Sensibility analysis of homogeneity tests of *in vitro* gas production curves by Monte Carlo simulation

O. Jay[1], Verena Torres[2], Yoandra Marrero[2] and J. P. Torres[3]

[1]*Centro de Estudios de Especies Menores, Universidad de Guantánamo, Carretera al Salvador km 6½, Guantánamo, Cuba*
[2]*Instituto de Ciencia Animal, Apartado Postal 24, San José de las Lajas, Mayabeque, Cuba*
[3]*Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara, Las Agujas, Zapopan, Jalisco, México*
Email: ojay@ica.co.cu

In order to improve the ruminant feed evaluation techniques through a statistical method permitting the comparison of treatments in experiments of *in vitro* gas production, based on regression curves, data were collected from an experiment of *in vitro* gas production. The Gompertz function was fitted and the Monte Carlo simulation method was used, with the aim of analyzing the sensibility of four homogeneity tests of non-linear regression models. Thirteen new treatments were obtained, with 300 repetitions each and 280 points simulated per each repetition. The homogeneity tests were: extra sum of squares, Bayesian information criterion (BIC), Akaike information criterion (AIC), and Akaike information criterion corrected (AICC). The comparison criteria were assessed with modifications to the parameters, on the order of 1, 3, 5, and 10 %. The methodology for the comparison of the curves was based on the definition of the complete and the reduced models. The sensibility of each method was established according to the probability of fulfillment of the null hypothesis. The AICC showed higher sensibility, followed by the test of extra sum of squares. The AIC and the BIC were the least sensible. The parameter that affected the most the differences between treatments was the asymptotic coefficient (A).

Key words: *extra sum of squares, information criterion, non-linear regression, comparison of treatments*.

In vitro gas production techniques in ruminant feed evaluation are of great interest to animal nutrition research, due to their low cost and because they are invasive. Besides, they admit large amount of experimental units (Rodríguez 2010). The development of statistical software has permitted the incorporation of non-linear models in the analysis of the results (López *et al*. 2007). These models permit the theoretical interpretation of the physiological process by means of the relationship of the mathematical equations with the experimental data. At the same time, they manage to quantify the influence of the parameters that characterize the degradability of the feed in the rumen.

Among the potentials for the use of modeling, there is the statistical comparison of the treatments with the corresponding experimental designs, taking as criterion of comparison the dynamics of feed degradation described by the regression curve. However, despite the advances in this branch of statistics, no deep research has been conducted about the homogeneity tests of the regression curves applicable to the experiments.

It is of interest the obtainment of a correct statistical method for these comparisons, because it would permit the conduction of new researches to find out new forms of profitable and sustainable management in ruminant feeding. As a result, the available resources would be better used, with the consequent increase in milk and meat yields.

The object of this research was to assess the sensibility of four homogeneity tests of models of non-linear regression through the simulation of *in vitro* gas production curves, with the application of the Monte Carlo method modifying the values of the parameters. Also, it was analyzed the influence of each parameter on the performance of the curve and on the determination of the significant differences between the treatments.

Materials and Methods

The original data came from an experiment of *in vitro* gas production, performed by the Physiology Department in the Institute of Animal Science in Cuba and the Laboratories of Molecular Microbiology and Chemistry of the Program of Animal Physiology and Nutrition of the Colombian Corporation of Agricultural Research (Marrero 2005).

The gas measurements were performed for 24 h, at intervals of 20.4 min, with four repetitions and 70 measurements per repetition. Also, it was considered, that at the start of the measurements (t=0), the gas production (GP) was null, for a total of 284 experimental points. The model of Thornley and France (2007) was considered for the non-linear regression (figure 1). Thus, the substrate was a compartment with two fractions, one degradable and the other non-degradable. The other compartment was the gas produced, at a production rate inversely proportional to the gas produced previously.

The graphic was represented mathematically through the differential equation:

$$\frac{dPG}{dt} = -\mu\,(t)\,PG \qquad (1):$$

whose solution was the function of Gompertz:

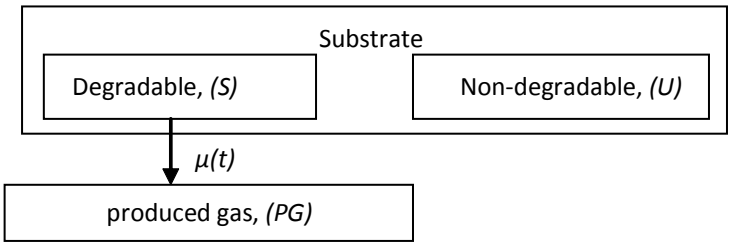

Figure 1. Bicompartimental layout of gas production

$PG = A\ exp\ (-be^{-ct})$ (2)

Where,

GP: gas production

A, = Asymptotic parameter, characterizes the maximum gas amount produced in the experiment.

b and c = Parameters related to the curve of the model. They characterize the speed of gas production and thus, the degradation of the substrate.

t = Time.

For the verification of the model adequacy, the significance of the regression and of the parameters, the coefficient of determination, and the value of the residual mean square were considered.

The simulation was performed according to Rubinstein and Kroese (2008):

1. The non-linear regression of the experimental data was fitted to obtain the estimators of the parameters and the original residues.
2. Simulated functions were constructed with the same original regression equation. It varied the value of a parameter and the rest was left constant.
3. The simulated function was assessed in each point of the variable, regardless the original function, and the value of a residue of the original regression was added through random selection.
4. With these simulated points the regression was adjusted again to obtain the simulated curve.

The order of the variation of the simulation parameters (1, 3, 5, and 10 %) was performed according to the results of Calabro *et al*. (2005) and García-Rodríguez *et al*. (2005), where the interval of confidence of the estimation of the parameters of regression was between 1 and 10 % of the absolute value of the parameter for experiments of *in vitro* gas production. These parameters were established in a way that, for the first four simulations, the value of the parameter A varied, on the order of 1, 3, 5, and 10 %. The rest of the parameters was left constant. For the following four, the parameter b was varied on the same order, and in the last four the value of the parameter c was modified on the same order.

The function for the simulation with the variations in the parameters was:

$PG^*_{k,i} = A^*_k\ exp\ (-b^*_k\ e^{-c_k t_i}) + \varepsilon^*_{k,i}$, (3)

Where,

k = Amount of simulated treatments (k=1, 2,…,12)

i = Amount of points per curve (i=1, 2,…, 84)

$GP^*_{k,i}$ = i-th point of the k-th simulation.

ti = Time corresponding to the i-th observation in the original experiment.

$\varepsilon^*_{k,i}$ = i-th component of the vector of simulated residues, obtained from a random selection with replacement of the original residues.

A^*_k, b^*_k, c^*_k: k = -th parameters of simulation, which were modifications of the originals, as shown in table 1.

The comparisons were performed between the treatment zero and the rest. Three hundred repetitions were conducted for each comparison between treatments.

For the comparison, the method of the complete and the reduced models was used, as described by Schabenberger *et al*. (1999) and Motulski and Christopoulos (2003).

Being the function of the regression model for the treatment k, F_k and for h, F_h, the hypothesis test would be as follows:

$$\begin{cases} H_0: F_k = F_h \\ H_1: F_k \neq F_h \end{cases}, \ k \neq h \quad (5)$$

The reduced model was declared for the fulfillment of the null hypothesis, considering that all the points belong to only one curve:

$M_{Reduced} = f(A_0, b_0, c_0, t) + \varepsilon$ (6)

The other expression, corresponding to the fulfillment of the alternative hypothesis, was denominated complete model. In it, the points were adjusted, as belonging to different curves for each treatment. A compound function

Table 1: Layout of calulation of the simulation parameters per treatment

SP	Treatments												
	0	1	2	3	4	5	6	7	8	9	10	11	12
A*	A	1.01A	1.03A	1.05A	1.1A	A	A	A	A	A	A	A	A
b*	b	b	b	b	b	1.01b	1.03b	1.05b	1.1b	b	b	b	b
c*	c	c	c	c	c	c	c	c	c	1.01c	1.03c	1.05c	1.1c

S.P. simulation parameters

was obtained as a result.

$$M_{Complete} = \begin{matrix} f(A_k, b_k, c_k, t) + \varepsilon \text{ Treatment } k \\ F(A_h, b_h, c_h, t) + \varepsilon \text{ Treatment } i \end{matrix} \quad (7)$$

The goodness of fit of the complete and the reduced models constituted a measurement of the fulfillment of one hypothesis or the other. There are statistical tests that permit the researcher to reject the null hypothesis or not.

Homogeneity tests of the curves:

1. Test of the extra sum of squares (Schabenberger *et al*. 1999), performed through the F-test, as follows:

$$F = \frac{SC_{Reduced} - SC_{Complete} / df_{Reduced} - df_{Complete}}{SC_{Complete} / df_{Reduced}} \quad (8)$$

Where,

SC = Sum of residual squares of the regression of the complete and reduced models.

df = Degrees of freedom of the corresponding sums of residual squares.

The calculated F was compared with the test of Fisher, with the degrees of freedom in the numerator and in the denominator dfReduced-dfComplete and dfComplete, respectively.

2. The Bayesian information criterion (BIC) (Konishi *et al*. 2004), was expressed as:

$$BIC = N \ln \left(\frac{SC}{N}\right) + K \ln (N) \quad (11)$$

Where,

N = Number of data

K = Number of parameters of the equation of fit to the regression plus one

SQ = Sum of squares of the residues

3. The Akaike information criterion (AIC), given by Motulsky and Christopoulos (2003), was defined by the expression:

$$AIC = N \ln \left(\frac{SC}{N}\right) + 2K \quad (9)$$

4. The Akaike information criterion corrected (AICC), with the correction of Hurvish and Tsai (1989), proposed originally for situations with few experimental points for the comparison of the models included the number of points in the correction:

$$AICC + \frac{AIC + 2K(K+1)}{N - K - 1} \quad (10)$$

The difference between the values of the information criteria, corresponding to the complete and the reduced models, was denoted by the Greek letter Δ:

Δ= *information criterion of the complete model – information criterion of the reduced model* (12)

The probability of fulfillment of the null hypothesis was established by the equation:

$$P = \frac{e^{-\frac{\Delta}{2}}}{1 + e^{-\frac{\Delta}{2}}} \quad (13)$$

The value of Δ and P were calculated in all the information criteria.

The theoretical decision criterion for the rejection of H_0 was given by:

$$P = \frac{e^{-\frac{\Delta}{2}}}{1 - e^{-\frac{\Delta}{2}}} < 0.05 \Rightarrow \Delta_{0.05} > 6.09 \quad (14)$$

The correspondence was analyzed between the border of the theoretical decision of each homogeneity test and that estimated through the simulation.

Sensibility test.

The sensibility of the homogeneity test of the regression curves was defined as the capacity of detecting significant differences ($P < 0.05$) in comparisons between treatments with values closed to the parameters. Thus, the comparison was made between the simulated curves of the treatment 0 (with the parameters of simulation A_0, b_0, c_0) and the rest.

It was considered that the method was efficacious in the comparison, when there were significant differences in more than 285 instances in the 300 repetitions. All the statistical analyses were conducted with the software SPSS 15.0 (2006) and Microsoft Excel 2007 (2008).

Results and Discussion

According to Calabro *et al*. (2005) and López *et al*. (2007), the Gompertz function, applied to *in vitro* gas production experiments, generally, brings about good criteria of fit. This was corroborated in this experiment, where the regression was significant, at a level of $P < 0.01$. The residual mean square had value of 0.715, with 280 degrees of freedom. The value of the coefficient of determination (R^2) was of 0.970. The estimated parameters were also significant and are shown in table 2.

Table 3 shows the parameters used in the modeling (4):

F-test of the extra sum of squares. Table 4 reports the performance of the test of the extra sum of squares for each of the comparisons.

With this test, there were significant differences in more than 285 instances, when the simulation parameters A and c varied in 5 % or more (A vs 1.05A and A vs 1.10A; c vs 1.05c and c vs 1.10c). For the parameter b, there were significant differences in more than 285 cases, only when the parameter varied in 10 % (b vs 1.10b). In these comparisons, the varied parameter was higher than the upper limit of estimation of the original parameters.

An interesting case occured in the comparison, where the parameter A varied on the order of 3 %. The calculated F-test was next to the decision border (2.62 and 2.15 respectively), but not enough. This was in correspondence with the number of differences detected, which represented 93.3 % of the total, not permitting in this comparison to reject the null hypothesis.

Figure 2 shows the performance of the value of the F-test, in respect to the increase in the differences between the

Table 2. Estimations of the parameters of the original regression

Parameter	Estimation	S. E.(±)	Interval of confidence 95%	
			Lower limit	Upper limit
A	17.970	0.414	17.155	18.786
b	3.212	0.088	3.038	3.386
c	0.130	0.005	0.119	0.140

Table 3. Values of the parameters for the simulation.

S.P.	Treatments												
	0	1	2	3	4	5	6	7	8	9	10	11	12
A*	17.97	18.15	18.51	18.87	19.77	17.97	17.97	17.97	17.97	17.97	17.97	17.97	17.97
b*	3.21	3.21	3.21	3.21	3.21	3.24	3.31	3.37	3.53	3.21	3.21	3.21	3.21
c*	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.131	0.134	0.137	0.143

S.P. Simulation parameters

Table 4. Performance of the test of extra sum of squares for the simulations

Parameter of simulation that varies	Variation of the parameter	Number of significant differences (P < 0.05)
A	1.01	71
	1.03	280
	1.05	300*
	1.10	300*
b	1.01	13
	1.03	96
	1.05	184
	1.10	300*
c	1.01	36
	1.03	169
	1.05	299*
	1.10	300*

*Indicates the efficacy of the test (P < 0.05).

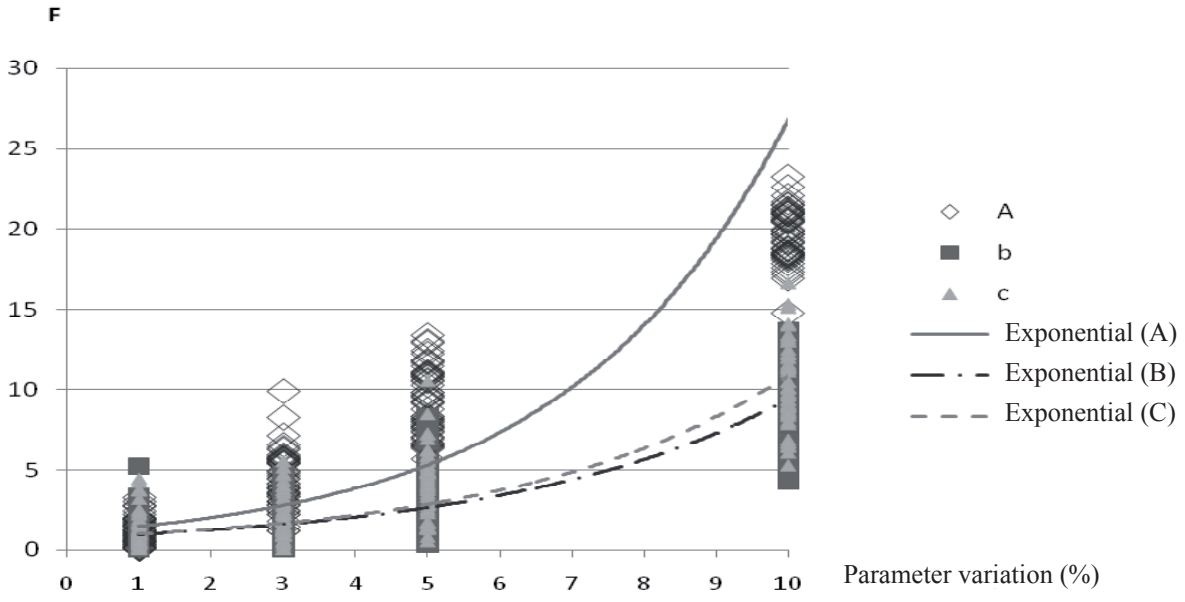

Figure 2. Values of the F-test and their tendency for each comparison.

compared treatments, given in percentage of variation of the simulation parameter. This had, approximately, an exponential behavior. For the parameter A, the rate of variation of the F-test was much higher than that of b and c, which was in relation to the significance of the parameter in the regression function. As the difference between the parameters of the regression function was increased, it increased the dispersion of the values of the test.

Table 5 shows the results corresponding to the BIC. For the parameter A, the test was able of finding significant differences in more than 285 cases, starting from the variation of the 5 % of the comparison parameter (A vs 1.05A). For the parameters b and c, it only reached this value for variations starting from 10 % (b vs 1.10b and c vs 1.10c). When the comparison was performed between treatments, with values next to the simulation parameters, the method did not detect almost any significant difference. This performance was more evident for the parameter b, for which it was fully inefficacious in the first three comparisons.

The dependence of the values of Δ, in respect to the variation of the parameters, was around that of a polynomial of second order (figure 3). In this instance, as for F, the dispersion of the values was increased as the difference between the parameters was incremented.

For differences up to 1 % in the estimations of A, up to 3 % in b, and up to 5 % in c, there were values of Δ lower than zero. According to the relationship expressed in 10, the value of Δ was lower than zero, when the inequality was fulfilled:

$$(SCR_{reduced})^N < (SCR_{complete})^N N^k \qquad (15)$$

When the values of the sums of squares of the complete and the reduced models were next one to the other, the difference between the BIC was negative.

These results showed that the sensibility of the Bayesian information criterion was inferior to the test of the extra sum of squares.

Table 5. Performance of the BIC for the parameters variations

Parameter of simulation that varies	Variation of the parameter	No. of significant differences (P<0.05)
A	1.01	1
	1.03	76
	1.05	292*
	1.10	300*
b	1.01	1
	1.03	1
	1.05	6
	1.10	295*
c	1.01	1
	1.03	9
	1.05	194
	1.10	300*

*Indicates the efficacy of the test (P < 0.05).

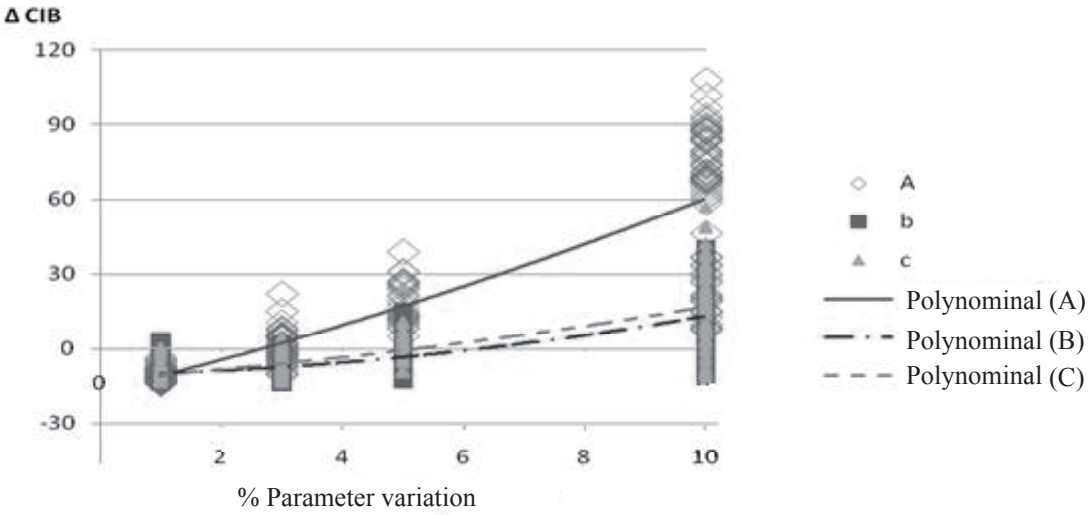

Figure 3. Values of ΔBIC and their trend for each comparison

Table 6. Performance of the criteria of information of Akaike and Akaike corrected for the variations of the parameters

Parameters	Differences. (%)	Variation of the parameter	No. of significant differences (P<0.05)	
			CIA	CIAC
A	1	1.01	17	104
	3	1.03	238	294*
	5	1.05	300*	300*
	10	1.10	300*	300*
b	1	1.01	3	30
	3	1.03	29	144
	5	1.05	98	214
	10	1.10	300*	300*
c	1	1.01	11	64
	3	1.03	107	225
	5	1.05	293*	299*
	10	1.10	300*	300*

*Indicates the efficacy of the test (P < 0.05).

Table 6 provides simultaneously the performance of the AIC and AICC. The two criteria were shown at once due to their similarity as to their mathematical nature, with emphasis on the correction by the number of experimental points.

The measurements of the Δ corresponding to the AIC and the AICC in each comparison were higher than the BIC, due to the factor accompanying the logarithm of the sum of residual squares was numerically higher. Nevertheless, there were negative values of Δ, corresponding to the fulfillment of the inequality derived from the equation (10):

$$\text{Ln}\left(\frac{SCR_{complete}}{SCR_{reduced}}\right) < \frac{2k}{N} \qquad (16)$$

The ΔAIC detected significant differences in more than 285 instances, comparisons in which the variation in the parameter A was of 5 % or higher, for variations of 10% of the parameter b, and of 5 % of the parameter c.

Significant differences were detected for variations in the simulation parameters, similar to those of the test of extra sum of squares. Therefore, it can not be stated that the sensibility of this test was higher.

The performance of the dependence of the values of ΔAIC in respect to the variation of the parameters was approximately that of polynomial of second order, just as in the ΔBIC. There was also, for this criterion, a higher dispersion of the values as the difference between the parameters of the simulation was increased, as shown in figure 4.

The number of differences found by the ΔAICC was higher than 285, as compared with the treatments where the simulation parameter A varied in 3 % or more, in respect to the variation of 10 % of the parameter b, and of 5% of the c. Likewise, the values of the means of Δ were slightly superior to the value of the theoretical

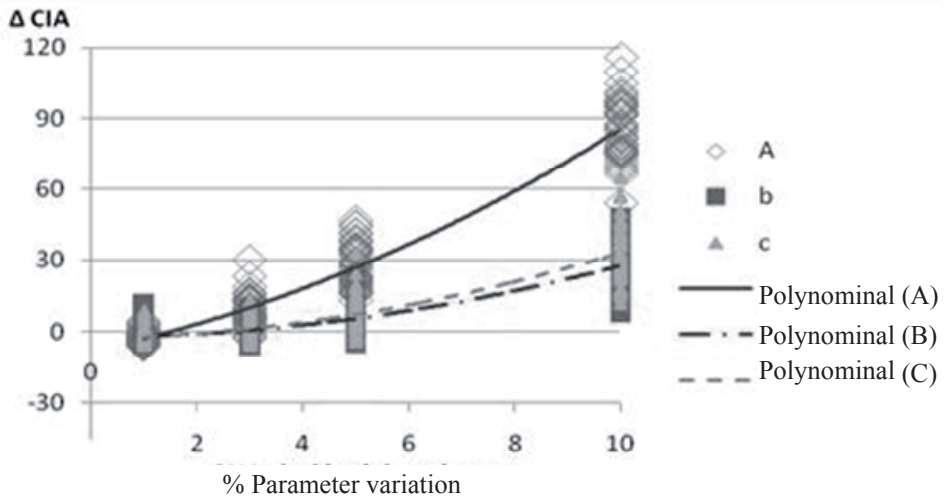

Figure 4. Values of ΔAIC and their trend for each comparison.

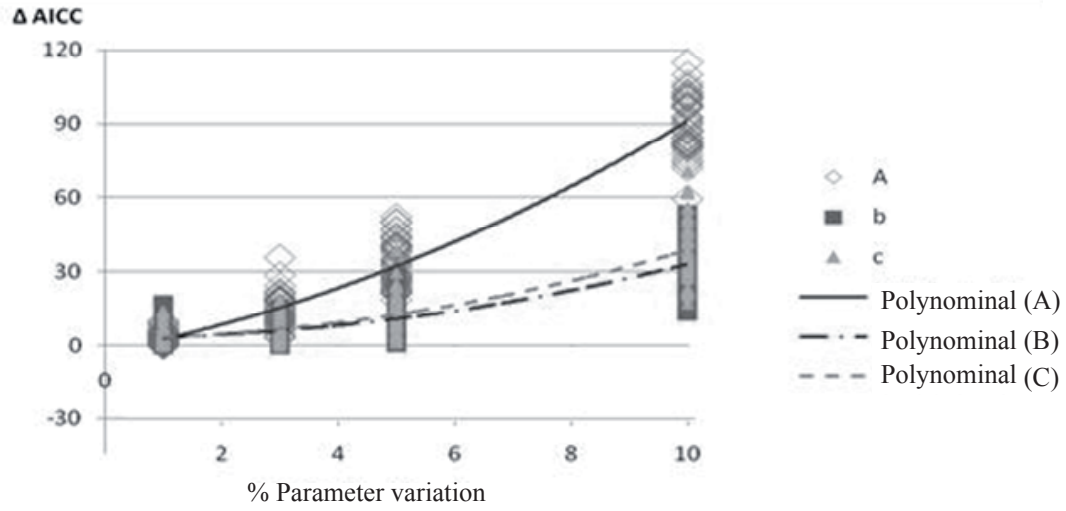

Figure 5. Values of ΔAICC and their trend for each comparison

decision border, even where the number of significant differences superior to 285 was not attained. This was due to the inclusion of the correction term (Hu 2007).

The performance of this criterion showed that the correction term was important in the comparison of the models, even when the number of experimental points was relatively large (Burham and Anderson 2004).

In all the comparisons, the number of detections of the AICC was higher than that of the F-test, the BIC, and the AIC. The test performance in respect to the variations in the treatments was quadratic, similar to the rest of the information criteria. The variation rate was higher, as a consequence from being the test of higher sensibility. Also, the dispersion was the lowest of all the information criteria, as the difference between treatments was increased.

The parameter A was the one affecting the most the sensibility, which represented the curve asymptote and showed the total amount of gas produced in the experiment as a result from the linear dependence of the gas production with this parameter. The other parameter of greatest influence was the c, indicator of the gas production rate, also as a consequence from the position occupied by this parameter in the Gompertz function. The case of the b, was the one with the lowest influence on the sensibility, because in respect to this parameter, the dependence of the gas production was logarithmical.

Considering the results, it was concluded that the most sensible test for the comparison of the treatments was the AICC, followed by the F-test of the extra sum of squares, and the AIC.

Although according to Schabenberger and Pierce (2002), the test of the extra sum of squares should have the same sensibility as to the information criteria, the fact of not being the most sensible can be explained due to the high dependence of the test on the strict fulfillment of the assumptions in the regression. As the experimental data got farther from the ideal distribution, the F-test may lose sensibility (Seber and Wild 2003).

It was concluded that the Akaike information criterion corrected is an efficacious tool for the comparison of treatments, taking as basis the *in vitro* gas production curves. This method constitutes a powerful tool to be applied in the different experimental designs used in research on this branch of science.

References

Burnham, K.P. & Anderson, D.R. 2004. Multimodel inference. Understanding AIC and BIC in model selection. Soc. Methods Res. 33:261

Calabro, S., Lopez, S., Piccolo, V., Dijkstra, J., Dhanoa, M.S. & France, J. 2005. Comparative analysis of gas production profiles obtained with buffalo and sheep ruminal fluid as the source of inoculum. Animal Feed Sci. Tech. Pp. 123-124

García-Rodríguez, A., Mandaluniz, N., Flores, G. & Oregui, L.M. 2005. A gas production technique as a tool to predict organic matter digestibility of grass and maize silage. Animal Feed Sci. Tech. 123:267

Hu, S. 2007. Akaike Information Criterion. Conferencia impartida en North Carolina State University. Center Res. Sci. Computing

Hurvish, C.M. & Tsai, C. L. 1989. Regression and time series model selection in small samples. Biometrika. 76: 297

Konishi, S., Ando, T. & Imoto, S. 2004. Bayesian information criteria and smoothing parameter selection in radial basis function networks. Biometrika 91:27

López, S., Dhanoa, M.S., Dijkstra, J., Bannink, A., Kebreab, E. & France, J. 2007. Some methodological and analytical considerations regarding application of the gas production tecnique. Anim. Feed Sci. Tech. 135:139

Marrero, Y. 2005. Las levaduras como mejoradores de la fermentación ruminal de dietas con alto contenido de fibra. PhD Thesis. La Habana. 151 pp.

MicroSoft_Corporation. 2008. Microsoft Office Excel 2007 SP 2.

Motulsky, H.J. & Christopoulos, A. 2003. Fitting models to biological data using linear and nonlinear regression. Graph Pad Software. San Diego. 351

Rodríguez, R. 2010. Estudio *in vitro* del valor nutritivo y de los efectos antinutricionales de cuatro leguminosas arbóreas tropicales con potencialidades como suplementos del

Pennisetum purpureum. PhD Thesis in Animal Production and Feed Scince. Zaragoza. 110

Rubinstein, R.Y. & Kroese, D.P. 2008. Simulation and the Monte Carlo Method. John Wiley & Sons Inc. New Jersey. 377

Schabenberger, O. & Pierce, F.J. 2002. Contemporary statistical models. CRC Press. Boca Raton. 195

Schabenberger, O., Tharp, B.E., Kells, J.J. & Penner, D. 1999. Statistical test for hormesis and effective dosages in herbicide dose response. Agron. J. 91:713

Seber, G.A.F. & Wild, C.J. 2003. Nonlinear regression. John Wiley & Sons. New York. 512.

Thornley, J.H.M. & France, J. 2007. Mathematical models in agriculture. CAB International. Oxon. 889.

Received: September 7, 2011