*Volume 59, January-December 2025, ISSN: 2079-3480*

BIOMATHEMATICS

CUBAN JOURNAL OF AGRICULTURAL SCIENCE

Journal site: https://cjascience.com Cu-ID: https://cu-id.com/1996/v59e08

# ALTERNATIVE MODELS FOR DESCRIBING BIOLOGICAL VARIABLES WITH NON-NORMALITY AND NON-LINEAR PERFORMANCE. STUDY CASE

## MODELOS ALTERNATIVOS PARA DESCRIBIR VARIABLES BIOLÓGICAS CON FALTA DE NORMALIDAD Y COMPORTAMIENTO NO LINEAL. ESTUDIO DE CASO

YANEILYS GARCÍA ÁVILA*, MAGALY HERRERA VILLAFRANCA, R. RODRÍGUEZ HERNÁNDEZ, MILDREY TORRES MARTÍNEZ, YOLAINE MEDINA MESA

[1]*Instituto de Ciencia Animal, C. Central, km 47½, San José de las Lajas, Mayabeque, Cuba*

**Email: feygarciaa@ica.edu.cu*

Alternative models, which do not follow a normal distribution or have linear performance, were examined to describe biological variables such as *in vitro* gas production. Two analyses were performed: the first ignored the lack of normality of the response variable, while the second considered the non-compliance with this assumption. Parameter estimation was performed using the NLMIXED procedure. The following statistical criteria were used to select the model with the best goodness of fit: CME, $R^2$aj, AIC, BIC, and parameter significance. The fulfillment of the hypotheses of independence and randomness of the residual component was also considered. The tests to select the probability distribution of *in vitro* gas production showed that it has an exponential distribution (P>0.05). It is concluded that the logistic model with a "log" link function to estimate the population mean of *in vitro* gas production did not show adequate results. However, the segmental linear model was the one that best described this performance, revealing the best $R^2$aj, CME, AIC, BIC and visually random residuals. Also, with the segmental linear model similar results were obtained, regardless of the normality of the response variable. The results showed that in certain cases parametric procedures can be used with data that do not fulfill with normality. However, the consequences of such violations must be taken into account. A segmental linear model is proposed as an alternative to describe *in vitro* gas production when the data do not fulfill with normality.

***Key words:*** *data distribution, segmental linear model, statistical criteria*

Se examinaron modelos alternativos, que no siguen una distribución normal ni tienen un comportamiento lineal, para describir variables biológicas como es la producción de gas *in vitro*. Se realizaron dos análisis: el primero ignoró la falta de normalidad de la variable respuesta, mientras que el segundo consideró el incumplimiento de dicho supuesto. La estimación de los parámetros se realizó con el proc NLMIXED. Para la selección del modelo con mejor bondad de ajuste, se utilizaron los criterios estadísticos: CME, $R^2$aj, AIC, BIC y significación de los parámetros. También se consideró el cumplimento de las hipótesis de independencia y aleatoriedad de la componente residual. Las pruebas para seleccionar la distribución de probabilidad de la producción de gas *in vitro* indicaron que tiene distribución exponencial (P>0.05). Se concluye que el modelo logístico con función de enlace "log" para estimar la media poblacional de la producción de gas *in vitro* no mostró resultados adecuados. Sin embargo, el modelo lineal por segmento fue el que mejor describió dicho comportamiento, al dejar ver los mejores $R^2$aj, CME, AIC, BIC y residuos visualmente aleatorios. Además, con el modelo lineal por segmentos se obtuvieron resultados similares, sin importar la normalidad de la variable respuesta. Los resultados evidenciaron que en determinados casos se pueden utilizar procedimientos paramétricos con datos que no cumplen la normalidad. No obstante, se deben tener en cuenta las consecuencias de dichas violaciones. Se propone como alternativa un modelo lineal por segmentos para describir la producción de gas *in vitro*, cuando los datos no cumplen con la normalidad.

***Palabras clave:*** *criterios estadísticos, distribución de datos, modelo lineal por segmentos*

Received: November 25, 2024
Accepted: February 22, 2025
**Conflict of interest:** There is no conflict of interest between the authors in this study.
**CRediT Authorship Contribution Statement:** Yaneilys García Avila: Research, Conceptualization, Methodology, Supervision, Writing original draft. Mildrey Torres Martínez: Methodology, Supervision, Writing original draft. Magaly Herrera Villafranca: Methodology, Supervision, Writing original draft. R. Rodríguez Hernández: Writing original draft. Yolaine Medina Mesa: Writing original draft

## Introduction

In the scientific community, procedures for modeling linear and non-normal data are known. However, sometimes the variable under study may have non-linear performance and not be normally distributed. This case has occurred with biological growth variables, which are characterized by having three phases: 1) acceleration, 2) deceleration and 3) linear or asymptotic (Ortega-Monsalve *et al.* 2021). This performance can be described with non-linear models. However, another possibility could be the use of segmented linear regressions, where each segment is related to a growth phase.

Whitlock and Schluter (2009) suggested three possible alternatives to analyze biological variables that do not fulfill with normality, but they assume randomness and independence: 1) ignore the non-compliance with the premises, 2) transform the data and 3) use non-parametric methods. These include those that use the probability density function of the data distribution: the maximum likelihood and restricted maximum likelihood methods (Gomez-Mejia 2021). For its application, a function that links the population mean with the linear predictor of the observations is needed. The link function can be non-linear and varies depending on the probabilistic distribution to which the response variable is fitted, which must belong to the exponential family (Mesa-Fúquen *et al.* 2021).

Programs such as R and SAS allow studying models with variables that do not follow a normal distribution (Hernández *et al.* 2021). The SAS includes the GENMOD, GLIMMIX, and NLMIXED procedures. The latter is designed to handle functions dependent on general conditional means, whether they contain a linear component or not. However, the appropriate way to treat non-normal and non-linear variables is to use link functions that relate the population mean with the non-linear predictors (Bono *et al.* 2023). But will link functions, used to model generalized linear models (GLMs), be useful if the predictor is non-linear? Is it appropriate to model non-linear data using a segmented linear model that allows using the GLM procedures? Violating the assumption of normality can be an option? These questions lead to look for alternatives for modeling non-normal and non-linear biological variables through a case study. Hence, the objective of this study.

## Materials and Methods

*Experimental procedure*: Data were selected from an experiment conducted at the Instituto de Ciencia Animal (ICA) in 2018. The IVGP production technique proposed by Theodorou *et al.* (1994) was used. The IVGP was measured at 3, 6, 9, 12, 15, 18, 21, 24, 29, 48, 72, 77, and 144 h. The IVGP data of silage were used, with 50% OM-22, 50% moringa and *Lactobacillus pentosus*.

*Statistical analysis:* For the non-parametric statistical analysis, three elements proposed by Bandera and Pérez (2018) were used:

1. distribution function of the variable, each result of the dependent variable "Y" is generated from a particular distribution of the exponential family (normal, binomial, Poisson, gamma, among others)
2. a predictor η=(Xβ), which can be linear or non-linear, where Xβ is the predictor, a linear or non-linear combination of unknown parameters.
3. a link function g, such that E (Y) = μ = $g^{-1}$ (η), where:

   E (Y): expectation of the dependent variable.

   μ: mean of the dependent variable.

   $g^{-1}$ (η): inverse of the link function evaluated at the "linear" or "non-linear" predictor.

Table 1 shows the equations used to fit the experimental data. The logistic and linear models were used, as they are the most widely used in the agricultural field (García Ávila *et al.* 2022). Two analyses were performed: the first ignored the lack of normality of the IVGP variable, while the other considered the non-compliance with this assumption

**Table 1.** Mathematical models used to describe the kinetics of IVGP

| Models | Mathematical equation |
|---|---|
| Logistic Schofield *et al.* (1994) | $\text{IVGP(t)} = \frac{b}{1 + e^{(2 - 4c(t - L))}}$ |
| Segmentally linear | $\text{IVGP}(\text{for } t \leq T_1) = V_1 t + F_1;$ (Phase1) |
| | $\text{IVGP}(for \quad T_1 < t \leq T_2) = V_2 t + F_2;$ (Phase2) |
| | $\text{IVGP}(for \quad t > T_2) = b;$ (Phase3) |

IVGP: *in vitro* gas production at time t (mL $g^{-1}$OM incubated), c: IVGP rate ($h^{-1}$), t: fermentation time (h), L: Lag phase (h), b: asymptote when t -> ∞ (ml $g^{-1}$OM incubated), $t_1$, $t_2$ critical points that mark the beginning of the second and third phases of IVGP, respectively. $V_1$ and $V_2$ average velocities of the first and second phases; $F_1$ and $F_2$: approximate IVGP at the beginning of the first and second phases.

Taking into account the performance of the IVGP over time, a break point was considered equivalent to an inflection point. When a curve reaches an inflection point, it is because its concavity has changed. The IVGP is the point where the curve transitioned from one phase to another. The ProGas v1.1 program was used to determine these inflection points (García *et al.* 2022).

Statistical analyses were mainly performed using the SAS 9.3 (2013). The (PROC) UNIVARIATE procedure was used to select the most suitable probability distribution for the data. This procedure evaluated the normal, exponential, and Weibull distributions. The Cramer-von Mises and Anderson-Darling goodness-of-fit tests were used to select the distribution (Zetina-Moguel *et al.* 2018). The estimation of the parameters in the generalized models was performed with PROC NLMIXED. This method uses likelihood estimation techniques. The "log" link function was applied because it is the most commonly used with exponential distributions (Bandera and Pérez 2018). The presence of independence or autocorrelation was tested using the Durbin-Watson (DW) test, according to Rozo (2017). The non-parametric Streaks test for the residuals was performed using the IBM-SPSS statistical program, version 22. This test allowed contrasting the hypothesis of a random ordering *versus* a trend alternative.

The selection of the model with the best fit was made based on the mean square error (MSE), the adjusted coefficient of determination ($R^2$aj), the Akaike and Bayesian information criteria (AIC and BIC, respectively) (Montoya and Quiroz 2021) and the significance of the parameters. Models with lower values MSE, AIC, and BIC were considered better fit, while higher values for $R^2$aj were preferred. The fulfillment of independence and randomness hypothesis of the residual component were also considered.

## Results and Discussion

When evaluating IVGP data with Cramer-von Mises and Anderson-Darling tests, it was observed that the data can be exponentially distributed (table 2). Authors such as Zetina-Moguel *et al.* (2018) used the Kolmogorov-Smirnof, Cramer-von Mises and Anderson Darling goodness of fit tests. In addition, the AIC and BIC information criteria to select the best distribution that their data followed. Finally, they considered that one of the best selection criteria is Anderson Darling's test. In this study, the two tests had similar results.

To fit the segmental linear model, the inflection points were calculated. Using the ProGas v1.1 program, it was determined that the IVGP acceleration phase developed during the first 18 h ($t_1$=18 h), time the maximum IVGP speed was reached. After this time (t>18 h), the deceleration phase began, which lasted until 48 h ($t_2$=48 h) to begin the linear or asymptotic phase, in which the GPIV stabilized.

Table 3 shows the results after modeling the data considering normal and exponential distribution with a "log" link. The fermentation of IVGP produced an asymptotic value of IVGP, which ranged between 121.03 and 139.06 mL.g$^{-1}$OMinc. In models assuming normal IVGP, all parameters were significant (P<0.05). However, when considering that the IVGP had an exponential distribution, the Lag phase (L) and the mean deceleration velocity ($V_2$) were not significant (P>0.05). The nonparametric fit made difficult to identify the Lag and deceleration phases.

The experimental data showed that after three hours of incubation there was little IVGP, with values below 2 mL.g$^{-1}$OMinc, which indicated the existence of a lag phase, where the microorganisms colonize or hydrate the substrate. Also, after 18 h, all models estimated deceleration rate $V_2$<1.58 mL.g$^{-1}$OMinc/h. Authors such as Solís *et al.* (2023) found that the Lag phase of whole-grain sweet potato mixtures was 2.5 to three hours, while the deceleration phase began after 12 hours.

The evaluation of the logistic model showed problems of residual independence in both analyses (table 3). Autocorrelation, especially in longitudinal data, could be due to that variables that really affect the model have not been included or that the appropriate model has not been chosen (Gómez and Agüero 2020). Failure to comply with the assumption of independence of errors is a limitation that can lead to biased estimates (Pérez Pelea 2018). In this case, the Runs test for randomness showed that the residuals were random, which was contradicted with the residuals graph (figure 1), which showed a trend and lack of randomness. The correlation of the residuals with the nonparametric logistic model may be due to that the "log" link function was not adequate. It is necessary to search for new link functions for nonlinear predictors. However, little is discussed in the literature because no published information on link functions for nonlinear models was found.

**Table 2.** Goodness-of-fit tests to select the probability distribution that best describes the IVGP variable

| Probability distribution function for IVGP | Cramer-von Mises test (P value) | Anderson-Darling test (P value) |
|---|---|---|
| Normal | P=0.024 | P=0.006 |
| Exponential | P=0.056 | P=0.050 |
| Weibull | P<0.010 | P<0.010 |

**Table 3.** Modeling of the data obtained from the experiment where the nutritional value of silage with 50 % OM-22, 50 % moringa and *Lactobacillus pentosus* was evaluated

| Model | Parameters | $R^2$aj % | CME | AIC | BIC | DW | Runs P value |
|---|---|---|---|---|---|---|---|
| Logistic | b= 133.11±2.24, P<0.0001 | 92.8 | 54.38 | 275 | 281 | 1.2 | 0.004 |
| (Normal distribution) | c=0.04±0.003, P<0.0001 | | | | | | |
| | L=9.37±0.78, P<0.0001 | | | | | | |
| Segmentally linear | $V_1$=3.2±0.23, P<0.0001 | 98.9 | 32.06 | 250 | 260 | 1.7 | 0.37 |
| $t_1$=18 | $V_2$=1.58±0.14, P<0.0001 | | | | | | |
| $t_2$=48 | $F_1$=-14.0±2.74, P<0.0001 | | | | | | |
| (Normal distribution) | $F_2$=40.72±4.53, P<0.0001 | | | | | | |
| | b=139.06±5.92, P<0.0001 | | | | | | |
| Logistic | b=121.03±1.29, P<0.0001 | 86.4 | 73.79 | 377 | 389 | 0.62 | 0.0005 |
| (Exponential distribution and log link function) | c= 1.05±0.01, P=0.0029 | | | | | | |
| | L=1.7±20.01, P=0.8601 | | | | | | |
| Segmentally linear | $V_1$=1.26±1.05, P<0.0001 | 98.8 | 42.09 | 372 | 380 | 1.5 | 0.0981 |
| $t_1$=18 | $V_2$=1.02±1.03, P=0.54 | | | | | | |
| $t_2$=48 | $F_1$=1.00±1.79, P=0.99 | | | | | | |
| (Exponential distribution) | $F_2$=51.84±2.46, P<0.0001 | | | | | | |
| | b=139.04±1.39, P<0.0001 | | | | | | |

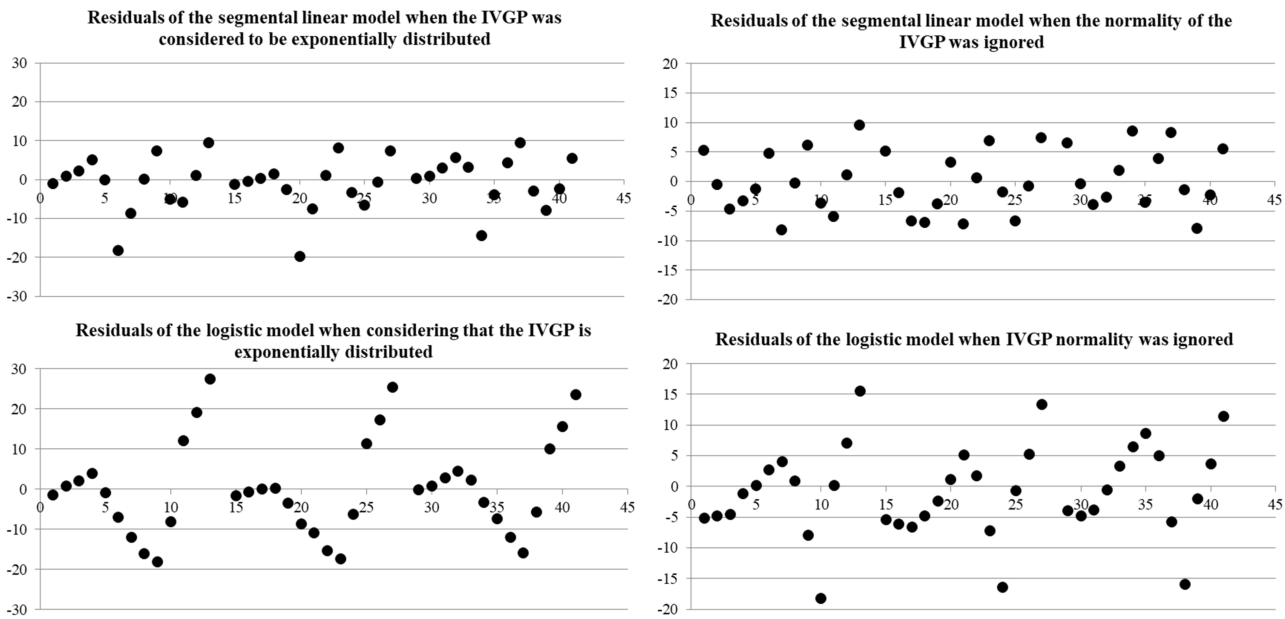


**Figure 1.** Residues of the models when the IVGP of silage was evaluated with 50 % OM-22, 50 % moringa and *Lactobacillus pentosus* and the assumption of normality was violated or not

The same table shows problems with residual independence. Gómez and Agüero (2020) suggest that, in longitudinal data, autocorrelation could be due to that the model has not included variables that actually affect it or that the appropriate model has not been selected. Pérez Pelea (2018) refers that the non-compliance of the assumption of independence of errors is a limitation that can lead to biased estimates.

The Streak test showed the randomness of the residuals with probability values above 0.05. However, the results in figure 1 contradicted this test by showing a trend and lack of randomness. The results obtained with the non-parametric logistic model could be due to that the proposed "log" link function was not adequate. It is necessary to search for new link functions for nonlinear predictors. It should be noted that little is reported in the scientific literature about other link functions that relate to different distributions. In the review carried out, there was not information on link functions for the case of nonlinear models.

When normality and the "log" link function were ignored for exponentially distributed data, the results were similar (table 3). However, with segmental linear models, the $R^2aj$ were higher and the CMEs were lower. The DW test did not have any problems in fulfilling the assumption of residual independence for these models.Also, the segmental models produced favorable residuals with a random point cloud and no pattern (figure 1). The Streaks test rejected the hypothesis of random performance for the residuals. It is necessary to highlight that the Wald and Wolfowitz (1940) Runs test takes into account the number of runs and their length, ignoring that the runs can be concentrated in specific intervals of the sample, as was the case with the logistic model fitted in this study (figure 1). In this case, it was assumed to have lost power.

The segmental linear model is an option that allows knowing the average speed of phases 1 and 2 of IVGP. In addition, it can be used to estimate the Lag phase by setting the first stage equation to zero: $0 = V_1 t + F_1$; where $Lag = F_1/V_1$. Linear equations also estimate the IVGP with which the second phase begins. Knowing the relation between $V_1/V_2$ is useful for researchers because it allows them to better understand the performance of the substrate under study. Using the phase 3 equation, the asymptotic IVGP can be determined.

Figure 2 shows the performance of the models fitted under the parametric and non-parametric approaches. It was observed that the maximum IVGP estimated by both linear models was 140 mL.g$^{-1}$OMInc at approximately 80 h. The segmental linear model best described the IVGP of silage with 50 % OM-22, 50 % moringa and *Lactobacillus pentosus*. It is incorrect to assume that data are normal when statistical tests show opposite. According to Pérez Pelea (2018), a significant deviation from the premises can seriously increase the researcher's chances of committing a type I or type II error, depending on the nature of the analysis, which implies inaccurate results and incorrect interpretations in statistical tests (Zhou Kimbeng 2010). Since there are no definitive instructions on how to act in each case, a study of the variable must be done before performing any test. When distributions are very far from normal, have outliers, or the distributions of the groups to be compared are very different and have very heterogeneous variances, the failure to meet the assumptions should not be ignored. In these cases, an alternative approach should be used, such as transforming the scale of the variable, use a non-parametric simulation method, or a generalized linear model.

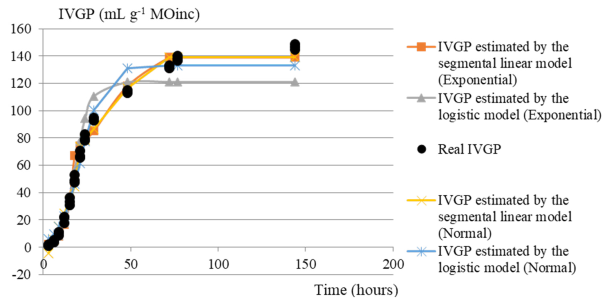


**Figure 2.** Contrast between the actual IVGP and that estimated by the models when the IVGP of silage was evaluated with 50 % OM-22, 50 % moringa and *Lactobacillus pentosus*

## Conclusions

It is concluded that the logistic model with the "log" link function to estimate the population mean of the IVGP did not show adequate results. However, the segmental linear model was the one that best described this performance, showing the best $R^2aj$, CME, AIC, BIC, and visually random residuals. In addition, with the segmental linear model, similar results were obtained regardless of the normality of the response variable. The results showed that in certain cases parametric procedures can be used with data that do not comply with normality. However, the consequences of such violations must be taken into account. A segmental linear model is proposed as an alternative to describe the IVGP when the data do not comply with normality.

## References

Bandera, E. & Pérez, L. 2018. Los modelos lineales generalizados mixtos. Su aplicación en el mejoramiento de plantas. *Cultivos tropicales*, 39(1): 127-133, ISSN: 1819-4087. https://ediciones.inca.edu.cu/index.php/ediciones/article/view/1437/2302.

Bono, R., Alarcón, R., Arnau, J., García-Castro, F.J. & Blanca, M.J. 2023. Robustez de los Modelos Lineales Mixtos Generalizados para diseños Split-Plot con datos binarios. *Anales de Psicología*, 39(2): 332-343. ISSN: 1695-2294. https://doi.org/10.6018/analesps.527421.

García, Y., Torres, M. & Rodríguez, R. 2022. ProGas v1.1: Programa para el pre-procesamiento y análisis de datos de producción de gas *in* vitro de alimentos para rumiantes. Nota técnica. *Cuban Journal of Agricultural Science*, 56(2): 105-109, ISSN: 2079-3480. https://www.cjascience.com/index.php/CJAS/article/view/1050/.

García Avila, Y., Herrera Villafranca, M., Rodríguez Hernández, R., & Ontivero Vasallo, Y. 2022. Evaluación de modelos no lineales y no lineales mixtos para describir la cinética de producción de gas *in vitro* de alimentos para rumiantes. *Cuban Journal of Agricultural Science*, 56(1): 1-9, ISSN: 2079-3480. https://www.cjascience.com/index.php/CJAS/article/view/1040/.

Gómez, M.A. & Agüero, Y. 2020. Ajuste de modelos mixtos no lineales para la descripción de curvas de lactación bovina bajo pastoreo en El Mantaro, Junín, Perú. *Revista Investigaciones Veterinarias del Perú*, 31(4): 19027, ISSN: 1609-9117. http://dx.doi.org/10.15381/rivep.v31i4.19027.

Gomez-Mejia, A. 2021. Modelo de máxima verosimilitud. *Libre Empresa*, 17(2): 121-138, ISSN: 1657-2815. https://doi.org/10.18041/1657-2815/libreempresa.2020v17n2.8027.

Hernández, A.Á., García-Munguía, C.A., García-Munguía, A.M., Valencia-Posadas, M., Ruiz, J.H. & Velázquez-Madrazo, P.A. 2021. Tipificación y caracterización del sistema de producción del cerdo criollo de la Región Centro, México. *Ecosistemas y Recursos Agropecuarios*, 8(2): 37, ISSN: 2007-901X. https://doi.org/10.19136/era.a8n3.2777.

Ortega-Monsalve, M., Velásquez-Henao, A.M., Ortiz-Acevedo, A., Galeano-Vasco, L.F. & Medina-Sierra, M. 2021. Ajuste a un modelo matemático, comparación de las curvas de crecimiento y características morfológicas de cuatro *Urochloas* de una colección *in vivo* establecida en Antioquía, Colombia. *Revista de Investigaciones Veterinarias del Perú*, 32(5): 1-7, ISSN: 1609-9117. https://doi.org/10.15381/rivep.v32i5.19678.

Mesa-Fúquen, E., Hernández, J.S. & Camperos, J.E. 2021. Uso de modelos lineales generalizados en el conteo de Leptopharsagibbicarina (Hemiptera: Tingidae) en palma de aceite. *Revista Colombiana de Entomología*, 47(1): 2-5, ISSN: 2665-4385. https://doi.org/10.25100/socolen.v47i1.7661.

Montoya, E.A.F. & Quiroz, A.B. 2021. Un Enfoque Bayesiano en Modelos Heterocedásticos de Series de Tiempo y su Aplicación en la Volatilidad de Activos Financieros.*Pesquimat*, 24(2): 1-12, ISSN: 1609-8439. https://doi.org/10.15381/pesquimat.v24i2.21152.

Pérez Pelea, L. 2018. ¿Cómo proceder ante el incumplimiento de las premisas de los métodos paramétricos? o ¿cómo trabajar con variables biológicas no normales? *Revista del Jardín Botánico Nacional,* 39: 1-12, ISSN: 2410-5546. https://www.researchgate.net/publication/327752027.

Rozo, A.J. 2017. La educación secundaria y sus dos dimensiones. Efectos del barrio y del colegio sobre los resultados saber 11. *Revista de Economía del Rosario*, 20(1): 33-69, ISSN: 2145-454X. http://dx.doi.org/10.12804/revistas.urosario.edu.co/economia/a.6148.

SAS Institute Inc. 2013.SAS/IML 9.3 User's Guide.SAS Institute Inc., Cary, NC. URL http://www.sas.com/

Schofield, P., Pitt, R.E. & Pell, A.N. 1994.Kinetics of fiber digestion from *in vitro* gas production. *Journal of Animal Science*, 72(11): 2980-2991, ISSN: 1525-3163. https://doi.org/10.2527/1994.72112980x.

Solís, C., Ruiloba, M.H., Rodríguez, R. & Marrero, Y. 2023. Dinámica de la fermentación ruminal *in vitro* de la mezcla integral de camote (*Ipomoea batata, l.*) presecada y ensilada. *Revista Investigaciones Agropecuarias*, 5(2): 88-96, ISSN: 2644-3856. https://revistas.up.ac.pa/index.php/investigaciones_agropecuarias/article/view/3899.

Theodorou, M.K., Williams, B.A., Dhanoa, M.S., McAllan, A.B. & France, J. 1994. A simple gas production method using a pressure transducer to determine the fermentation kinetics of ruminant feeds. *Animal Feed Science and Technology*, 48: 185-197, ISSN: 0377-8401. https://doi.org/10.1016/0377-8401.

Wald, A. & Wolfowitz, J. 1940. On a Test Whether Two Samples are from the Same Population. *The Annals of Mathematical Statistics*, 11(2): 147-162, ISSN: 0003-4851. http://dx.doi.org/10.1214/aoms/1177731909.

Wang, T., He, P., Ahn, K.W., Wang, X., Ghosh, S. & Laud, P. 2015. A re-formulation of generalized linear mixed models to fit family data in genetic association studies. *Frontiers in Genetics*, 6: 120, ISSN: 1664-8021. http://dx.doi.org/10.3389/fgene.2015.00120.

Whitlock, M.C. & Schluter, D. 2009. The Analysis of Biological Data. Roberts and Company Publishers, Greendwood Village, Colorado, USA.

Zetina-Moguel, C.E., Sánchez y Pinto, I., González-Herrera, R., Osorio-Rodríguez, J.H., Barceló- Quintal, I.D. & Méndez-Novelo, R.I. 2018. Modelación estocástica del nivel freático en pozos de la ciudad de Mérida, Yucatán, México. *Ingeniería*, 22(2): 25-35, ISSN: 1665-529X. https://www.redalyc.org/journal/467/46758579003/html/.

Zhou, M.M. & Kimbeng, C.A. 2010. Multivariate repeated measures: A statistical approach for analyzing data derived from sugarcane breeding variety trials. *Proceeding sof the South African Sugar Technologists' Association*, 83: 92-105, ISSN: 0370-1816. https://www.cabidigitallibrary.org/doi/pdf/10.5555/20113349222.