Comparison of Nonlinear Models for Dry Matter Yield of Brachiaria Hybrid cv Cayman in Drought Season by Bootstrapping

1Danny Villegas Rivas, 2Nora Valbuena Torres, 3Manuel Milla Pino, 4Zadith Garrido Campaña, 5Martín Grados Vasquez, 6Erick Delgado Bazan, 7Cesar Osorio Carrera, 8Ydalia Velasquez Casana and 9Wilfredo Ruiz Camacho

1Facultad de Ingeniería Forestal y Ambiental, Universidad Nacional de Jaén, Cajamarca, Perú
2Programa de Ciencias del Agro y del Mar, Universidad Nacional Experimental de los Llanos Occidentales Ezequiel Zamora, Guanare, Venezuela
3Facultad de Ingeniería Civil, Universidad Nacional de Jaén, Cajamarca, Perú
4Escuela de Postgrado, Universidad César Vallejo, Perú

Abstract: The dynamics of biomass accumulation at different seasons of the year is of vital importance in the production of a forage species. In this study nonlinear models (logistic and gompertz) to describe growth dynamics of plants based on dry matter yield of Brachiaria hybrid Cayman at different interval between cuts in drought season were used. Results for logistic and gompertz model showed an excellent fit to the data set of accumulated dry matter with a R² close to 1.0. In the dry season the best fitted model to estimate the cumulative dry matter yield was gompertz model. It is possible suggest logistic model use in drought conditons. The empirical distribution of the estimated intrinsic growth rate (K) was symmetric and leptokurtic for the logistic and gompertz model in the intervals (-0.0141; 0.2071 Kg MS ha⁻¹) and -0.00086; 0.724986 Kg MS ha⁻¹), respectively and the empirical distribution of the asymptotic accumulated dry matter yield (B) was asymmetric for the logistic and gompertz models in the intervals (4373; 16980 Kg MS ha⁻¹) and (9840; 26654 Kg MS ha⁻¹), respectively. The empirical distribution of these two parameters reflects the fact that the accumulated dry matter production of this grass is strongly related to the distribution of rainfall, especially in the tropical dry forest. The nitrogen fertilization, interval between cuts, intrinsic growth rate of B. hybrid Cayman as well as other factors affect this species development determine dry matter yield sustainability.

Keywords: Pasture, Growth Curve, Entropy, Bootstrapping

Introduction

Knowledge of the biomass accumulation dynamics of a forage species at different seasons is a useful methodology for better planning and harvesting of the crop to obtain the highest yields and plant material of good nutritional quality (Montes Cruz et al., 2016). For forage yield have been used various species and varieties of creeping or erect growth habit, including Brachiaria brizantha it is of great interest to know their growth dynamics in different edaphoclimatic conditions in order to establish mechanisms that help to its best use and management. In this way, obtaining non-linear models can be very helpful to explain and predict pastures and forages behavior in relationship certain factors such as cutting age. Studies that is limited in the geographical area (Verdecia et al., 2012). Gompertz models (Laird, 1965), logistic (Nelder, 1961; Richards, 1959; Von Bertalanffy, 1957; Brody, 1945) are the most frequently used to describe the plants growth, animals and organisms. In this sense, knowledge and control of growth and development of the crops are great utility parameters for researchers, since their characterization allows they can be seen besides enabling the management programs to be designed for growth inherent to each species. The main objective of this work is to study the dynamics of dry

© 2021 Danny Villegas Rivas, Nora Valbuena Torres, Manuel Milla Pino, Zadith Garrido Campana, Martin Grados Vasquez, Erick Delgado Bazan, Cesar Osorio Carrera, Ydalia Velasquez Casana and Wilfredo Ruiz Camacho. This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.
matter yield of *Brachiaria* hybrid cv. Cayman in drought using non-linear models.

**Materials and Methods**

**Study Data**

Data was obtained from a trial carried out to study the grasses management with different interval between cuts in a farm located in Potrguesa, Venezuela located between 450570.307 west longitude 985346.992 north latitude, with an annual rain average of 1.847.3 mm. Plant material used is certified commercially embedded seed of *Brachiaria* hibrido cv. Cayman. The treatments consisted interval between cuts: 21, 28, 35 and 42 days under scheme of a completely randomized block design. The response variate to be considered for the model is the total biomass. The indicator describes this variate is kilograms of dry matter per hectare (kg MS * ha−1). The covariate of model was interval between cuts. Two non-linear models were considered; gompertz (Laird, 1965) and logistic (Nelder, 1961) to study the relationship between cutting age and yield. The statistical analyzes were carried out in the R programming environment.

According to (Seber and Wild, 1988), the Gompertz model is defined as follows:

\[ f(x) = c + (d − c) \times \exp(-\exp(b(x−e))) \]

It is a response/growth curve across the true axis, that is, it is not limited to non-negative values even though this is the range for most response and growth data.

If *b*<0 the mean function increases, while it decreases for *b*>0.

In practice, several reparametrizations of the model have been carried out.

According to (Bruce and Versteeg, 1992) the logistic model is defined as follows:

\[ f(x) = c + \frac{d − c}{1 + \exp(b(x−e))} \]

**Selection Criteria based on Information Measures**

In this research, in addition to the widely known criteria to evaluate the goodness of fit of the models, such as; the coefficient of determination \((R^2)\) and the residual standard Error (EER), the information criteria or entropy indices.

**Akaike Information Criteria (AIC)**

This criterion is detailed in (Garay-Martínez et al., 2018), who point out that if the problem consists of selecting the coefficients *β* that are as close as possible to the vector \(\beta^*\), the distance between the distributions \(f(Y/\beta^*)\) and \(f(Y/\beta)\) can be characterized by an entropy measure of the form (Akaike, 1978):

\[
D(\beta^*, \beta) = \int \left[ f(y/\beta^*) \ln f(y/\beta) - f(y/\beta^*) \ln f(y/\beta^*) \right] dy
\]

(where the first addend of the second member represents the ability to fit of \(f(Y/\beta)\) with respect to \(f(Y/\beta^*)\) and the second addend, for a given function \(f(Y/\beta^*)\), is a constant).

The minimization of the entropy measure implies the minimization of the information criterion (Kullback, 1959):

\[
KL(\beta^*, \beta) = -D(\beta^*, \beta)
\]

\[
= \int \left[ \ln f(y/\beta^*) - \ln f(y/\beta) \right] f(y/\beta^*) dy
\]

Assuming that \(\beta = \beta^* + \Delta \beta\) (where \(\Delta \beta = [\Delta \beta_1, \Delta \beta_2, ..., \Delta \beta_k]^T\) is an arbitrary norm vector small), then the criterion \(KL(\beta^*, \beta)\) admits a Taylor series expansion of the form:

\[
KL(\beta^*, \beta + \Delta \beta) \approx \sum_{k} \frac{\partial^2 \log f(y/\beta^*)}{\partial \beta_k^2} \Delta \beta_k^2 + ... + \sum_{k, l} \frac{\partial^2 \log f(y/\beta^*)}{\partial \beta_k \partial \beta_l} \Delta \beta_k \Delta \beta_l
\]

If \(f(y/\beta)\) is a regular function, the first term of the second member of this expression vanishes and, consequently, it follows that \(KL(\beta^*, \beta + \Delta \beta) \approx \frac{1}{2} [\Delta \beta]^T \Sigma [\Delta \beta]\) (where \(\Sigma = \frac{\partial^2 \log f(y/\beta^*)}{\partial \beta \partial \beta} \) is the information matrix of Fisher). Suppose that \(\beta\) is included in a *s*-dimensional space \(\Theta(1,2, ..., k-1)\), while the vector of the true values of the coefficients, \(\beta^*\), is included in a *k*-dimensional space (\(k > s\)). Denoting by \(\beta^*_k\) the projection of \(\beta^*\) on \(\Theta_0\) in the sense of the Euclidean norm, it is shown that \(2KL(\beta^*, \beta) = [\beta_k - \beta^*_k] + [\beta - \beta^*_k]^T\) (where \(\beta_k \in \Theta_0\) and it is verified that \(\beta_k \approx \beta^*_k\)).

Replacing \(\beta_k\) by the vector of random variables \(\hat{\beta}_k\), formed by the restricted maximum-likelihood estimators of \(\beta^*_k\) in \(\Theta_0\) and, taking into account that, for values of \(n\) that are sufficiently large, \(n[\beta_k - \beta^*_k]^T d \chi^2_n\), it is verified that \(2E[KL(\beta^*, \hat{\beta})] \approx [\beta_k - \beta^*_k]^T + \frac{\chi^2_n}{n}\). This expression constitutes a measure of the deviations of \(\hat{\beta}_k\) with respect to the vector \(\beta^*_k\) and allows to conclude that the expected value of this deviation includes a component that represents the error related to the selection of a coefficient space approximated by \(\hat{\beta}_k\) and another which
represents the error due to the estimation of the vector of the coefficients. Akaike showed that, under certain conditions of regularity, the likelihood ratio:

\[ LR(Y) = -2\sum_{i=1}^{n} \log \left[ \frac{f(y_i, \hat{\beta})}{f(y_i, \hat{\beta}_{MV})} \right] d\hat{\beta}^{2}\sum_{i=1}^{n} (\hat{\beta}_i - \beta_i)^2 \]

And therefore, that \( \frac{1}{n} LR(Y) + 2s - k \) is an unbiased estimator of the measure \( E\left[ KL(\beta - \hat{\beta}) \right] \). The Akaike Information Criterion (AIC) consists of minimizing the logarithm of the likelihood function \(-2L_{\theta}(Y, \hat{\beta}) + 2s (s = 1, 2, ... , k-1)\) in which the first term represents the measure of the error due to the lack of capacity to adapt to the approximation and the second term defines the penalty factor. Under the assumption of normality of the assumed true model, its density function assumes the form:

\[ f(Y) = \frac{1}{(\sigma_n \sqrt{2\pi})^{n-p}} \exp \left[ -\frac{1}{2\sigma_n^2} \sum_{i=1}^{n} (Y_i - \hat{\beta}_i)^2 \right] \]

and the likelihood function of the candidate model \( \hat{Y}_c \) will be of the form. Therefore, the Kullback-Leibler distance will assume the form:

\[ f(\hat{Y}) = \frac{1}{(\sigma_n \sqrt{2\pi})^{n-p}} \exp \left[ -\frac{1}{2\sigma_n^2} \sum_{i=1}^{n} (\hat{Y}_i - \hat{\beta}_i)^2 \right] \]

Therefore, the Kullback-Leibler distance will assume the form:

\[ KL = \frac{2}{n-p} E \left[ \ln \left( \frac{f(Y)}{f(Y')} \right) \right] \]

\[ = \ln \left( \frac{\sigma_{\hat{\beta}}}{\sigma_{\hat{\beta}'}^2} \right) + \frac{1}{n-p} \sum_{i=1}^{n} \left[ m(\hat{Y}_i - \hat{\beta}_i) - \hat{\beta}_i - \cdots - \hat{\beta}_i \right] - 1. \]

Thus, substituting in this expression the coefficients \( \hat{\beta}_i \) and \( \sigma^2_{\hat{\beta}} \) by their maximum-likelihood estimators, we obtain that:

\[ KL = \ln \left( \frac{\sigma_{\hat{\beta}}^{MV12}}{\sigma_{\hat{\beta}}^{MV2}} \right) + \frac{\sigma_{\hat{\beta}}^{MV2}}{\sigma_{\hat{\beta}}^{MV12}} - \frac{L_2}{\sigma_{\hat{\beta}}^{MV12}} - 1 \]

From this definition the following selection criteria results:

\[ AIC(p) = \ln \left( \sigma_{\hat{\beta}}^{MV12} \right) + \frac{2(p+1)}{n-p} \]

which allows obtaining an asymptotically efficient estimator \( \hat{p} = \min AIC(p) \).

**Bootstrapping Estimation**

In addition to the maximum likelihood estimators of the parameters of the nonlinear models considered in this investigation, the estimation was performed using the bootstrap method proposed by (Efron, 1979), which is one of the simplest methods used to obtain an estimator of a parameter \( \hat{\beta} = \beta(P) \) where \( P \) is the postulated statistical model. Alonso (2001) presents the bootstrap method in a general situation.

Let be \( Z = (Z_1, Z_2, ..., Z_n) \) a data set generated by the statistical model \( P \) and let be \( T(Z) \) the statistic whose distribution \( L(T; P) \) we wish to estimate. The bootstrap method proposes as an estimator of \( L(T; P) \) the distribution \( L(T^*; n P^*) \) of the statistic \( T^* = T(Z^*) \), where \( Z^* \) is a data set generated by the estimated model \( \hat{P} \). Note that if \( \hat{P} = P \), then the distributions \( L(T; P) \) and \( L(T^*; \hat{P}) \) coincide. Then if we have a good estimator of \( P \), it is logical to suppose that \( L(T^*; \hat{P}) \) it will approach \( L(T; P) \).

The models described above, their estimators (EMV and Bootstrap) and the model selection criteria (AIC and EER) were determined in the R environment, using the “drc” package and the “boot” package (R Core Team, 2020). For details see Appendix 1.

**Results and Discussion**

The results of the fit two non-linear models to cumulative dry matter (Kg MS ha\(^{-1}\)) in Brachiaria hybrid cv Cayman pastures at different times between cuts in dry season (Table 1, 2 and Fig. 1) showed an excellent fit to the data set of accumulated dry matter with a \( R^2 \) close to 1.0 for the logistic and gompertz models. However, gompertz model shows the lower value of AIC (99.712) compared to the obtained values by the logistic model (AIC = 180.958, Table 1). Results indicated above suggest in the dry season (Table 2) the best fit model to estimate the cumulative dry matter yield (Kg MS ha\(^{-1}\)) in Brachiaria hybrid cv Cayman at different intervals between cuts is gompertz model. These results match with was indicated by (Rodríguez et al., 2011) in a trial to estimate the growth dynamics of Pennisetum purpureum cv. Cuba CT-169, who report that growth classic models gompertz and logistic for the rainy season are those of better fit for dry matter accumulation and plant height. Likewise, (Torres and Herrera, 2010; Rodríguez et al., 2011; 2013) in the modeling of cumulative dry matter yield of King grass or its clones concluded gompertz and logistic models were the best fit. In those papers are considered functions allowed to
estimate the biomass production as a time function. Likewise, these results coincide with reported by (Villegas et al., 2019) in a trial to estimate dry matter production in Brachiaria brizantha at different interval between cuts and under nitrogen fertilization.

On the other hand, in Table 2 it is observed that highest value of the inflection point ordinate (75 days), which is obtained for a cutting age of 42 days. This suggests that under this experimental condition the accumulated dry matter yield of Brachiaria hybrid cv Cayman is in a constant manner for a longer time period than the rest of the experimental conditions in this trial. In this sense, determining optimal moments of cut or harvest based on values obtained in the growth dynamics of the species, in a particular place and climate allows to maximize forage yield greater, obtaining forage of higher nutritional quality in comparison to the cuts that are made when the forage is already dry (Castro et al., 2017).

Table 1: Fit of a logistic model to estimate accumulated dry matter yield (Kg MS ha⁻¹) in Brachiaria hybrid cv Cayman at different interval between cuts (days) during the dry season

<table>
<thead>
<tr>
<th>Interval between cuts (days)</th>
<th>Age at the inflection point (t) (days)</th>
<th>Autocorrelation</th>
<th>Model fitting</th>
<th>Estimated parameters of the model and its significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dw</td>
<td>P value</td>
<td>R²</td>
<td>AIC</td>
</tr>
<tr>
<td>21</td>
<td>61</td>
<td></td>
<td>99.30</td>
<td>246.72</td>
</tr>
<tr>
<td>28</td>
<td>64</td>
<td></td>
<td>99.49</td>
<td>205.14</td>
</tr>
<tr>
<td>35</td>
<td>69</td>
<td></td>
<td>99.78</td>
<td>149.83</td>
</tr>
<tr>
<td>42</td>
<td>80</td>
<td></td>
<td>99.88</td>
<td>151.05</td>
</tr>
<tr>
<td>Mean</td>
<td>2.6052</td>
<td>99.61</td>
<td>188.15</td>
<td>8843</td>
</tr>
</tbody>
</table>

Table 2: Fit of a gompertz model to estimate accumulated dry matter yield (Kg MS ha⁻¹) in Brachiaria hybrid cv Cayman at different interval between cuts (days) during the dry season.

<table>
<thead>
<tr>
<th>Interval between cuts (days)</th>
<th>Age at the inflection point (t) (days)</th>
<th>Autocorrelation</th>
<th>Model fitting</th>
<th>Estimated parameters of the model and its significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dw</td>
<td>P value</td>
<td>R²</td>
<td>AIC</td>
</tr>
<tr>
<td>21</td>
<td>54</td>
<td></td>
<td>99.45</td>
<td>241.83</td>
</tr>
<tr>
<td>28</td>
<td>56</td>
<td></td>
<td>99.61</td>
<td>200.89</td>
</tr>
<tr>
<td>35</td>
<td>66</td>
<td></td>
<td>99.78</td>
<td>205.56</td>
</tr>
<tr>
<td>42</td>
<td>75</td>
<td></td>
<td>99.88</td>
<td>200.89</td>
</tr>
<tr>
<td>Mean</td>
<td>2.6053</td>
<td>99.68</td>
<td>199.58</td>
<td>11152</td>
</tr>
</tbody>
</table>

Fig. 1: Fit of two nonlinear models for estimate accumulated dry matter yield (Kg MS ha⁻¹) in Brachiaria hybrid cv Cayman at different interval between cuts (days) during the dry season.
Therefore, this result suggests that the cut age, the intrinsic rate of growth of *Brachiaria* hybrid cv Cayman as well as other factors that affect the development of this forage species determine sustainability in dry matter yield of this species, which coincides with what was pointed out by Bello, 2014 in a theoretical description of nonlinear models, especially of the gompertz and logistic models in relation to factors that affect the ordinate \( X \) of the inflection point. Moreover, the growth dynamics determine the phenological behavior of the crop at different year seasons which varies depending on the environmental conditions that arise (Montes Cruz et al., 2016) so that the growth and quality of the pastures it can vary considerably according to the management to which they are subjected, with favorable or non-favorable effects according to the plant species and the edaphoclimatic conditions (del Pozo Rodríguez, 1998). This coincides with that reported by (Cruz-Hernández et al., 2017), who in a trial with *B*. hybrid cv Mulato at different frequency and intensity of grazing concluded that during rainy season the pasture presents greater forage accumulation of forage when harvested in rest periods of 28 days.

In Table 3 and Fig. 2 and 3 show that the empirical distribution of the estimated intrinsic growth rate \( K \) is symmetric and leptokurtic for the logistic and gompertz models in the intervals \((-0.0141; 0.2071 \text{ Kg MS ha}^{-1})\) and \((-0.00086; 0.724986 \text{ Kg MS ha}^{-1})\), respectively. However, the empirical distribution of the asymptotic accumulated dry matter yield \( B \) is asymmetric for the logistic and gompertz models in the intervals \((4373; 16980 \text{ Kg MS ha}^{-1})\) and \((9840; 26654 \text{ Kg MS ha}^{-1})\), respectively. This behavior of the empirical distribution of these two parameters reflects the fact that the accumulated dry matter production of this grass is strongly related to the distribution of rainfall, especially in the tropical dry forest. In this sense, (Baruch and Fisher, 1991) suggest that the most variable climatic factor in the tropical area is precipitation and its distribution throughout the year, which has a significant impact on the annual and seasonal production of forage biomass. Hence, annual forage production occurs during the season of highest rainfall (Garay-Martínez et al., 2018). However, any analysis of the effect of moisture availability on the growth of grasses should not be isolated from the type of soil and the genetic potential of the plants, since among the latter there is great variability regarding their tolerance to drought stress.

**Table 3:** Bootstraping estimation of accumulated dry matter yield (Kg MS ha\(^{-1}\)) in *Brachiaria* hybrid cv Cayman at interval between cuts (42 days) during the dry season by a gompertz and logistic model

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Bootstrap estimated coefficient</th>
<th>Bootstrap confidence interval Lower limit</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>( K )</td>
<td>0.0266</td>
<td>-0.0141</td>
<td>0.2071</td>
</tr>
<tr>
<td></td>
<td>( A )</td>
<td>13260</td>
<td>4373</td>
<td>16980</td>
</tr>
<tr>
<td></td>
<td>( B )</td>
<td>8.446</td>
<td>-20.7269</td>
<td>30.195</td>
</tr>
<tr>
<td>Gompertz</td>
<td>( K )</td>
<td>0.0136</td>
<td>-0.00086</td>
<td>0.724986</td>
</tr>
<tr>
<td></td>
<td>( A )</td>
<td>16860</td>
<td>9840</td>
<td>26654</td>
</tr>
<tr>
<td></td>
<td>( B )</td>
<td>2.772</td>
<td>0.34568</td>
<td>6.22579</td>
</tr>
</tbody>
</table>

**Fig. 2:** Empirical bootstrapping distribution of estimated intrinsic growth rate \( K \) and the asymptotic accumulated dry matter yield \( A \) in *Brachiaria* hybrid cv Cayman at interval between cuts (42 days) during the dry season by a logistic model
Fig. 3: Empirical bootstrapping distribution of estimated intrinsic growth rate (K) and the asymptotic accumulated dry matter yield (A) in Brachiaria hybrid cv Cayman at interval between cuts (42 days) during the dry season by a gompertz model

**Conclusion**

In the drought period, the best fit model to estimate the accumulated dry matter yield in Brachiaria hybrid cv Cayman at different interval between cuts was the gompertz model. Both models, logistic and gompertz tend to underestimate the initial dry matter yield. The gompertz model presented a higher growth speed than the logistic model. The empirical distribution of the estimated intrinsic growth rate is symmetric and leptokurtic for the logistic and gompertz models. The empirical distribution of the asymptotic accumulated dry matter yield is asymmetric for the logistic and gompertz models. Finally, the inflection point in logistic and the gompertz models are conditioned by the intrinsic growth rate, the initial dry matter production and by the factors that affect grass growth.

**Appendix 1**

R code to fit nonlinear models to data from accumulated dry matter yield in Brachiaria hybrid cv Cayman at different intervals between cuts during the dry season.

```r
library(bootstrap)
library(ISLR)
library(drc)
library(stats)
library(kableExtra)
tablamo <- data.frame(Modelo = c("Logistico"), Parametro.fct = c("L.3("))
kable(tablamo, caption = "Tabla 1. Códigos para modelos en fct")

data.frame(dc,kgms)

result.Log <- drm(kgms~dc, data = dataset1, fct = L.3())
summary(result.Log)
AIC(result.Log)
set.seed(1)
library(ISLR)
indices.train <- sample(x = nrow(dataset1), size = 0.5*(nrow(dataset1)), replace = FALSE)
datos.entrenamiento <- dataset1[indices.train,]
datos.test <- dataset1[-indices.train,]
n = nrow(dataset1)
head(result.Log)
result.Log$fit$par[3]
result.Log$fit$par
meanstar1 = mean(dataset1$kgms)
sdstar1 = sd(dataset1$kgms)
R = 1000
Fstar = numeric(R)
for (i in 1:R) {
simkgms = rnorm(n, mean = meanstar1, sd = sdstar1)
simtiempo = dataset1$dc
simdata = data.frame(simkgms,simtiempo)
result.Log <- drm(simkgms~simtiempo, data = simdata,fct = L.3())
Fstar[i]=result.Log$fit$par[2]
}
Fstar
mean(Fstar)
hist(Fstar,main = "",ylab = " Kg MS ha^{-1}\),breaks = 90,freq = FALSE)
confint(object = result.Log, level = 0.95 )
```

**Authors Contributions**

Danny Villegas Rivas, Nora Valbuena Torres and Manuel Milla Pino: Design the research plan and
organized the study, participated in all experiments, coordinated the data-analysis and contributed to the writing of the manuscript.

Zadith Garrido Campaña, Martín Grados Vasquez, Erick Delgado Bazan, Cesar Osorio Carrera and Ydalía Velasquez Casana: Participated in all experiments and contributed to the writing of the manuscript.

Wilfredo Ruiz Camacho: Coordinated the mouse work and participated in all experiments and contributed to the writing of the manuscript.

Conflict of Interest

The authors declare that there are no conflicts of interest.

References


https://doi.org/10.22267/rcia.193601.96

https://doi.org/10.1086/401873