Acta Scientiarum
http://periodicos.uem.br/ojs
ISSN on-line: 1807-8664
Doi: 10.4025/actascitechnol.v46i1.63428

STATISTICS

Accuracy of goodness-of-fit criteria for nonlinear regression: a study via Monte Carlo simulation

Walleff da Silva e Silva*, Felipe Augusto Fernandes, Tales Jesus Fernandes and Joel Augusto Muniz

Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, Departamento de Estatística, Universidade Federal de Lavras, Trevo Rotatório Professor Edmir Sá Santos, s/n, Cx. Postal 3037, 37203-202, Lavras, Minas Gerais, Brazil. *Author for correspondence. E-mail: Walleff100@hotmail.com

ABSTRACT. Nonlinear regression models are widespread in the literature, and one of their main applications is in the study of sigmoidal growth curves. There are several models for this purpose, and the most used are the logistic, Gompertz, von Bertalanffy, and Brody models. Each one of them has its own characteristics and is more suitable for a given curve shape. There are several criteria in the literature for selecting the most appropriate model, but there is no consensus on the best criteria. Thus, the present study aims to evaluate the accuracy of the main selection criteria via Monte Carlo simulation, considering the logistic, Gompertz, von Bertalanffy, and Brody nonlinear regression models. Four simulation scenarios are used, each simulated with ideal curves of the logistic, Gompertz, von Bertalanffy and Brody models. Next, the 4 models are adjusted for each of the scenarios, and the main quality criteria found in the literature are calculated to assess the ability of the criteria to identify the most appropriate model for each scenario. The results show that the criteria asymptotic index, mean absolute error and coefficient of determination choose the correct model more often than the other criteria studied. Although the measures of the Battes and Watts curvature and box bias are important for the evaluation of the goodness-of-fit of the models, they are not indicated for the selection of the best model.

Keywords: Growth curves; Coefficient of determination (R^2); Asymptotic Index (AI); Mean Absolute Error (MAE); nonlinearity measures.

Received on June 02, 2022.
Accepted on February 19, 2024.

Introduction

Regression models are widely used, with applications in many areas. The main objective of regression analysis is to identify and quantify some functional relationships between two or more variables. Prediction and estimation of parameters for subsequent inference are other objectives of regression analysis, and these models can be classified in three ways: linear, nonlinear, and linearizable.

Nonlinear regression models are often used by researchers in various fields of knowledge, such as agriculture, biology, econometrics, engineering, and chemistry. (Mischan & Pinho, 2014). These models are deduced from theoretical assumptions inherent to the phenomenon of interest, and the resulting parameters are interpretable. Growth curves are an application of nonlinear regression models. A growth curve can be characterized by the description of the development of some variables, such as weight, height, and length as a function of age. Mischan and Pinho (2014) state that the growth of living beings shows distinct behaviour; it starts slowly, moves to an exponential phase, and tends to stabilize at the end, which is also known as sigmoidal growth.

Thus, when the response variable has a sigmoidal aspect of development, nonlinear regression models with normal errors are one of the most commonly used methods for describing its behaviour over time, and the most common models are the logistic, Gompertz, von Bertalanffy, and Brody models (Fernandes, Muniz, Pereira, Muniz, & Muianga, 2015; Diel et al., 2019; Fernandes, Fernandes, Pereira, Meirelles, & Costa, 2019; Jane et al., 2020; Prado, Savian, & Muniz, 2020; Silva, Fernandes, Muniz, Muniz, & Fernandes, 2021; Souza et al., 2017; Teixeira et al., 2021).

In practice, different models are fitted to estimate the parameters of the growth curves. When several models are fitted to a dataset, it is important to determine which of the models has the best descriptive fit as well as the best ability to predict the response variable. It is necessary to be aware that there is no way to define true models; appropriate models are chosen to explain the phenomenon with the least possible loss of information. In the choice of the most appropriate model, criteria for evaluating the goodness-of-fit are important; however, there is no consensus on which quality criteria are more efficient in this selection.

Aiming at a lower measurement error, the study of situations controlled by simulations is a viable option. Monte Carlo simulation is a way to study situations in computationally controlled environments and refers to the use of artificial models to represent real data-generation processes to obtain a greater understanding of such processes (Barreto & Howland, 2005).

Therefore, the present study aims to evaluate the accuracy of the main selection criteria by using Monte Carlo simulation and considering the logistic, Gompertz, von Bertalanffy, and Brody nonlinear regression models to describe sigmoidal growth curves.

Material and methods

Simulation

The data analysed were selected by sampling from datasets simulated by the Monte Carlo method using nonlinear regression model equations according to the flowchart shown in Figure 1.

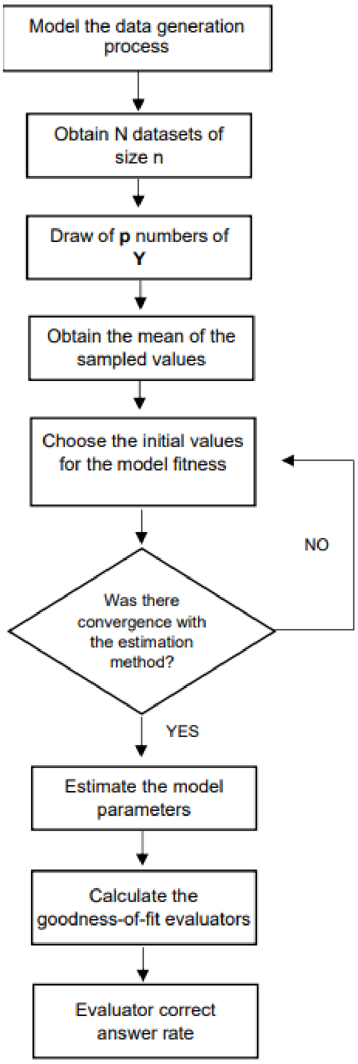

Figure 1. Structure of the Monte Carlo simulation process

The idea for modelling the data-generation process was to use the structure of some nonlinear regression model (logistic, Gompertz, von Bertalanffy, or Brody) for the construction of the simulation model. Random error (ε_i) generated through a process with pseudorandom numbers assuming a normal distribution with zero mean and variance σ^2 (constant and known) was added to the results obtained by this structure.

The random error terms generated 1,000 values following a normal distribution with a mean of zero and a standard deviation proportional to the mean, defined by $\tau \times \bar{X}_i$. The value τ was defined after a survey of bibliographic references on growth curves in different scenarios for the modelling of animal, plant, and fruit growth curves, which provided standard deviation information for the response variable. From the observed deviation values in each study, the mean proportion of the standard deviations in relation to the means at each age was calculated. After the survey, a value of 0.15 was obtained as an estimate of $\bar{\tau}$.

The study population was represented by the 1,000 simulated values for the dependent variable and 15 different ages. Subsequently, a random sampling of 20 values of Y_i in the population for each of the 15 points of the axis X_i was performed, and then the mean of this sampling was calculated. The result formed the final set of values of Y_i (15 means) for the estimation of the model parameters; subsequently, the selection of the appropriate fit was made by the goodness-of-fit evaluators. As the "original model" was known, the previous procedure was repeated 100 times to determine the percentage of correct answers for each evaluator for the selection of the correct model.

For illustrative purposes, suppose that we wish to study the growth of cattle; there are 1,000 cows in the herd, 20 of them are randomly selected, and they are weighed every 8 days, totalling 15 longitudinal measurements. The mean of the 20 weights at each age are calculated, and the 4 models studied are adjusted to this mean weight. We analyse which model had the best fit based on the quality evaluators.

Nonlinear regression models

Data were generated in four scenarios, and the values observed in each scenario come from the simulations performed using the equations from the following models: logistic, Gompertz, von Bertalanffy, and Brody.

$$Y_i = \frac{\alpha}{1 + e^{k(\beta - x_i)}} + \epsilon_i$$

$$Y_i = \alpha e^{-e^{k(\beta - x_i)}} + \epsilon_i$$

$$Y_i = \alpha \left[1 - \frac{e^{k(\beta - x_i)}}{3}\right]^3 + \epsilon_i$$

$$Y_i = \alpha\left[1 - \beta e^{(-k x_i)}\right] + \epsilon_i$$

where i = 1, 2..., n; Y_i is the i-th observation of the dependent variable, x_i is the i-th observation of the independent variable; α is the asymptotic value, i.e., the expected value for the maximum growth of the object under study; β is the abscissa of the inflection point (except in the Brody model, which does not have an inflection point), that is, from where the growth decelerates; k is an index of maturity or precocity and is associated with growth – the higher its value, the less time it takes for the object under study to reach the asymptotic value (α); and ϵ_i is the random error associated with the i-th observation, which is assumed to be independent and identically distributed following a normal distribution of zero mean and constant variance, that is, $\epsilon_i \sim N(0, \sigma^2)$ (Fernandes et al., 2015; Teixeira et al., 2021).

Goodness-of-fit criteria

According to Navarro and Myung (2004), there are several factors to consider when evaluating a model. In general, statistical methods can be used to measure the descriptive sufficiency of a model (by fitting it to the data and testing these adjustments), as well as its generalization ability and simplicity (using model selection tools).

The quality of a model also depends on its interpretability, its consistency with others, and its overall plausibility. This implies inherently subjective judgements but is no less important, explained in Table 1.

Table 1. The main selection criteria presented in the literature were used to compare and evaluate the model fit.

Criterion	Equation	Interpretation
Coefficient of determination	$R^2 = 1 - \frac{SSE}{SQTotal}$	The model with the highest R^2 is considered the most appropriate fit. SSE is sum of squares of residuals and $SQTotal$ is the total sum of squares.
Adjusted coefficient of determination	$R^2_{ajs} = 1 \frac{(n-1)}{n-p} \times (1 - R^2)$	The model with the highest adjusted R^2 is considered the most appropriate fit. Where n is the number of data points, p is the number

Criterion	Equation	Interpretation
		of parameters and R^2 is the Coefficient of determination.
Mean square error	$MSE = \dfrac{SSE}{(n-p-1)}$	For the selection of models, the model with the lowest MSE is desired. SSE is sum of squares of residuals, n is the number of data points and p is the number of parameters.
C_p of Mallows	$C_p = \dfrac{SSE}{S^2} - (n - 2p)$	The best model is desired to have the least biased estimates, so C_p close to p is preferable. SSE is sum of squares of residuals, S^2 is the estimation of residuals variance, n is the number of data points and p is the number of parameters.
Akaike information criterion (AIC)	$AIC = -2\log L(\hat{\theta}) + 2(k)$	Models with lower AIC values are classified as better. Log $L(\hat{\theta})$ is the log-likelihood estimate and k is the number of parameters.
Corrected Akaike information criterion (AICc)	$AIC = -2\log L(\hat{\theta}) + 2(k) + 2\dfrac{k(k+1)}{n-k-1}$	The selection of the most appropriate model is similar to that of the AIC. The terms Log $L(\hat{\theta})$ and k are equal to AIC and n is the number of data points.
Bayesian information criterion (BIC)	$BIC = -2L(\hat{\theta}) + k\ln(n)$	The model that minimizes the BIC value is the best model for the data. $L(\hat{\theta})$ is the maximized values of the likelihood function, n is the number of data points and k is the number of parameters.
Root mean square error (RMSE)	$RMSD = \sqrt{\dfrac{1}{N}\sum_{n=1}^{n}(\hat{y}_i - y_i)^2}$	The closer the RMSE is to zero, the higher the quality of prediction in relation to the observed data. N is the number of data points, y_i is the i-th measurement, and $\hat{y}_i$ is its corresponding prediction.
Mean absolute error (MAE)	$MAE = \dfrac{\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert}{n}$	A model with a lower MAE is preferable. The terms are the equal to RMSE.
Asymptotic index (AI)	$AI = (AAD - MAE) - R^2_{ajs}$	The model with AI closest to zero is considered the best fit model. (AAD – Average Absolute Devation)
Mean prediction error (PE)	$PE_M = 100\left(\dfrac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}\right)$	For the selection of adjustments, models with low PE_M are chosen. The terms are the equal to RMSE.
Box bias	$\%B(\hat{\theta}) = \dfrac{100 \times B(\hat{\theta})}{\hat{\theta}}$	An acceptable bias is one in which $\%B(\hat{\theta}) <$ 1%. $B(\hat{\theta})$ represents the discrepancy between the estimates of the parameters and the true values.
Bates–Watts curvature	$\dfrac{1}{2}\sqrt{F}$	Lower values of intrinsic nonlinearity (IN) and effect of parameters (EP) indicate a greater linear approximation of the model. 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.

It should be noted that, when choosing a model, parsimony must be considered, models with a large number of parameters, in general, present lower residual sum of squares values. When making decisions in various scenarios, it is best to stick with the simplest explanation possible.

Methods

The estimate of the model parameters was performed by using the Gauss–Newton convergence algorithm or the linearization method, which consists of using an expansion of the Taylor series to approximate the nonlinear regression model in linear terms and then applying the least squares method and finding the estimate of the parameters (Mazucheli, Achcar, 2002).

The evaluators used the criteria described above to select the most appropriate model, and the one selected by the most evaluators was chosen as the best fit of the four fitted models (logistic, Gompertz, von Bertalanffy, and Brody). Ideally, the model selected by the evaluators should be the model whose equation was used as a basis for the simulation of the data sampled in the study scenario.

The correct answer rate was calculated for each of the goodness-of-fit evaluators who correctly selected the fit that generated the sampled data of the 100 replicates of the simulation process. The most efficient evaluation criteria were those with the highest percentages of correct model selection.

All analyses and computations were performed in the statistical *software* R (THE R DEVELOPMENT CORE TEAM 2021) by using the packages rlang, tidyverse, dplyr, qpcR, lmtest, nlme, and IPEC, which are available for parameter estimation, statistical tests, and preparation of graphs.

Results

The results presented here were obtained following the simulation reasoning presented in the previous section. Table 2 shows the values of the parameters chosen for each simulation scenario (logistic, Gompertz, von Bertalanffy, and Brody) and the initial values assigned to the start of the iterative method:

Table 2. Values for the parameters chosen in each equation of the simulated model and the initial values used in the subsequent adjustment.

Simulated models (α; β; k)	Initial values of the fitted models (α_0; β_0; k_0)
Logistic (300; 60; 0.1)	(250; 40; 0.02)
Gompertz (300; 60; 0.05)	(200; 40; 0.03)
von Bertalanffy (300; 1; 0.04)	(250; 0.03; 0.02)
Brody (300; 0.3; 0.06)	(250; 0.3; 0.04)

The parameter α, which represents the asymptote (highest point on the curve), was set at 300 to better explain the results in all simulations of the models performed, while the other parameters (β, k) were altered according to the curve shape characteristics for each model.

The parameter k is the most sensitive parameter in model estimation, mainly due to the possible association with another parameter, depending on the model parameterization, as exemplified by Fernandes et al. (2015), where it can be correlated with parameter β. Therefore, the selection of the initial value of k has a significant influence on the convergence of the model compared to the initial values of the other parameters. The convergence space of each parameter in the analyzed models is conditioned by the values present in the observed data set, with emphasis on the parameter k, whose variation is generally between 0 and 1, as evidenced in the literature.

All 4 simulations of the nonlinear regression models were performed, and the means were obtained for each scenario (logistic, Gompertz, von Bertalanffy, and Brody). The same conditions were replicated 100 times, goodness-of-fit was calculated for all models within these scenarios, and the percentage of the "correct model" choice was calculated. This is summarized in Figures 2 to 9.

There was no simulation convergence using the equations of the logistic and Gompertz models when trying to model the curve of the Brody model. A possible explanation is the sigmoidal shape of these models; the Brody model has a curve shape with an already decreasing growth rate and is therefore nonsigmoidal, different from the other models studied.

It is important to note that the initial values of the adjustments were chosen to be well below the initial simulated values. In all scenarios, this was done to verify how accurately the model estimated its parameters, as shown in Table 1.

After all models were adjusted for each scenario in all 100 replicates, the evaluation of the most appropriate model was performed by the goodness-of-fit criteria. Figure 2 shows the percentage of choice of the appropriate model for the simulation considering the equation of the logistic model in its 100 replications.

Each column discriminated by colour in Figure 2 represents the model fitted to the simulated data and the respective percentages of the goodness-of-fit choice to the appropriate model for that simulation performed by each evaluator.

In this scenario, there were no discrepancies regarding the choice of the most appropriate model among the evaluators shown in Figure 2 considering the simulation through the equation of the logistic model.

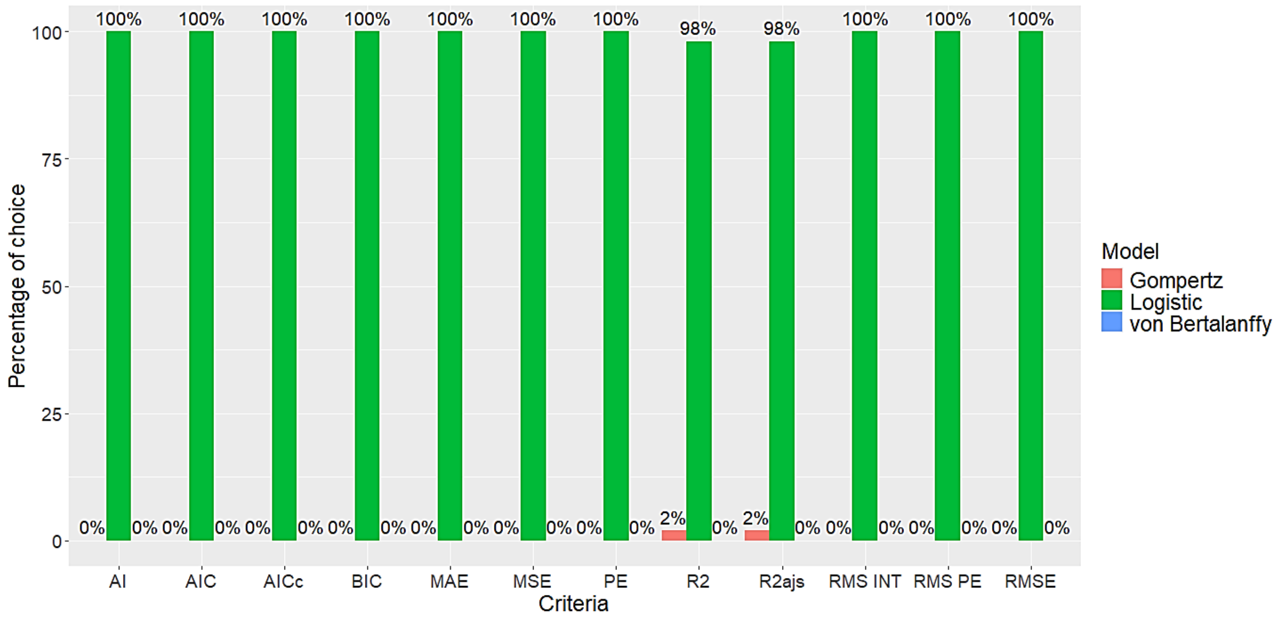

Figure 2. Percentage of choice of the evaluators for the model adjustments considering the simulated logistic model.

Overall, the evaluators chose the fit of the logistic model as the most appropriate to represent scenario 1 of simulation, which used the equation of the logistic model to generate the data. This represented 100% of the choices of the listed criteria, except for the coefficient of determination (R^2) and adjusted coefficient of determination (R^2_{ajs}). In 98% of the replicates, these evaluators considered the logistic model as preferable, and the complementary percentage was allocated to the Gompertz model, which shows a behaviour closer to the logistic model.

Evaluators such as R^2 and MAE are used in many studies in the literature. Deprá, Lopes, Noal, Reiniger, and Cocco (2016), found that through these evaluators, the fit of the logistic model explained most of the variability in the variables plant height and number of leaves per plant for the genotypes of creole corn cultivars and maternal half-sib progenies according to the thermal sum. Diel et al. (2020) studied strawberry production and selected the logistic model as the most appropriate based on the AIC, R^2, and Battes and Watts curvature measures.

The results of the box bias are presented in a separate graph. Figure 3 shows the percentage of choice information for the smallest biases in each parameter of the fitted models.

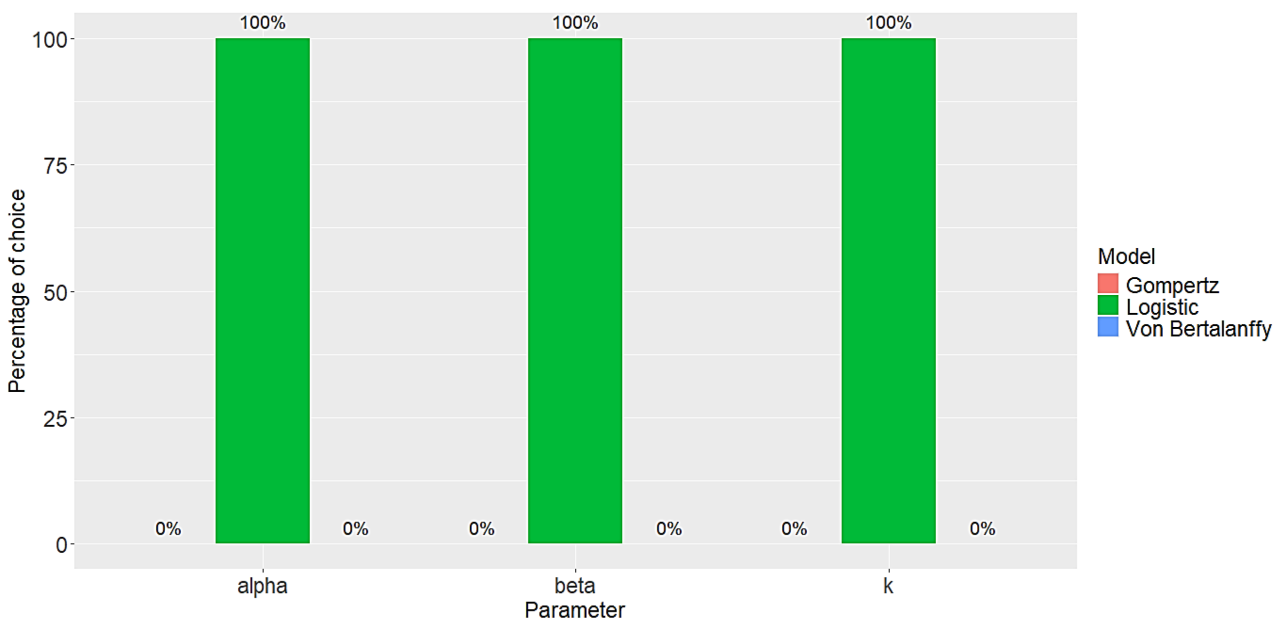

Figure 3. Percentage of box bias for the choice of the fit and models considering the simulated logistic model.

Once again, the agreement of the most appropriate fit in relation to the parameters and box bias was noticeable. All evaluators selected the fit of the logistic model as the most appropriate, and this fit was chosen 100% of the time for the repetitions. Thus, the fit of the logistic model in the simulation scenario for this same model showed lower box biases for all parameters.

In contrast, Figure 4 shows some changes observed for scenario 2 using the Gompertz model equation in relation to the results presented in the simulation of the previous scenario.

Figure 4 shows that the fit of the Gompertz model was the most appropriate for representing the simulation of scenario 2, and it reflects the high percentages of choice of this model for most of the evaluators presented here. A total of 11 out of the 12 evaluators (Figure 4) had higher percentages of choice for this model, all with percentages greater than 85% choice.

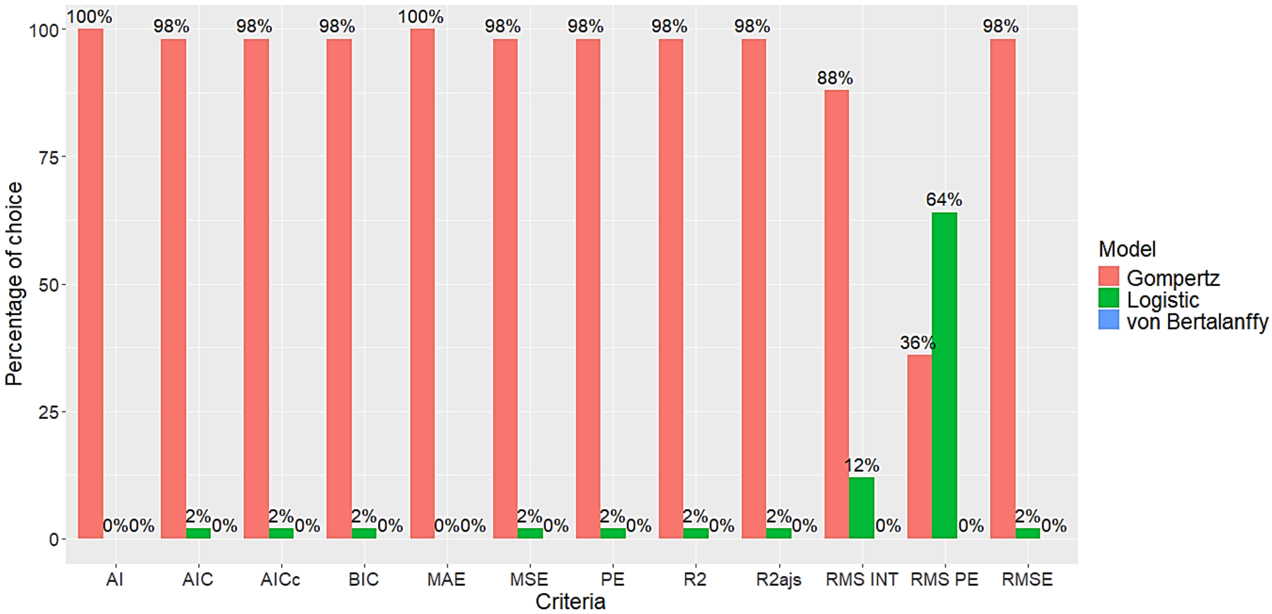

Figure 4. Percentage of choice of evaluators for model fitting considering the simulated Gompertz model

It can be noted that the evaluators asymptotic index (AI) and mean absolute error (MAE) were slightly more accurate than the others for this scenario. In 100% of the replicates, they selected the fit of the Gompertz model as the "original equation" for data simulation, while the Akaike information criterion (AIC), corrected Akaike information criterion (AICc), Bayesian information criterion (BIC), mean prediction error (PE_M), mean square error (MSE), R^2, R^2_{ajs}, and root mean square deviation (RMSE) selected the same model in only 98% of the replicates.

Drumond et al. (2013) studied meat quail growth and explained that considering the criteria for the model goodness-of-fit, R^2 was not a good indicator for choosing the models, as all of the values were high and similar. Thus, the models that best fit the data were selected through the AI. Lower AI values were observed for the Gompertz model in male quails and for the logistic model in females.

The Batts and Watts curves (intrinsic and parametric) did not lead to the same conclusion regarding the choice of adequate fit. The intrinsic curvature (RMS INT) led to the selection of the correct ideal model in 88% of the repetitions, while the parametric curvature (RMS PE) did not lead to the same conclusion in accepting the Gompertz model as the original equation. The parametric curvature evaluations led to the conclusion that the logistic fit would be the ideal model for 64% of the repetitions and that the Gompertz model fit would be satisfactory in only 36% of the replications.

Considering the choice of the appropriate model with the box bias, Figure 5 shows that the fit of the Gompertz model obtained higher percentages of the lowest bias for all parameters. Considering the smallest biases, the parameter k, which represents a relationship between the first derivative calculated at the inflection point and the asymptotic value, presented the highest percentage of choice for the Gompertz model – 98% of the cases. In turn, the parameter β considered the fit of the Gompertz model to have the least bias in 55% of the fits, while 58% of the adjustments for the asymptote, parameter α, showed lower biases for the fit of the original data equation – the Gompertz model.

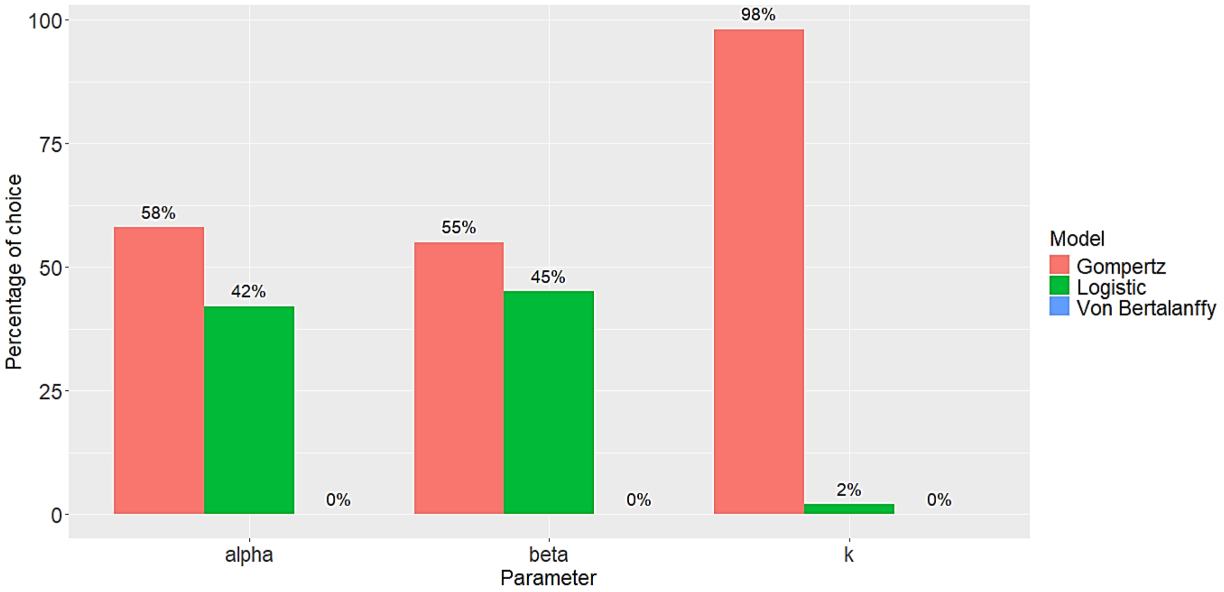

Figure 5. Percentage of box bias for the choice of the fit and models considering the simulated Gompertz model.

Figure 6 shows results for the fit quality evaluators by simulating scenario 3 using the von Bertalanffy model equation and then fitting the 100 simulation replicates. For the most part, the fit quality evaluators for this scenario agreed that the von Bertalanffy model was the appropriate model, which simulated the initial data.

Despite the high correct answer rate of most evaluators in selecting the ideal fit from the equation that simulated the data in this scenario, the for the MAE and AI evaluators had clearer accuracy according to their high percentage of "correct model" selection – both evaluators agreed in 89% of the simulations that the von Bertalanffy model generated the data, which was a percentage higher than the percentage of all the other evaluators discussed here.

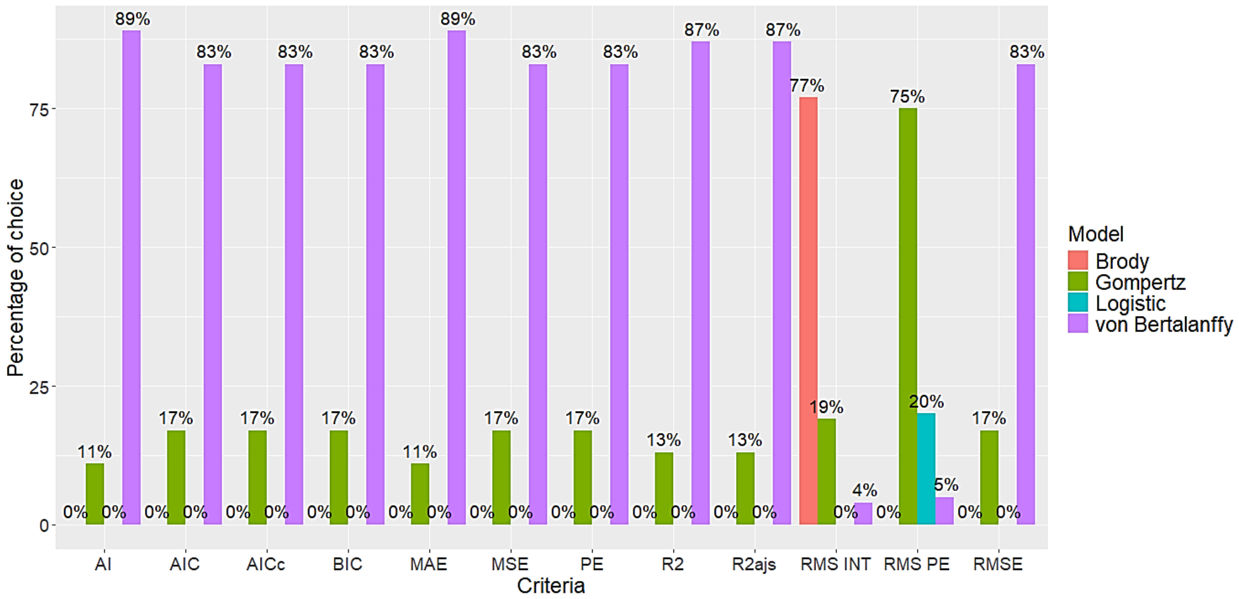

Figure 6. Percentage of choice of the evaluators for the model adjustments considering the simulated von Bertalanffy model.

The use of the AI in growth curve studies and its effectiveness in choosing the appropriate model was emphasized by Veloso et al. (2015), who used the AI to select the appropriate model because it is a more complete evaluation criterion. Thus, the AI indicated that the best fit was presented by the curve proposed by von Bertalanffy, which is recommended for describing the growth pattern of birds of the free-range chicken genotypes evaluated in the study.

The coefficient of determination and its adjusted form (R^2 and R^2_{ajs}) demonstrate 87% accuracy in the choice of the von Bertalanffy model fit. The evaluators AIC, AICc, BIC, PE, MSE, and RMSE formed a group with similar behaviour regarding the selection of the appropriate fit, where they considered the von Bertalanffy model equation to be the simulated data generation matrix in 83% of the replicates.

It is noteworthy how inaccurate the Battes and Watts curvatures were in selecting the correct model that simulated the data from this scenario. According to the intrinsic and parametric curvature measurements, the data came from a von Bertalanffy model equation in only 4 and 5%, respectively, of the 100 fitted replicates. The accuracy of these evaluators in choosing the correct equation that generated the data in this scenario was the lowest of all the simulations. According to these evaluators, the intrinsic curvature indicates that the data showed behaviour characteristic of the Brody model in 77% of the replicates and that the parametric curvature shows behaviour similar to that of the Gompertz model in 75% of the repetitions.

Figure 7 shows the percentages for the adjustments with lower biases by parameter. In this scenario, parameter k had the highest correct answer rate considering the lowest bias; the von Bertalanffy model was considered ideal to represent the simulated data in 61% of the fits. Only 29% of the simulations presented lower box bias results for α for the initially simulated model, while 69% of the fits show lower biases for the Gompertz model for this same parameter.

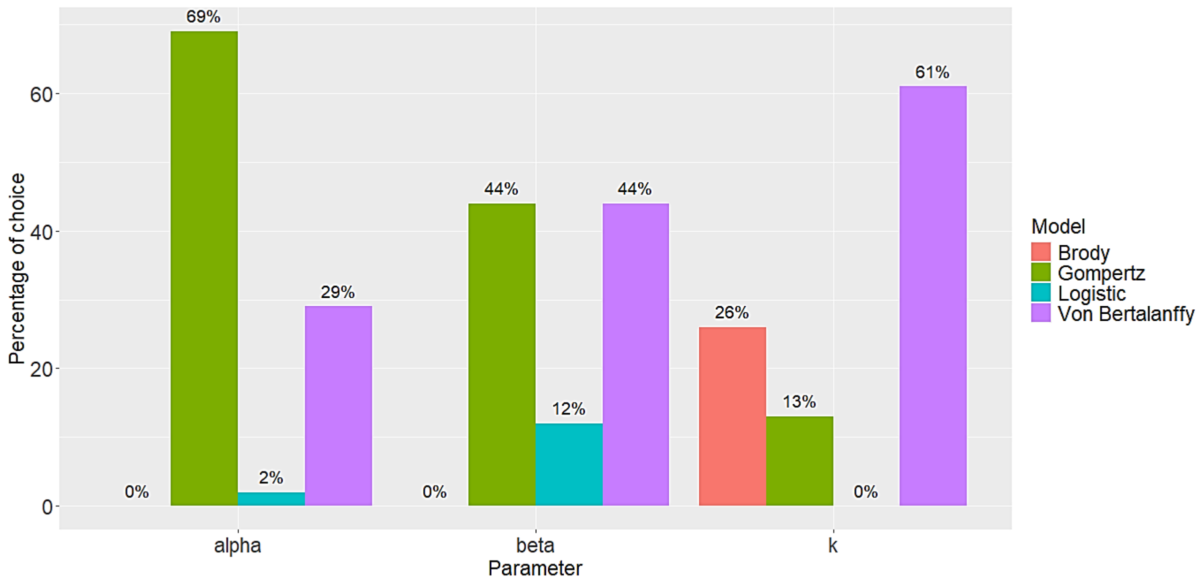

Figure 7. Percentage of box bias for the choice of the fit and models considering the simulated von Bertalanffy model.

Finally, there was no fit preference for the parameter β considering the smallest bias, and the selection of the "real model" – von Bertalanffy – was responsible for 44% of the adjusted replicates, and it tied with the Gompertz model. Thus, the results show that the box bias was not efficient in selecting the appropriate fit in this simulation scenario.

Simulation scenario 4 using the Brody model equation led to key results for the evaluation of the fit quality criteria and definition of the objective of this study because it presents a growth pattern different from the patterns of others (Figure 8).

According to the results obtained by the goodness-of-fit evaluators in this scenario, most were mistaken in relation to the selection of the initial simulation model. According to the evaluators and the highest percentages of choice, the fit of the logistic model would be ideal, while the Brody model that generated the simulated data was correctly selected only by the evaluators MAE, AI, R^2, and R^2_{ajs}.

The group composed of AIC, AICc, BIC, PE, MSE, and RMSE showed greater choice for the logistic model in terms of percentage. In 50% of the replicates, these evaluators indicated that the data came from the logistic model equation, while these evaluators were correct regarding the appropriate model only in 46% of the total number of simulations, where the Brody model was selected as correct. R^2 and R^2_{ajs} chose the Brody model as appropriate in this scenario in 47% of the repetitions.

Veloso et al. (2016) recommended using goodness-of-fit evaluators other than R^2_{ajs} because the differences between the R^2_{ajs} values of the different models used in the growth curve study were negligible. The best adjustment groups for the growth of broiler chickens were those that, in general, showed simultaneously lower values of MSE, AIC, BIC, and AI.

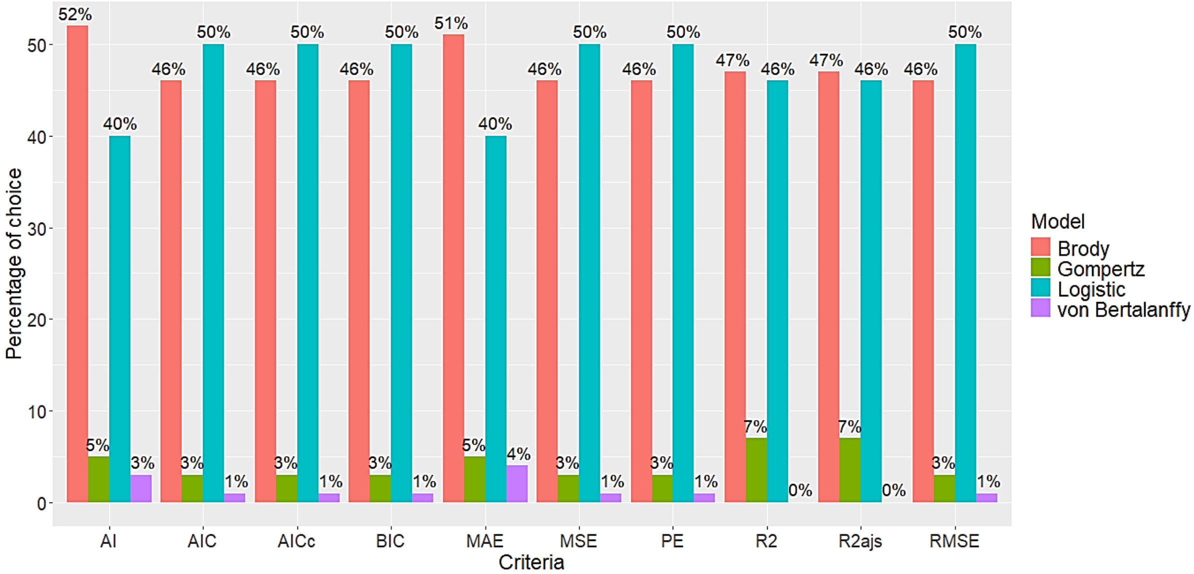

Figure 8. Percentage of choice of the evaluators for the model adjustments considering the simulated Brody model.

Regarding the accuracy of the selection, the MAE selected the Brody model as adequate in 51% of the simulations, and the difference between the Brody model and the second most chosen model (logistic) was 11%; the MAE evaluator was more assertive than the other indicators.

The AI was even more precise regarding choosing the Brody model for this scenario, as it initially generated the simulated data. This evaluator obtained correct answers in 52% of the adjusted repetitions, a 12% difference from the logistic model, which was the second most appropriate fit. Thus, this evaluator obtained the highest accuracy of the evaluators studied.

For this scenario the Battes and Watts curvature did not show convergence of results; thus, it was not possible to evaluate the behaviour of this criterion for the selection of the appropriate model. The results for the box bias converged only for the fits of the von Bertalanffy and Brody models, and their percentages are shown in Figure 9.

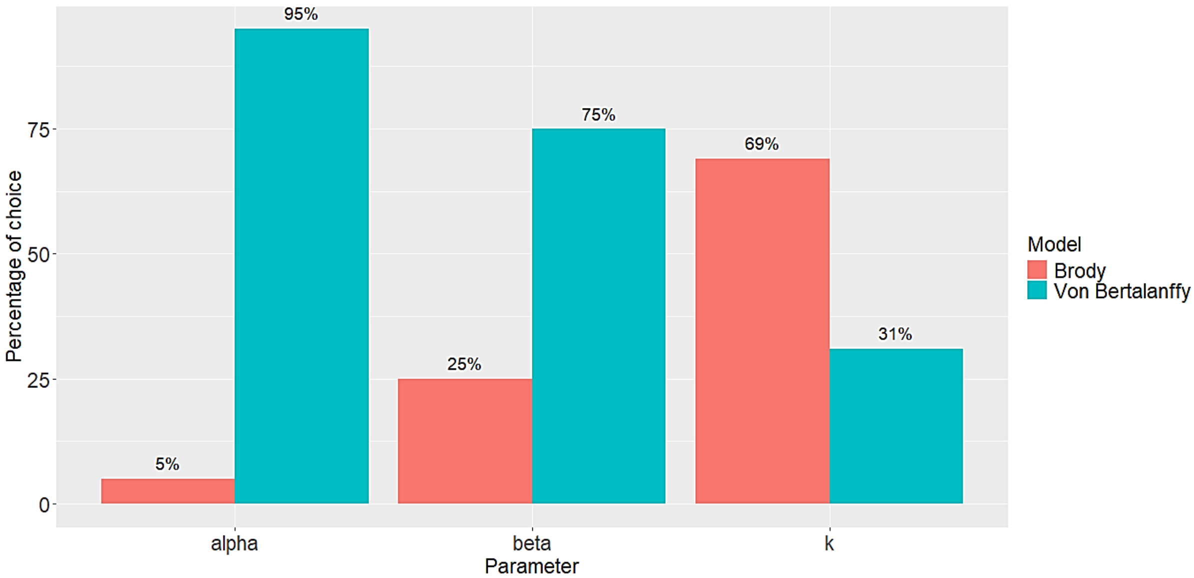

Figure 9. Percentage of box bias for the choice of the fit and models considering the simulated Brody model.

The adjustments of the von Bertalanffy model showed lower biases for the parameters α and β in 95% and 75% of the replicates, respectively, in this simulation scenario. The choice for the Brody model was satisfactory only for the parameter k, which showed lower biases in 69% of the replicates.

In general, it can be concluded through the graphs presented (Figures 2 to 9) that most of the evaluators studied chose correctly in some of the scenarios. However, some evaluators were more efficient than others in selecting the most appropriate fit equation, the equation that initially simulated the fitted data. The accuracy or efficiency of the evaluators R^2, R^2_{ajs}, AIC, AICc, BIC, PE, MSE, RMSE, MAE and AI are very close for small variations in the curve shape, such as in the case of logistic and Gompertz models.

In addition, AIC, AICc, and BIC did not differ regarding the ability to select the appropriate model; thus, if authors wish to use any of these evaluators, they can choose only one. The same was observed for the evaluators PE, MSE, and RMSE, and therefore, it was not necessary to use more than one evaluator from each of these groups.

The effectiveness of the coefficient of determination in choosing the appropriate model is questioned by some authors in a study of growth curves. Spiess, and Neumeyer, (2010) used a Monte Carlo simulation to evaluate the results of R^2 in adjustments of nonlinear models in pharmacy and biomedicine, and they found that the R^2 values decreased and the AICc values increased with the highest error; however, the greatest criticism of the authors regarding the use of R^2 is due to its inefficiency in the selection of models in the simulated scenario (many ties), where the differences between the R^2 of the various models occurred only in the fourth decimal place (ten thousandths).

Considering all the studied scenarios, the evaluators MAE and AI stand out for their high efficiency in choosing the appropriate model. It is noteworthy that the percentage of correct answers of these evaluators was always followed by those of the evaluators R^2 and R^2_{ajs}, which were also efficient and useful as evaluators for the selection of nonlinear regression models, mainly because they provide a simple and clear idea of how much the data variation is explained by the model under study.

Mallows's Cp is not shown in the figures because it did not present significant results in this study. Its value remained constant for all models and different simulation scenarios, and this evaluator was not efficient in classifying these nonlinear models.

The nonlinearity measures, although necessary to evaluate the suitability of the linear approximations at the time of parameter estimation, showed that the Battes and Watts curvatures as well as the box bias were not effective in selecting the most appropriate model. Therefore, it is suggested that researchers use these nonlinearity measures only to verify the suitability of the model and to verify whether the Battes and Watts curvatures are significant and whether the biases are small, for example. However, they should not be used as selection criteria (the lower the better), as commonly found in studies in the literature such as those by Fernandes et al. (2015), Sari et al. (2018), Diel et al. (2019), Silva, and Savian, (2019) and Diel et al. (2020).

Conclusion

By considering the simulation scenarios of growth curves with the different shapes and by evaluating the fit of the logistic, Gompertz, von Bertalanffy, and Brody models, it is concluded that the following evaluators (in this order) should be used for the selection of the best model to decide which model is the most appropriate to describe the data: the asymptotic index (AI), mean absolute error (MAE), and coefficient of determination (R^2) (or R^2_{ajs} if there are models with different numbers of parameters). If more than one evaluator is deemed necessary, only one of the information criterion – either the Akaike information criterion (AIC), corrected Akaike information criterion (AICc), or Bayesian information criterion (BIC) – and only one of the following need to be used: the mean prediction error (PE), mean square error (MSE), or root of the mean square error (RMSE). The measures of Battes and Watts curvature and box bias, although important for the evaluation of the goodness-of-fit of the models, are not indicated for the selection of the best model.

Acknowledgements

The present study was conducted with the financial support of the National Council for Scientific and Technological Development (CNPq).

References

Barreto, H., Howland, F. (2005). *Introductory econometrics. Using Monte Carlo simulation with Microsoft Excel*. Cambridge, UK: Cambridge University Press Cambridge.

Deprá, M. S., Lopes, S. J., Noal, G., Reiniger, L. R. S., & Cocco, D. T. (2016). Modelo logístico de crescimento de cultivares crioulas de milho e de progênies de meios-irmãos maternos em função da soma térmica. *Ciência Rural, 46*(1), 36-43. DOI: https://doi.org/10.1590/0103-8478cr20140897

Diel, M. I., Sari, B. G., Krysczun, D. K., Olivoto, T., Pinheiro, M. V. M., Meira, M., ... Lúcio, A. D. (2019). Nonlinear regression for description of strawberry (*Fragaria x ananassa*) production. *The Journal of Horticultural Science and Biotechnology, 94*(2), 259-273. DOI: https://doi.org/10.1080/14620316.2018.1472045

Diel, M. I., Lúcio, A. D., Valera, O. V. S., Sari, B. G., Olivoto, T., Pinheiro, M. V. M., ... Shmidt, D. (2020). Production of biquinho pepper in different growing seasons characterized by the logistic model and its critical points. *Ciência Rural, 50*(4). DOI: https://doi.org/10.1590/0103-8478cr20190477

Drumond, E. S. C., Gonçalves, F. M., Veloso, R. C., Amaral, J. M., Balotin, L. V., Pires, A. V., & Moreira, J. (2013). Curvas de crescimento para codornas de corte. *Ciência Rural, 43*(10), 1872-1877. DOI: https://doi.org/10.1590/S0103-84782013001000023

Fernandes, T. J., Muniz, J. A., Pereira, A. A., Muniz, F. R., & Muianga, C. A. (2015). Parameterization effects in nonlinear models to describe growth curves. *Acta Scientiarum Technology, 37*(4), 397–402. DOI: https://doi.org/10.4025/actascitechnol.v37i4.27855

Fernandes, F. A., Fernandes, T. J., Pereira, A. A., Meirelles, S. L. C., & Costa, A. C. (2019). Growth curves of meat-producing mammals by von Bertalanffy's model. *Pesquisa Agropecuária Brasileira, 54*. DOI: https://doi.org/10.1590/S1678-3921.pab2019.v54.01162

Jane, S. A., Fernandes, F. A., Silva, E. M., Muniz, J. A., Fernandes, T. J., & Pimentel, G. V. (2020). Adjusting the growth curve of sugarcane varieties using nonlinear models. *Ciência Rural, 50*(3). DOI: https://doi.org/10.1590/0103-8478cr20190408

Mazucheli, J., & Achcar, J. A. (2002) Algumas considerações em regressão não-linear. *Acta Scientiarum, 24*(6), 1761-1770. DOI: http://dx.doi.org/10.4025/actascitechnol.v24i0.2551

Mischan, M. M., & Pinho, S. Z. (2014). *Modelos não lineares [recurso eletrônico]: funções assintóticas de crescimento*. São Paulo: Cultura Acadêmica, 184 p.

Navarro, D. J., & Myung, In Jae. (2004). *Model evaluation and selection*. Columbus, USA: Ohio State University, Department of Psychology, 6p.

Prado, T. K. L., Savian, T. V., & Muniz, J. A. (2013). Ajuste dos modelos gompertz e logístico aos dados de crescimento de frutos de coqueiro anão verde. *Ciência Rural, 43*(5), 803–809. DOI: https://doi.org/10.1590/S0103-84782013005000044

Sari, B., Olivoto, T., Diel, M., Krysczun, D., Lúcio, A., & Savian, T. (2018). Nonlinear Modeling for Analyzing Data from Multiple Harvest Crops. *Agronomy journal, 110*(6), 2331-2342. DOI: https://doi.org/10.2134/agronj2018.05.0307

Silva, P. V., & Savian, T. V. (2019). Chanter model: nonlinear modeling of the fruit growth of cocoa/Modelo Chanter: modelagem não linear do crescimento de frutos do cacaueiro. *Ciencia Rural, 49*(11). DOI: https://doi.org/10.1590/0103-8478cr20190409

Silva, W. S., Fernandes, F. A., Muniz, F. R., Muniz, J. A., & Fernandes, T. J. (2021). Eucalyptus grandis x Eucalyptus Urophylla growth curve in different site classifications, considering residual autocorrelation. *Brazilian Journal of Biometrics, 39*(1), 122-138. DOI: https://doi.org/10.28951/rbb.v39i1.511

Souza, F., Fernandes, T., Moura, R., Meirelles, S., Ribeiro, R., Cunha, F., & Muniz, J. (2017). Nonlinear modeling growth body weight of Mangalarga Marchador horses. *Ciência Rural, 47*(4). DOI: https://doi.org/10.1590/0103-8478cr20160636

Spiess, A., & Neumeyer, N. (2010). An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach. *BMC pharmacology, 10*(6). DOI: https://doi.org/10.1186/1471-2210-10-6

Teixeira, G. L., Fernandes, T. J., Muniz, J. A., Souza, F. A. Ca., Moura, R. S., & Melo, R. M. P. S. (2021). Growth curves of campolina horses using nonlinear models. *Livestock Science, 251*. DOI: https://doi.org/10.1016/j.livsci.2021.104631

The R Development Core Team. (2021). *R: A language and environment for statistical computing*. Retrieved from https://www.R-project.org/

Veloso, R. C., Pires, A.V., Torres Filho, R. A., Drumond, E. C. S., Costa, L. S., Amaral, J. M., & Pereira, I. G. (2015) Crescimento de genótipos de frangos tipo caipira. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 67*(5), 1361-1371. DOI: https://doi.org/10.1590/1678-4162-6816

Veloso, R. C., Winkelstroter, L. K., Silva, M. T. P., Pires, A. V., Torres Filho, R. A., Pinheiro, S. R. F., ... Amaral, J. M. (2016) Seleção e classificação multivariada de modelos não lineares para frangos de corte. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 68*(1), 191-200. DOI: https://doi.org/10.1590/1678-4162-7894