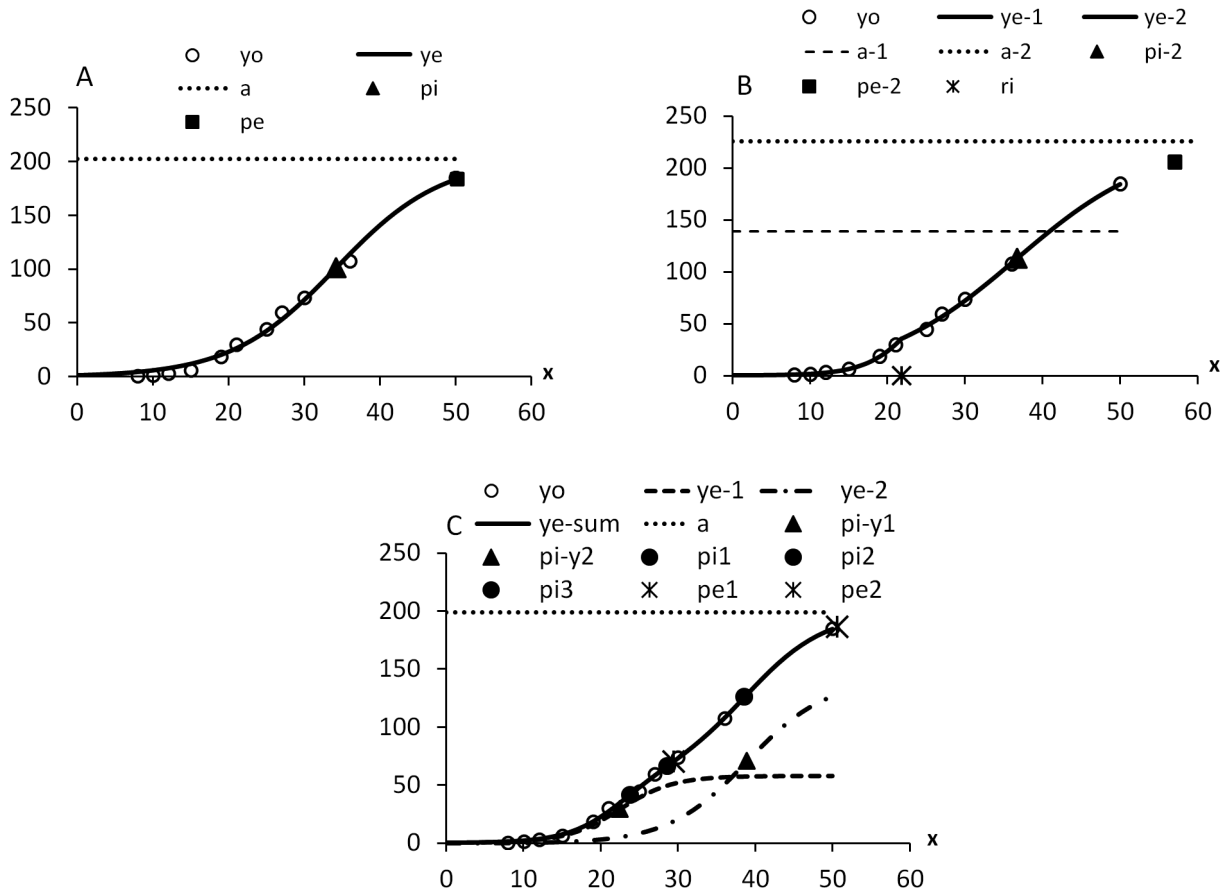


Figure 1 – Logistic models fitted to data from a *Eucalyptus grandis* plant. (A) Monophasic logistic, (B) diphasic segmented logistic, (C) diphasic logistic sum and its components in phases 1 and 2; yo = observed values, ye = fitted values, a = asymptote, ri = abscissa of the intersection point between the phases in model II, pi = inflection points and pe = stability points. In model III, pi-y1 and pi-y2 = the inflection points in phases 1 and 2 curves, respectively; pi1, pi2 and pi3 are inflection points, pe1 and pe2 are stability points in the logistic sum curve. Indices 1 and 2 refer to phases 1 and 2.

Table 3 – Mean (standard deviation) of the x values that are roots of the second order derivative functions (abscissas of the inflection points) and fourth order (abscissas of the stability points) of models I - monophasic, II - diphasic segmented and III - diphasic sum. (n = number of trees).

Models (n)	Abcissas of inflection points			Abcissas of stability points	
	phase 1	inter-phases	phase 2	phase 1	phase 2
Model I (29)	33.46 _(1.902)	..	48.36 _(2.088)
Model II (26)	22.07 _(2.104)	..	34.47 _(1.560)	29.07 _(3.583)	50.46 _(3.594)
Model III (26)	21.26 _(1.392) [20.90 _(1.271)]	26.49 _(2.295)	36.60 _(0.950) [36.80 _(1.062)]	27.16 _(2.314) [27.48 _(2.626)]	47.41 _(2.828) [47.44 _(2.843)]

In brackets, abscissas determined by the functions at each phase of model III.

in the second phase. The abscissas values of the latter inflection point in the model III are higher in all plants, comparing to the values determined by the monophasic model. In this model III, the inflection points of the double logistic for all plants, are within the interval (t_1 , t_2), where t_1 is the abscissa of the inflection point of the first logistic and t_2 , of the second, as demonstrated in this work. Table 3 shows $20.90 < 21.26 < 26.49 < 36.60 < 36.80$ months.

Figure 2 shows the graphs of the derivatives of the first and second orders, with the location of the abscissas of the inflection points for two plants. In the interval $(-\infty; \infty)$, in (B) the derivative of second order $y'' = y''_1 + y''_2$ intercepts the x-axis three times, and in (D) just once. The x value that defines the interphase, 26.49 months in the model III, is quite close to the estimate of the abscissa of the intersection point between the two logistics in model II, 25.59 months. These are points that

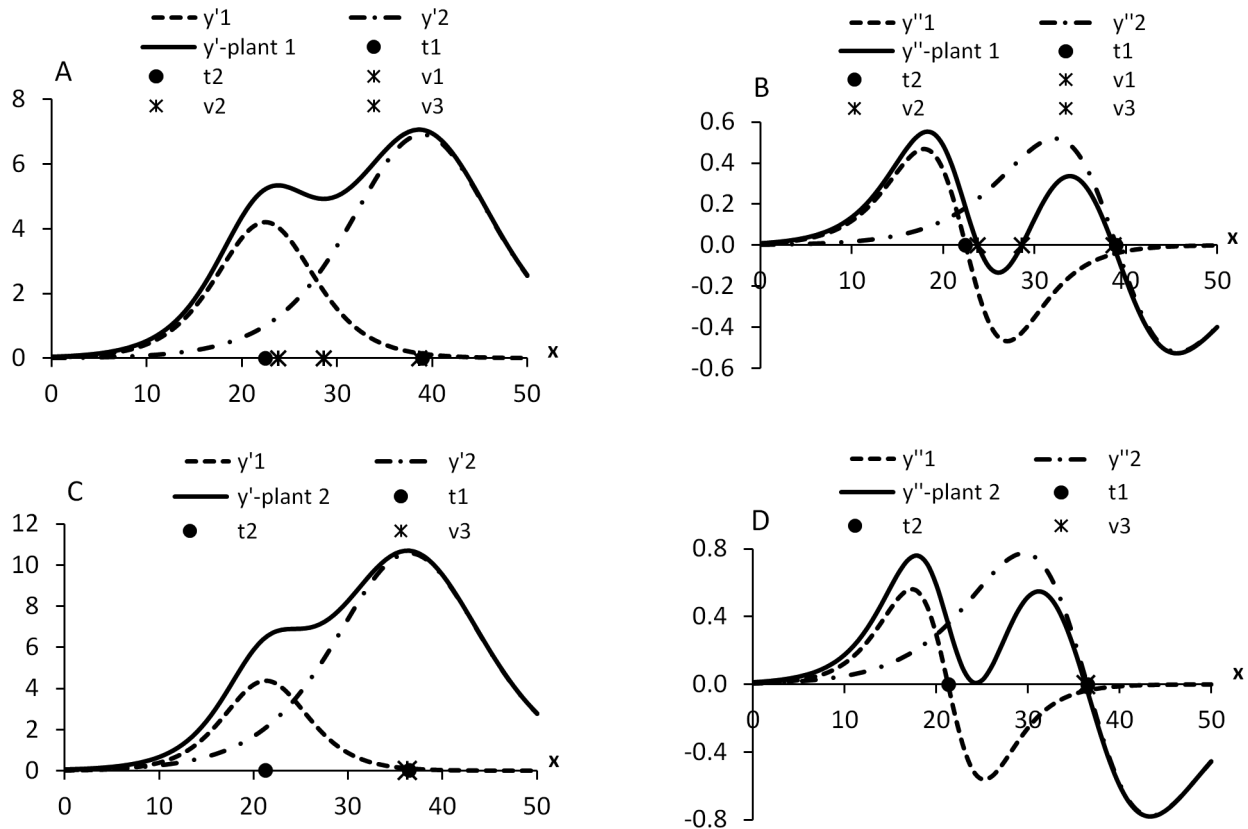


Figure 2 – Inflection points in the diphasic logistic sum in an example of two plants of *Eucalyptus grandis*. In (A) first order derivatives (y') and in (B) second order (y'') for plant 1: $a_1 = 58.057$, $b_1 = -6.506$, $c_1 = 0.2898$, $a_2 = 141.104$, $b_2 = -7.632$, $c_2 = 0.1959$, $t_1 = 22.45$, $t_2 = 38.95$, $v_1 = 23.81$, $v_2 = 28.61$, $v_3 = 38.62$; in (C) and (D) for plant 2: $a_1 = 52.603$, $b_1 = -7.096$, $c_1 = 0.3333$, $a_2 = 222.031$, $b_2 = -6.952$, $c_2 = 0.1906$, $t_1 = 21.29$, $t_2 = 36.48$, $v = 36.28$ months. The abscissas t_1 and t_2 are estimates of τ_1 and τ_2 ; v_1 , v_2 , v_3 , of υ_1 , υ_2 , υ_3 .

can define the separation between the two phases of growth of the organism.

In determining the points that are the roots of the fourth order derivative in model III, the diphasic sum, there were five solutions for (16), considering as stability points the third solution, x_{e_1} = the abscissa of the stability point of y in phase 1, and the fifth solution x_{e_2} = the abscissa of the stability point of y in growth phase 2. These solutions are different from those obtained when determining stability points considering a function for each phase. Table 3 presents the averages of the abscissas of the stability points: in model I they are the third solution of the equation $y^{(4)} = 0$; in model II, the third solution for phase 1 and the sixth for phase 2; in model III, the third solution for phase 1 and the fifth for phase 2.

The abscissas of the stability points in phase two of models II and III are quite similar to the values found for model I, for all plants. The abscissas determined by the logistics at each stage in model III (Table 3 in brackets) are very similar, but not identical to those that are the roots of the fourth order derivative of the sum function; this can be observed in all plants.

Conclusions

The use of a diphasic logistic to represent eucalyptus growth up to 50 months was more effective than a monophasic logistic. The critical points of inflection and stability are determined in the diphasic segmented model by the known formulas for determining these points in monophasic models, but in the diphasic model sum, the points that are the roots of second and fourth orders derivatives cannot be determined explicitly, and their values are different from those determined for the individual logistics components of this model.

References

- Beyene, S.; Ramakrishnan, V. 2013. A numerical method for estimating the variance of age at maximum growth rate in growth models. *Communications in Statistics: Theory and Methods* 42: 1464-1475.
- Bock, R.D.; Wainer, H.; Petersen, A.; Thissen, D.; Murray, J.; Roche, A. 1973. A parameterization for individual human growth curves. *Human Biology* 45: 63-80.

- Breusch, T.S.; Pagan, A.R. 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica* 47: 1287-1294.
- Brody, S.; Ragsdale, A.C. 1921. The rate of growth of the dairy cow. *The Journal of General Physiology* 3: 623-633.
- Draper, N.R.; Smith, H. 1998. *Applied Regression Analysis*. 3ed. John Wiley, New York, NY, USA.
- Fenner, T.; Levene, M.; Loizou, G. 2013. A bi-logistic growth model for conference registration with an early bird deadline. *Central European Journal of Physics* 11: 904-909.
- Grossman, M.; Koops, W.J. 1988. Multiphasic analysis of growth curves in chickens. *Poultry Science* 67: 33-42.
- Koops, W.J. 1986. Multiphasic growth analysis. *Growth* 50: 169-177.
- Koops, W.J.; Grossman, M.; Michalska, E. 1987. Multiphasic growth curve analysis in mice. *Growth* 51: 372-382.
- Koops, W.J.; Grossman, M. 1991. Applications of a multiphasic growth function to body composition in pigs. *Journal of Animal Science* 69: 3265-3273.
- Kurnianto, E.; Shinjo, A.; Suga, D. 1999. Multiphasic analysis of growth curve body weight in mice. *Asian-Australasian Journal of Animal Sciences* 12: 331-335.
- Mendes, P.N.; Muniz, J.A.; Silva, F.F.; Mazzini, A.R.A. 2008. Difasics logistic model in the study of the growth of Hereford breed females. *Ciência Rural* 38: 1984-1990 (in Portuguese, with abstract in English).
- Mendes, P.N.; Muniz, J.A.; Silva, F.F.; Mazzini, A.R.A.; Silva, N.A.M. 2009. Analysis of the difasics growth curve of Hereford females by the Gompertz non-linear function. *Ciência Animal Brasileira* 10: 454-461 (in Portuguese, with abstract in English).
- 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.
- Narinc, D.; Karaman, E.; Firat, M.Z.; Aksoy, T. 2010. Comparison of non-linear growth models to describe the growth in Japanese Quail. *Journal of Animal and Veterinary Advances* 9: 1961-1966.
- Nešetřilová, H. 2005. Multiphasic growth models for cattle. *Czech Journal of Animal Science* 50: 347-354.
- Özkan, M. 2004. Diphasic analysis of growth in Japanese quail. *Asian-Australasian Journal of Animal Sciences* 17: 1281-1285.
- Passos, J.R.S.; Pinho, S.Z.; Carvalho, L.R.; Mischan, M.M. 2012. Critical points in logistic growth curves and treatment comparisons. *Scientia Agricola* 69: 308-312.
- 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.
- Robbins, K.R.; Saxton, A.M.; Southern, L.L. 2006. Estimation of nutrient requirements using broken-line regression analysis. *Journal of Animal Science* 84: E155-E165.
- Schinckel, A.P.; Craig, B.A. 2002. Evaluation of alternative nonlinear mixed effects models of swine growth. *The Professional Animal Scientist* 18: 219-226.
- Seber, G.A.F.; Wild, C.J. 1989. *Nonlinear Regression*. John Wiley, New York, NY, USA.
- Sheehy, J.E.; Mitchell, P.L.; Ferrer, A.B. 2004. Bi-phasic growth patterns in rice. *Annals of Botany* 94: 811-817.