Research Article

Open Access CrossMark

Description of the strawberry production cycle via non-linear quantile regression models

Abstract

The description of the production cycle of crops with multiple harvests, such as strawberries, can be conducted using non-linear regression models. When there is a lack of homogeneity of variances and non-normality of errors, an alternative is to use non-linear quantile regression. Thus, in this study, the objective was to propose more robust estimates of the parameters of the growth models via a non-linear quantile regression model, with greater precision and parsimoniousness, using the Logistic, Gompertz, von Bertalanffy, and Brody models with the variables cumulative number and cumulative mass of strawberry fruits. The work was conducted using data from two experiments conducted in Santa Maria, RS, and Brazil. In Experiment I, the treatments were four strawberry cultivars: LBR F, LBR, Albion, and Estiva; in Experiment II, there were three strawberry cultivars: Albion, Estiva, and LBR. The non-linear regression models Logistic, Gompertz, and von Bertalanffy for the variables cumulative number of strawberry fruits and cumulative mass of strawberry fruits, both by the ordinary least squares method and by the non-linear quantile regression method, were those that presented the best results in the quality of fits. The Logistic model showed high accuracy in Experiment I (in the central quantiles) and maintained a dominant performance in Experiment II (between 0.44 and 0.70), even with greater data variability. The use of non-linear quantile regression is recommended as an alternative for fitting the Logistic, Gompertz, and von Bertalanffy models in data with a sigmoidal distribution when the assumptions are not met. The NLQR showed a drastic reduction in residual deviations compared to the OLS, being between 35 and 2,450 times more accurate in the experiments.

Keywords: assumptions violation, multi-harvest crops, quantile modeling

Volume 15 Issue 1 - 2026

Valdecir José dos Santos,[1] Alessandro Dal'Col Lúcio,[2] Dilson Antônio Bisognin,[2] Gabriel de Araujo Lopes[2]

[1]Integrated Secretariat for Postgraduate Courses, Federal University of Santa Maria, Brazil
[2]Crop Science Department, Rural Science Center, Federal University of Santa Maria, Brazil

Correspondence: Valdecir José dos Santos, Integrated Secretariat for Postgraduate Courses, Frederico Westphalen Campus, Federal University of Santa Maria, Frederico Westphalen 98400-000, Brazil, Tel +55 55 9 9901-3455

Received: February 12, 2026 | **Published:** February 23, 2026

Introduction

The strawberry plant (*Fragaria x ananassa* Dusch.) is a crop with a high production capacity in small cultivation areas, great economic and social importance, and expressive production. It is one of the most significant in the horticulture sector in family farming, universally attractive due to its sensory qualities, benefits for human health, and profitability.[1,2] It is a crop of great importance in south and southeast regions of Brazil, and it serves as a good alternative for diversifying production and improving income on family farms. The cultivation areas of this fruit are generally small, ranging from 0.2 to 2.0 hectares, due to its high demand for labor.[3] Worldwide strawberry production in 2022 was 9.57 million tons, with China and the United States being the largest producers, at 3.35 and 1.26 million tons, respectively. Brazil ranked tenth, with 183,922.5 tons of strawberries.[4]

The strawberry plant is a crop with multiple harvests on the same plant, and the cumulative values of the productive variables in each harvest show that production begins slowly and undergoes exponential growth, then decreases until it stabilizes. This type of sigmoidal response is typical of non-linear regression models known as growth models.[5]

Biological non-linear growth models can be used to extract information from a dataset. They provide the reality of the production cycle in each experimental treatment, allowing for inferences and interpretations that are not obtained in analyses of variance or complementary statistical tests, such as comparisons of treatment means or linear regression analyses.[6–8] Data related to strawberry production throughout its production cycle follow a sigmoidal response and, at times, with a lack of homogeneity of residual variances and non-normality of errors.[6,9] In this way, it is possible to use non-linear regression models as a statistical tool for data analysis.[6,7,10,11] There are many non-linear regression models used to describe growth curves, including the Logistic, Brody, Von Bertalanffy, Gompertz, Richards, and Santos models.[12]

In the process of estimating the parameters of non-linear regression models, it is necessary to meet their assumptions, the most problematic of which are the homogeneity of residual variances and the normality of errors. If these assumptions are not met, some alternatives to overcome this situation are presented, and non-linear quantile regression (NLQR) is one of them.

NLQR is a statistical analysis technique that models the quantile (τ) of a response variable in relation to the dependent variable.[13] It can be efficient in describing data with asymmetry, heterogeneous variances, or outliers.[14] NLQR is a method that locates a non-linear model, suitable for describing growth curves, as it has parameters with practical biological interpretation, between the dependent variable and a set of independent.[9]

Thus, this study is justified by the lack of adequate models for evaluating productivity data in multiple-harvest crops, such as strawberries, due to the heterogeneity of variances and non-normality of errors. Thus, the purpose of this study is to obtain more robust estimates of the parameters of growth models using a non-linear quantile regression model, which offers greater precision and parsimony.

Material and methods

Data

Strawberry productivity data were obtained by Lopes[15] from two experiments conducted in a highly randomized design in Santa Maria, RS, Brazil, in 2021, as described in Table 1.

Experiment I was installed on area of the Plant Breeding and Vegetative Propagation Center (29°43'24"S, 53°43'11"W, at an altitude of 99 m), belonging to the Department of Plant Science of the Federal University of Santa Maria (UFSM) and Experiment II, was conducted in a protected system on a rural property (29°40'02"S, 53°41'11"W, at an altitude of 97 m). The region has a Cfa climate, according to the Köppen-Geiger classification.[16] The minimum and maximum daily air temperatures were measured at a meteorological station belonging to the 8th Meteorological District of the National Institute of Meteorology (DISME/INMET) located in the Department of Plant Science at UFSM.

The fruits were harvested twice a week from 75% of their epidermis being red. In each harvest, the number of fruits was counted and the production mass was measured (g plant^{-1}). Crop management and treatments followed the technical recommendations for the crop, which are detailed in Lopes.[15]

The average daily air temperature (Tavg) was calculated by the arithmetic mean of the minimum and maximum daily air temperatures. The daily thermal sum (DTS, °C day) was calculated using the method: DTS = [(Tmax + Tmin)/2 - Bt] or DTS = (Tavg-Bt), if Tavg<Bt then Tavg=Bt. Where Bt is the lower basal temperature. The DTS was accumulated from planting, resulting in the accumulated thermal sum (ATS), that is: ATS = ΣDTS. The lower basal temperature (Bt) used was 7 °C.[9,17]

The fruit mass (g plant^{-1}) and the number of fruits per plant obtained in each harvest were cumulative consecutively for each experimental plot (H1, H1 + H2, H1 + H2 + H3, ..., H1 + H2 + …Hn), obtaining the cumulative fruit mass and the cumulative number of fruits.

Models

The non-linear regression models fitted to the strawberry productivity data were Logistic (1), Gompertz (2), von Bertalanffy (3), and Brody (4):

$$Y_i(x) = \beta_1 / \left(1 + exp^{(\beta_2 - \beta_3 * x_i)}\right) + \varepsilon_i \qquad (1)$$

$$Y_i(x) = \beta_1 exp^{\left[-\beta_2 exp^{(-\beta_3 x_i)}\right]} + \varepsilon_i \qquad (2)$$

$$Y_i(x) = \beta_1 \left[1 - \beta_2 exp^{(-\beta_3 x_i)}\right]^3 + \varepsilon_i \qquad (3)$$

$$Y_i(x) = \beta_1 \left[1 - \beta_2 exp^{(-\beta_3 x_i)}\right] + \varepsilon_i \qquad (4)$$

in which Y_i is the fruit mass or number of fruits (dependent variable); x_i is the accumulated thermal sum (ATS), in degree-days, elapsed from the moment of seedling transplantation until harvest (independent variable), β_1 and β_3 are biological interpretation parameters, β_2 is a mathematical constant and ε_i is the random error assumed to be independent and identically distributed, following a normal distribution with mean 0 (zero) and variance σ^2.[18]

Parameter estimation and residual analysis

The assumptions of the residuals were verified by the Shapiro-Wilk test[19] for normality of errors; Durbin-Watson statistic[20] for independence of errors, and Breusch-Pagan test[21] for homogeneity of residual variances. Parameter estimates were obtained using the ordinary least squares method with the Gauss-Newton algorithm, using the *nls* function and the quantile regression method, using the *nlrq* function in the R software.[22] For the quantile regression estimates, the quantiles τ = 0.1; τ = 0.25; τ = 0.5; τ = 0.75 and τ = 0.9 for each studied model.

Statistics for assessing goodness of fit

To assess the quality of the model's fit, the Akaike information criterion (AIC) was used, as defined by the equation[23]: $AIC = -2\ln(L) + 2p$, where L is the maximum of the likelihood function and p is the number of parameters fitted. The model with the lowest AIC value was considered the one that best fit the data. However, comparisons are only possible within each estimation method (OLS or NLQR) and not between methods. The coefficients of determination *(R^2)* of each fitted model. Mean squared residuals (MSRs), which represent the mean squared differences between the observed values and the values predicted by a model, smaller values of the mean squared residual indicate a better fit of the model. The average absolute deviation (AAD), calculated as the sum of the deviations between the observed and estimated values, divided by the number of observations, indicates that lower values of the average absolute deviation of the residuals indicate a better fit of the model to the data. And the values of residual deviances, which quantify the difference between the model fit and a model perfectly fitted to the data, a high residual deviance indicates that the model does not fit the data well, while a low residual deviance suggests a good fit.[24-29]

Results and discussions

Assessment of statistical assumptions

The assessment of the assumptions of normality, homoscedasticity, and independence was not fully met by the non-linear regression models used ($p < 0.05$). Similar results were also found in studies with strawberry productivity data by Diel et al.,[6,9] with the risk of biased estimates in the fitness of the growth curve using the ordinary least squares (OLS) method.[30,31] Therefore, the use of the non-linear quantile regression method (NLQR) is an alternative to overcome this problem, as it does not depend on the normality of errors, highlighting the robustness of quantile regression in scenarios with asymmetry or heterogeneity.[32–34]

The fitness of the Logistic model

The fitness of the Logistic model used the NLQR method and the OLS method, in the four treatments studied of experiment I for the variable cumulative number of fruits (Figure 1) and for cumulative mass of strawberry fruits (Figure 2) were shown to be possible since there were significant equation fitness for both OLS and NLQR, in the five quantiles studied (τ = 0.10, τ = 0.25, τ = 0.50, τ = 0.75, and τ = 0.90), with well-defined curves and in the sigmoid shape typical of plant growth curves.[9,28,34–37] This was also the non-linear regression model based on biological criteria that best described the response of pumpkin (*Cucurbita pepo*) and bell pepper (*Capiscum annuum*) fruit production[10] as a function of days after harvest.

In all quantiles (τ = 0.10, τ = 0.25, τ = 0.50, τ = 0.75, and τ = 0.90), good fitness of the equations were obtained, with a coefficient of determination in quantiles $R^2_{(\tau)}$ higher than 0.80 (Table 2), thus, there was a fitness of equations very similar to that found by the least squares method, but with the advantage of overcoming the lack of normality of the data and uniformity of the residual variances, presented in these data.[9] A similar response was found in studies with the Logistic model for dry matter accumulation in garlic plants, with non-linear quantile regression showing greater efficiency for model fitness than ordinary least squares regression[38,39] the Logistic model using quantile regression was effective in predicting heterogeneous growth in *Oreochromis niloticus* tilapia, highlighting its ability to predict growth patterns and reduce uncertainties in intensive aquaculture systems. In a study using these quantile regression quantiles to assess tree diameter and total height, Lanssanova et al.[40] concluded that the fitted logistic model was accurate in estimating the total heights of the stands and the quantile regression allowed a more complete view of the relationship between response height and diameter at breast height, recommending this methodology for the calibration of hypsometric models in *Tectona grandis* stands.

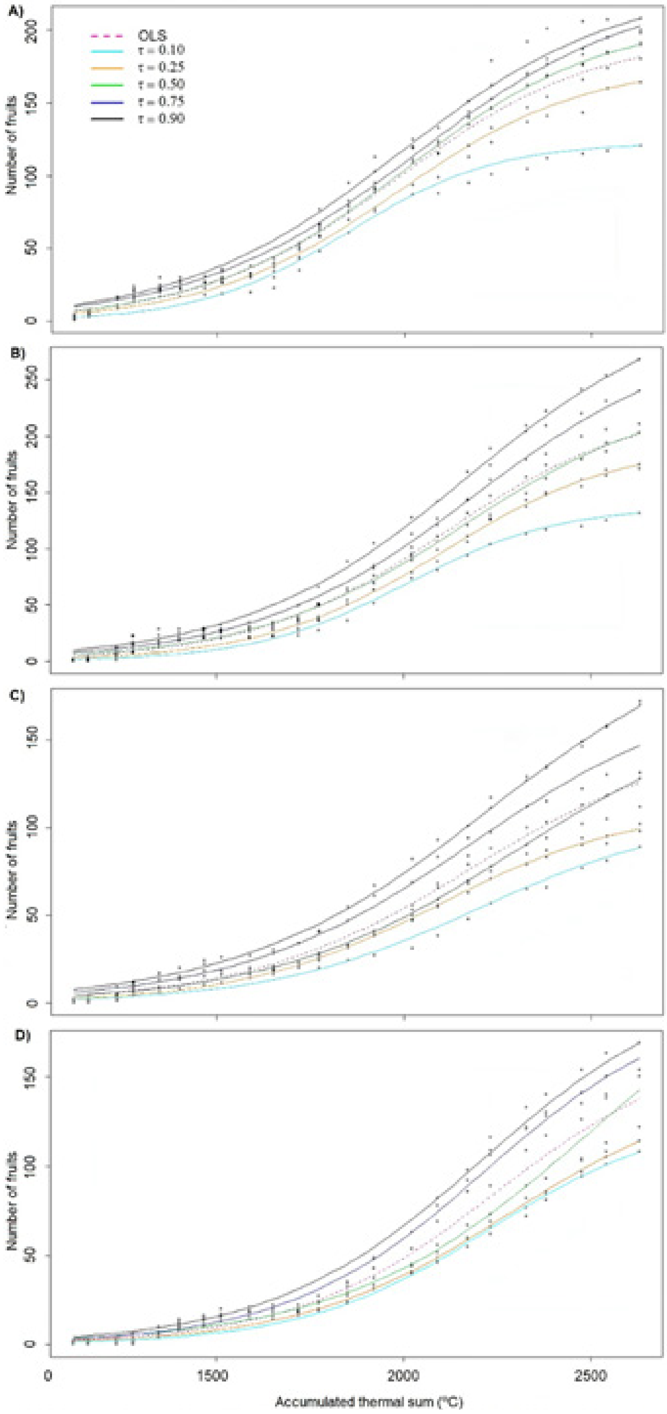

Figure 1 Fitness of the non-linear logistic model estimated via OLS and NLQR for the number of strawberry fruits as a function of the accumulated thermal sum in experiment I. A: LBRF, B: LBR, C: Albion, and D: Estiva.

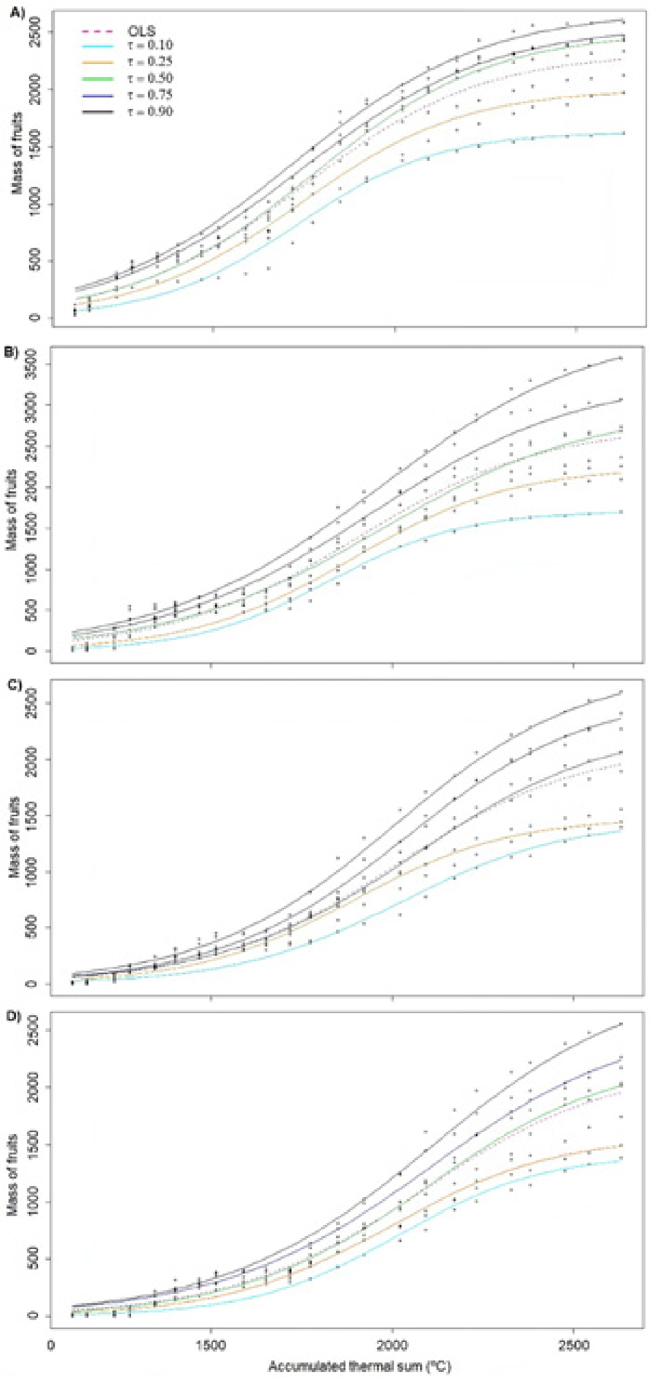

Figure 2 Fitness of the non-linear logistic model estimated via OLS and NLQR for the mass of strawberry fruits (g plant^{-1}) as a function of the accumulated thermal sum in experiment I. A: LBRF, B: LBR, C: Albion, and D: Estiva.

In the equations fitted by the quantile regression model, both for the cumulative number of fruits (Figure 1) and for the cumulative mass of fruits (Figure 2), it was possible to differentiate the most productive plants (τ=0.90)from the least productive ones τ =0.1. Those that obtained the best fitness, with lower values of the Akaike criterion (AIC), were the curves close to the mean (τ =0.25, τ =0.5and τ =0.75). This same methodology was employed by Puiatti et al.[28] to classify 30 garlic accessions based on their agricultural interests, with groupings determined by the quantile of the closest estimates.

In Experiment II, with greater data variability, the fitness of the Logistic model for the cumulative number of fruits variable was also efficient when using the NLQR in the four treatments studied (Figure 3 & 4). The fitness of the quantiles presented a coefficient of determination $R^2_{(\tau)}$ ranging from 0.44 to 0.70, with evidence of a better fit at quantile $\tau = 0.50$, as indicated by the lowest values of AIC and average absolute deviation (AAD); these estimates are lower than those observed for experiment I.

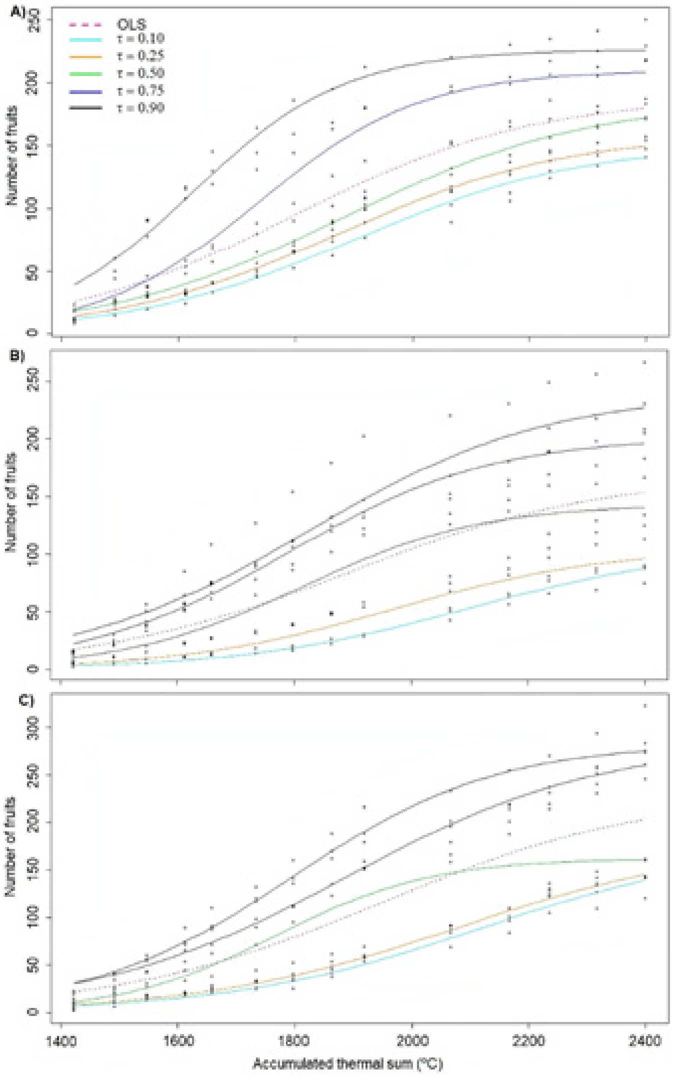

Figure 3 Fitness of the non-linear logistic model estimated via OLS and NLQR for the number of strawberry fruits as a function of the accumulated thermal sum in experiment II. A: LBR, B: Albion, and C: Estiva.

The quantile $\tau = 0.50$ allowed for better fitness than those found by OLS. In studies on larch forest stands (*Larix principis-rupprechtii* Mayr.) developed by Xu et al.[41] in northern China, it was concluded that the NLQR had the highest fitting accuracy, especially for quantile $\tau = 0.5$, providing a scientific basis for sustainable forest management. For site classification and growth models for spruce (*Picea sitchensis*) plantations in Ireland, Lekwadi et al.[24] concluded that the NLQR was efficient for modeling, highlighting its robustness against outliers and variability. Another study using NLQR was applied to modeling the crown profile of *Pinus sylvestris* var. *Mongolica*, where it was suitable to describe the external profile of the crown,[42] this methodology is efficient in estimating the regression model used.

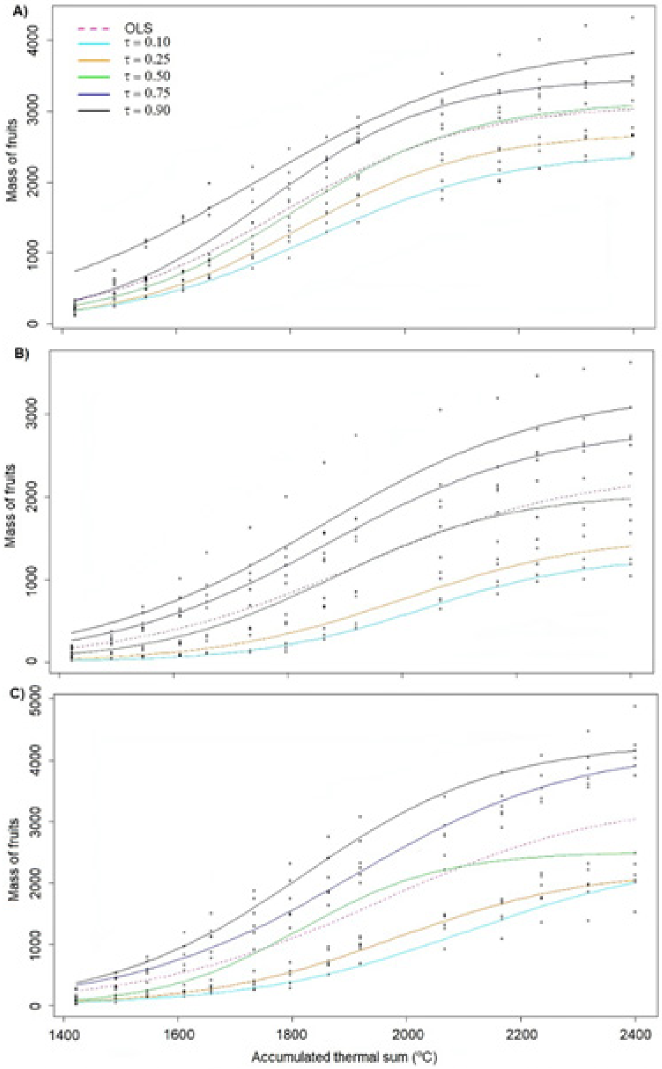

Figure 4 Fitness of the non-linear logistic model estimated via OLS and NLQR for the mass of strawberry fruits (g plant^{-1}) as a function of the accumulated thermal sum in experiment II. A: LBR, B: Albion, and C: Estiva.

At quantiles = 0.10 and $\tau = 0.90$, despite there being a significant fitness of the equations obtained by NLQR, these presented lower $R^2_{(\tau)}$. Similarly to Experiment I, in Experiment II there is increasing variation for productivity β_1 from quantile $\tau = 0.10$ to $\tau = 0.90$ for the variables number and cumulative mass of strawberry fruits (Tables 2 & 3), while β_2 and β_3 were very similar for both variables, with average values of 8.0 and 0.005, respectively. Similar results were found by Silva et al.[34] with corn plant growth data.

Regarding the quality of the fitness of the Logistic model, in Experiment I and II, for the variables cumulative number of fruits and cumulative mass of fruits (Tables 2 and 3), significant and quite high values of $R^2_{(\tau)}$ were observed, close to 0.9, in the quantiles $\tau = 0.25$, $\tau = 0.50$, and $\tau = 0.75$, $R^2_{(\tau)}$ was similar to R^2 obtained by the OLS, with a better fitness for the quantile $\tau = 0.50$ where the AAD and the AIC presented the lowest values. These results corroborate those found by Puiatti et al,[28] who reported better fitness with the reduction in the Akaike criterion when using quantile regression, in logistic models, for quantile $\tau = 0.5$, in garlic accession growth data. For quantiles $\tau = 0.10$ and $\tau = 0.90$, the fitness was slightly worse, with slightly lower $R^2_{(\tau)}$, but with quite high values for the mean squared residuals (MSR), higher AAD, and higher AIC, indicating lower quality in the fitness of the Logistic model for these quantiles.

By fitting the equations of the two experiments, presented in Figures 1-4, it was possible to verify that the Logistic model using NLQR, in all quantiles, is indicated for production data (number and cumulative mass) of strawberry fruits. The use of non-linear logistic regression and von Bertalanffy models was also proposed by Oliveira et al.[43] for the fruit length and width traits of pepper genotypes. In this study, the NLQR was more efficient in fitting models with these variables, compared to the OLS. Silva et al.[34] found that the NLQR, Logistic, Gompertz, and Chanter models exhibited good fitness to corn plant height data, even without adhering to the normal distribution of errors. Additionally, for determining the environmental factors affecting the growth and survival of Greek fir (*Abies cephalonica* Loudon) seedlings, Detsis et al.[44] employed both the quantile regression model and the logistic regression model, achieving successful results.

The fitness of the Gompertz model

The fitness of the Gompertz model using the NLQR method and the OLS, in Experiment I for cumulative number of fruits (Figure 5) and cumulative mass of fruits (Figure 6), presented good fitness, similar to that of the Logistic model.

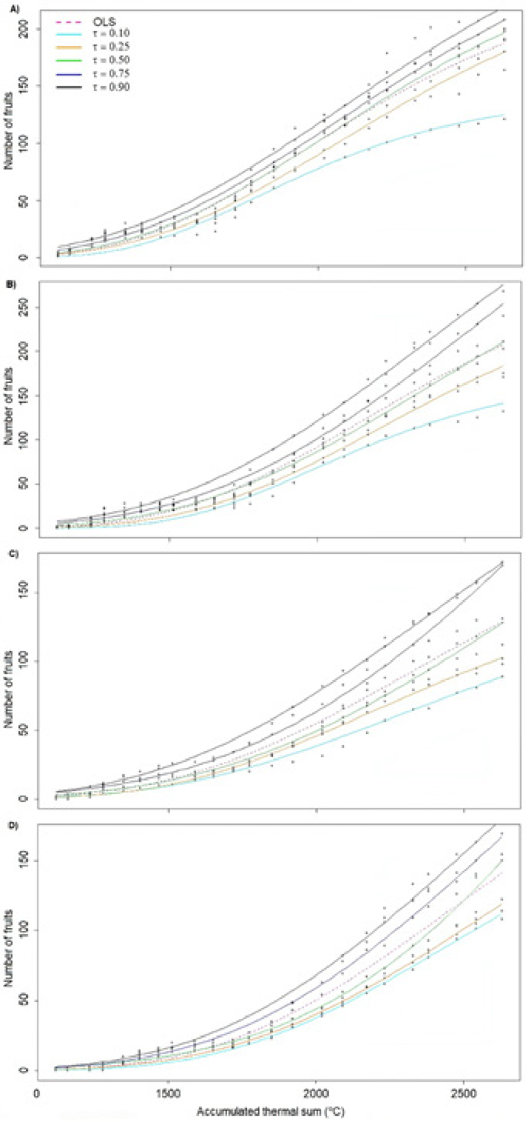

Figure 5 Fit of the non-linear Gompertz model estimated via OLS and NLQR for the number of strawberry fruits as a function of the accumulated thermal sum in experiment I. A: LBRF, B: LBR, C: Albion, and D: Estiva.

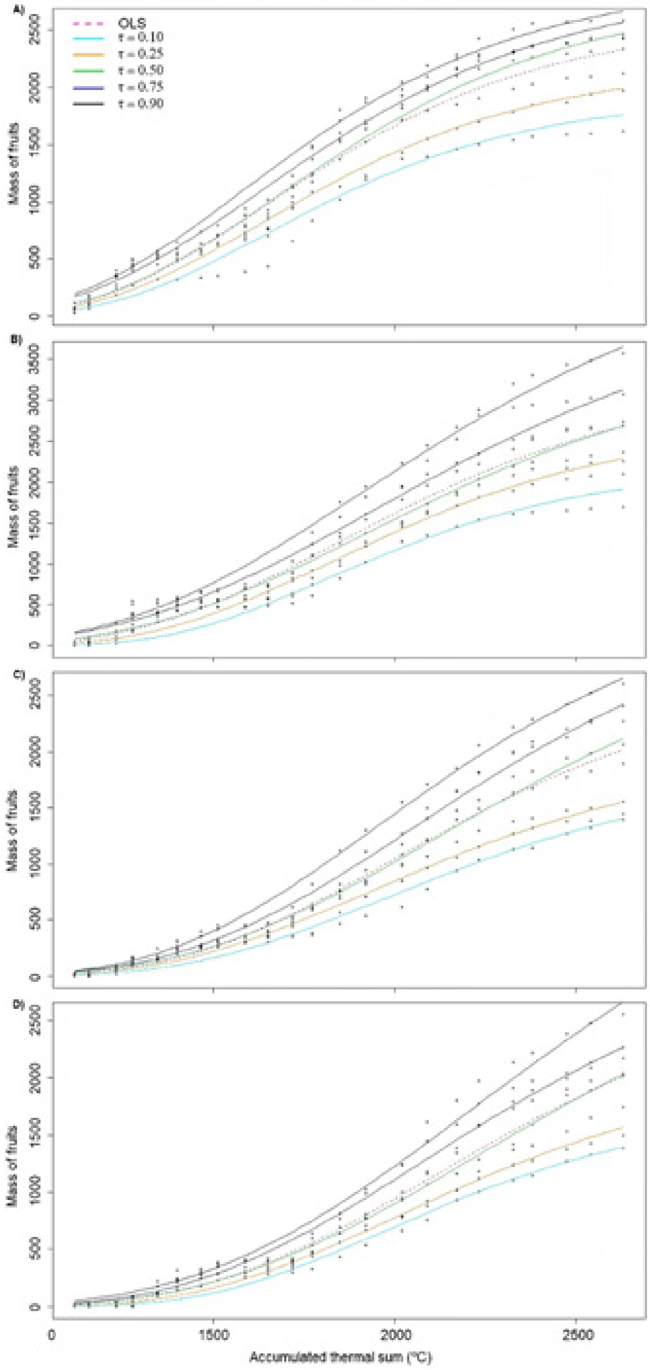

Figure 6 Fit of the non-linear Gompertz model estimated via OLS and NLQR for the mass of strawberry fruits (g plant^{-1}) as a function of the accumulated thermal sum in experiment I. A: LBRF, B: LBR, C: Albion, and D: Estiva.

Using the NLQR and OLS methods, fitted with the Gompertz model on the data from experiment II for cumulative number of fruits (Figure 7) and cumulative mass of fruits (Figure 8), we verified the effect of the fitness of this methodology on data with greater dispersion. Even with the high variability of the data, the fitness of non-linear equations using the Gompertz model was observed for all quantiles and for the OLS in all treatments.

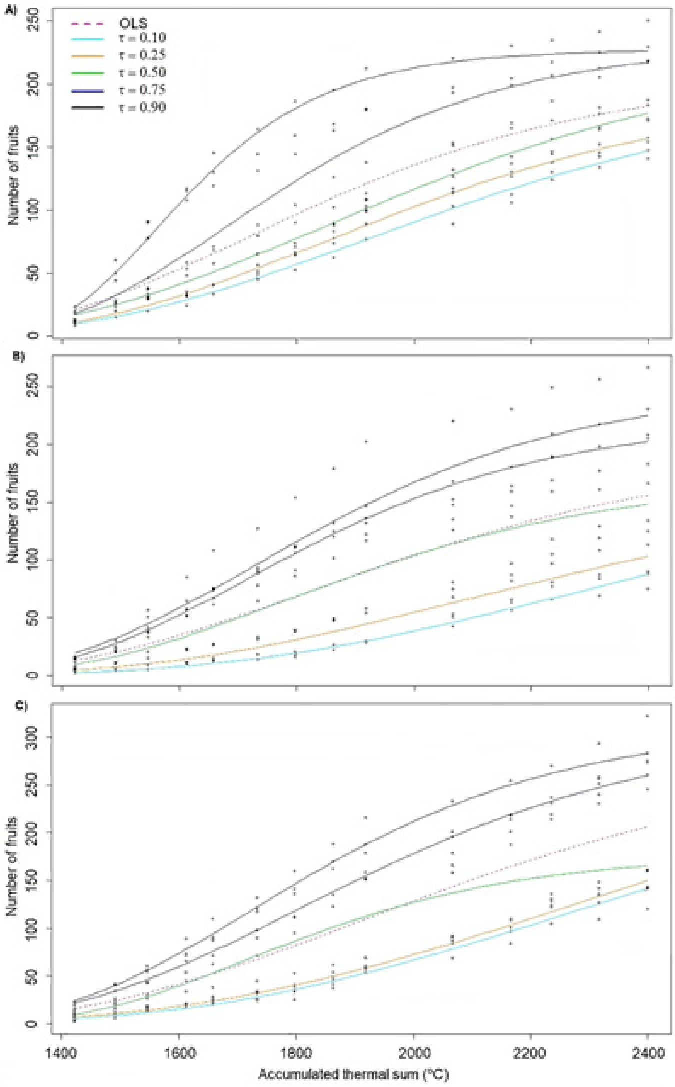

Figure 7 Fit of the non-linear Gompertz model estimated via OLS and NLQR for the number of strawberry fruits as a function of the accumulated thermal sum in experiment II. A: LBR, B: Albion, and C: Estiva.

In the evaluation of the quality of the fitness of the Gompertz model, significant and quite high $R^2_{(\tau)}$ values are observed, close to 0.90 for experiment I (Table 4) and around 0.50 to 0.80 for experiment II (Table 5), with coefficients of determination in the higher quantiles for quantiles $\tau = 0.25$, $\tau = 0.50$, and $\tau = 0.75$ being those that presented the best fitness, where the AAD presented the lowest values, confirmed by the lowest AIC values. Very high values are also observed for MSR, higher AAD, and higher AIC in quantiles $\tau = 0.10$ and $\tau = 0.90$, showing lower quality in the fitness of the Gompertz model for these quantiles.

The Gompertz non-linear regression model has been used in other studies with various plant and animal species and is also effective in fitting their growth. In a study conducted by Fuentes-Andraca et al.,[39] the Logistic model was the best-fitted model; however, the Gompertz model presented consistent results, estimating growth differences and adequately representing the decrease in growth rates as the initial tilapia density increases. The non-linear Gompertz model was also the most versatile for fitting the growth curve data of Alpine goats.[45] Evaluating two populations of sage-grouse (*Centrocercus urophasianus*) from Wyoming, USA, and Gunnison sage-grouse (*Centrocercus minimus*) from Crawford, USA, Cade et al.[46] used quantile regressions for the Gompertz model in both locations, providing more efficient estimates from quantile regression when there are discrepant growth rates.

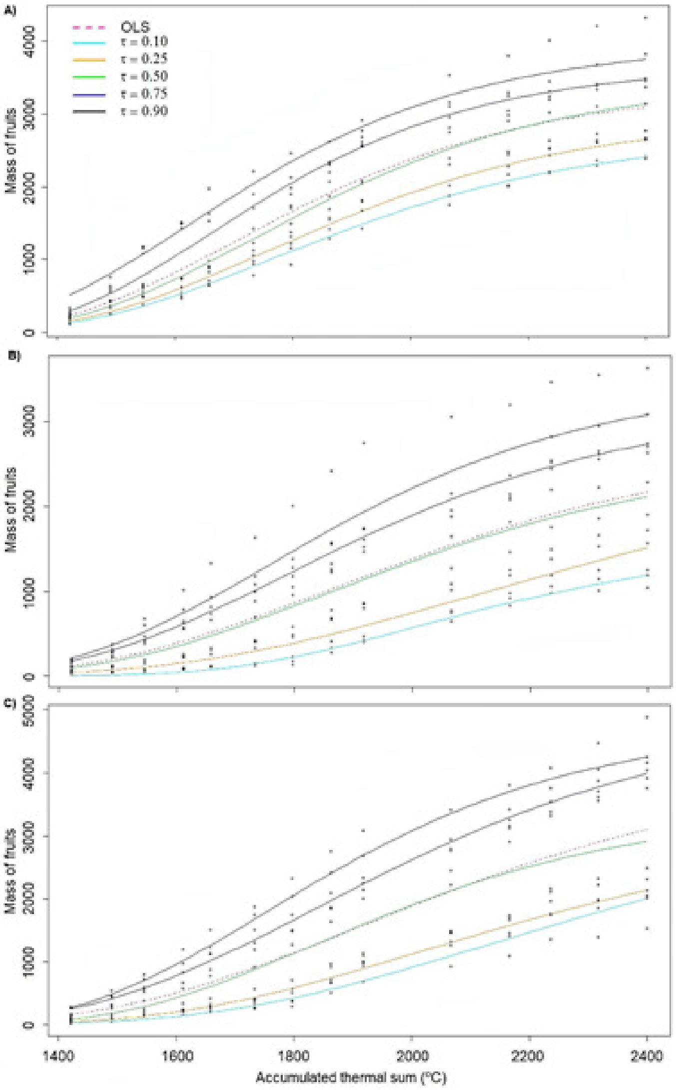

Figure 8 Fit of the non-linear Gompertz model estimated via OLS and NLQR for the mass of strawberry fruits (g plant^{-1}) as a function of the accumulated thermal sum in experiment II. A: LBR, B: Albion, and C: Estiva.

The fitness of the von Bertalanffy model

The assessment of the fit of the von Bertalanffy model for the variables under study (Figures 9&10), related to experiment I, also had fitness of quantile regression equations for all quantiles and treatments studied, as well as for experiment II (Figures 11&12) where there was greater variability, showing good fitness of this model also to the strawberry fruit productivity data.

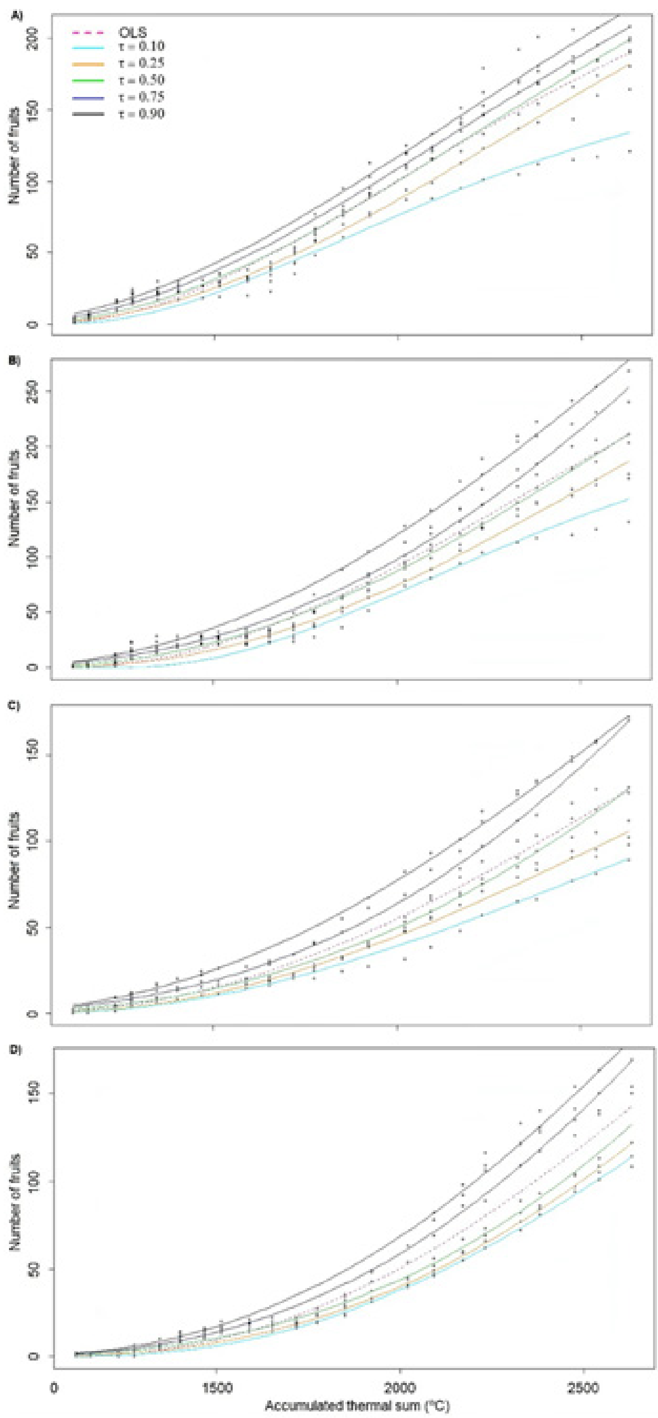

Figure 9 Fitness of the non-linear von Bertalanffy model estimated via OLS and NLQR for the number of strawberry fruits as a function of the accumulated thermal sum in experiment I. A: LBRF, B: LBR, C: Albion, and D: Estiva.

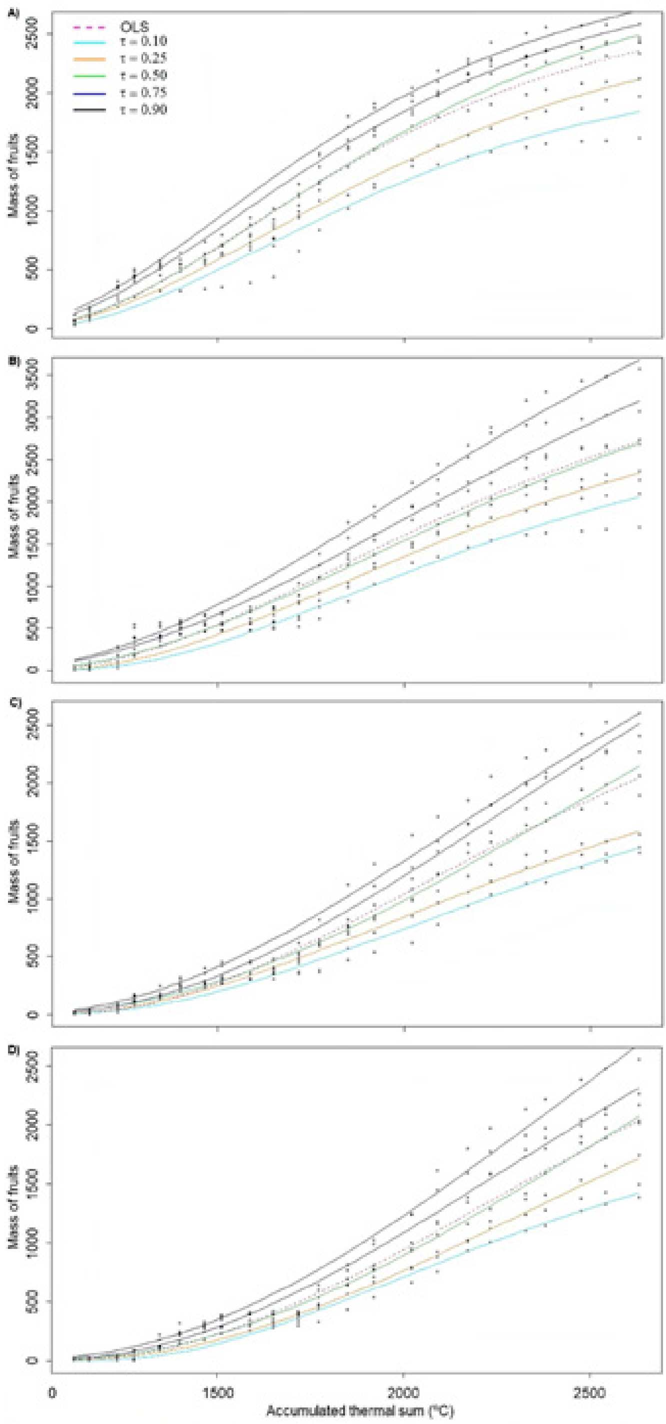

Figure 10 Fitness of the non-linear von Bertalanffy model estimated via OLS and NLQR for the mass of strawberry fruits (g $plant^{-1}$) as a function of the accumulated thermal sum in experiment I. A: LBRF, B: LBR, C: Albion, and D: Estiva.

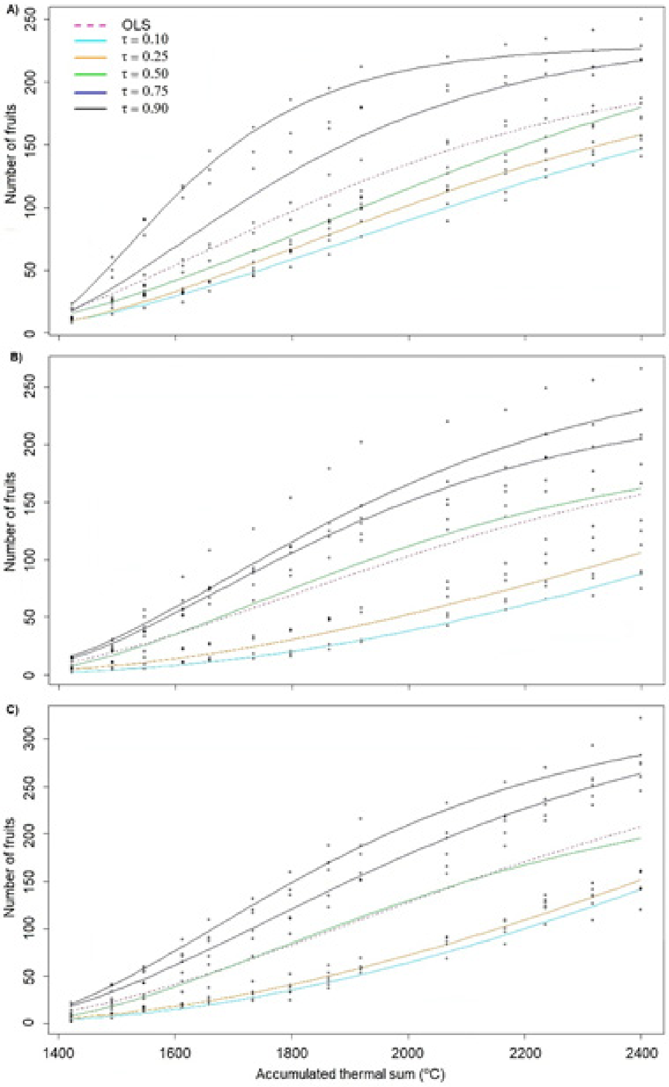

Figure 11 Fitness of the non-linear von Bertalanffy model estimated via OLS and NLQR for the number of strawberry fruits as a function of the accumulated thermal sum in experiment II. A: LBR, B: Albion, and C: Estiva.

The von Bertalanffy non-linear regression model was also proposed by Oliveira A.C.R. et al.[43] for the fruit length and width traits of pepper genotypes, with NLQR being more efficient for fitting models with these variables compared to OLS. Nonlinear logistic and von Bertalanffy growth regression models satisfactorily described the response of the mean fruit weight, the mean number of fruits, and the mean number of harvested bunches of tomato (*Lycopersicon esculentum* var. *cerasiforme*) grown in a plastic greenhouse.[47]

The assessment of the quality of the fitness of the von Bertalanffy model for the variables under study (Tables 6 &7), high values of the coefficient of determination are observed, close to $\tau = 0.90$, when the models are fitted via NLQR in quantiles $\tau = 0.25$, $\tau = 0.50$, and $\tau = 0.75$ with evidence of better fitness for quantile $\tau = 0.50$ where the AAD presented the lowest values, confirmed by the lowest AIC. Similar to the other models studied, there was a reduction in the quality of fit at quantiles $\tau = 0.10$ and $= 0.90$, both for the cumulative number of fruits and cumulative mass of fruits variables, with higher values for MSR, AAD, and AIC, evidencing a worse fit of the von Bertalanffy model for these quantiles, especially for experiment II. There are numerous citations in the literature regarding the application of the von Bertalanffy model to describe growth and productivity variables; however, few utilize quantile regressions. Some deserve to be highlighted: A study with non-linear logistic and von Bertalanffy regression models concluded that NLQR was efficient in fitting models to model the fruit length and width traits of pepper genotypes.[43] Through quantile regression, the von Bertalanffy model allowed the prediction of different growth patterns in intensive tilapia farming, successfully assessing size heterogeneity throughout the production cycle.[39]

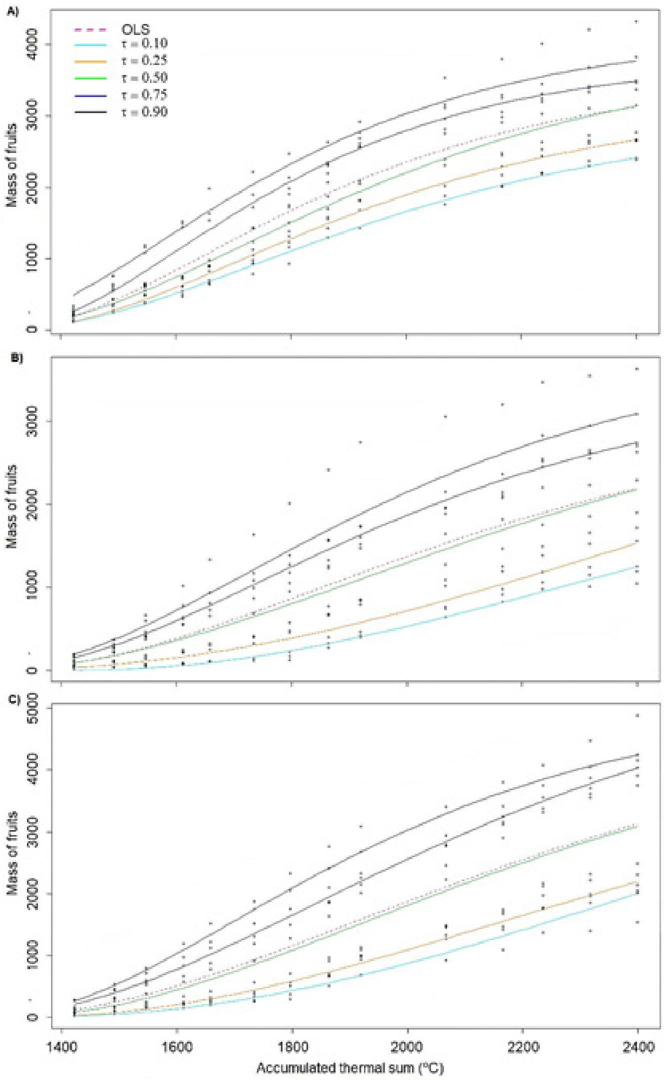

Figure 12 Fitness of the non-linear von Bertalanffy model estimated via OLS and NLQR for the mass of strawberry fruits (g plant^{-1}) as a function of the accumulated thermal sum in experiment II. A: LBR, B: Albion, and C: Estiva.

Other studies used the von Bertalanffy model to describe the growth curve of the biquinho-type pepper, where the fitness of the Logistic, Gompertz, and von Bertalanffy models were indicated to the height data of the biquinho-type pepper under water stress, with the von Bertalanffy model being the most appropriate, considering the AIC and BIC statistics.[48] With the parameters of the Logistic, von Bertalanffy, and Gompertz models, fitted for the number and weight of strawberry fruits, it showed good quality in the fitness for the models, considering the AIC.[49]

The fitness of the Brody model

The fitness of the NLQR and OLS curves for Experiment I with the Brody model in the variables number and cumulative mass of fruits (Figures 13 &14) shows an inversion of the Brody curve, with many negative values of β_1 and β_3 (Table 8). Despite the good fit observed for this model with high coefficients of determination, low AAD, and AIC, this model is not recommended for data similar to these, due to the lack of biological interpretations for negative values. When this model was evaluated in Experiment II (Figure 15 & 16), it also did not present a good fit, with low coefficients of determination for the Albion and Estiva cultivars, and a lack of fitness of the equations when using OLS (Table 9). Furthermore, in both experiments, it was possible to observe many equations tending towards linearity, with non-significant fitted parameters. In a study with non-linear models in cedar (*Cedrela fissilis* Vell.) in a seasonally dry tropical forest, the Brody model proved to be the best for describing diametric growth over time.[50]

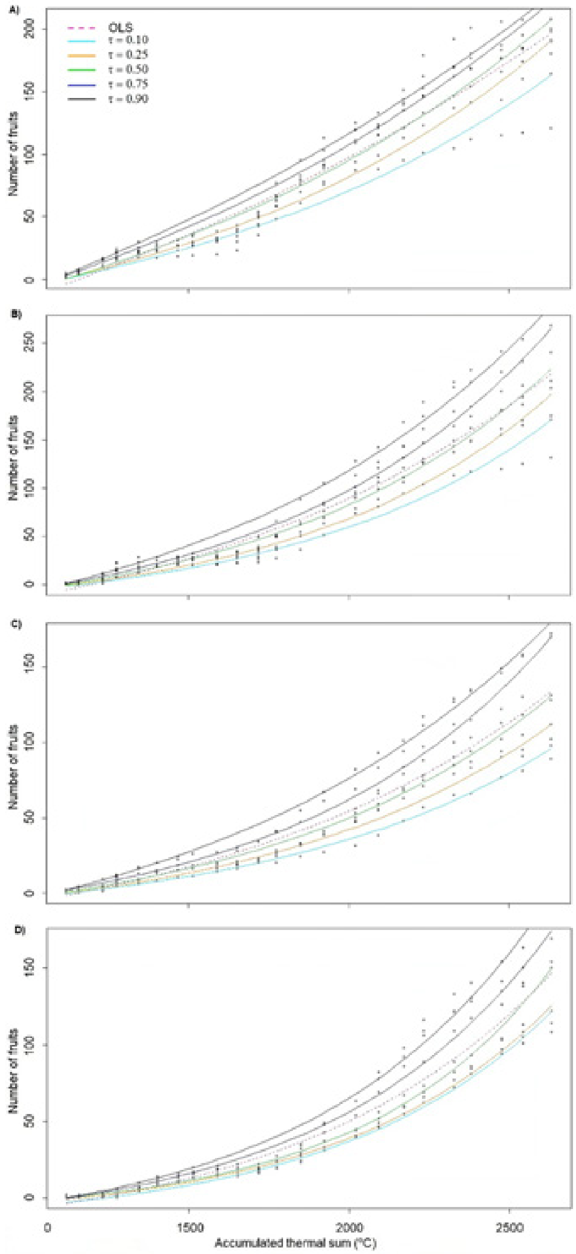

Figure 13 Fitness of the non-linear Brody model estimated via OLS and NLQR for the number of strawberry fruits as a function of the accumulated thermal sum in experiment I. A: LBRF, B: LBR, C: Albion, and D: Estiva.

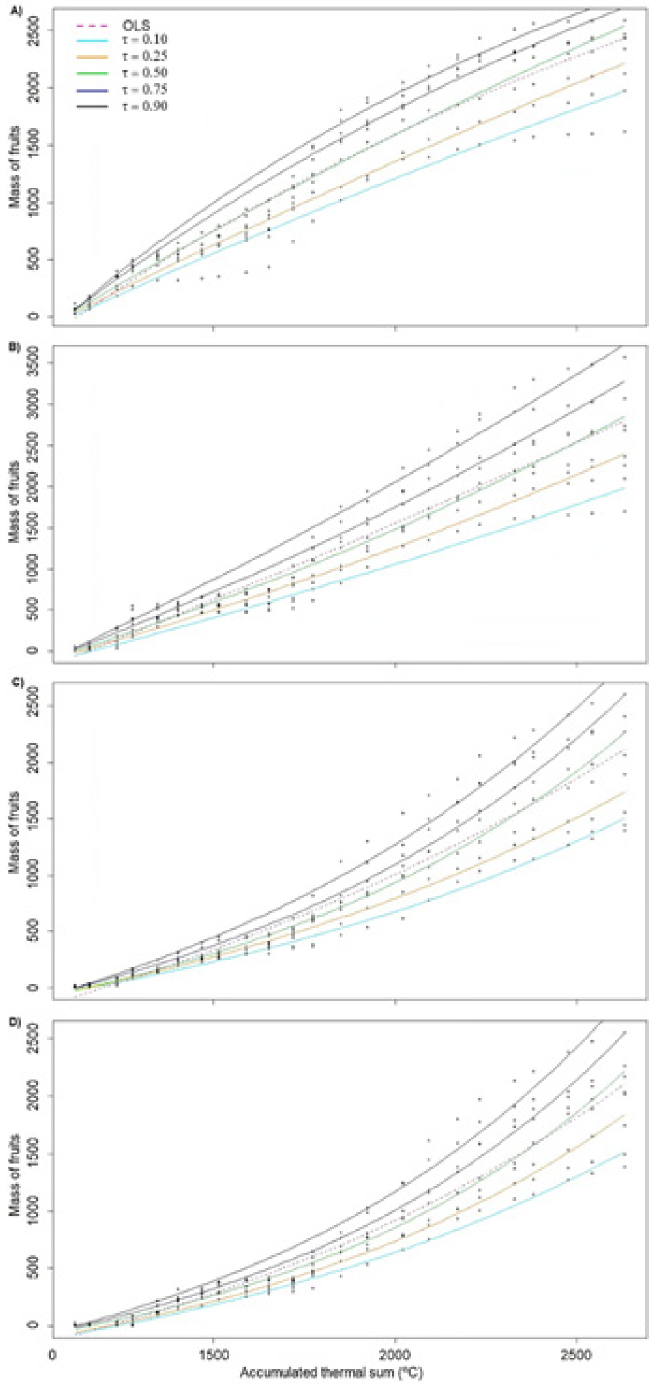

Figure 14 Fitness of the non-linear Brody model estimated via OLS and NLQR for the mass of strawberry fruits (g plant^{-1}) as a function of the accumulated thermal sum in experiment I. A: LBRF, B: LBR, C: Albion, and D: Estiva.

Citation: Santos VJ, Lúcio ADC, Bisognin DA, et al. Description of the strawberry production cycle via non-linear quantile regression models . *Biom Biostat Int J.* 2026;15(1):13‒24. DOI: 10.15406/bbij.2026.15.00444

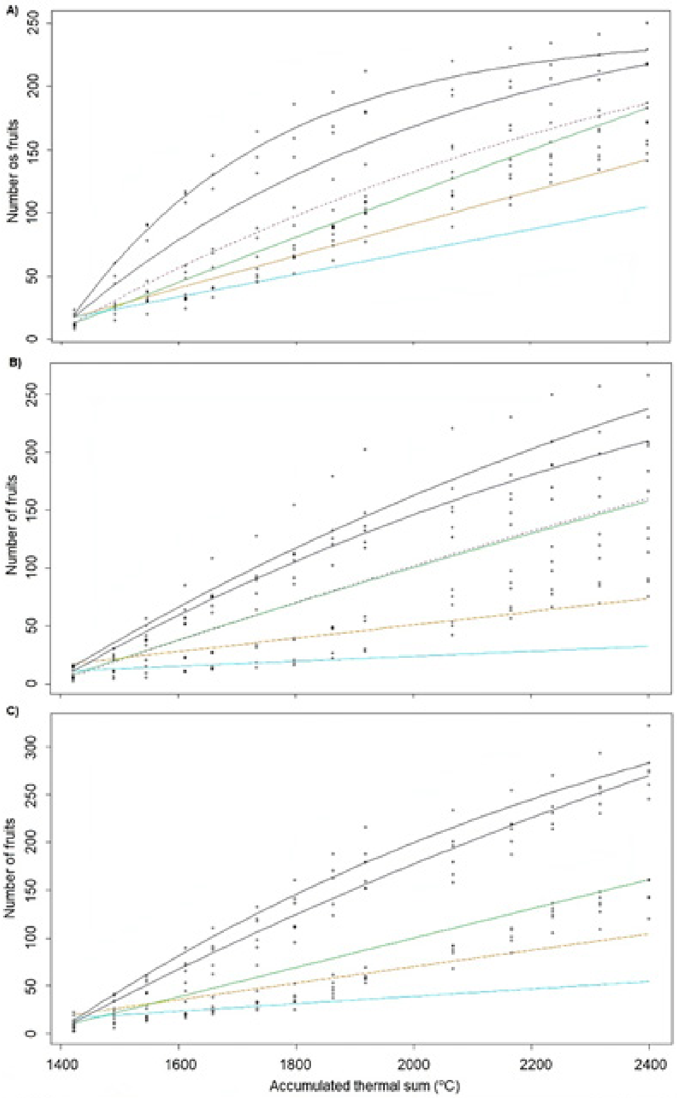

Figure 15 Fitness of the non-linear Brody model estimated via OLS and NLQR for the number of strawberry fruits as a function of the accumulated thermal sum in experiment II. A: LBR, B: Albion, and C: Estiva.

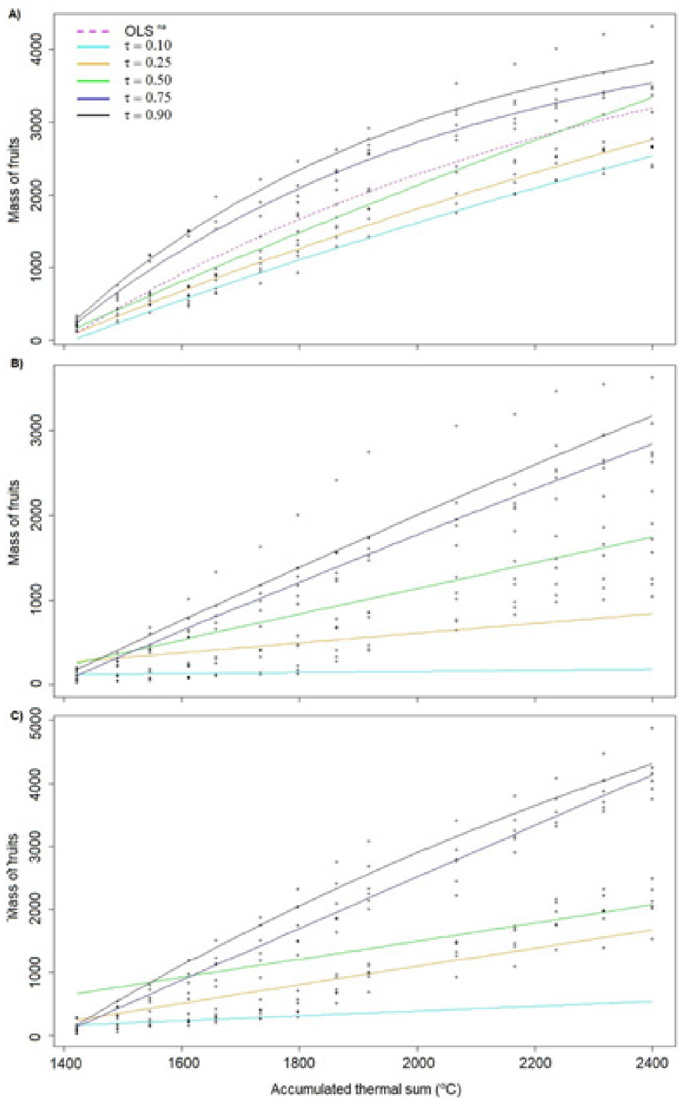

Figure 16 Fitness of the non-linear Brody model estimated via OLS and NLQR for the mass of strawberry fruits (g plant^{-1}) as a function of the accumulated thermal sum in experiment II. A: LBR, B: Albion, C: Estiva, and ^{na} = not adjusted.

Comparison of nonlinear models

The results presented are added to the residual deviance values for Experiment I (Table 10) and Experiment II (Table 11), where the lowest values were found for the Logistic model, followed by the Gompertz model, then the von Bertalanffy model, and finally the Brody model with the highest deviance values, therefore a clear classification of these models in this sequence. With greater magnitude, the comparison between the OLS and the NLQR is presented in Table 10, with deviance values for the variable cumulative number of fruits ranging from 221.81 to 1325.5 for the NLQR, while the deviance values for the OLS were from 28298 to 88370, that is, from 35 to 160 times higher, therefore with lower quality of fit. These proportions are even higher for the cumulative mass of strawberry fruits, where the deviance for the OLS is 70 to 2450 times higher. In a study by Lekwadi et al.,[24] the deviance ranged from 10,639.6 for the NLQR (τ= 0.5), being considerably lower than that of 71,736.3 for the comparable model of mean height-age growth by OLS.

Therefore, despite the Logistic, Gompertz, and von Bertalanffy models obtaining good fitness by the OLS, with the lack of normality of the data and heterogeneity of the residual variances verified, there is the possibility of bias in the interpretations.[28,49] Using the NLQR method, which is more robust and independent of data normality and heterogeneity of residual variances, more parsimonious and higher-quality fitness was achieved for the quantile regression equations.

Therefore, it is recommended to use the NLQR to fit the Logistic, Gompertz, and von Bertalanffy models in data with a sigmoidal distribution when the assumptions are not met.

Conclusion

The non-linear regression models Logistic, Gompertz, and von Bertalanffy are recommended for the fitness of non-linear models based on growth for the variables cumulative number of strawberry fruits and cumulative mass of strawberry fruits, both by the ordinary least squares method (OLS) and by the quantile regression method (NLQR).

For data with a sigmoidal distribution, which do not meet the assumptions of homogeneity of residual variances and normality of errors, the use of non-linear quantile regression is an important alternative for fitting the Logistic, Gompertz, and von Bertalanffy models.

Data availability

Data will be made available on request.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code CAPES/PROEX/88881.844984/2023-01

Conflicts of interest

The Authors declares that there are no conflicts of interest.

References

1. Whitaker VM, Knapp SJ, Hardigan MA, et al. A roadmap for research in octoploid strawberry. *Hortic Res*. 2020;7:e33.
2. Negi YK, Sajwan P, Uniyal S, et al. Enhancement in yield and nutritive qualities of strawberry fruits by the application of organic manures and biofertilizers. *Sci Hortic*. 2021;283:110038.
3. Antunes LEC, Peres NA. Strawberry production in Brazil and South America. *Int J Fruit Sci*. 2013;13(1-2):156–161.
4. FAOSTAT. *Food and agriculture organization corporate statistical database*. Published 2024.
5. Mischan MM, Pinho SZ, Carvalho LR. Determination of a point sufficiently close to the asymptote in nonlinear growth functions. *Sci Agric*. 2011;68(1):109–114.
6. Diel MI, Sari BG, Krysczun DK, et al. Nonlinear regression for description of strawberry *(Fragaria × ananassa)* production. *J Hortic Sci Biotechnol*. 2019;94(2):259–273.
7. Diel MI, Lúcio ADC, Valera OVS, et al. Production of biquinho pepper in different growing seasons characterized by the logistic model and its critical points. *Cienc Rural*. 2020;50:e20190477.
8. Sari BG, Lúcio ADC, Santana CS, et al. Nonlinear growth models: An alternative to ANOVA in tomato trials evaluation. *Eur J Agron*. 2019;104:21–36.
9. Diel MI, Lúcio ADC, Bisognin DA, et al. Nonlinear logistic model for describing strawberry fruit production. *Agronomy*. 2024;14(9):1884.
10. Lúcio ADC, Nunes LF, Rego F. Nonlinear models to describe production of fruit in *Cucurbita pepo* and *Capsicum annuum*. *Sci Hortic*. 2015;193:286–293.
11. Lúcio ADC, Sari BG, Rodrigues M, et al. Nonlinear models for estimating cherry tomato production. *Cienc Rural*. 2016;46(2):233–241.
12. Amaral LS, Moreira GR, Gomes-Silva F, et al. Fitting nonlinear models to the growth of New Zealand White rabbits. *Semin Cienc Agrar*. 2024;45(6):1765–1784.
13. Haupt H, Lösel F, Stemmler M. Quantile regression analysis and other alternatives to ordinary least squares regression. *Methodology*. 2014;10(3):81–91.
14. Ibrahim A, Yahaya A. Analysis of quantile regression as alternative to ordinary least squares. *Int J Adv Stat Probab*. 2015;3(2):138–145.
15. Lopes GA. *Productivity and fruit quality of strawberry clones in a closed soilless cultivation system [master's thesis]*. Santa Maria, Brazil: Federal University of Santa Maria; 2023.
16. Alvares CA, Stape JL, Sentelhas PC, et al. Köppen's climate classification map for Brazil. *Meteorol Z.* 2013;22(6):711–728.
17. Mendonça HFC, Calvete EO, Nienow AA, et al. Estimation of the phyllochron of strawberry in intercropped and monocropped systems in protected environments. *Rev Bras Frutic*. 2012;34(1):15–23.
18. Mazzini ARA, Muniz JA, Aquino LH, et al. Analysis of the growth curve of Hereford males. *Cienc Agrotec*. 2003;27(5):1105–1112.
19. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965;52(3-4):591–611.
20. Durbin J, Watson GS. Testing for serial correlation in least squares regression: I. *Biometrika*. 1950;37(3-4):409–428.
21. Breusch TS, Pagan AR. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*. 1979;47(5):1287–1294.
22. R Core Team. *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing; 2024.
23. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974;19(6):215–222.
24. Lekwadi SO, Nemesova A, Lynch T, et al. Site classification and growth models for Sitka spruce plantations in Ireland. *For Ecol Manag*. 2012;283:56–65.
25. Sarmento JLR, Regazzi AJ, Sousa WH, et al. Analysis of the growth curve of Santa Ines sheep. *Rev Bras Zootec*. 2006;35(2):435–442.
26. Maia E, Siqueira DL, Silva FF, Peternelli LA, et al. Method of comparison of models non-linear regression in banana trees. *Cienc Rural*. 2009;39(5):1380–1386.
27. Reis RM, Cecon PR, Puiatti M, et al. Nonlinear regression models applied to clusters of garlic accessions. *Hortic Bras*. 2014;32(2):178–183.
28. Puiatti GA, Cecon PR, Nascimento M, et al. Quantile regression of nonlinear models to describe different levels of dry matter accumulation in garlic plants. *Cienc Rural.* 2018;48:e20170322.
29. Zuur G, Fryer RJ, Ferro RST, et al. Modelling the size selectivities of a trawl codend and an associated square mesh panel. *ICES J Mar Sci*. 2001;58(3):657–671.
30. Azevedo AT, Soares S, Oliveira ARL, et al. Interior point methods applied to multicommodity flow problems with additional constraints. Trends Comput Appl Math. 2002;3(1):41–50.
31. Miranda LF, Frühauf AC, Lima KP, et al. Nonlinear models to describe the growth of *Jatropha curcas* L. *Rev Cienc Agron.* 2021;52:e20207602.
32. Koenker R, Park BJ. An interior point algorithm for nonlinear quantile regression. *J Econom*. 1996;71(1-2):265–283.
33. Koenker R, Hallock KF. Quantile regression. *J Econ Perspect*. 2001;15(4):143–156.
34. Silva PV. *Use of nonlinear quantile regression in the description of growth data [doctoral thesis]*. São Paulo, Brazil: University of São Paulo; 2022.
35. Muianga CA, Muniz JA, Nascimento MS, et al. Description of the growth curve of cashew fruits using nonlinear models. Rev Bras Frutic. 2016;38(1):22–32.
36. Vanegas DM, Ramírez ME. Correlation of the growth of Pseudomonas fluorescens with the production of medium-chain-length polyhydroxyalkanoates using Gompertz, logistic, and Baranyi primary models. *Inf Tecnol*. 2016;27(2):87–96.
37. Santos ALP, Ferreira TAE, Brito CCR, et al. Proposal for a new non-linear model to describe growth curves. *Biosci J*. 2024;40:e40011.

38. Puiatti GA, Cecon PR, Nascimento M, et al. Nonlinear quantile regression to describe the dry matter accumulation of garlic plants. *Cienc Rural*. 2020;50:e20180385.

39. Fuentes-Andraca VH, Araneda-Padilla ME, Domínguez-May R, et al. Modeling Nile tilapia heterogeneous growth under different stocking densities during pre-grow-out stage. *Aquac Res*. 2023;2023:9347654.

40. Lanssanova LR, Silva FA, Machado SA, et al. Hypsometric relationship in *Tectona grandis* L. f. stands using quantile regression. *Sci For.* 2021;49:e3559.

41. Xu A, Wang D, Liu Q, et al. Incorporating stand density effects and regression techniques for stem taper modeling of a *Larix principis-rupprechtii* plantation. *Front Plant Sci*. 2022;13:902325.

42. Sun Y, Gao H, Li F. Using linear mixed-effects models with quantile regression to simulate the crown profile of planted *Pinus sylvestris* var. *mongolica* trees. *Forests*. 2017;8(11):446.

43. Oliveira ACR, Cecon PR, Puiatti GA, et al. Nonlinear models based on quantiles in the fitting of growth curves of pepper genotypes. *Rev Bras Biom*. 2021;39(3):447–459.

44. Detsis V, Efthimiou G, Theodoropoulou O, et al. Determination of the environmental factors that affect the growth and survival of Greek fir seedlings. *Land.* 2020;9(4):100.

45. Oliveira JA, Pires LC, Silva LP, et al. Growth curves in Alpine goats raised in the northeastern semiarid region. In: Delke CA, Moraes G, Galati RL, eds. Animal Science: *Contemporary Research and Practices*. Científica Digital; 2021:64–82.

46. Cade BS, Edmunds DR, Ouren DS. Quantile regression estimates of animal population trends. *J Wildl Manage*. 2022;86(3):e22228.

47. Lúcio ADC, Nunes LF, Rego F. Nonlinear regression and plot size to estimate green beans production. *Hortic Bras*. 2016;34(4):507–513.

48. Mangueira RAF, Silva VF, Martins WA. Description of the growth curve of the biquinho-type pepper plant under water stress. *Cienc Nat.* 2022;44:e17.

49. Diel MI, Pinheiro MVM, Thiesen LA, et al. Cultivation of strawberry in substrate: Productivity and fruit quality are affected by the cultivar origin and substrates. *Cienc Agrotec*. 2018;42(2):229–239.

50. Frühauf AC, Pereira GA, Barbosa ACMC, et al. Nonlinear models in the study of the cedar diametric growth in a seasonally dry tropical forest. *Rev Bras Cienc Agrar*. 2020;15:e8558.