Semina Ciências Exatas e Tecnológicas

e-ISSN 1679-0375
BIOMATHEMATICS

ORIGINAL ARTICLE DOI 10.5433/1679-0375.2025.v46.53256

iThenticate

Quantile Nonlinear Regression for Modeling Maize Growth Data

Regressão Não Linear Quantílica para Modelagem de Dados de Crescimento de Milho

Pollyane Vieira da Silva[1], Taciana Vilella Savian[2]

Received: August 10, 2025 *Received in revised form: October 15, 2025* *Accepted: October 27, 2025* *Available online: December 2, 2025*

ABSTRACT

Statistical modeling is fundamental in plant growth studies as it guides management practices across different developmental stages. Among existing approaches, nonlinear regression models stand out for describing growth patterns over time through estimated parameters, typically obtained via least squares methods. However, this approach is limited to average data analysis and remains sensitive to outliers and variance heterogeneity. Quantile regression emerges as a robust alternative, enabling estimation across different quantiles without requiring normality assumptions for errors. This study aimed to analyze corn plant growth over time using quantile nonlinear regression and to investigate critical points in the applied models. We evaluated logistic, Gompertz, and Chanter models, assessing their goodness-of-fit through the Akaike Information Criterion (AIC), intrinsic nonlinearity measures, and two novel statistics proposed here: quantile correlation and the weighted sum of squared deviations (WSSD). Results demonstrated that the Chanter model provided the best fit for the analyzed dataset, with strong performance across different quantiles.

keywords nonlinear models, quantile regression, intrinsic nonlinearity measure, corn culture, growth curve analysis

RESUMO

O embasamento estatístico em estudos de crescimento vegetal é fundamental, pois orienta o manejo em diferentes fases do desenvolvimento das plantas. Dentre as abordagens utilizadas, destacam-se os modelos de regressão não linear, que descrevem o crescimento ao longo do tempo com base em parâmetros estimados, usualmente, pelo método dos mínimos quadrados. No entanto, essa abordagem se limita à análise média dos dados, sendo sensível a valores extremos e à heterogeneidade de variâncias. A regressão quantílica surge como alternativa robusta, permitindo estimativas em diferentes quantis sem pressupor normalidade dos erros. Este estudo teve como objetivo analisar o crescimento de plantas de milho ao longo do tempo por meio da regressão não linear quantílica, além de investigar os pontos críticos dos modelos utilizados. Foram aplicados os modelos logístico, Gompertz e Chanter. Para avaliar a qualidade dos ajustes, utilizaram-se o critério de informação de Akaike (AIC), a medida de não linearidade intrínseca, bem como duas estatísticas propostas neste trabalho: a correlação dos quantis e a soma ponderada dos quadrados dos desvios. Os resultados indicaram que o modelo de Chanter proporcionou melhor ajuste para o conjunto de dados analisado, com bom desempenho em diferentes quantis.

palavras-chave modelos não lineares, modelagem quantílica, medida de não linearidade, cultura do milho, curvas de crescimento

[1]Prof. Dr., Department of Mathematics and Statistics, UFPel, Pelotas, RS, Brazil. pollyane.silva@ufpel.edu.br
[2]Prof. Dr., Department of Exact Sciences, USP, Piracicaba, SP, Brazil. tvsavian@usp.br

Introduction

Corn is one of the world's most produced cereals due to its extensive use in human and animal nutrition. Understanding plant growth is fundamental for proper crop management. Studies evaluating plant growth trajectories are essential for developing improved cultivation, harvesting, and conservation techniques, as well as for detecting developmental issues (Souza & Macêdo, 2009). Brazil ranks as the third-largest corn producer globally, with an estimated yield of 5.40 tons per hectare and a total production of 94.50 million tons across 17.50 million hectares (United States Department of Agriculture [USDA], 2025).

Regression analysis has been widely employed to predict the behavior of a primary variable based on multiple explanatory variables in corn research. For instance, Cargnelutti and Toebe (2020) used multiple regression to model grain yield as a function of ear length and diameter. Nonlinear regression models are particularly suited for describing growth curves, as their formulations incorporate theoretical considerations specific to the biological processes being modeled (Mazucheli & Achcar, 2002). Such models have been successfully applied in growth studies for cassava (Gray, 2000), ryegrass (Migault et al., 2012), and other crops.

In plant growth studies, researchers often seek to distinguish between subpopulations with varying productivity levels after fitting a model. Quantile regression offers a robust alternative for characterizing predictor-response relationships across different quantiles of the distribution, eliminating the need for subsampling. Unlike traditional methods relying on conditional means, this approach models conditional quantiles, revealing relationships between time and specific percentiles of the dependent variable. This provides richer information about distributional characteristics, including tail behavior, while offering greater robustness against heteroskedasticity (Hao & Naiman, 2007; Koenker, 2005). Moreover, it is possible to compute confidence intervals and perform hypothesis testing within this framework, most commonly through bootstrap methods that avoid complex asymptotic estimations; the key advantage lies in the ability to compare quantiles and uncover heterogeneous effects, although limitations remain regarding interpretation, computational cost, and the need for more robust inference procedures.

The application of nonlinear quantile regression to plant growth analysis has demonstrated significant potential. Sorrel et al. (2012) effectively employed this method to differentiate phenotypic responses to water depth from non-phenotypic variations in wetland plants. Here, nonlinear quantile regression is applied to analyze correlated experimental data. Previous work in this domain includes Amador (2010), who modeled castor bean plant height using nonlinear mixed-effects models.

While growth curves typically lack local extrema, certain physiologically critical points hold particular significance. These are identified through derivatives of the growth equations with respect to time (Mischan & Pinho, 2014). Silva et al. (2021) demonstrated that the coordinates of these critical points provide actionable insights, enabling farmers to optimize harvest timing for green coconuts.

This article aims to visualize critical points in growth models through derivative analysis while implementing nonlinear and nonlinear quantile regression techniques. Additionally, two novel statistics are proposed to assess model fit: quantile correlation and the weighted sum of squared deviations (WSSD). These metrics are designed to evaluate the performance of nonlinear quantile regression models when applied to empirical corn plant height data, with the further objective of comparing the alternatives and indicating the most suitable nonlinear model.

Materials and methods

The data were adapted from Crosariol Netto (2013) and refer to the plant height of the transgenic hybrid corn 30F35 Y (Yieldgard). The experiment was conducted in Votuporanga, São Paulo State, at the experimental field of the Northwest Regional Pole of APTA (São Paulo State Agribusiness Technology Agency) during the 2011/2012 growing season. Each experimental plot consisted of four 4-meter-long rows spaced 0.90 m apart, with a planting density of seven plants per linear meter. Ten plants were selected from each center row, tagged, and monitored throughout the crop cycle.

Plant height measurements (in centimeters) were taken for 58 plants on each evaluation day using a ruler, measuring from the plant base (soil level) to the apex of the last fully expanded whorl leaf. Initial height measurements were taken 15 days after sowing (DAS), with subsequent measurements at 30, 40, 50, and 122 DAS.

Three nonlinear models were employed in this study: Logistic, Gompertz, and Chanter. Following Fernandes and Pereira (2017), we adopted the following parameterizations for these models:

Logistic model:

$$y_i(x) = \frac{\beta_1}{1 + \exp\left(\frac{\beta_2 - x_i}{\beta_3}\right)} + \varepsilon_i. \tag{1}$$

Gompertz model:

$$y_i(x) = \beta_1 \exp\left[-\exp\left(\beta_2(\beta_3 - x_i)\right)\right] + \varepsilon_i. \tag{2}$$

In equations (1) and (2), the index i represents longitudinal observations $i = 1, 2, \ldots, 58$, where y_i corresponds to the i-th measurement of corn plant height and x_i denotes the plant age in days counted from 15 days after sowing. The parameter β_1 estimates the asymptotic plant height at maturity. In the logistic model, β_2 corresponds to the abscissa of the inflection point of the curve, whereas in the Gompertz model it represents a growth rate. The parameter β_3, in the logistic model, indicates the growth rate, while in the Gompertz model it expresses a value proportional to the slope of the tangent line to the curve at the inflection point. The random error term ϵ_i follows an independent and identically distributed normal distribution with mean zero and constant variance σ^2 ($\epsilon_i \sim N(0, \sigma^2)$).

The Chanter model represents a hybrid formulation that combines features of both the logistic and Gompertz models, with parameters that admit analogous biological interpretations, being:

Chanter model:

$$y_i(x) = \frac{\beta_1\beta_2}{\beta_1 + (\beta_2 - \beta_1)\exp\left[-\left(\frac{\beta_3}{\beta_4}\right)(1 - \exp(-\beta_4 x_i))\right]} + \varepsilon_i, \tag{3}$$

where i indexes the longitudinal observations $i = 1, 2, \ldots, 58$, with y_i denoting the i-th measurement of corn plant height and x_i representing the observation time in days (measured from 15 days post-sowing). The parameters β_1 through β_4 govern the model's dynamic behavior, while ϵ_i represents the random error term, assumed to be independently and identically distributed (iid) following a normal distribution with mean zero and constant variance σ^2 ($\epsilon_i \sim N(0, \sigma^2)$).

The Chanter model equation is derived as the solution to the following differential equation:

$$\frac{dy(x)}{dx} = \beta_3\, y(x) \left(1 + \frac{y(x)}{\beta_2}\right) e^{-\beta_4 x} \tag{4}$$

The detailed derivation of this equation is presented in Silva and Savian (2019). The differential equation fundamentally states that the relative growth rate $\frac{dy(x)}{dx}$ is proportional to:

i) $y(x)$ itself,

ii) the remaining quantity for the function to reach parameter β_2 (characteristic shared with the logistic model),

iii) a negative exponential function of x (characteristic shared with the Gompertz model).

This dual proportionality makes the Chanter model a true hybrid between the logistic and Gompertz formulations.

Parameter β_1 represents the curve's intercept (the $y(x)$ value when $x = 0$), while β_2 defines the upper horizontal asymptote when β_4 is negative. The remaining parameters lack immediate practical interpretation.

Parameter estimates for the nonlinear models were obtained through least squares estimation, with the Gauss–Newton algorithm solving the system of nonlinear normal equations. For quantile nonlinear regression, parameter estimates were derived via weighted absolute error minimization using an interior-point algorithm (Silva, 2022).

Quantile regression provides more comprehensive information than classical regression by enabling analysis of changes in the τ-th quantile of the response variable for any $\tau \in (0,1)$, rather than solely detecting mean shifts (Koenker & Basset, 1978). As Younesi et al. (2019) demonstrate, the mean does not always represent an adequate fit, particularly when the conditional distributions of the response variable deviate from normality.

In quantile regression, the following model can be specified for any $\tau \in (0,1)$:

$$y_i = f(x_i, \beta(\tau)) + \epsilon_i, \quad i = 1, 2, \ldots, 58 \tag{5}$$

where x_i is the covariate vector (in this study: days after sowing), $\beta(\tau)$ represents unknown parameters at the τ-th quantile, and ϵ_i is the random error term requiring no distributional assumptions.

We analyzed quantiles at $\tau = \{0.1, 0.25, 0.5, 0.75, 0.9\}$. Parameter estimation for $\beta(\tau)$ was achieved through minimization of the weighted sum of absolute residuals, given by:

$$\hat{\beta}(\tau) = \arg\min \sum_{i=1}^{n} \rho_\tau\big(y_i - f(x_i, \beta(\tau))\big) = \arg\min \sum_{i=1}^{n} \rho_\tau(\epsilon_i) \tag{6}$$

where ρ_τ is called the check function or loss function, defined as:

$$\rho_\tau(\epsilon) = \begin{cases} \epsilon\tau, & \text{if } \epsilon \geq 0, \\ \epsilon(\tau - 1), & \text{if } \epsilon < 0. \end{cases} \tag{7}$$

Quantile regression employs an asymmetric loss function that differentially weights positive and negative residuals using τ and $\tau - 1$ coefficients respectively. Due to the computational complexity of direct minimization, the optimization problem is typically reformulated as a linear programming task solved through interior-point algorithms. For model implementation, we utilized both the nonlinear least squares (`nls`) and nonlinear quantile regression (`nlrq`) functions available in the `R` statistical software (R Core Team, 2021).

Model evaluation incorporated multiple criteria. The Akaike Information Criterion (AIC) served as our primary model selection tool, calculated as

$$AIC = -2\ln(L) + 2p, \tag{8}$$

where L represents the maximized likelihood value (equivalent to minimized residual sum of squares under normality assumptions) and p denotes the number of estimated parameters. Following the foundational work of Akaike (2003), models with lower AIC values are preferred.

We complemented this with Bates and Watts' intrinsic curvature measure to assess nonlinearity characteristics, where models exhibiting more linear-like behavior are generally favored (Ratkowsky, 1983; Zeviani et al., 2012).

To enhance quantile regression assessment, we propose two novel goodness-of-fit measures: *quantile correlation* and *weighted sum of squared deviations (WSSD)*. These metrics, which will be formally derived and applied in subsequent analysis, provide robust tools for evaluating nonlinear quantile models across different conditional distributions.

Results and discussion

As demonstrated by Mischan and Pinho (2014), the logistic function allows for practical interpretation of its parameters. By evaluating the limits of the logistic function as the independent variable x approaches positive and negative infinity, we obtain

$$\lim_{x \to \infty} y(x) = \beta_1 \quad \text{and} \quad \lim_{x \to -\infty} y(x) = 0, \tag{9}$$

respectively. This demonstrates that the logistic function has two horizontal asymptotes: a right asymptote at $y = \beta_1$ and a left asymptote at $y = 0$.

The inflection point of the logistic curve can be determined by setting its second derivative to zero ($y'' = 0$), yielding $x = \beta_3$. Thus, the coordinates of the inflection point are given by

$$PI = (\beta_3, \frac{\beta_1}{2}), \tag{10}$$

corresponding to 50% of the model's asymptotic value (Silva,2022). As discussed by Mischan and Pinho (2014), the second derivative represents growth acceleration, which remains positive until the organism reaches its inflection point (maximum growth rate) and becomes negative thereafter. During its positive phase, the acceleration function increases to a maximum point (PAM) before decreasing, while in its negative phase, it reaches a minimum point (PDM).

By setting the third derivative to zero ($y''' = 0$), we obtain the x-coordinates of the maximum acceleration and deceleration points. For the logistic model, the point of maximum acceleration (PAM) is located at

$$x = \frac{\beta_2\beta_3 - \ln(2 + \sqrt{3})}{\beta_2}, \tag{11}$$

with a y-coordinate of

$$y = \frac{\beta_1}{3 + \sqrt{3}}, \tag{12}$$

corresponding to approximately 21% of the asymptotic value. Conversely, the point of maximum deceleration (PDM) occurs at

$$x = \frac{\beta_2\beta_3 - \ln(2 - \sqrt{3})}{\beta_2}, \tag{13}$$

with a y-coordinate of

$$y = \frac{\beta_1}{3 - \sqrt{3}}, \tag{14}$$

corresponding to approximately 79% of the asymptotic value. Graphical representation of growth functions and their derivatives are presented in Figure 1, where Critical points derived from second- and third-order derivatives are annotated, showing inflection points (IP), points of maximum acceleration (PMA), and points of maximum deceleration (PMD). Red horizontal dashed lines indicate asymptotic maturity thresholds.

These critical points of the logistic model are shown in Figure 1(a).

The Gompertz model, originally proposed by Benjamin Gompertz in 1825 for actuarial tables describing population survival and mortality probabilities, has since been widely applied in animal and plant growth studies. Similar to the logistic model, the Gompertz curve exhibits an initial exponential growth phase followed by an asymptotic plateau (Mischan & Pinho, 2014).

Analogous to the logistic function, the critical points of the Gompertz model can be derived from its higher-order derivatives (Figure 1b). Setting $y'' = 0$ yields the inflection point at

$$PI = (\beta_3, \frac{\beta_1}{e}). \tag{15}$$

The third derivative ($y''' = 0$) provides the coordinates for the maximum acceleration (PAM) and deceleration (PDM) points. Specifically, PAM is located at

$$x = \frac{\beta_2\beta_3 - \ln\left(\frac{-3+\sqrt{5}}{2}\right)}{\beta_2}, \tag{16}$$

whith

$$y = exp\left(\frac{-3-\sqrt{5}}{2}\ \beta_1\right), \tag{17}$$

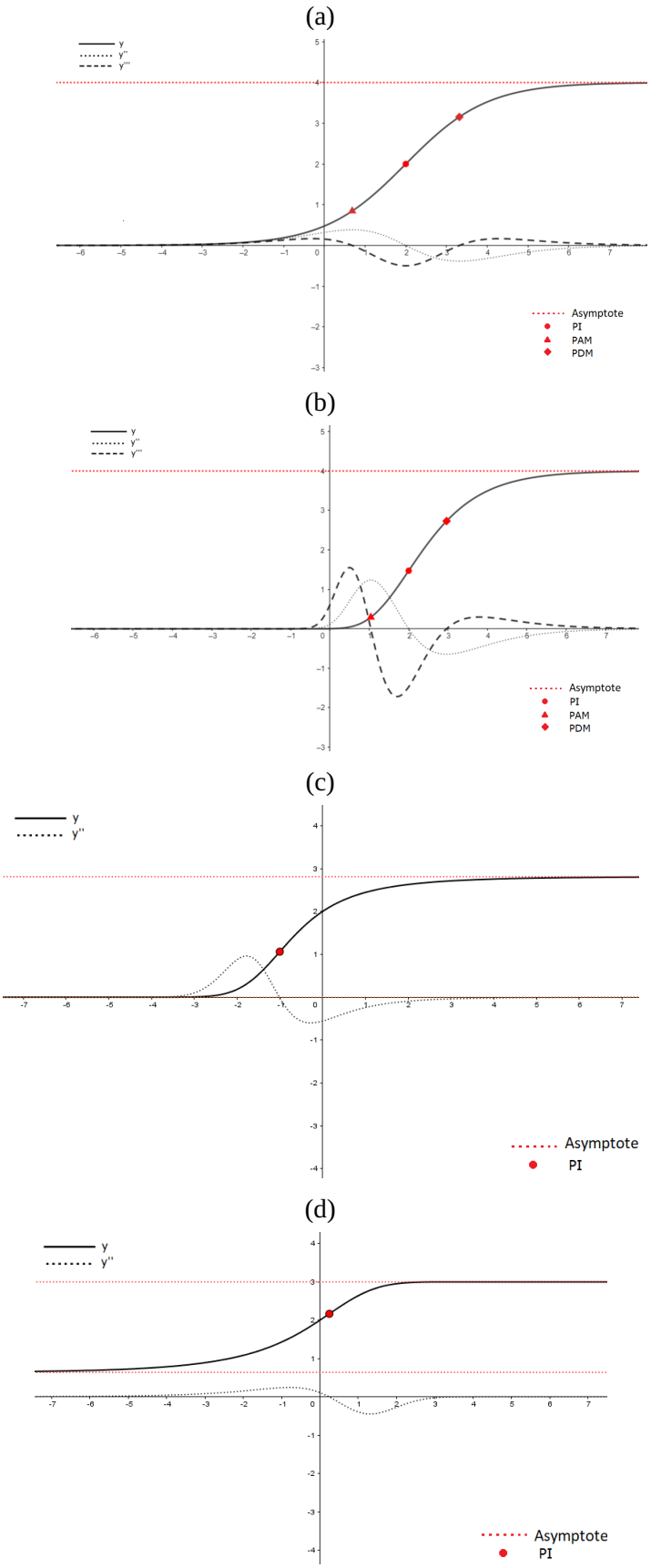
while PDM occurs at

$$x = \frac{\beta_2\beta_3 - \ln\left(\frac{3-\sqrt{5}}{2}\right)}{\beta_2}, \tag{18}$$

whith

$$y = exp\left(\frac{-3+\sqrt{5}}{2}\ \beta_1\right). \tag{19}$$

Figure 1 - Graphical representation of growth functions and their derivatives: (a) Logistic model, (b) Gompertz model, and (c)-(d) Chanter model.



The logistic and Gompertz models share similarities but differ in their parametric interpretations. In both models, β_1 represents the upper right asymptote. However, the logistic model is symmetric about its inflection point (occurring at 50% of asymptotic growth), whereas the Gompertz model is asymmetric, reaching its inflection point at approximately 37% of the asymptotic value β_1/e.

This distinction has practical implications for agricultural management, as the timing of growth deceleration varies between crop varieties. Certain crops may require later harvesting periods than others, making the Gompertz model particularly useful for determining optimal harvest times (Silva, 2018).

The Chanter model, originally developed by Dennis Osborne Chanter in 1976, represents a four-parameter growth function with distinct mathematical properties (Silva, 2022). As demonstrated by Silva (2018), the model's parameters must satisfy specific conditions: $\beta_1, \beta_2, \beta_3 \in \mathbb{R}^+$, $\beta_4 \in \mathbb{R}^*$ (non-zero real numbers), with the additional constraint $\beta_2 > \beta_1$.

The model's asymptotic behavior shows an interesting dependence on the sign of β_4. When β_4 is positive, the function approaches

$$y = \frac{\beta_1\beta_2}{\beta_1 + (\beta_2 - \beta_1)e^{-\frac{\beta_3}{\beta_4}}} \quad \text{as} \quad x \to \infty, \tag{20}$$

and

$$y = 0 \quad \text{as} \quad x \to -\infty. \tag{21}$$

Conversely, for negative β_4 values, the right asymptote becomes

$$y = \beta_2, \tag{22}$$

while the left asymptote follows

$$y = \frac{\beta_1\beta_2}{\beta_1 + (\beta_2 - \beta_1)e^{-\frac{\beta_3}{\beta_4}}}. \tag{23}$$

A notable feature of the Chanter model is its consistent y-intercept at β_1, ensuring that the curve always passes through the point $(0, \beta_1)$. However, as Silva (2018) established, the inflection point presents mathematical complexity—lacking a closed-form solution and requiring numerical methods for full characterization.

Our analysis using Maple software confirmed that the second derivative exhibits multiple roots. This behavior is visually demonstrated in Figures 1(c)–1(d), which contrast the model's behavior for positive versus negative β_4 values while highlighting the relationship between the function, its second derivative, and critical points.

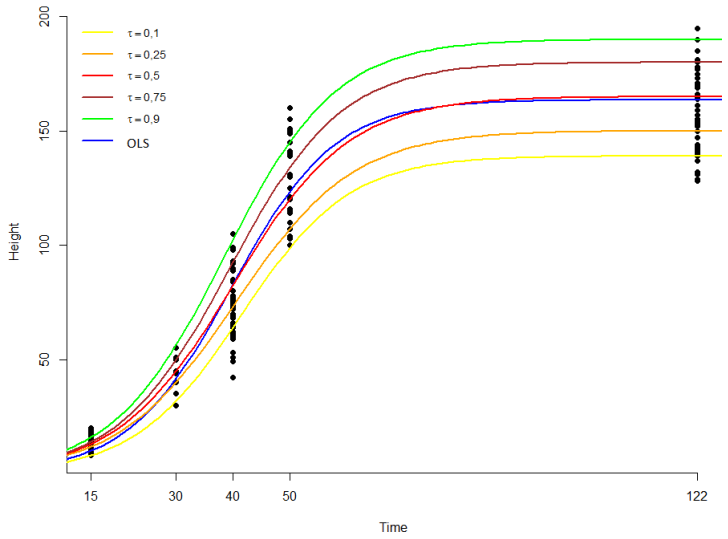
From an applied perspective, the parameters β_1 and β_2 admit clear biological interpretations: β_1 represents the initial measurement value at time zero, while β_2 serves as the right asymptotic limit when β_4 is negative, indicating the stabilization value for growth processes. The remaining parameters β_3 and β_4, while mathematically essential, currently lack direct physiological interpretation.

When applied to our corn growth dataset (plant height in cm versus days after sowing), all three models: logistic, Gompertz, and Chanter, were fit across multiple quantiles ($\tau = \{0.1, 0.25, 0.5, 0.75, 0.9\}$) alongside conventional OLS estimation. The corresponding parameter estimates for all models are detailed in Table 1.

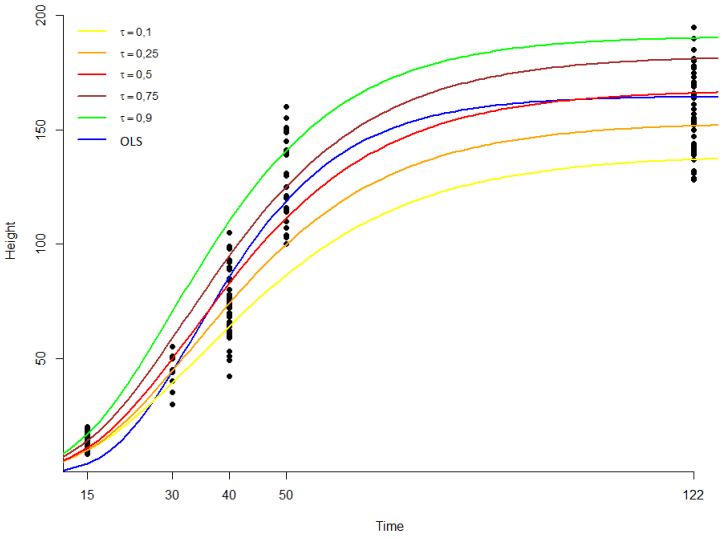
The logistic model revealed consistent patterns: β_1 (asymptotic height) increased monotonically across quantiles, while β_2 (abscissa of the inflection point) remained stable around 40, consistently indicating a growth phase transition at approximately 40 days after sowing (DAS). Parameter β_3 represented a stable growth rate, close to 10. The Gompertz model showed asymptotic increases with higher quantiles, as evidenced by β_1, along with a growth rate clustered around 6 for β_2 and values close to 1 for β_3. For the Chanter model, parameter estimates fell within biologically plausible ranges (β_1: 5.16–8.76; $\beta_3 \approx 0.04$; $\beta_4 \approx -0.02$), with β_2 maintaining its role as the asymptotic parameter. Notably, ordinary least squares estimates consistently approximated the median quantile ($\tau = 0.5$) results across all models, supporting the robustness of our quantile regression approach. Figure 2 presents the fitted growth curves for the nonlinear regression models, showing quantile-specific fits at $\tau = \{0.1, 0.25, 0.5, 0.75, 0.9\}$ alongside ordinary least squares (OLS) estimates.

Figure 2 - Fitted growth curves of (a) logistic, (b) Gompertz, and (c) Chanter models at different quantiles for corn plant height data.

(a)



(b)



(c)

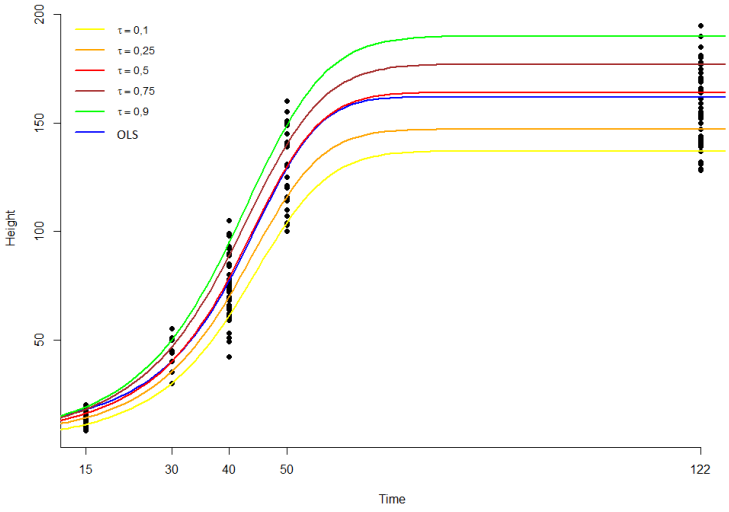


Table 1 - Parameter estimates (standard errors) for growth models fitted to corn plant height data across different quantiles (τ) and ordinary least squares (OLS).

Parameter	Quantiles (τ)					OLS
	0.1	**0.25**	**0.5**	**0.75**	**0.9**	
Logistic model						
β_1	139.03 (3.21)	148.81 (4.25)	165.05 (3.81)	180.04 (4.73)	190.04 (3.04)	1163.58 (1.82)
β_2	41.51 (0.82)	40.36 (0.76)	40.01 (0.67)	39.45 (0.80)	38.49 (0.58)	39.78 (0.37)
β_3	9.48 (0.67)	10.35 (0.49)	10.19 (0.35)	9.88 (0.45)	9.84 (0.28)	9.12 (0.33)
Gompertz model						
β_1	138.94 (3.60)	151.98 (4.40)	167.38 (3.89)	182.33 (4.92)	194.58 (3.02)	164.70 (2.08)
β_2	5.48 (0.56)	6.05 (0.49)	6.13 (0.55)	5.84 (0.29)	5.86 (0.21)	10.52 (1.14)
β_3	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.95 (0.00)	0.94 (0.00)	0.93 (0.00)
Chanter model						
β_1	5.16 (1.40)	7.24 (0.97)	8.36 (1.11)	8.54 (0.94)	8.76 (0.74)	10.67 (1.84)
β_2	137.00 (3.81)	147.00 (4.53)	164.00 (3.85)	177.00 (4.22)	190.00 (4.01)	161.96 (1.72)
β_3	0.04 (0.01)	0.03 (0.01)	0.04 (0.01)	0.04 (0.01)	0.05 (0.00)	0.03 (0.00)
β_4	-0.02 (0.01)	-0.03 (0.01)	-0.03 (0.01)	-0.02 (0.01)	-0.02 (0.00)	-0.04 (0.00)

The graphical analysis presented in Figures 2(a), 2(b) and 2(c) shows that all three models (logistic, Gompertz, and Chanter) provided good fits to the corn plant height data. However, visual inspection alone proved insufficient to determine which model offered the best fit to the observed growth patterns. To address this limitation, we performed a thorough residual analysis and model diagnostics, complemented by quantitative goodness-of-fit assessments.

The residual analysis included normality verification through normal probability plots with simulated envelopes. These plots revealed that for all quantiles and across all models, numerous residual points fell outside the expected envelope boundaries. This graphical evidence suggested potential deviations from normality in the residual distributions. We confirmed this observation statistically using Shapiro-Wilk tests, which consistently rejected the null hypothesis of normally distributed residuals at the 5% significance level for all model-quantile combinations.

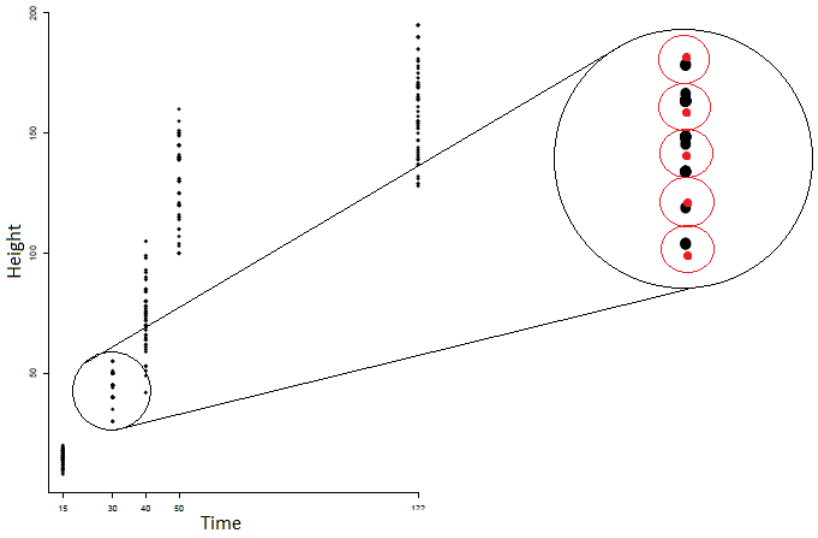
To objectively evaluate model performance, we employed four distinct criteria: the Akaike Information Criterion (AIC), Bates and Watts' intrinsic nonlinearity measure (IN), quantile correlation, and the weighted sum of squared deviations (WSSD). The comparative analysis revealed consistent superiority of the Chanter model, which demonstrated the lowest AIC values across all examined quantiles. The logistic model ranked second in performance, followed by the Gompertz model. Notably, the Chanter model also showed the smallest values for Bates and Watts' intrinsic nonlinearity measure at all quantiles, indicating its behavior most closely approximated linear model characteristics. These quantitative results are comprehensively presented in Table 2, which summarizes all comparative metrics for the three candidate models.

We propose quantile correlation as a novel goodness-of-fit measure for nonlinear models, calculated through four sequential steps: First, the nonlinear model is fitted to the data for specified quantiles ($\tau = 0.1, 0.25, 0.5, 0.75$, and 0.9). Second, predicted values are computed for all sampling days across all quantiles. Figure 3 illustrates this through a scatter plot showing observed values (black dots) versus predicted values (red dots) for one sampling day, with predicted values arranged vertically from bottom to top corresponding to increasing quantiles.

Table 2 - Comparison of goodness-of-fit measures for logistic, Gompertz, and Chanter models fitted to corn plant height data across different quantiles (τ): Akaike Information Criterion (AIC), Bates and Watts intrinsic curvature (IN), quantile correlation (r_τ), and weighted sum of squared deviations ($WSSD_\tau$).

Parameter	Quantiles				
	0.1	**0.25**	**0.5**	**0.75**	**0.9**
Logistic model					
AIC	2444.31	2373.53	2336.05	2364.68	2399.91
IN	0.0050	0.0056	0.0057	0.0048	0.0032
r_τ	**0.9899**	**0.9988**	**0.9981**	**0.9993**	**0.9983**
$WSSD_\tau$	117.03	36.42	39.51	35.99	61.55
Gompertz model					
AIC	2482.83	2425.55	2422.54	2495.22	2534.25
IN	0.0233	0.0213	0.0177	0.0056	0.0034
r_τ	**0.9899**	**0.9985**	**0.9988**	**0.9986**	**0.9979**
$WSSD_\tau$	119.04	29.07	33.11	64.85	78.16
Chanter model					
AIC	2413.82	2350.76	2309.29	2312.61	2324.40
IN	0.0037	0.0043	0.0043	0.0035	0.0023
r_τ	**0.9949**	**0.9981**	**0.9990**	**0.9880**	**0.9985**
$WSSD_\tau$	47.09	53.73	33.50	33.69	62.53

Figure 3 - Graphical representation of observation classification for the quantile correlation goodness-of-fit metric.



The third step involves classifying observed values into quantile-specific groups for each sampling day. A value belongs to quantile τ's group if its Euclidean distance to τ's predicted value is smaller than its distance to other quantiles predictions. The Euclidean distance between an observed value (x_j, y_{ij}) and predicted value $(x_j, \hat{y}_{j\tau})$ simplifies to

$$d[(x_j, y_{ij}), (x_j, \hat{y}_{j\tau})] = \sqrt{(x_j - x_j)^2 + (y_{ij} - \hat{y}_{j\tau})^2} = \sqrt{(y_{ij} - \hat{y}_{j\tau})^2} = |y_{ij} - \hat{y}_{j\tau}|, \qquad (24)$$

where x_j denotes the j-th day of data collection, y_{ij} represents the i-th observed plant height measurement on day j, and $\hat{y}_{j\tau}$ corresponds to the predicted plant height value for day j at quantile level τ.

Fourth, we calculate Pearson's correlation coefficient (r_τ) between the vector of observed values in quantile τ's group across all days and the corresponding vector of predicted values. This quantile-specific Pearson correlation (ranging from -1 to 1) evaluates model fit quality across the response distribution. Higher absolute values indicate stronger agreement between predicted and observed values for each quantile.

Building on this framework, we propose a second metric: the Weighted Sum of Squared Deviations (WSSD_τ), computed as

$$\text{WSSD}_\tau = \sum_{j=1}^{k} \sum_{i=1}^{n_{j\tau}} \left(\frac{y_{ij} - \hat{y}_{j\tau}}{n_{j\tau}} \right)^2, \tag{25}$$

where $n_{j\tau}$ represents the count of observations in quantile τ's group for day j. Lower WSSD values indicate better model fit.

Table 2 presents quantile correlation results for all models. The Chanter model achieved the highest correlations for all quantiles except $\tau = 0.25$ and 0.75, though all models showed strong positive correlations throughout. For WSSD metrics, the Chanter model demonstrated superior performance at $\tau = 0.1$, with significantly lower values than competitors. The Gompertz model showed slight advantages at $\tau = 0.25$ and 0.5, while the logistic and Chanter models performed comparably at $\tau = 0.75$, and the logistic model excelled at $\tau = 0.9$.

Both statistics confirmed the Chanter model's exceptional fit for the 0.1 quantile, evidenced by its high correlation and markedly lower WSSD value. While all models showed strong overall performance (high quantile correlations), the WSSD metric revealed more nuanced differences in quantile-specific accuracy, particularly for extreme quantiles. These results suggest that model selection should consider both global fit statistics and quantile-specific performance measures.

Conclusions

The quantile nonlinear regression models (logistic, Gompertz, and Chanter) demonstrated good fit to the plant height data of transgenic hybrid 30F35 (Yieldgard) corn, despite the non-normal distribution of residuals. The proposed models successfully estimated growth curve parameters across different quantiles of the plant height distribution.

Among the evaluated models, the Chanter model showed superior performance according to both the Akaike Information Criterion (AIC) and Bates and Watts' intrinsic curvature measure of nonlinearity. For the quantile-specific goodness-of-fit statistics (quantile correlation and WSSD), the Chanter model performed particularly well at the 0.1 quantile, while all models showed comparable results at other quantiles. The asymptotic height estimates ranged from 137 cm to 190 cm (a 53 cm range), representing the mature plant height at physiological maturity. The model intercepts, corresponding to plant height at 15 days after sowing, varied between 5.16 cm and 8.76 cm across quantiles.

This analysis provides robust evidence that quantile regression approaches can effectively model corn growth patterns across different quantiles of the population, with the Chanter model being particularly suitable for this crop hybrid.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil – Finance Code 001.

Author Contributions

P. V. da Silva participated in: conceptualization, formal analysis, programming, methodology, and writing. **T. V. Savian** participated in: supervision, and validation.

Conflicts of Interest

The authors declare that they have no competing financial or commercial interests in relation to the work described in this manuscript.

References

Akaike, H. (2003). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Amador, J. P. (2010). *Modelos mistos no ajuste de curvas de crescimento de Ricinus communis L.* [Tese de doutorado, Universidade Federal de Santa Maria]. Repositório institucional. https://repositorio.ufsm.br/handle/1/3192

Cargnelutti, A. Filho & Toebe, M. (2020). Reference sample size for multiple regression in corn. *Pesquisa Agropecuária Brasileira*, *55*, e01400. https://doi.org/10.1590/S1678-3921.pab2020.v55.01400

Crosariol Netto, J. (2013). *Infestação de dados e Dichelops melacanthus (Dallas, 1851) (Heteroptera: Pentatomidae) em híbridos transgênicos e convencionais de milho, submetidos ao controle químico* [Dissertação de mestrado, Universidade Estadual Paulista "Júlio de Mesquita Filho"]. Repositório institucional. http://hdl.handle.net/11449/91328

Fernandes, T. J., & Pereira, A. A. M. J. A. (2017). Double sigmoidal models describing the growth of coffee berries. *Ciência Rural*, *47*(8), 1–8. https://doi.org/10.1590/0103-8478cr20160646

Gray, V. M. (2000). A comparison of two approaches for modelling cassava (Manihot esculenta Crantz.) crop growth. *Annals of Botany*, *85*(1), 77–90. https://doi.org/10.1006/anbo.1999.0999

Hao, L., & Naiman, D. Q. (2007). *Quantile Regression*. Sage Publications.

Koenker, R. (2005). *Quantile Regression* (Vol. 1). Cambridge University Press.

Koenker, R., & Basset, G. J. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society*, *46*(1), 33–50. https://doi.org/10.2307/1913643

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

Migault, V., Combes, D., Barre, P., Gueye, B., Louarn, G., & Escobar-Gutiérrez, A. J. (2012). Improved modelling of ryegrass foliar growth. *Proceedings of the Plant Growth Modeling, Simulation, Visualization and Applications* [Proceesding]. 4th International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications, Shanghai, China. https://doi.org/10.1109/PMA.2012.6524847

Mischan, M. M., & Pinho, S. Z. (2014). *Modelos não lineares*. Cultura Acadêmica.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ratkowsky, D. A. (1983). *Nonlinear Regression Modeling*. Dekker.

Silva, E. M., Fruhauf, A. C., Silva, E. M., Muniz, J. A., Fernandes, T. J., & Silva, V. F. (2021). Evaluation of the critical points of the most adequate nonlinear model in adjusting growth data of ‘green dwarf’ coconut fruits. *Revista Brasileira de Fruticultura*, *43*(1), e–726. https://doi.org/10.1590/0100-29452021726

Silva, P. V. (2018). *Modelo não linear Chanter: uma aplicação aos dados de crescimento de frutos do cacaueiro* [Dissertação de mestrado, Universidade de São Paulo]. Repositório institucional. https://doi.org/10.11606/D.11.2018.tde-02072018-103309

Silva, P. V. (2022). *Uso da regressão não linear quantílica na descrição de dados de crescimento* [Tese de doutorado, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo]. Repositório institucional. https://doi.org/10.11606/T.11.2022.tde-10052022-155257

Silva, P. V., & Savian, T. V. (2019). Chanter model: nonlinear modeling of the fruit growth of cocoa. *Ciência Rural*, *49*(11), e20190409. https://doi.org/10.1590/0103-8478cr20190409

Sorrel, B. K., Tanner, C. C., & Brix, H. (2012). Regression analysis of growth responses to water depth in three wetland plant species. *AoB Plants*, pls043. https://doi.org/10.1093/aobpla/pls043

Souza, R. J., & Macêdo, F. S. (2009). *Cultura do alho: Tecnologias modernas de produção*. Editora UFLA.

United States Department of Agriculture. (2025). *World Agricultural Production* (Circular Series WAP 09-25). https://apps.fas.usda.gov/psdonline/circulars/production.pdf

Younesi, H. N., Shariati, M. M., Zerehdaran, S., Nooghabi, M., & Lovendahl, P. (2019). Using quantile regression for fitting lactation curve in dairy cows. *Journal of Dairy Research*, *86*(1), 19–24. https://doi.org/10.1017/S0022029919000013

Zeviani, W. M., Silva, A. C., Carneiro, W. J. O., & Muniz, J. A. (2012). Modelos não-lineares para a liberação de potássio de estercos animais em latossolos. *Ciência Rural*, *42*(10), 1789–1796. https://doi.org/10.1590/S0103-84782012001000012