

Statistical-Mathematical Modeling in Agrarian Processes. An application in Agricultural Engineering



Modelación Estadístico-Matemática en Procesos Agrarios. Una aplicación en la Ingeniería Agrícola

<http://opn.to/a/xisqZ>

Dr.C. Lucía Fernández-Chuairey ^{1*}, MSc. Lazara Rangel-Montes de Oca ¹, Dr.C. Caridad Walkiria Guerra-Bustillo ¹, Dr. MVZ Jany del Pozo-Fernández ¹

¹Universidad Agraria de La Habana, San José de las Lajas, Mayabeque, Cuba.

ABSTRACT: In the agrarian sector, there is an ever-growing demand to describe processes that require Statistical-Mathematical Modeling, as well as work methodologies that allow the efficient development of research projects. In this sense, the Agrarian University of Havana (UNAH) has multidisciplinary groups that develop the use of Mathematical Models, statistical tools and the use of differential calculus in search of optimal solutions and productions. Due to the interest of this topic and its repercussion in the research, the present review was developed to establish criteria and assessments in the analysis and application of models that describe Agrarian Processes, on a mathematical - statistical basis. It includes an application in studies of post-harvest curves for the weight loss (g) in pineapple crop. In addition, the Logistics and VonBertalaffy models were tested and a comparative analysis in which the Logistic Model allowed a better description and explanation of the kinetics of weight loss as well as indicators of the speed and moment, where the maximum speed of weight loss was reached. It is concluded that the development of Statistical-Mathematical Modeling allows responding to several current problems of research in the agricultural sector and to the teaching and scientific-research challenges of the new University.

Keywords: Mathematical modeling, agrarian processes, post-harvest curves, pineapple cultivation.

RESUMEN: En el sector agrario existe una demanda siempre creciente de describir procesos que requieren de la Modelación Estadístico-Matemática, así como de metodologías de trabajo que permitan el desarrollo eficiente de los proyectos de investigación. En tal sentido la Universidad Agraria de La Habana (UNAH) cuenta con grupos multidisciplinarios que desarrollan el uso de Modelos Matemáticos, herramientas estadísticas y el empleo del cálculo diferencial en busca de soluciones y producciones óptimas. Por el interés de esta temática y su repercusión en la investigación, se desarrolló la presente revisión para establecer criterios y valoraciones en el análisis y aplicación de modelos que describen Procesos Agrarios, sobre bases matemático - estadísticas. Se incluye una aplicación en estudios de curvas pos cosecha para la pérdida de peso (g) en el cultivo de la Piña, se probaron los Modelos Logístico y Von Bertalaffy y se realiza un análisis comparativo en el que el Modelo Logístico permitió mejor descripción y explicación de la cinética de pérdida de peso e indicadores de la velocidad y momento donde se alcanzó la máxima velocidad de pérdida de peso. Se concluye que el desarrollo de la Modelación Estadístico-Matemática, permite dar respuesta a diversos problemas vigentes de investigación en el sector agrario y a retos docentes y científico-investigativo de la nueva Universidad.

Palabras clave: modelación matemática, procesos agrarios, curvas poscosecha, cultivo de piña.

*Author for correspondence: Lucía Fernández-Chuairey, E-mail: lucia@unah.edu.cu

Received: 25/07/2018

Accepted: 25/02/2019

INTRODUCTION

Historically, the study of the dynamics of agrarian processes has been accompanied by algebraic expressions, but because they are very complex systems by nature, at first they could not be solved by the traditional methods of Mathematical Analysis. In practice, these phenomena have evolved, and in many cases, they were represented through Mathematical Statistical Models, which allow describing processes, perform detailed quantitative analyses, predict the behavior of objects in different conditions, and develop techniques that allow establishing work strategies to achieve optimal solutions and productions.

Statistical-Mathematical Modeling and its applications in Agrarian Sciences, have been present for more than 40 years at the Agrarian University of Havana and researching in collaboration with Institute of Animal Science (ICA), National Center for Agricultural Health (CENSA) and National Institute of Agricultural Sciences (INCA). They have worked biometry and modeling of agrarian and biological processes. This is corroborated in the works of [Menchaca \(1978, 1990\)](#), in studies of lactation curves in dairy cows, in which he used a multiplicative model with effects of these curves, in parallel. [Guerra \(1980\)](#), used the Response Surface Methodology (MSR) in crops of sugarcane, citrus fruits, pastures and fodder, to explore the optimal conditions of fertilization.

[Menchaca \(1990\)](#), used models called staged or by steps, to describe animal growth curves, and in the work of [Del Pozo and Herrera \(1995\)](#), the use of multiplicative models with control of growth curves and environmental effects in star grass (*Cynodon nlemfumsis*) growth studies is reported.

[Fernández \(1996, 2004\)](#), conducted a study associated with Mathematical Models that describe the dynamics of biological processes in agricultural sciences, with emphasis on nonlinear regressions and the use of iterative methods for the estimation of parameters and not by the linearization of the model as it had traditionally been worked. The development of computer science and statistical softwares facilitated the work of non-linear regressions in this way. More

recently, [Fernández et al. \(2018\)](#), exposed the evolution in nonlinear regressions and the proposal of initial parameters from mathematical-biological interpretations of them; she also reported examples ranging from animal and plant growth models, allometric growth and lactation curves, among other processes.

Likewise, models associated with genetic evaluations and improvements in cattle (with the use of the BLUP (Best Linear Unbiased Predictor) and the use of weighing in the control day), as well as pasture growth models (King Grass) have been approached for the determination of the optimum cutting moment ([Rodríguez, 2015](#)). Other researches in the area of Agricultural Sciences have been reported by [Fernández \(1996a, 1996b, 2004\)](#), [Del Valle \(2000\)](#), [Del Valle and Guerra \(2012\)](#) and [Vázquez et al. \(2014\)](#), among others.

In Agricultural Technical Sciences, Mathematical Modeling has been applied to the prediction of quality properties of [Rangel fruits \(2015\)](#), in which linear and non-linear models are involved in the characterization of these processes. The theoretical basis of the most used models are addressed in this work. For the interest of this topic and its impact on research, the present review was developed to establish criteria and assessments in the analysis and application of models that describe agrarian processes, on mathematical - statistical bases.

METHODS

Theoretical fundament

Linear models

The best-known case of Mathematical Models is the linear model, which is expressed in a general way as:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_i x_{pi} + e_i$$

Where:

Y_i - Dependent variable or response.

β_j - Parameters of the model ($j = 0, 1, 2 \dots p$)

x_{ji} - i -th value of the independent variable ($j = 1, 2, 3, \dots, p$).

e_i - independent random error, with normal distribution with zero mean and variance σ^2

The concept of linearity and non-linearity is referred to the parameters. In this case, (x) is considered as a constant, and the dependence of

(Y) with the parameters is a combination of addition and subtraction. [Menchaca \(1990\)](#), when referring to linear models, points out that they present limitations from the practical and biological points of view, but they are extremely useful if the objective corresponds more to the field of estimation than to biological interpretation. In addition, these are very used by the simple mathematical method they utilize to estimate their parameters (method of ordinary least squares), which is accurate, and the solution is unique.

The general model can be taken to specific forms. Some of them are the first order polynomial, second order polynomial, and the like. Among the linear models that have shown good level of adjustment in applications within agricultural technical sciences, the cubic and polynomial logarithmic polynomial can be mentioned, whose fundamental characteristics and functional form are shown in [Table 1](#).

Non-Linear Models

Although linear models are suitable for many situations, some variables are not connected to each other by a simple relationship, such as the algebraic representation of the growth dynamics of individuals in animal species or the kinetics of postharvest weight loss. This has led to the search and construction of mathematical models, which have turned out to be non-linear in many cases.

The disadvantage of these models is that for the calculation of their parameters, iterative methods are used, and this leads to very complex calculations and can even fall into errors when estimating the parameters. Many authors use transformations to linearize the model and apply the minimum method square, but in practice, this

procedure is not advisable. At present, there are specific programs for non-linear adjustment, which allow overcoming this disadvantage.

Within the non-linear models frequently used are those that are asymptotic without inflection points (Models of Brody, Modified Exponential) and those that are asymptotic with inflection points (Models of Gompertz, Von-Bertalanffy, and Logistic, among others). [Table 2](#) shows the algebraic expressions of some of these models, as well as indicators associated with the calculation of speed and / or the rate of profit (first derivative), point of inflection.

In the Brody model, the function is concave and grows until its value stabilizes as the value of t increases; however, the velocities or rates (f l (x)) decrease as the value of t increases. ([Figure 1](#))

The shape of the rest of the models is sigmoid. The difference between them is the location of the inflection point. The speeds or rates in these models is flared; this increases until it reaches its maximum value, and then decreases to zero. ([Figure 1](#))

Application of These Models in Studies of Post-Harvest Curves of Pineapple (variety Cayena Lisa)

The study was carried out in areas of the company of several crops located within the Havana-Matanzas Plain, with a range of average annual temperature between 25 and 32 °C and high humidity. Weight Loss (PP) was performed through the weighing of fruits with the use of the electronic balance, during the first nine days after harvesting. Average values (g) were calculated on the days observed. The Logistics and Von Bertalanffy models (described above) were

TABLE 1. Summary of useful aspects of linear models

Model	Functional Form	First Derivative	Inflection Point
1	$Y_i = \beta_0 + \beta_1x_i + \beta_2x_i^2 + \beta_3x_i^3 + e_i$	$Y_i' = \beta_1 + 2\beta_2x_i + 3\beta_3x_i^2$	$x = -\frac{\beta_2}{3\beta_3} \pm \sqrt{\frac{\beta_2^2}{9\beta_3^2} - \frac{\beta_3}{\beta_3}}$
2	$y = \beta_0 + \beta_1x + \beta_2x^2 + \beta_3\ln(x) + e_i$	$y' = \beta_1 + 2\beta_2x + \frac{\beta_3}{x}$	$x = \sqrt{\frac{\beta_3}{2\beta_2}} \pm \sqrt{\frac{\beta_3}{2\beta_2}}$

1- Cubic polynomial 2-polynomial logarithmic

Y_i- Dependent variable.

β_j - Parameters of the model (j = 0,1,2, ..., p)

x_i - i-th value of the independent variable.

e_i - independent random error, with normal distribution, with zero mean and variance σ²

TABLE 2. Functional form and aspects related to non-linear models

	Brody Model	Logistic Model	Von Bertalanffy Model
Functional form	$f(t) = A(1 - be^{-kt})$	$f(t) = \frac{A}{(1 + be^{-kt})}$	$f(t) = A(1 - be^{-kt})^3$
Description	Asymptotic model without inflection point	Asymptotic model with inflection point	Asymptotic model with inflection point
Asymptote	A	A	A
Profit rate or first rate	$f'(t) = Abke^{-kt}$	$f'(t) = \frac{Abke^{-kt}}{(1 + be^{-kt})^2}$	$f'(t) = 3A(1 - be^{-kt})^2 bke^{-kt}$
Second derivative features	$f''(t) = -Ak^2be^{-kt}$	$f''(t) = \frac{Abk^2e^{-kt}(be^{-kt} - 1)}{(1 + be^{-kt})^3}$	$f''(t) = 3Abke^{-kt}(1 - be^{-kt})^2 + 2(3Abe^{-kt}(1 - be^{-kt})(bke^{-kt}))$
Inflection point	-	$t = \frac{\ln b}{k}$	$t = \frac{\ln 3b}{k}$

f (t): Dependent random variable, t: Independent variable, controlled or predefined, A, b, k: parameters to be estimated.

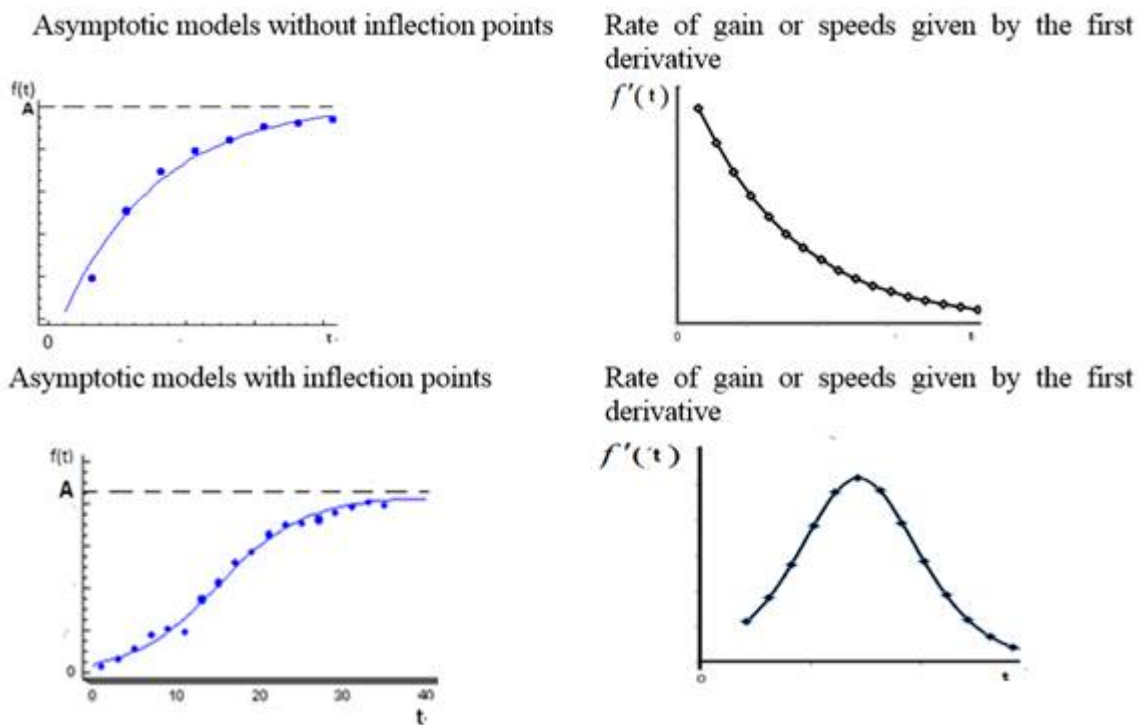


FIGURE 1. Representation of non-linear models and their corresponding rates or speeds tested. For the goodness of fit and discrimination between models, the coefficient of determination, sum of the square of the error, significance of the model and the parameters, distribution of the residuals and Durbin Watson's hypothesis test were taken into account. The loss process is described from the best model. The speed with which the weight loss occurs at each instant of the period analyzed is estimated from the derivative of the function, and the search of the moment where the maximum speed of weight

loss occurred was calculated by the inflection point.

RESULTS OBTAINED

The average values of weight loss (g) and the corresponding days allowed us to estimate the parameters of the adjusted curves (Table 3), which showed good adjustment with coefficient of determination above 97%. This indicates that both models are an alternative for the description of the process and for the prediction within the range of values studied.

The discrimination between the models allowed selecting the Logistic Model as the best one, which explains 99.27% of the total variability. All its parameters are significant and it achieved lower standard error of estimation, absolute mean of the error and sum of squared of the error than the model of Von - Bertalaffy (Table 3).

The Logistic model selected (Figure 2a) shows the increase in weight loss as the post-harvest day's increase, with abrupt behavior during the first days and a tendency to stabilize this value from the seventh day. This trend described by the model is considered adequate and is associated, among other aspects, with the irreversible physiological changes that occur in the fruit.

Another indicator of interest is associated to the behavior of the speed with which this weight loss occurs (Figure 2b), which is in flared form. This increases until reaching its maximum value,

approximately on the fourth day, and then decreases to zero.

Table 4 shows the inflection point (3.8, 2.24) which is where the maximum speed of weight loss is reached. It was of 1.07 g. It also shows the values reached around this point.

These indicators associated with the loss of weight, are of special interest to save resources, time and efforts, which allows tracing postharvest handling strategies and guarantee the quality of the product.

Currently, there are valuable results regarding the use of these and other types of models that have been used to predict the production of green and dry grass biomass (Maralfalfa). They have been studied with several doses of nitrogen fertilizer; or in the description and characterization of the physical and chemical properties of the pineapple (variety Cayenne Lisa) stored at room temperature. Moreover, they

TABLE 3. Result of the adjustment of the models

	Logistic Model	Von-Bertalanffy Model
Coefficient of Determination (R ²) %	99.27%	97.64%
Significance of the model	significant	significant
Adjusted model (estimated parameters)	$f(t) = \frac{4.57}{(1 + 36.25e^{-0.93t})}$	$f(t) = 4.98 (1 - 1.12 e^{-0.43t})^3$
Significance of the parameters	All significant	All significant
Standard estimation error	0.20	0.36
Sum of error square	0.16	0.52
Durbin Watson	2.38	2.19

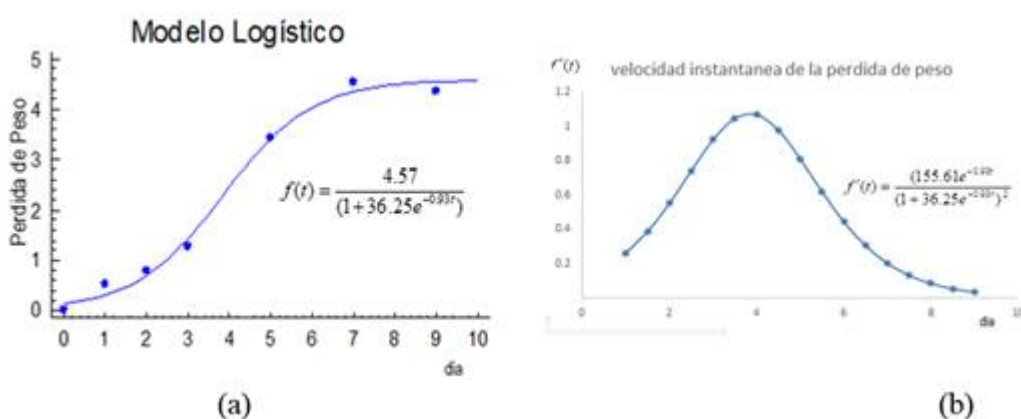


FIGURE 2. Model of best fit for weight loss (2a) and speed with which it occurs (2b).

TABLE 4. Behavior of the weight loss and the speed with which the weight loss occurs around the inflection point

Day	3.5	3.6	3.7	3.8	3.9	4	4.1	4.2
Weight Lossing (g)	1.927	2.032	2.138	2.245	2.352	2.459	2.565	2.669
Speed of PP (g)	1.046	1.060	1.068	1.073	1.072	1.067	1.058	1.044

have been used in the modeling and simulation of the star grass (*C. nlemfuensis*) performance under different management conditions and climatic scenarios (Fernández *et al.*, 2011; Rangel, 2015; López, 2016 and others).

CONCLUSIONS

- To deal in detail with the analysis of agrarian processes, knowledge of a group of models is required with the application of Differential Calculus and Mathematical Statistics, among other aspects.
- The graph of different models, as well as their absolute rates, help to clarify the analysis of the dynamics of different agrarian phenomena.
- Mathematical modeling allows deepening the laws that govern each of the phenomena, facilitates the study of the process and creates a methodological basis for the processing of information
- It is concluded that the development of Statistical-Mathematical Modeling allows responding to current problems of research in the agricultural sector and to the teaching and scientific-research challenges of the new University.

REFERENCES

- DEL POZO, P.P.; HERRERA, R.S.: "Modelado del crecimiento del pasto estrella (*Cynodon nlemfuensis*). 1. Modelo multiplicativo con control de la curva de crecimiento y los efectos ambientales", *Pastos y Forrajes*, 18(2): 171, 1995, ISSN: 0864-0394.
- DEL VALLE, J.: *La Multicolinealidad en Modelos de Regresión Lineal Múltiple. Propuesta de solución*, Universidad Agraria de La Habana (UNAH), PhD Thesis, San José de las Lajas, La Habana, Cuba, 2000.
- DEL VALLE, M.J.; GUERRA, B.W.: "La multicolinealidad en modelos de regresión lineal múltiple", *Revista Ciencias Técnicas Agropecuarias*, 21(4): 80-83, 2012, ISSN: 1010-2760, E-ISSN: 2071-0054.
- FERNÁNDEZ, C.L.; GUERRA, B.C.W.; DE CALZADILLA, P.J.; CHANG, L.N.U.: "Desarrollo de la modelación estadístico-matemática en las ciencias agrarias. Retos y perspectivas", *Investigación Operacional*, 38(5): 462-467, 2018.
- FERNÁNDEZ, L.: *Modelos que describen la dinámica de los procesos biológicos en las ciencias*, Universidad Agraria de La Habana, Cuba, MSc. Thesis, San José de las Lajas, La Habana, Cuba, 60 p., 1996a.
- FERNÁNDEZ, L.: "Modelos que describen la dinámica de los procesos biológicos en las Ciencias Agropecuarias", *La Habana, Cuba: Universidad Agraria de la Habana*, 1996b.
- FERNÁNDEZ, L.: *Modelos estadísticos-matemáticos en el análisis de la curva de lactancia y factores que la afectan en el genotipo Siboney de Cuba*, Universidad Agraria de La Habana, Instituto de Ciencia Animal, Veterinary Doctor Thesis., San José de las Lajas, La Habana, Cuba, 100 p., 2004.
- FERNÁNDEZ, L.; TONHATI, H.; ALBUQUERQUE, L.G.; ASPILCUETA-BORQUIS, R.R.; MENÉNDEZ BUXADERA, A.: "Modelos de regresiones aleatorias para la estimación de parámetros genéticos y estudios de curvas de lactancia del Holstein en Cuba", *Revista Cubana de Ciencia Agrícola*, 45(1), 2011, ISSN: 0034-7485.
- GUERRA, B.C.W.: *Relación Modelo-Diseño de tratamientos en la determinación de las dosis óptimas de fertilizantes con experimentos de Campo en Cuba*, Instituto de Suelos y Agroquímica "Nikola Pushkarov", PhD. Thesis, Sofia, Bulgaria, 184 p., 1980.
- LÓPEZ, J.L.: *Modelación y simulación del rendimiento del pasto estrella (*C. nlemfuensis*) bajo diferentes condiciones de manejo y escenarios climáticos*, Universidad Agraria de La Habana, Cuba, MSc. Thesis, San José de las Lajas, Mayabeque, Cuba, 79 p., 2016.
- MENCHACA, M.A.: *Modelo multiplicativo. Efecto de curva de lactancia controlado para el análisis estadístico de experimentos con vacas lecheras*, Instituto Superior de Ciencias Agropecuarias-Instituto de Ciencia Animal, PhD Thesis, San José de las Lajas, La Habana, Cuba, 1978.
- MENCHACA, M.A.: "El uso de modelos estáticos para describir las curvas de

crecimiento animal”, *Revista Cubana Ciencias Agrícolas*, 24(1): 29-34, 1990, ISSN: 0034-7485.

RANGEL, M.L.: *Variabilidad temporal de las propiedades físicas, químicas y mecánicas de la piña, variedad Cayena Lisa, almacenada a temperatura ambiente*, Universidad Agraria de La Habana, Cuba, MSc. Thesis, San José de las Lajas, Mayabeque, Cuba, 99 p., 2015.

RODRÍGUEZ, L.: *Modelación y simulación de la producción de biomasa de Pennisetum purpureum Schum vs. King Grass y su aplicación en la alimentación animal*, Universidad Agraria de La Habana, PhD Thesis, San José de las Lajas, Mayabeque, Cuba, 2015.

VÁZQUEZ, A.Y.; GUERRA, B.C.W.; SÁNCHEZ, L.O.E.: “Modelación estadístico-computacional para el estudio de la sostenibilidad socioeconómica de la empresa pecuaria Valle del Perú”, *Investigación Operacional*, 35(2): 121-130, 2014.

Lucía Fernández-Chuairéy, Profesor Titular, Universidad Agraria de La Habana (UNAH), Departamento de Matemática y Física, e-mail: lucia@unah.edu.cu

Lazara Rangel-Montes de Oca, Profesor Asistente, (UNAH), Departamento de Ingeniería Agrícola, e-mail: lazarar@unah.edu.cu

Caridad Walkiria Guerra-Bustillo, Profesor Titular, (UNAH), Departamento de Matemática y Física, e-mail: luzmi@infomed.sld.cu

Jany del Pozo-Fernández. Adiestrada, Facultad de Medicina Veterinaria, e-mail: janydelpozo@gmail.com

The authors of this work declare no conflict of interest.

This article is under license [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/)

The mention of commercial equipment marks, instruments or specific materials obeys identification purposes, there is not any promotional commitment related to them, neither for the authors nor for the editor.