

DOI: <http://dx.doi.org/10.1590/1807-1929/agriambi.v27n10p820-827>

Nonlinear modeling of liming reaction and extractable base curves¹

Ajuste não linear de curvas de reação de calcário e bases trocáveis

João H. M. Viana^{2*}, Manoel R. de Albuquerque Filho²,
Flávia C. dos Santos² & Daniela de A. Ladeira³

¹ Research developed at Correntina, BA, Brazil

² Embrapa Milho e Sorgo/Centro Nacional de Pesquisa de Milho e Sorgo, Sete Lagoas, Brazil

³ Universidade Federal de São João del Rey/Departamento de Ciências Agrárias, Sete Lagoas, Brazil

HIGHLIGHTS:

Statistically nonlinear models were proposed for pH as a function of lime doses.

Statistically nonlinear models were proposed for exchangeable $Ca^{+2} + Mg^{+2}$ as a function of lime doses.

The use of a statistical index combined with constraints on extrapolation values enabled the selection of the best-fitting models.

ABSTRACT: Modeling the response of soils to liming is important for understanding neutralization reactions and predicting lime residual effects. Models based on simple or quadratic polynomial equations are the most used due to their simplicity and ease of fitting; however, they fail to reproduce a realistic soil response to liming, indicating a decrease in pH as the lime dose is increased after reaching a maximum point. Thus, several nonlinear functions were tested and compared to polynomial models, using a dataset from a liming test conducted on a sandy clay loam soil in a farm. The best-fitting models for pH data were the Mitscherlich, three-parameter logistic, and Morgan-Mercer-Flodin models. The best-fitting models for exchangeable $Ca^{+2} + Mg^{+2}$ data were Skaggs et al., Gompertz, and Morgan-Mercer-Flodin. The use of the proposed T index, which ranks models based on their residual standard error and Akaike information criterion values, combined with constraints on extrapolation values, was useful for selecting models that are statistically robust and empirically coherent.

Key words: pH correction, nonlinear functions, liming effect, soil correction

RESUMO: A modelagem da resposta de solos à calagem é importante para a compreensão das reações de neutralização e para a predição de seu efeito residual. Os modelos baseados em equações polinomiais simples ou múltiplas são os mais usados, em função de sua simplicidade e facilidade de ajuste, mas falham em reproduzir uma resposta realística à calagem, indicando um decréscimo do pH com a dose, após o máximo valor ter sido alcançado. Neste trabalho, uma série de funções não lineares foi comparada à polinomial, usando uma base de dados de um teste de calagem em uma fazenda com solo franco-argilo-arenoso. Os melhores modelos foram os de Mitscherlich, o logístico de três parâmetros e o modelo de Morgan, Mercer e Flodin. Para os dados de $Ca^{+2} + Mg^{+2}$ trocáveis, os melhores modelos foram os Skaggs et al., o de Gompertz e o modelo de Morgan Mercer e Flodin. O uso conjunto do proposto índice T, baseado no ranqueamento dos modelos baseado em seus valores de erro padrão residual e do critério de Akaike, associados às restrições dos valores extrapolados, ajudaram na escolha de modelos que são estatisticamente robustos e empiricamente coerentes.

Palavras-chave: correção de pH, funções não lineares, efeito de calagem, correção de solo



INTRODUCTION

The modeling of responses of soils to liming is important for understanding neutralization reactions and predicting lime residual effects. The use of liming is a well-established agricultural practice that has been extensively studied, however, many details still require clarification. The soil reaction time to different lime sources and rates and the short-term effects on crop yield still require detailed information.

Soil preparation for new agricultural fields in the Cerrado region of western Bahia, Brazil, usually involves the use of high lime rates, which exceed the current recommended rates for any textural soil group (Albuquerque Filho et al., 2011). The soil preparation process includes applying lime rates of 6 Mg ha⁻¹ in the first year, followed by 2 Mg ha⁻¹ in the second year plus 1 Mg ha⁻¹ of gypsum, and finally 4 Mg ha⁻¹ of lime and 1 Mg ha⁻¹ of gypsum.

Fitting neutralization curves with mathematical models is used to study the kinetics of lime dissolution and estimate the unreacted lime. Models based on simple or quadratic polynomial equations are commonly used for this purpose due to their simplicity and ease of fitting.

However, these models fail to reproduce a realistic soil response to liming, as they indicate a decrease in pH as the lime rate is increased after reaching a maximum point. Thus, they cannot be used for extrapolation, not even for values slightly above the highest ones used for curve fitting. The pH levels should increase continuously as the lime rate is increased until they reach the theoretical limit defined by the plateau, which is determined by the limestone abrasion pH (just above eight); thus, equations from the growth function family, such as the monomolecular or Mitscherlich types, should be included to describe the lime reaction over time.

The aim of this work was to test several nonlinear functions and compare them to polynomial models currently used.

MATERIAL AND METHODS

The soils were sampled in agricultural plots of the Xanxere farm (13°47'01"S; 46°00' 13"W; altitude of 930 m) (Figure 1). The mean chemical properties of 20 soil samples collected from the experimental area are shown in Table 1. The farm is on the Sao Francisco plateau, characterized by a flat to gently undulating terrain with deep, well-drained, acidic,

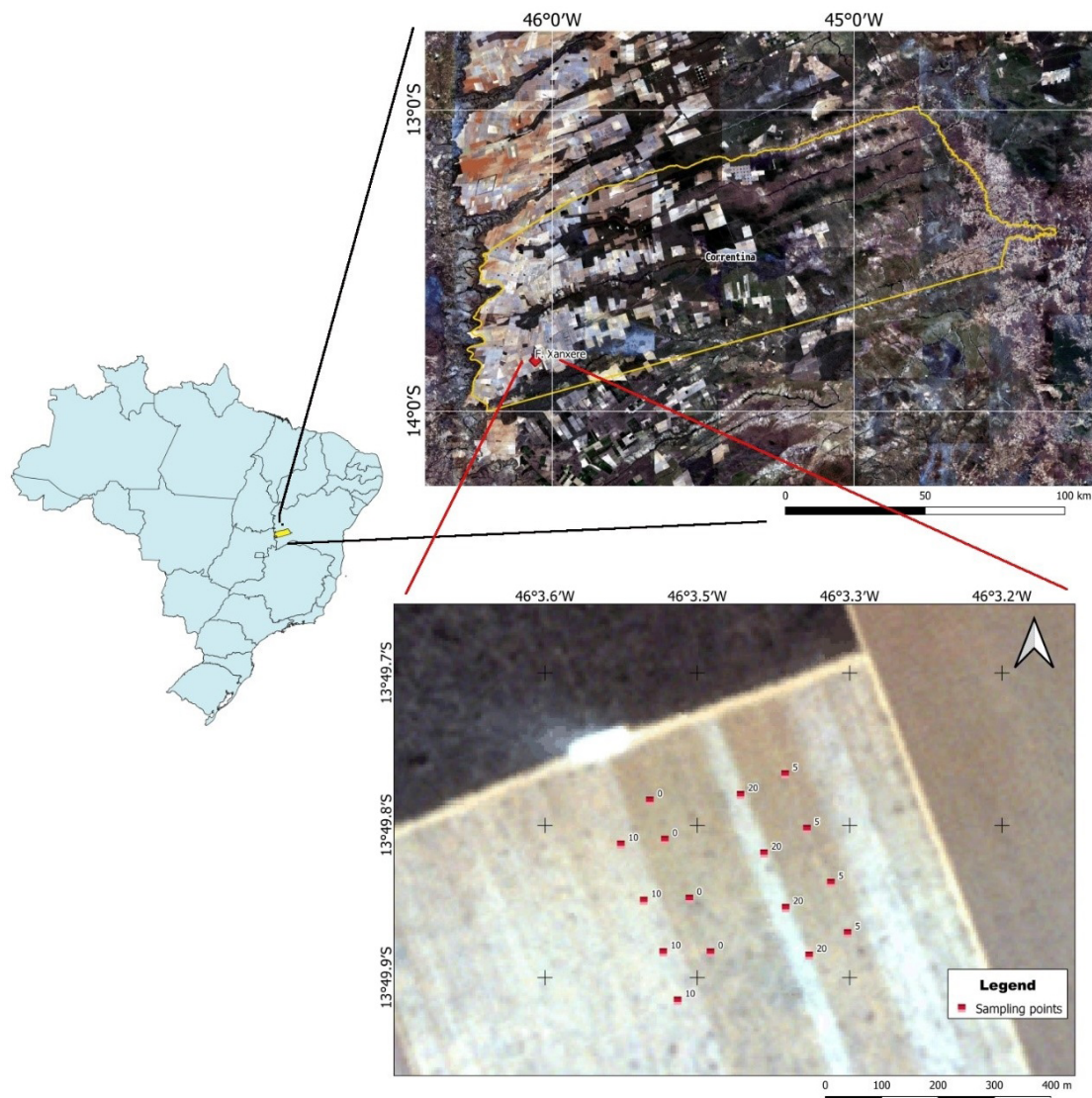


Figure 1. Map of the location of the experimental area and soil sampling points

Table 1. Soil chemical attributes of 20 samples from the experimental area

OM (g kg ⁻¹)	pH (H ₂ O)	P (mg dm ⁻³)	K (mg dm ⁻³)	Ca (cmol _c dm ⁻³)	Mg (cmol _c dm ⁻³)
10.43 (2.9)	5.37 (0.9)	9.88 (12)	25.15 (18.3)	0.59 (0.6)	0.59 (0.4)
Al ⁺³ (cmol _c dm ⁻³)	H + Al (cmol _c dm ⁻³)	CEC (T)	Sand (dag kg ⁻¹)	Silt	Clay
0.16 (0.1)	2.32 (1.5)	3.55 (1.3)	75.25 (3.9)	1.31 (0.4)	23.44 (3.8)

OM - Organic matter; (°) Means and standard deviation (in parentheses)

and low-fertility soils. The soils were classified as Oxisols or Psammentes (Latossolos Vermelho-Amarelos, Latossolos Amarelos, or Neossolos Quartzarenicos; Santos et al., 2018) of sandy clay loam texture. The mean annual rainfall depth in the region is 1,300 mm, concentrated from November to March. The water deficit period ranges from April/May to September/October. The mean air relative humidity is 64%, decreasing to 34% from June to September.

A liming test was conducted in the Xanxere farm, where 50 × 200 m plots were subjected to liming using dolomitic limestone at rates ranging from zero to 20 Mg ha⁻¹ (0, 5, 10, and 20 Mg ha⁻¹). Four points were sampled in the middle of the plots, equally spaced in a straight line, in two periods (one and two and a half years after liming) for each of the evaluated rates. A composite soil sample was obtained by collecting three samples of the 0-20 cm and 20-40 cm layers for each sampling point, using an auger. Soybean crops were grown in the area for two consecutive years and soil samples were collected in the first and second crop seasons.

The soil samples were air-dried and sent for standard fertility analyses, according to the procedures of Teixeira et al. (2017).

The soil experimental data (pH in water and sum of Ca⁺² and Mg⁺² as a function of the applied lime rates) were fitted to a set of nonlinear models (Table 2) and statistically compared using the R software (R Core Team, 2022). The nonlinear models listed in Table 1 were selected from the reference literature and have been applied for various purposes (Lima & Silva, 2003).

The performance of the models was evaluated with a ranking system that uses the values of residual standard error (RSE) and Akaike information criterion (AIC), as well as an

Table 2. Nonlinear models applied in this work

Model	Formula(°)
Four-parameter log-logistic	$Y = b + (c-b)/(1 + \exp(a \cdot \log(x)-d))$
Gompertz	$Y = 8.3 \cdot \exp(-\exp(-a \cdot (x-b)))$
Haverkamp & Parlange	$Y = 10/(1 + (a/x) \wedge b) \wedge c$
Lima & Silva	$Y = a + ((10-a)/(1 + (b/x) \wedge c) \wedge d)$
Morgan-Mercer-Flodin (MMF)	$Y = 10 - ((10-a)/(1 + (b \cdot x) \wedge c))$
Quadratic polynomial	$Y = a + (b \cdot x) + (c \cdot (x \wedge 2))$
Richards	$Y = 100 \cdot (1 + (a-1) \cdot \exp(-b \cdot (x-c))) \wedge (1/(1-a))$
Skaggs et al.	$Y = 10/(1 + ((10/a)-1) \cdot \exp(-b \cdot (x \wedge c)))$
Tangent hyperbolic	$Y = 50 \cdot (1 + \tanh((x-a)/b))$
Three-parameter log-logistic	$Y = b/(1 + \exp(a \cdot \log(x)-c))$
Three-parameter logistic	$Y = b/(1 + \exp((a \cdot x)-c))$
Weibull - 3 parameters	$Y = a \cdot \exp(-\exp(-(b \cdot \log(x)-c)))$
Weibull - 4 parameters	$Y = a + (b-a) \cdot \exp(-\exp(-(c \cdot \log(x)-d)))$
Mitscherlich	$Y = a - (a-b) \cdot \exp(-c \cdot x)$
Second-order polynomial exponential	$Y = a \cdot \exp((b \wedge 2) \cdot x \cdot (c-0.5 \cdot x))$

(°) a, b, c, d, fitted parameters

additional criterion based on the result of the extrapolation of the models outside the range of the data used for modeling. This proposed criterion compares the estimated values generated by the model for values above or below the range of the data used for its fit with the expected values based on soil chemical theory. This criterion was selected to avoid models that produce unrealistic results outside the range of the data used for model fitting, such as pH values above the maximum of the lime (8.3) or below the untreated soil pH, or declining pH values with increasing lime rates. This undesired trend is found in second-order polynomial models fitted to neutralization curves.

The ranking index show the average position of each model for each of soil layer and sampling year in a classification list, according to Eq. 1, 2, and 3. The highest values indicate the best-fitting models.

$$\text{Index } T_m = \frac{\text{RSE}_{m_{ij}} + \text{AIC}_{m_{ij}}}{2} \quad (1)$$

where:

Index T_m - final score index of the model m.

$$\text{RSE}_{m_{ij}} = \frac{\sum_{i=1}^I \sum_{j=1}^J \text{RSE}_m}{\left(\sum_{a=1}^N x_a \right) \cdot i \cdot j} \quad (2)$$

where:

$\text{RSE}_{m_{ij}}$ - score index of the model m by the RSE from year 1 to year i, and from soil layer 1 to layer j;

N - number of models to be tested (15 models; and sum of x_a , 120);

I - number of sampling years to be tested (2 years);

J - number of soil layers to be tested (2 layers);

RSE_m - position in the classification list of the model m (1 ton models in the list); and,

1st = worst and 15th = the best model, with the lowest RSE.

$$\text{AIC}_{m_{ij}} = \frac{\sum_{i=1}^I \sum_{j=1}^J \text{AIC}_m}{\left(\sum_{a=1}^N x_a \right) \cdot i \cdot j} \quad (3)$$

where:

$\text{AIC}_{m_{ij}}$ - score index of the model m by the RSE from year 1 to year i, and from soil layer 1 to layer j;

N - number of models to be tested (15 models; and sum of x_a , 120);

I - number of sampling years to be tested (2 years);

j - number of soil layers to be tested (2 layers);

AIC_m - position in the classification list of the model m (1 ton models in the list); and,

1st = worst and 15th = the best model, with the lowest AIC.

RESULTS AND DISCUSSION

Soil pH as a function of lime rates presented the expected increases in both years evaluated, but with a decreasing effect

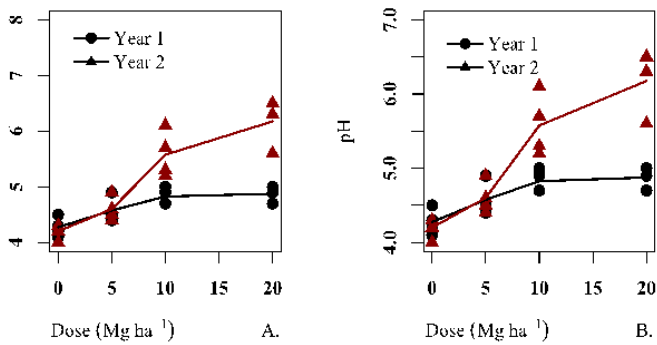


Figure 2. Soil pH as a function of the applied lime doses, a) 0-20 cm soil layer, b) 20-40 cm soil layer, in two years

of lime rates (Figures 2A and B), despite the increase in lime rates at double rates (from 5 to 10 to 20 Mg ha⁻¹). The mean pH was 7 for the highest lime rate in the first year, and 6.6 in the year 2. The decrease in pH of the soil surface layer from the first to the second year is connected to the effects of the crop grown in the area, such as root acidification and nutrient extraction, among others. The flattening trend observed in the reaction curves over time, indicated by the measured pH levels, has been previously reported in other studies (Nogaroli & Fonseca, 2020; Cavalli et al. 2021).

The increase in pH of the soil subsurface layer from the first to the second year is due to vertical percolation and/or biological activity, decreasing the liming effects. Long-term experiments using high lime rates of up to 20 Mg ha⁻¹ have reported a pH plateau in the range of 7 to 7.5 (Goulding, 2016). Although the decrease in soil surface pH over time was small, it may have been affected by soil spatial variability, sampling procedures, and irregular distribution of lime throughout the area. However, even small changes in pH can impact model predictions, which will be discussed later. Decreases in lime solubility as pH increases may reduce the effects of higher lime rates, as already found in previous studies.

Tables 3 and 4 present the statistical evaluation of model fitting to pH data, while Table 5 presents the evaluation for exchangeable Ca⁺² + Mg⁺² data. Some models performed poorly, with higher values of residual standard error (RSE) and Akaike and Bayesian information criteria (AIC and BIC). The Haverkamp & Parlange, Weibull (three parameters), and Three-parameter log-logistic models did not adequately fit the data, mainly for values close to zero (Figures 3 and 4), resulting in lower ranks in the classification by the T index based on RSE and AIC (Tables 6 and 7). Therefore, these models were discarded.

The others models had mixed performance, with a narrow range of RSE and AIC values. Considering the 15 models, 10 of them [Gompertz; Lima & Silva; Mitscherlich; Four-parameter log-logistic; Morgan-Mercer-Flodin, Richards, Quadratic polynomial, Second-order polynomial exponential, Skaggs et al.; and Weibull (4 parameters)] were among the three best models for each of the four sets of pH data, denoting the relative flexibility of these models to fit the data. Additionally, this shows the practical difficulty of choosing an appropriate function model based solely on empirical fit quality.

The three best-fitting models, based on T index classification, were the Four-parameter log-logistic, Quadratic

Table 3. Statistics for comparison of model fit to pH data in the first year

Model	RSE	Convergence	AIC	BIC	RDF
0-20 cm soil layer (*)					
Haverkamp & Parlange	2.700	false	81.8	84.9	13
Weibull - 3 parameters	2.690	false	81.8	84.9	13
Three-parameter log-logistic	2.690	false	81.8	84.9	13
Tangent hyperbolic	0.484	true	26.1	28.4	14
Second-order polynomial exponential	0.426	true	22.8	25.9	13
Morgan-Mercer-Flodin (MMF)	0.422	true	22.5	25.5	13
Gompertz	0.421	true	21.6	23.9	14
Quadratic polynomial	0.419	true	22.3	25.3	13
Weibull - 4 parameters	0.412	false	22.4	26.3	12
Three-parameter logistic	0.412	true	21.7	24.8	13
Lima & Silva	0.411	false	22.3	26.2	12
Four-parameter log-logistic	0.411	false	22.3	26.2	12
Richards	0.411	true	21.7	24.7	13
Skaggs et al.	0.411	true	21.6	24.7	13
Mitscherlich	0.407	true	21.4	24.4	13
20-40 cm soil layer (*)					
Weibull - 3 parameters	2.380	true	77.8	80.9	13
Haverkamp & Parlange	2.380	true	77.8	80.9	13
Three-parameter log-logistic	2.380	true	77.8	80.9	13
Tangent hyperbolic	0.195	true	-3.1	-0.8	14
Gompertz	0.191	true	-3.8	-1.4	14
Lima & Silva	0.173	false	-5.3	-1.5	12
Weibull - 4 parameters	0.171	true	-5.8	-1.9	12
Four-parameter log-logistic	0.171	true	-5.8	-1.9	12
Skaggs et al.	0.169	true	-6.7	-3.6	13
Richards	0.169	true	-6.8	-3.7	13
Mitscherlich	0.168	true	-7.0	-3.9	13
Three-parameter logistic	0.168	true	-7.1	-4.0	13
Morgan-Mercer-Flodin (MMF)	0.167	true	-7.2	-4.1	13
Quadratic polynomial	0.165	true	-7.6	-4.5	13
Second-order polynomial exponential	0.165	true	-7.7	-4.6	13

(*)RSE = residual standard error; Convergence = whether the model converged after 50 iterations; AIC = Akaike information criterion; BIC = Bayesian information criterion; RDF = residual degrees of freedom

polynomial, and Second-order polynomial exponential models. However, the latter two models failed to predict acceptable pH levels beyond the range of the model data, as they predicted a decrease in pH levels as lime rates increased, which is unrealistic (Tables 8 and 9). The four-parameter log-logistic model also failed to predict an acceptable pH level (9.3) for the 0-20 cm soil layer in the first year for a lime rate of 100 Mg ha⁻¹, which exceeds the theoretical limit (8.3) achievable by liming. Other models, such as Weibull (3 parameters), Three-parameter log-logistic, Tangent hyperbolic, and Weibull (4 parameters), overestimated pH levels for higher lime rates (Tables 8 and 9). Thus, according to the combined use of T index and constraints on extrapolation values, the best models to fit the pH data were the Mitscherlich, Three-parameter logistic, and Morgan-Mercer-Flodin models.

The statistical evaluation of model fitting to exchangeable Ca⁺² + Mg⁺² data in the second year (Table 5) resulted in the ranking presented in Table 7, in which the best models were the Second-order polynomial exponential, Gompertz, and Morgan-Mercer-Flodin models. Most of the models produced acceptable results within the range of fitting. However, as expected, the polynomial models predicted a decrease in pH levels beyond that range, which is not realistic or acceptable in real-world situations (Table 7 and Figures 5A and B). Although the theoretical upper limit for exchangeable Ca⁺² + Mg⁺² in soils may be high, such as in soils developed from limestones, stoichiometric calculations constrain the increase

Table 4. Statistics for comparison of model fit to pH data in the second year

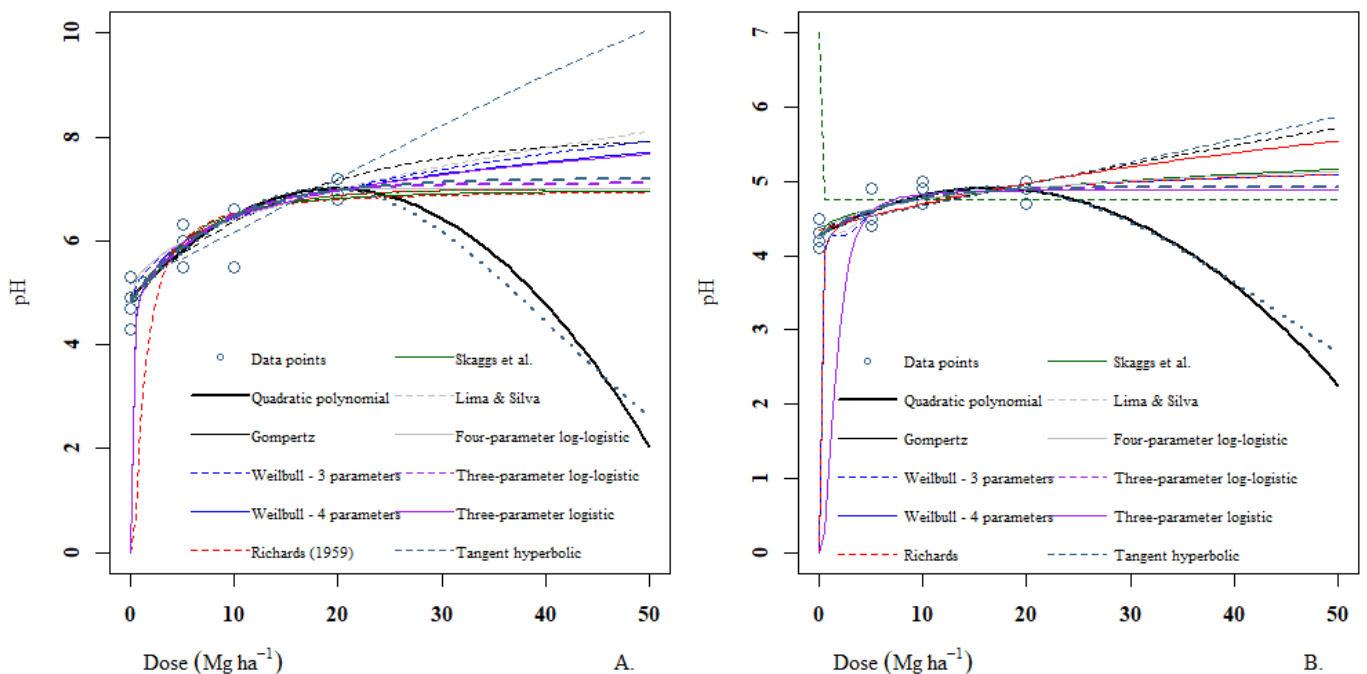
Model	RSE	Convergence	AIC	BIC	RDF
0-20 cm soil layer (°)					
Weibull - 3 parameters	2.580	true	80.4	83.4	13
Haverkamp & Parlange	2.580	false	80.4	83.5	13
Three-parameter log-logistic	2.580	true	80.4	83.4	13
Tangent hyperbolic	0.379	true	18.2	20.6	14
Mitscherlich	0.303	true	11.9	15.0	13
Richards	0.295	true	11.0	14.1	13
Three-parameter logistic	0.295	true	11.0	14.1	13
Skaggs et al.	0.292	true	10.7	13.8	13
Gompertz	0.288	true	9.5	11.8	14
Quadratic polynomial	0.287	true	10.1	13.2	13
Second-order polynomial exponential	0.275	true	8.8	11.9	13
Morgan-Mercer-Flodin (MMF)	0.258	true	6.8	9.9	13
Lima & Silva	0.254	false	6.9	10.8	12
Weibull - 4 parameters	0.235	true	4.5	8.3	12
Four-parameter log-logistic	0.235	true	4.5	8.3	12
20-40 cm soil layer (°)					
Weibull - 3 parameters	2.270	true	148.0	154	29
Haverkamp & Parlange	2.270	false	148.0	154	29
Three-parameter log-logistic	2.270	true	148.0	154	29
Tangent hyperbolic	0.474	true	47.0	51.4	30
Lima & Silva	0.470	false	48.2	55.6	28
Skaggs et al.	0.468	true	47.1	53.0	29
Richards	0.467	true	46.9	52.8	29
Mitscherlich	0.466	true	46.8	52.6	29
Morgan-Mercer-Flodin (MMF)	0.466	true	46.7	52.6	29
Weibull - 4 parameters	0.465	true	47.5	54.8	28
Four-parameter log-logistic	0.465	true	47.5	54.8	28
Three-parameter logistic	0.465	true	46.6	52.5	29
Quadratic polynomial	0.464	true	46.5	52.4	29
Second-order polynomial exponential	0.463	true	46.4	52.3	29
Gompertz	0.462	true	45.3	49.7	30

(*)RSE = residual standard error; Convergence = whether the model converged after 50 iterations; AIC = Akaike information criterion; BIC = Bayesian information criterion; RDF = residual degrees of freedom

Table 5. Statistics for comparison of model fit to exchangeable Ca⁺² + Mg⁺² data in the second year

Model	RSE	Convergence	AIC	BIC	RDF
0-20 cm soil layer (°)					
Tangent hyperbolic	0.508	true	27.6	29.9	14
Haverkamp & Parlange	0.399	false	20.7	23.8	13
Weibull - 4 parameters	0.398	true	21.4	25.2	12
Lima & Silva	0.398	true	21.4	25.2	12
Four-parameter log-logistic	0.398	true	21.4	25.2	12
Weibull - 3 parameters	0.395	true	20.4	23.4	13
Three-parameter log-logistic	0.395	true	20.4	23.4	13
Mitscherlich	0.391	true	20.0	23.1	13
Quadratic polynomial	0.389	true	19.9	23.0	13
Gompertz	0.389	true	19.1	21.4	14
Second order polynomial exponential	0.388	true	19.8	22.9	13
Richards	0.387	true	19.7	22.8	13
Three-parameter logistic	0.386	true	19.7	22.8	13
Skaggs et al.	0.386	true	19.6	22.7	13
Morgan-Mercer-Flodin (MMF)	0.383	true	19.4	22.5	13
20-40 cm soil layer (°)					
Three-parameter logistic	0.280	false	9.37	12.5	13
Tangent hyperbolic	0.271	true	7.45	9.77	14
Mitscherlich	0.247	false	5.32	8.41	13
Quadratic polynomial	0.247	true	5.31	8.41	13
Skaggs et al.	0.240	false	4.44	7.53	13
Richards	0.240	true	4.36	7.45	13
Lima & Silva	0.238	false	4.88	8.74	12
Weibull - 4 parameters	0.238	true	4.82	8.68	12
Four-parameter log-logistic	0.238	true	4.82	8.68	12
Haverkamp & Parlange	0.235	false	3.74	6.83	13
Weibull - 3 parameters	0.235	true	3.73	6.82	13
Three-parameter log-logistic	0.235	true	3.73	6.82	13
Morgan-Mercer-Flodin (MMF)	0.234	true	3.64	6.73	13
Gompertz	0.233	true	2.72	5.04	14
Second order polynomial exponential	0.230	true	3.06	6.15	13

(*)RSE = residual standard error; Convergence = whether the model converged after 50 iterations; AIC = Akaike information criterion; BIC = Bayesian information criterion; RDF = residual degrees of freedom

**Figure 3.** Soil pH as a function of applied lime doses, with extrapolation beyond the evaluated doses (A) 0-20 cm soil layer, (B) 20-40 cm soil layer, in the first year

in exchangeable Ca⁺² + Mg⁺² to be proportional to the applied rate; moreover, empirical evidence indicates that only a portion of it will be available for measurements due to different soil reactions. Therefore, models that predict Ca⁺² + Mg⁺² contents

far beyond those usually found in soils are unrealistic and should be discarded. Furthermore, this is coherent with the expected pH response. The three-parameter logistic model and Tangent hyperbolic model predicted pH levels that were one to

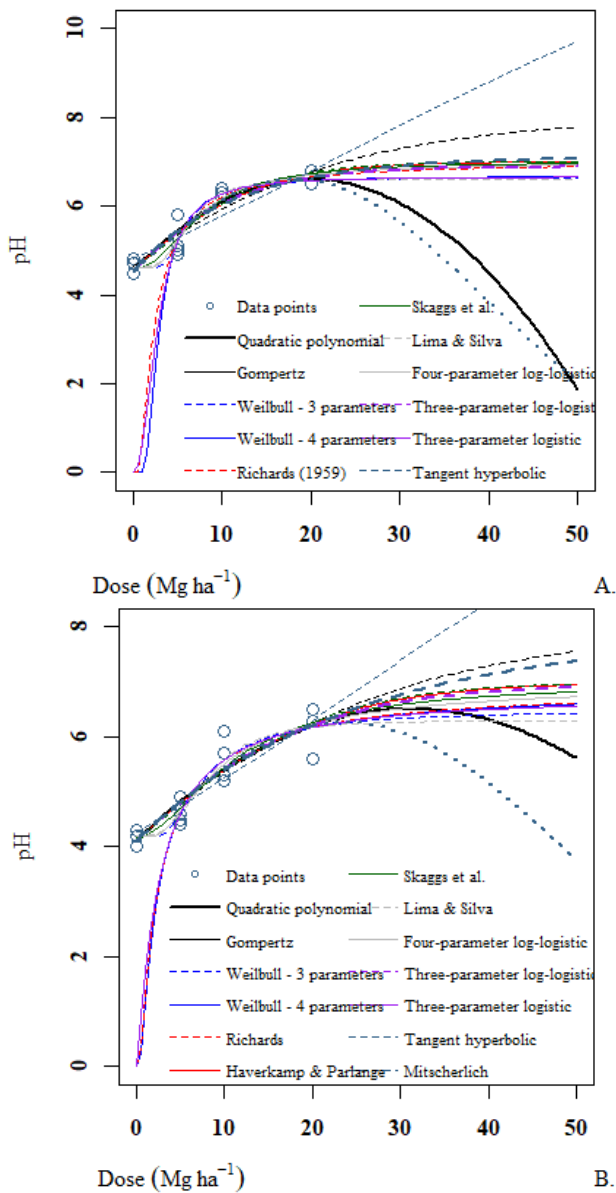


Figure 4. Model prediction extrapolated beyond the fitted data (pH data as a function of applied lime doses), (A) 0-20 cm soil layer, (B) 20-40 cm soil layer, in the second year

Table 6. Ranking of the models according to results of Tables 3 and 4 (model fit to pH data in both sampling years). The higher values of T index indicate the best-fitting models

pH Depth Model	RSE		AIC		T index
	0-20	20-40	0-20	20-40	
Weilbull - 3 parameters	0.01	0.01	0.01	0.01	0.01
Haverkamp & Parlange	0.01	0.02	0.01	0.02	0.02
Three-parameter log-logistic	0.03	0.03	0.03	0.03	0.03
Tangent hyperbolic	0.03	0.03	0.03	0.05	0.04
Lima & Silva	0.10	0.05	0.09	0.04	0.07
Richards	0.08	0.07	0.08	0.08	0.08
Weilbull - 4 parameters	0.10	0.07	0.09	0.05	0.08
Skaggs	0.09	0.06	0.09	0.07	0.08
Gompertz	0.07	0.08	0.10	0.08	0.08
Mitscherlich	0.08	0.08	0.08	0.09	0.08
Three-parameter logistic	0.07	0.10	0.08	0.10	0.09
Morgan-Mercer-Flodin (MMF)	0.08	0.09	0.08	0.10	0.09
Four-parameter log-logistic	0.11	0.08	0.10	0.06	0.09
Quadratic polynomial	0.08	0.11	0.07	0.11	0.09
Second-order polynomial exponential	0.07	0.12	0.07	0.12	0.09

RSE = residual standard error; AIC = Akaike information criterion

Table 7. Ranking of the models according to results of the tables 3 and 4 (model fit to exchangeable Ca⁺² + Mg⁺² data in the second year). The higher values of T index indicate the best-fitting models

Ca ⁺² + Mg ⁺² Model	RSE	AIC	T index
	Index		
Tangent hyperbolic	0.013	0.013	0.013
Lima & Silva	0.046	0.033	0.040
Weilbull - 4 parameters	0.046	0.033	0.040
Mitscherlich	0.046	0.046	0.046
Four-parameter log-logistic	0.058	0.046	0.052
Quadratic polynomial	0.054	0.054	0.05
Haverkamp & Parlange	0.050	0.063	0.056
Three-parameter logistic	0.058	0.054	0.056
Weilbull - 3 parameters	0.071	0.071	0.071
Richards	0.075	0.083	0.079
Three-parameter log-logistic	0.079	0.079	0.079
Skaggs	0.079	0.088	0.083
Second-order polynomial exponential	0.108	0.100	0.104
Gompertz	0.100	0.125	0.113
Morgan-Mercer-Flodin (MMF)	0.117	0.113	0.115

RSE = residual standard error; AIC = Akaike information criterion

Table 8. Extrapolated estimated pH values by the models, in the first year, for lime doses beyond the evaluated doses

Model	Lime rate (Mg ha ⁻¹)		
	25	50	100
Layer - 0-20			
Quadratic polynomial	6.85	2.03	-27.97
Second-order polynomial exponential	6.74	2.61	0.01
Haverkamp & Parlange	6.84	6.93	6.97
Morgan-Mercer-Flodin (MMF)	6.87	6.96	6.99
Skaggs	6.93	7.00	7.00
Richards	6.93	7.00	7.00
Lima & Silva	7.00	7.00	7.00
Three-parameter logistic	7.04	7.10	7.10
Mitscherlich	7.07	7.19	7.19
Gompertz	7.41	7.90	8.00
Three-parameter log-logistic	7.13	7.66	8.21
Weilbull - 3 parameters	7.14	7.70	8.29
Weilbull - 4 parameters	7.19	7.92	8.77
Four-parameter log-logistic	7.22	8.10	9.27
Tangent hyperbolic	7.71	10.07	12.85
Layer - 20-40			
Quadratic polynomial	4.74	2.25	-11.59
Second-order polynomial exponential	4.72	2.68	0.12
Four-parameter log-logistic	4.88	4.88	4.88
Weilbull - 4 parameters	4.88	4.88	4.88
Three-parameter log-logistic	4.88	4.89	4.89
Weilbull - 3 parameters	4.88	4.89	4.89
Three-parameter logistic	4.91	4.92	4.92
Mitscherlich	4.91	4.93	4.93
Haverkamp & Parlange	4.92	4.96	4.98
Lima & Silva	4.93	4.97	4.99
Morgan-Mercer-Flodin (MMF)	4.94	4.98	4.99
Skaggs	4.95	4.99	5.00
Richards	4.95	4.99	5.00
Gompertz	5.10	5.65	6.35
Tangent hyperbolic	5.11	5.82	7.13

almost two orders of magnitude higher than those predicted by the other models. According to the joint use of T index and constrains on extrapolation values, Skaggs, Gompertz, and Morgan-Mercer-Flodin were the best models.

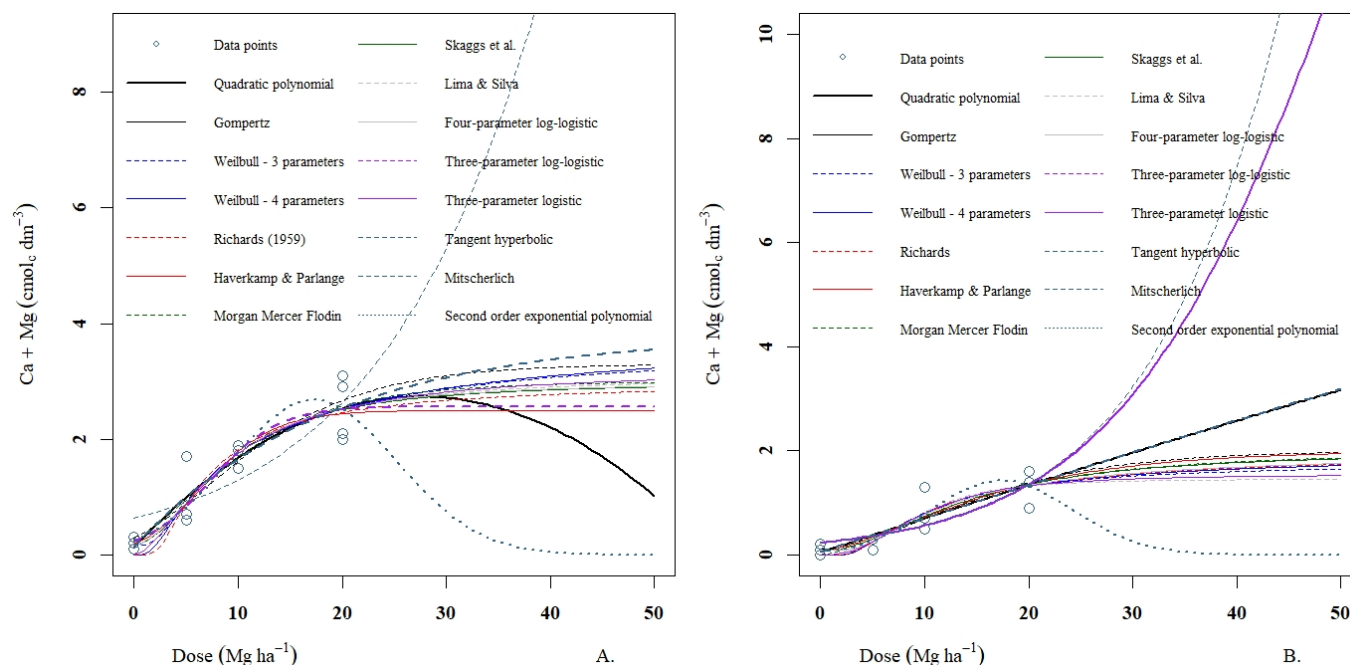
The results of the present study highlight two important but often not obvious points when fitting models. Firstly, the risk of extrapolating the model beyond the range of the data

Table 9. Extrapolated estimated pH values by the models, in the second year, for lime doses beyond the evaluated doses

Model	Lime dose (Mg ha ⁻¹)		
	25	50	100
Layer - 0-20			
Quadratic polynomial	6.47	1.87	-26.93
Second-order polynomial exponential	6.28	2.08	0.00
Four-parameter log-logistic	6.59	6.61	6.61
Weibull - 4 parameters	6.60	6.63	6.64
Three-parameter log-logistic	6.61	6.65	6.66
Weibull - 3 parameters	6.61	6.66	6.67
Three-parameter logistic	6.76	6.88	6.89
Haverkamp & Parlange	6.74	6.89	6.96
Lima & Silva	6.79	6.93	6.98
Morgan-Mercer-Flodin (MMF)	6.83	6.96	6.99
Gompertz	6.79	6.99	7.00
Richards	6.85	7.00	7.00
Skaggs	6.88	7.00	7.00
Mitscherlich	6.80	7.06	7.10
Tangent hyperbolic	7.31	9.72	12.70
Layer - 20-40			
Quadratic polynomial	5.58	3.93	-8.43
Second-order polynomial exponential	5.50	3.36	0.13
Four-parameter log-logistic	5.56	5.60	5.60
Weibull - 4 parameters	5.58	5.65	5.68
Three-parameter log-logistic	5.59	5.71	5.77
Weibull - 3 parameters	5.59	5.72	5.79
Three-parameter logistic	5.67	5.87	5.89
Mitscherlich	5.70	6.00	6.08
Haverkamp & Parlange	5.68	6.00	6.24
Lima & Silva	5.73	6.12	6.42
Morgan-Mercer-Flodin (MMF)	5.75	6.24	6.59
Richards	5.75	6.34	6.79
Skaggs	5.78	6.44	6.87
Gompertz	5.85	6.56	6.94
Tangent hyperbolic	6.00	8.24	15.07

Table 10. Extrapolated estimated values of exchangeable Ca²⁺ + Mg²⁺ by the models, in the second year, for lime doses beyond the evaluated doses

Model	Lime dose (Mg ha ⁻¹)		
	25	50	100
Layer - 0-20			
Quadratic polynomial	2.71	1.02	-15.15
Second-order polynomial exponential	1.67	0.00	0.00
Richards	2.48	2.50	2.50
Three-parameter logistic	2.56	2.57	2.57
Haverkamp & Parlange	2.59	2.83	2.93
Morgan-Mercer-Flodin (MMF)	2.66	2.90	2.97
Lima & Silva	2.68	2.91	2.98
Skaggs	2.75	2.98	3.00
Four-parameter log-logistic	2.69	2.96	3.05
Three-parameter log-logistic	2.70	3.02	3.15
Gompertz	2.94	3.28	3.30
Weibull - 4 parameters	2.73	3.19	3.46
Weibull - 3 parameters	2.73	3.23	3.54
Mitscherlich	2.84	3.55	3.78
Tangent hyperbolic	3.75	19.24	89.92
Layer - 20-40			
Second-order polynomial exponential	0.76	0.00	0.00
Four-parameter log-logistic	1.39	1.45	1.46
Three-parameter log-logistic	1.41	1.52	1.54
Weibull - 4 parameters	1.44	1.65	1.74
Weibull - 3 parameters	1.45	1.72	1.84
Haverkamp & Parlange	1.46	1.74	1.88
Lima & Silva	1.47	1.76	1.89
Morgan-Mercer-Flodin (MMF)	1.52	1.84	1.96
Skaggs	1.53	1.86	1.97
Richards	1.56	1.94	2.00
Gompertz	1.60	1.97	2.00
Quadratic polynomial	1.66	3.17	5.95
Mitscherlich	1.66	3.19	6.04
Three-parameter logistic	2.06	11.38	23.79
Tangent hyperbolic	2.10	16.26	94.06

**Figure 5.** Exchangeable Ca²⁺ + Mg²⁺ as a function of applied lime doses, with extrapolation beyond the evaluated lime doses (A) 0-20 cm soil layer, (B) 20-40 cm soil layer, in the second year

without proper evaluation of the results (Table 10). Secondly, the practical usefulness of adding more criteria, both theoretical and empirical, to aid in selecting the best model. Purely statistical criteria may be misleading if used blindly,

as model fitting is sensitive to the absolute values of the data, mainly when considering nonlinear data. Furthermore, the uncertainty of the data limits the reliability of the fitted models, no matter how accurate the fitting procedures are.

CONCLUSIONS

1. The best-fitting models for soil pH as a function of lime doses were the Mitscherlich, three-parameter logistic, and Morgan-Mercer-Flodin models.

2. The best-fitting models for exchangeable $\text{Ca}^{+2} + \text{Mg}^{+2}$ as a function of lime doses were the Skaggs, Gompertz, and Morgan-Mercer-Flodin models.

3. The use of the proposed T index, which ranks models based on their residual standard error and Akaike information criterion values, combined with analysis of estimated values through extrapolation beyond the limits of the data used for fitting, was useful for selecting models that are statistically robust and empirically coherent.

ACKNOWLEDGEMENTS

The authors thank the Agronomist Luiz Carlos Bergamaschi, the owner of Xanxere Farm, for his support during the fieldwork and for granting permission to conduct research on his farm.

LITERATURE CITED

- Albuquerque Filho, M.R. de; Egreja Filho, F. B.; Santos, F. C.; Ferreira, G. B. Manejo da calagem em solo de textura média da Bahia – produtividade de soja, alterações no solo e reação do calcário. Sociedade Brasileira de Ciência do Solo, 2011. <https://www.alice.cnptia.embrapa.br/alice/handle/doc/897741>
- Cavalli, E., Lange, A., Teixeira, L.A.J., Cavalli, C., Abreu, C.A. and Mellis, E.V. Efficiency of a mine by-product as agricultural liming material and source of zinc. *Bragantia*, v.80, p.1-13, 2021. <https://doi.org/10.1590/1678-4499.20200466>
- Goulding, K.W.T. Soil acidification and the importance of liming agricultural soils with particular reference to the United Kingdom. *Soil Use Manage*, v.32, p.390-399. 2016. <https://doi.org/10.1111/sum.12270>
- Lima, J. E. F. W.; Silva, E. M. Teste e comparação de modelos matemáticos para o traçado de curvas granulométricas. Embrapa Cerrados: Planaltina, 2003. 122p.
- Nogaroli, J. A. Fonseca, A. F. Acidity control after surface application of micronized liming material in a Typic Distrudept under no-tillage. *Brazilian Archives of Biology and Technology*, v.63, p.1-12, 2020. <https://doi.org/10.1590/1678-4324-2020190411>
- R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna: Austria. 2022. Available on: <<https://www.r-project.org/>>. Accessed on: Jun. 2022.
- Santos, H. G.; Jacomine, P. K. T.; Anjos, L. H. C. Dos; Oliveira, V. A.; Lumberras, J. F.; Coelho, M. R.; Almeida, J. A.; Cunha, T. J. F.; Oliveira, J. B. Sistema Brasileiro de Classificação de Solos. Brasília: Embrapa, 2018. 356p.
- Teixeira, P. C.; Donagemma, G. K.; Fontana, A.; Teixeira, W. G. Manual de métodos de análise de solo. Brasília: Embrapa. 2017. 573p.