Research Article
Open Access
CrossMark

Longevity evaluation of cattle Curraleiro Pé-Duro breed using the inverse Gaussian frailty model

Abstract

Cattle of the Curraleiro Pé-Duro cattle breed have a great ability to adapt to extreme climatic conditions in a natural environment, providing the minimization of losses in productive and reproductive performance, helping in disease resistance, consequently reducing mortality and increasing longevity when exposed to stressful conditions. The research aimed to use of the survival analysis technique, analyzing the effects of unobserved factors, in a study on the stay ability in the herd of cattle Curraleiro Pé-Duro breed using the model of inverse Gaussian fragility. Data from records of 102 cattle born between 2005 and 2014 in an experimental herd of Embrapa Meio-Norte located in São João do Piauí, in the semiarid region of Piauí, Brazil, were observed. The failure was considered to be the inactivity of the cattle caused by the death or sale and the censorship as the animal remaining active in the herd. The methodology of inverse Gaussian frailty models with log normal basis risk was used. Were considered significant ($P < 0.05$) the covariates season of birth, sex, weight at 365 days and the interaction weight at 365 days with sex. There was a predominance of birth in the dry season (July to December). It was observed that the cattle that more remained in the herd were born in the dry season and were male. The use of fragility models proved to be efficient to meet the proposed objectives, pointing out the potential of its use to contemplate unobserved heterogeneity, being a great tool for animal breeding. Frailty models allowed incorporating a term for the unobserved heterogeneity that affects the estimation of risk influencing longevity of the bovine. Thus, so these models can be used as a tool to help in animal improvement.

Keywords: Survival, univariate frailty, censorship, failure, inverse Gaussian

Volume 11 Issue 3 - 2022

KMM S Soares,[1] VLD Tomazella,[2] SCS Júnior,[3] GMC Carvalho,[4] CMM Lima,[1] KR Santos[5]

[1]Department of Statistics, Federal University of Piauí, Brazil
[2]Department of Statistics, Federal University of São Carlos, Brazil
[3]Department of Zootechnics, Federal University of Piauí, Brazil
[4]Department of Statistical Methods, Empresa Brasileira de Pesquisa Agropecuária Meio Norte, Brazil
[5]Department of Medicine, Federal University of the Parnaíba Delta, Parnaiba Campus, Av. São Sebastião, 2819, Parnaíba, Piauí, Brazil

Correspondence: V LD Tomazella, Department of Statistics, Federal University of São Carlos, São Paulo, Brazil, Tel +55 16 981671390; Email veratomazella@gmail.com

Received: August 22, 2022 | **Published:** September 09, 2022

Introduction

The permanence of the animal in the herd is linked to aspects of production, reproduction, nutrition and economics. The permanence of the matrix in the herd is influenced by factors such as the characteristics of animal reproduction, considering that the insufficient performance of the animal in terms of reproduction can determine a reduction in the permanence time. The productive characteristics are of fundamental importance for the functional permanence of the matrix in the herd. Maximizing longevity generates profit optimization by decreasing involuntary culling rates, allowing the producer to carry out a higher voluntary culling rate, increasing genetic gain.[1,2]

Survival Analysis, in turn, is considered the most appropriate statistical methodology to deal with data from time to the occurrence of an event of interest (time of failure), in the presence of censorship,[3] which is its main feature. In animal production, we can highlight Bonetti et al.[4] who estimated genetic parameters in a genetic evaluation for the longevity of Italian Brown-Swiss bulls, using the Weibull proportional hazards model. The authors considered the method satisfactory for the use and inclusion of bulls in breeding programs. However, Caetano et al.[5] proposed the age of the cow at the last calving as a measure to assess the cow's ability to remain in the herd. The authors concluded that the variable is relevant to assess the ability of cows to remain in the herd and that the survival analysis model estimate a greater proportion of genetic variability for the trait studied. A special feature associated with survival data is the possibility that, for some individuals, the complete time until the occurrence of the event of interest is not observed due to various causes. Failure to consider these individuals with incomplete information about their lifetimes can lead to biased or less efficient inferences.[3,6] Therefore, one can see the importance of introducing a variable in the analysis that indicates whether survival time was observed.[7] This variable is defined in the literature as a variable indicating censorship or simply "censorship".

In recent studies, there are situations in which the response variable, failure time, may be influenced by unobservable factors, called latent factors. Survival models with latent variables or frailty models are characterized by the inclusion of a random effect, that is, an unobservable random variable, which represents information that cannot or was not observed; such as environmental, genetic or information factors that for some reason were not considered in the planning. One of the ways found to incorporate this random effect, called the fragility variable, is to introduce it in the modeling of the risk function, with the objective of controlling the unobservable heterogeneity of the units under study.[8,9] The frailty can be inserted in the model in an additive or multiplicative way, with the objective of evaluating the heterogeneity among the units in the risk function or the dependence for multivariate data. In animal studies, associations appear due to shared genetic or environmental influences, and if ignored, incorrect inferences can be drawn.

The term frailty was introduced by Vaupel et al.[10] in survival models with univariate data. Due to the characteristics of frailty in the multiplicative frailty model, the natural candidates for the frailty distribution, supposedly continuous and not dependent on time, are the gamma, log-normal, inverse Gaussian and Weibull distributions. Hougaard[11] was one of the first authors to address the impact of using different distributions for the frailty variable.

Curraleiro Pé-Duro was the first bovine breed selected in Brazil from Portuguese breeds brought by colonizers from the 15th century onwards.[12] It was introduced in the region of the state of Piauí, from the São Francisco River by Domingos Afonso Mafrense, in the middle of the year 1674, later resulting in the adaptation of the cattle to the environmental conditions of the region.[13] It is indicated as a pure breed for the production of semen and embryos for use in reproduction and industrial crosses with specialized breeds for the production of milk and tender meat, marketed with a protected origin designation. The natural resistance to ecto and endoparasites and adaptation to our grasses and legumes are the great weapons of these cattle, which have been naturally selected for centuries to face local adversities. To all this comes the great thermal amplitude in which they can be bred and great longevity, living for more than 20 years. However, the great merit of this breed is to convert low-quality foods into noble foods and enable people to live together in semiarid regions.

In this context, the objective of this work was to analyze the effects of unobserved factors, in a study on the length of stay in the herd, of Curraleiro Pé-Duro cattle using the inverse Gaussian univariate frailty model.

Material and methods

The data for this study were provided by the Curraleiro Pé-duro (CPD) cattle conservation center belonging to Embrapa Meio-Norte, in Teresina-Piauí-Brazil, with an experimental field located on the Otávio Domingues farm, in São João do Piauí (between 8° 26' and 8° 54' South latitude and between 42° 19' and 42° 45' West longitude), in the semiarid region of Piauí belonging to an insitu conservation herd. The management of these cattle was carried out extensively with the supply of only salt, mineral and water, which justifies their low weight compared to CPD cattle raised in other regions and cattle of other breeds. It is worth highlighting the lack of food supplementation, the increased incidence of environmental effects with the presence of many toxic plants, ticks, babesia, worms, which contribute to the low weight of CPD.

102 cattle (58 males and 44 females) of the CPD breed were evaluated from birth to 550 days. Animal data were collected from 2005 to 2014. To model the survival time (in months), after the beginning of the reproductive life of the cattle until the occurrence of failure (inactivity caused by death or sale), in relation to the cattle remaining active (alive), time was considered as the response variable, and the calf's date of birth was the beginning of the study. The variable T (time) was obtained from the difference between the date of birth and the date of disposal. The date of the last disposal in the herd was considered as the final observation period for animals that had not been disposed yet (August 19, 2016). Failure was defined as cattle inactivity (death or sale), while censorship was defined for cattle that remained alive in the herd. The type of censorship used was the right one. For each animal observed, it was registered a corresponding indicator of censorship, called status (δ=1 if it failed and δ=0 if censored) indicating whether the animal is active or inactive in the herd.

In this study, the covariates considered as possible risk factors in the length of stay of CPD cattle in the herd were: SB: season of birth, S: sex, BW: birth weight, WW: weaning weight, W365: weight at 365 days, W550: weight at 550 days categorized according to the average weight in each phase[14] and described according to Table 1.

Table 1 Number and percentage of cattle that failed or were censored by variable

Variable	Category	n	Fail (%)	Censors (%)
Birth season	0 - rainy	35	31(88,6)	4(11,4)
	1 - dry	67	42(62,7)	25(37,3)
Sex	0 - male	58	48(82,7)	10(17,3)
	1 - female	44	25(56,8)	19(43,2)
Birth weight	0 - $<$ 20 kg	56	41(73,2)	15(26,8)
	1 - $\geq$ 20 kg	46	32(69,6)	14(30,4)
Weaning weight	0 - $<$ 66 kg	40	29(72,5)	11(27,5)
	1 - $\geq$ 66 kg	62	44(71,0)	18(29,0)
365-days weight	0 - $<$ 95 kg	34	29(85,3)	5(14,7)
	1 - $\geq$ 95 kg	68	44(64,7)	24(35,3)
550-days weight	0 - $<$ 131 kg	48	35(72,9)	13(27,1)
	1 - $\geq$ 131 kg	54	38(70,4)	16(29,6)

Model formulation

To analyze the longevity of cattle in the herd, survival analysis techniques were used. Due to the presence of censors in survival data, they are summarized with estimates of the survival function and the risk function.[15] To estimate these functions, the non-parametric method of Kaplan and Meier[16] was used. To analyze the influence of covariates on the longevity of cattle in the herd, the univariate frailty model was used, which is an extension of the Cox model. All statistical analyzes were performed using free statistical software[17] and the parfm package.[18]

Univariate frailty model

Frailty models for univariate survival data take into account that the population is non-homogeneous. Heterogeneity can be explained by covariates, but when important covariates are not incorporated into the model, this leads to unobserved heterogeneity.[3] The multiplicative frailty model is an extension of the Cox model,[19] where individual risk depends on an unobservable, non-negative random variable Z, which acts multiplicatively on the basis risk function. The risk function with the presence of covariates at time t for the i[th] individual is given by:

$$h(t|X,z_i) = z_i h_0(t) exp\{\boldsymbol{X}'\boldsymbol{\beta}\} \quad (1)$$

where $\boldsymbol{X}$ is the vector of covariates and $\boldsymbol{\beta}$ the vector of parameters associated with $\boldsymbol{X}$. As z_i represents a value of the unobservable random variable, the individual risk increases when $z_i > 1$, decreases if $z_i < 1$ and for $z_i = 1$ the model of frailty(1) reduces to the Cox proportional hazard model.[19] The fact that the frailty variable acts in a multiplicative way in the risk function implies, the higher the value of the frailty variable, the greater the chance of failure. Thus, the greater the z_i, the more "fragile" the observations belonging to individual i are about to fail, hence the name fragility. Therefore, the event of interest is expected to occur for the most "fragile" individuals.[8]

An important problem in frailty models is the choice of distribution for the random effect. Due to the way the frailty term acts in the risk function, the candidates for the frailty distribution are supposedly

Citation: Soares KS, Tomazella V, Júnior S, et al. Longevity evaluation of cattle Curraleiro Pé-Duro breed using the inverse Gaussian frailty model. *Biom Biostat Int J.* 2022;11(3):111‒117. DOI: 10.15406/bbij.2022.11.00363

non-negative, usually continuous and not time-dependent, such as gamma, lognormal, inverse Gaussian and weibull distributions.[9,20] Distributions traditionally used to represent lifetimes can be attributed to the basis risk function, such as exponential, lognormal, weibull, gamma, among others.[21]

Inverse Gaussian frailty distribution

The inverse Gaussian distribution was introduced by Hougaard.[11] Thus, let Z be a random variable that follows an inverse Gaussian frailty distribution with $E(Z)=1$ and $Var(Z)=\theta$. The probability density function is given by

$$f(z)=\frac{1}{\sqrt{2\pi\theta}}z^{-\frac{3}{2}}exp(-\frac{(z-1)^2}{2\theta z},\theta>0$$

A useful tool for frailty model analysis is the Laplace transform. Given a function *g(x)*, the Laplace transform considered as a real function with argument *s* is defined by P Hougaard.[21]

$$L(s)=\int_0^\infty g(s)e^{-sx}ds$$

The reason this is useful in our context is that the Laplace transform has exactly the same shape as the unconditional survival function. The unconditional survival function, integrating the frailty term, is given by:

$$S(t)=\int_0^\infty [S_0(t)]^z g(z)dz=\int_0^\infty e^{-H_0(t)z} g(z)dz=L[H_0(t)]$$

where *g(z)* is the probability density function of the frailty variable and $L[H_0(t)]$ denotes the Laplace transformation of the function *g(z)* considering the accumulated risk function, $H_0(t)$.

Consequently, the Laplace transform of the inverse Gaussian distribution is given by

$$L(z)=\exp\left[\frac{1}{\theta}\left(1-\sqrt{1+2\theta z}\right)\right],z\geq 0$$

The risk function and unconditional survival function of the inverse Gaussian frailty variable are given, respectively, by

$$h(t)=\frac{h_0(t)}{\left(1+2\sigma^2 H_0(t)\right)^{1/2}} \text{ and } S(t)=exp\left[\frac{1}{\sigma^2}\left(1-\sqrt{1+2\sigma^2 H_0(t)}\right)\right]$$

where h_0 and H_0 are the basis risk and basis accumulated risk functions.

In the presence of covariates, the risk and unconditional survival functions are given, respectively, by:

$$h(t,X)=\frac{h_0(t)e^{\boldsymbol{\beta X'}}}{(1+2\sigma^2\boldsymbol{H}_0(t)e^{\boldsymbol{\beta}'\boldsymbol{X}})^{1/2}} \quad (2)$$

and

$$S(t,X)=exp\left[\frac{1}{\sigma^2}\left(1-\sqrt{1+2\sigma^2\boldsymbol{H}_0(t)e^{\boldsymbol{\beta}'\boldsymbol{X}}}\right)\right] \quad (3)$$

where $\boldsymbol{X}$ is the vector of covariates and $\boldsymbol{\beta}$ the vector of parameters associated with $\boldsymbol{X}$.

LogNormal Inverse Gaussian fragility model

Different parametric forms can be assumed for the basis risk function $h_0(t)$ for example: Lognormal, Log-logistic, Gompertz and so on. Table 2 shows the probability density functions and the basis risk and survival functions for these distributions.

Table 2 Functions of density f(t), survival S_0 (t) and risk h_o(t)of the lognormal, log-logistic and Gompertz distributions

Distribution	$f(t)$	$h_o(t)$	$S_0(t)$
lognormal	$\frac{1}{\sqrt{2\Pi t\theta}}\exp\left[-\frac{1}{2}\left(\frac{\log(t)-\mu}{\sigma}\right)^2\right]$	$\frac{\frac{1}{\sqrt{2\Pi t\theta}}\exp\left[-\frac{1}{2}\left(\frac{\log(t)-\mu}{\sigma}\right)^2\right]}{\phi\left(\frac{-\log(t)+\mu}{\sigma}\right)}$	$\phi\left(\frac{-\log(t)+\mu}{\sigma}\right)$
log-logístic	$\frac{\gamma}{\lambda^\gamma}t^{\gamma-1}\left(1+\left(\frac{t}{\lambda}\right)^\gamma\right)^{-2}$	$\frac{\gamma\left(\frac{t}{\lambda}\right)^{\gamma-1}}{\lambda\left[1+\left(\frac{t}{\lambda}\right)^\gamma\right]}$	$\frac{1}{\left[1+\left(\frac{t}{\lambda}\right)^\gamma\right]}$
Gompertz	$\lambda\exp(\gamma t)\exp\left\{-\left(\frac{\lambda}{\gamma}\right)\left(e^{\gamma t}-1\right)\right\}$	$\lambda\exp(\gamma t)$	$\exp\left\{-\left(\frac{\lambda}{\gamma}\right)\left(e^{\gamma t}-1\right)\right\}$

In this study, the parametric approach was considered for the univariate inverse Gaussian frailty model, where the lifetimes of cattle at risk follow lognormal distribution. Thus, substituting the basis risk functions $h_0(t)$, survival basis $S_0(t)$, and cumulative basis risk $H_0=-log(S_0(t))$, from the lognormal distribution, respectively in (2) and (3) , with the presence of covariates, the lognormal inverse Gaussian univariate frailty model with unconditioned survival and risk function given, respectively, by:

$$S(t,X)=\exp\left[\frac{1}{\sigma^2}\left(1-\sqrt{1-2\sigma^2\log\left\{\phi\left(\frac{-\log(t)+\mu}{\sigma}\right)\right\}e^{\boldsymbol{\beta}'\boldsymbol{X}}}\right)\right] \quad (4)$$

and

$$h(t,X)=\left[\frac{\frac{1}{\sqrt{2\Pi t\sigma}}\exp\left[-\frac{1}{2}\left(\frac{\log(t)-\mu}{\sigma}\right)^2\right]}{\phi\left(\frac{-\log(t)+\mu}{\sigma}\right)}e^{\boldsymbol{\beta}'\boldsymbol{X}}\right]\cdot\left[\frac{1}{1-2\sigma^2\log\left(\phi\left(\frac{-\log(t)+\mu}{\sigma}\right)e^{\boldsymbol{\beta}'\boldsymbol{X}}\right)^{1/2}}\right] \quad (5)$$

where $\Phi(.)$ is the cumulative distribution function of a standard normal.

Estimation of model parameters

In survival analysis, the most widely used method to estimate the vector of parameters

$\tau = (\theta,\mu,\sigma,\beta)$ of the model (5) is the maximum likelihood method, since it can incorporate censored data. Assuming that the data are independent and identically distributed, the unconditional likelihood function, with censored data, is given by:

$$L(\tau) = \Pi_{i=1}^{n}[h_i(t)]^{\delta_i} S_i(t) \quad (6)$$

where δ_i is the censorship indicator, h_i and S_i is the risk and survival function of the frailty distribution. Substituting (2) and (3) into (6), we have the non-conditional likelihood function of inverse Gaussian frailty given by:

$$L(\tau) = \prod_{i=1}^{n}\left[\frac{h_0(t)e^{\beta' X}}{(1+2\sigma^2 H_0(t)e^{\beta' X})^{1/2}}\right]^{\delta_i} \exp\left[\frac{1}{\sigma^2}\left(1-\sqrt{1-2\sigma^2 H_0(t)e^{\beta' X}}\right)\right] \quad (7)$$

Consequently, substituting (4) and (5) in (7), we have for the inverse Gaussian frailty model with lognormal basis risk with parameters τ, the non-conditional likelihood function given by

$$L(\tau) = \prod_{i=1}^{n}\left[\left[\frac{\frac{1}{\sqrt{2\Pi}t\sigma}\exp\left[-\frac{1}{2}\left(\frac{\log(t)-\mu}{\sigma}\right)^2\right]}{\phi\left(\frac{-\log(t)+\mu}{\sigma}\right)}e^{\beta' X}\right]\left[\frac{1}{\left(1-2\theta\log\left(\phi\left(\frac{-\log(t)+\mu}{\sigma}\right)e^{\beta' X}\right)\right)^{1/2}}\right]\right]^{\delta_i} \times \exp\left[\frac{1}{\sigma^2}\left(1-\sqrt{1-2\theta\log\left(\phi\left(\frac{-\log(t)+\mu}{\sigma}\right)e^{\beta' X}\right)}\right)\right] \quad (8)$$

The maximum likelihood estimates are obtained by numerically maximizing the log-likelihood function described in (8). To estimate the parameters, the parfm package[18] of the R software[17] was used. The construction of confidence intervals for the model parameters are based on the asymptotic normality properties of the maximum likelihood estimators. If $\hat{\tau}$ denotes the maximum likelihood estimators of τ then the distribution of $\hat{\tau}-\tau$ is approximated by a q-varied normal distribution with zero mean and covariance matrix $I^{(-1)}(\hat{\tau})$ where $I(\tau)$ is the observed information matrix. Thus, an asymptotic confidence interval with $100(1-\alpha\%)$ for each parameter is:

$\hat{\tau} \pm z_{\alpha/2}\sqrt{\widehat{Var}(\tau)}$ where $\hat{\tau}$ is the element of the main diagonal of $I^{-1}(\hat{\tau})$ corresponding to each parameter and $z_{\alpha/2}$ is the quantile $(1-\alpha\%)$ of the standard normal distribution.

To test hypotheses related to the parameters $(\theta,\mu,\sigma,\beta)$ three tests were used: the Wald, the Likelihood Ratio and the Score.[3] Model selection criteria such as the Akaike information criterion (AIC) proposed by Akaike[22] and the Bayesian information criterion (BIC), proposed by Schwarz et al.[23], are often used to select models in different areas. The best models are considered those with lower AIC and BIC values. For the adequacy of the model, several methods are available in the literature and they are essentially based on Cox-Snell residues, which help to examine the global adjustment of the model, Schoenfeld's, which has a time-dependent coefficient, the one of martingale, which is given by the difference between the observed number of events for an individual and the expected one given the adjusted model, and the deviance, which facilitate the detection of atypical points (outliers).

Results

In this study, the herd consists of 102 cattle, where 57% is made up of males and 43% of females, of the CPD breed born in the period from 2005 to 2014. When considering the survival of these animals in the herd, it was found that 28% (29) were censored, that is, they remained alive until the final observation period, and 72% (73) failed, that is, died or were sold. It was observed for SB, a greater predominance (65%) of the dry period, where there is a lower incidence of rain. Regarding standard weights (birth, weaning, 365 days and 550 days), a frequency greater than 50% was found in relation to the average weight for weaning, at 365 days and at 550 days.

Table 1 shows the number and percentage of cattle that failed or were censored by variable. Regarding the SB covariate, there was a higher percentage of failures (88.6%) in calves born in the rainy season. In relation to S, male calves failed more (82.7%) compared to females (56.8%). Regarding the variables BW, WW, W365 and W550 all had a higher percentage of failures, for weights below the average.

The estimates for the parameters of the inverse Gaussian-lognormal univariate frailty model are described in Table 3, where only the estimates of the parameters of the covariates that were significant when considering p-value less than 5% were presented.

Among the analyzed variables (SB, S, BW, WW, W365 and W550) significance was observed in the variables SB, S, W365 and the interaction of S with W365 $(p-value < 0.05)$. The failure rate in dry SB was $exp(-1.310) = 0.269$ times compared to the rainy one, indicating that there is a decrease in the risk of cattle failing (death or sale), among those that manage to stay alive, demonstrating a cattle adaptability in the herd. For variable S, the failure rate for females was $exp(1.677) = 5.349$ times the failure rate for males, that is, female calves are 5.349 times more likely to fail (death or sale) than male calves. For variable W365, the failure rate was $exp(0.775) = 2.170$ times for animals weighing 95 kg or more. For the interaction of variable S with W365, a failure rate of $exp(-3.204) = 0.040$ was observed, indicating a decrease in the risk of cattle failing, that is, dying or leaving for sale. As for the estimated variance for θ frailty, it is equal to 0.155, revealing the presence of unobserved heterogeneity, such as genetic or environmental factors.

The survival (4) and unconditioned risk (5) functions of the inverse Gaussian frailty model with log-normal basis risk are given, respectively, by:

$$S(t,X) = \exp\left[\frac{1}{0,155}\left(1-\sqrt{1-0,31\log\left\{\phi\left(\frac{-\log(t)+3,924}{0,544}\right)\right\}g(\beta' X_{ij})}\right)\right]$$

and

$$h(t,X) = \left[\frac{\frac{1}{0,155\sqrt{2\Pi}t}\exp\left[-\frac{1}{2}\left(\frac{\log(t)-3,924}{0,544}\right)^2\right]}{\phi\left(\frac{-\log(t)+3.924}{0,544}\right)}g(\beta' X_{ij})\right].\left[\frac{1}{\left(1-0,31\log\left(\phi\left(\frac{-\log(t)+3,924}{0,544}\right)g(\beta' X_{ij})\right)\right)^{1/2}}\right]$$

where

$$g(\beta' X_{ij}) = exp\{-1.310\times \text{EN} + 1.677\times S + 0.775\times P365 - 3.204\times(S * P365)\}$$

According to Figure1, it was possible to observe that 50% of the calves had frailty varying between $[0.682; 0.985]$ and 50% ranging from $[1; 1.153]$, indicating the presence of relatively high unobserved heterogeneity in half of the herd.

To assess the goodness of fit of the model, in Figure 2 are graphs of Cox-Snell residuals, Martingale residuals and Deviance residuals.

In panels (a) and (b), it is observed that the Cox-Snell residuals approximately follow a standard exponential distribution, which indicates an acceptable global goodness-of-fit of the model. Panels (c) and (d) do not suggest the existence of outlier points.

A comparative study of the inverse Gaussian frailty model was carried out using the lognormal and log-logistic basis risk functions, described in Table 4, where the Akaike criterion (AIC) and Bayesian information criterion (BIC) were used for the selection of models, showing that the inverse Gaussian-lognormal frailty model is the most suitable for presenting the lowest value of AIC and BIC.

Discussion

Adaptability, or ability to adapt, can be assessed by the animal's ability to adjust to average environmental conditions as well as to climatic extremes. Well-adapted animals are characterized by maintenance or minimal reduction in productive performance, high reproductive efficiency, resistance to diseases, longevity and low mortality rate during exposure to stress.[24]

The small size of CPD cattle is the main reason alleged to nearly drive them to extinction. However, the extreme conditions to which animals of this breed are subjected are rarely mentioned. Improved breeds, adapted to good pastures, if kept in the same conditions offered to CPD cattle, will soon have their weight reduced and their reproductive performance negatively affected. Furthermore, it is important to add that the production capacity of a breed does not depend only on the individual weight of the animals, the birth and mortality rates being fundamental. A herd of heavier animals, but less prolific and with a higher percentage of deaths, results in less meat production. In addition, the same pasture can guarantee the feeding of a greater number of smaller animals. Therefore, a more adequate index to evaluate the performance of a cattle breed in pasture is its annual productivity per unit of area.[25]

The perspective addressed in this research was supported by the longevity of cattle and the departure from the herd caused by death or sale. There was a predominance of 66% of calves born in the dry season, indicating that more than half of the calves were born in a favorable period for their development, caused by the reduction of diseases in the environment and mortality of the calves, in addition to the supply of nutrients in the pasture in the pregnant season for the mother. That is, the calf born during the dry season, the mother during pregnancy had more pasture available, therefore more milk and consequently a fatter calf. Regarding the standard weights (birth, weaning, 365 days and 550 days), a higher frequency of cattle with weight above the average was found, indicating a better genetic concentration and good adaptability.

The univariate inverse Gaussian frailty model with lognormal basis risk was used, where the following covariates were significant (p-value < 5%) and the interaction of S with W365, as shown in Table 3. It was still observed through the failure rate ratio, that for the covariate SB, there was a decrease in the failure rate (death or sale) for cattle born in the dry season, which can be explained by having less diseases for the calves, in addition to the supply of nutrients in the pasture at the time of pregnancy for the mother, demonstrating greater longevity for the cattle. For variable S, a higher failure rate was found for females than that for males, which can be explained by the fact that as females give birth annually, in the absence of supplementation, their organism directs calcium from the bones to the milk during the breastfeeding, making the cow leaner and more subject to mortality. For variable W365, a higher failure rate was found for cattle weighing 95kg or more, which can be explained by heavier cattle leaving the herd for sale, slaughter or reproduction.

Table 3 Estimates of maximum likelihood (EMV), standard error (SE), confidence interval (CI 95%), p-value and failure rate ratio (FRR) for the parameters of the lognormal inverse Gaussian univariate frailty model

Parameters	EMV	EP	CI (95%)	*p*-valor	FRR
θ	0.155	0.593	-	-	-
μ	3.924	0.187	-	-	-
σ	0.544	0.134	-	-	-
β_1 (SB)	-1.310	0.319	[0.144; 0.504]	<0,001	0.269
β_2 (S)	1.677	0.527	[1.904;15.028]	0,001	5.349
β_3 (W365)	0.775	0.373	[1.044; 4.515]	0,038	2.170
β_4 (S*W365)	-3.204	0.738	[0.010; 0.173]	<0,001	0.040

Table 4 Estimates for the parameters of the lognormal inverse Gaussian and log-logistic inverse Gaussian fragility models

	lognormal		log-logistic	
Parameters	Estimates	Standard error	Estimates	Standard error
θ	0.155	0.593	0.014	0.322
μ	3.924	0.187	-	-
σ	0.544	0.134	-	-
α	-	-	-11.617	2.082
γ	-	-	2.962	0.581
β_1 (SB)	-1.310	0.319	-1.264	0.289
β_2 (S)	1.677	0.527	1.513	0.493
β_3 (W365)	0.775	0.373	0.695	0.355
β_4 (S*W365)	-3.204	0.738	-2.978	0.687
Log-Veross	-372.94		-374.79	
AIC	759.88		763.59	
BIC	778.25		781.97	

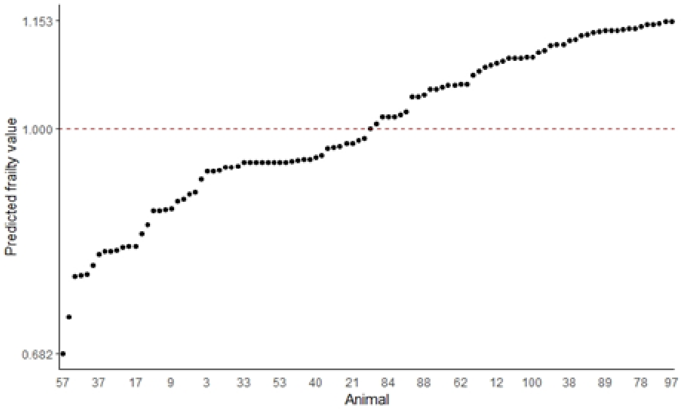

Figure 1 Univariate inverse Gaussian frailty with lognormal baseline risk.

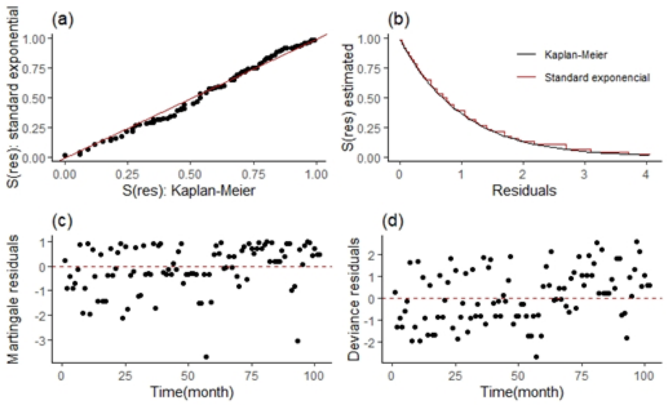

Figure 2 Residues of Cox-Snell, Martingale and Deviance from the lognormal inverse Gaussian fragility model for data from cattle of the Curraleiro Pé-Duro breed.

The methodology used in this study proved to be promising, with interesting results in the study to assess the permanence of animals in the herd, as it can be seen in Figure 1, where weaknesses inherent to environmental and genetic factors were detected, and in Figure 2 where the adjustment tests of models for effects of genetic nature as random were carried out, which include the indication of adequacy of the frailty model to meet the objectives proposed by the Cox-Snell, Martingale and Deviance tests.

In this study, modeling the length of stay of the calf in the herd with the inverse Gaussian - lognormal univariate frailty model, where the random effect was associated with each animal, the estimated variance for frailty, $\hat{\theta} = 0.155$, indicating the presence of unobserved heterogeneity, caused by genetic or environmental factors, which is related to factors such as: coat color, fur color, maternal ability, thermoneutrality zone, rearing condition, type of pasture, type of food, among others.[26]

Conclusion

The main conclusion of this article is that the use of frailty models proved to be an adequate selection criterion for the proposed situations. It also allowed incorporating a term for the unobserved heterogeneity that affects the estimation of risk, where it was possible to observe for the univariate frailty model, the statistical significance in the covariates SB, S, W365 and the interaction of S with W365 affecting longevity of the bovine. In this sense, frailty models can be used as a tool to help in animal improvement.

Acknowledgments

None.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Van Arendonk, JAM. Economic importance and possibilities for improvement of dairy cow herd life. *In:World congress of genetic applied to livestock production*'. Lincoln;1986.
2. Ferreira WJ. Estudo de tendência genética e de medidas de longevidade em bovinos da raça Holandesa no estado de Minas Gerais. 2003.
3. Colosimo EA, Giolo SR. Applied survival analysis. Editora Edgard Blucher. São Paulo, 2006.
4. Bonetti O, A Rossoni C, e Nicoletti. Genetic parameters estimation and genetic evaluation for longevity in italian brown swiss bulls. *Italian Journal of Animal Science*. 2009;8:30–32.
5. Caetano , Rosa G, Savegnago R, et al. Characterization of the variable cow's age at last calving as a measurement of longevity by using the kaplan–meier estimator and the cox model. *Animal*. 2013;7:540–546.
6. Bolfarine HRJAJ. 'Análise de Sobrevivência'. 2ª Escola de Modelos de Regressão, Rio de Janeiro,1991.
7. Louzada Neto F, Pereira BdB. Modelos em análise de sobrevivência. Caderno de saúde coletiva, *Rio Janeiro*. 2000;9–26.
8. Tomazella VLD. Recurring event data modeling via Poisson process with frailty term. Universidade de São Paulo;2003.
9. Calsavara VF. Fractional cure survival models using a generalized modified Weibull frailty and lifespan term. 2011.
10. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*. 1979;16:439–454.
11. Hougaard P. Life table methods for heterogeneous populations: distributions describing the heterogeneity. *Biometrika*.1984;71:75–83.
12. Salles P, Medeiros G, Costa RG, et al. Conservation and breeding program of brazilian curraleiro (pé duro) cattle breed. AICA-*Actas Iberoamericana de Conservacion animal*. 2011;1:453–456.
13. Azevedo DMRR, Alves AA, Feitosa FS. Adaptability of Pé-duro cattle to the climatic conditions of the semi-arid region of the State of Piauí.. *Arch Zootec*. 2008;57:5-11.
14. Carvalho GMC, Lima Neto A, Da Frota MNL, et al. The use of "curraleiro pé-duro" cattle in crossbreeding for the production of good quality meat in the hot tropics-phase 1. In 'Embrapa Meio Norte-Article in congress proceedings (ALICE)'. Northeastern Congress of Animal Production; 2015.
15. Moore DF. *Applied survival analysis using R*. Springer;2016.
16. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*. 1958; 53: 457–481.
17. R Core Team 2016. R: a language and environment for statistical computing. Retrieved on 9 Septembre 2019.
18. Munda M, Rotolo F. Parfm: Parametric Frailty Models in R. *Journal of Statistical Software*. 2012;51(11).
19. Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1972;34:187–202.
20. Hougaard P. Frailty models for survival data. Lifetime data analysis. 1995;1(3):255–273.

21. Wienke A. Frailty models in survival analysis. 1st ed. Chapman and Hall/CRC;2010.
22. Akaike H. A new look at the statistical model identification. *IEEE transactions on automatic control*. 1974;19:716–723.
23. Schwarz G. et al. Estimating the dimension of a model. *The annals of statistics*. 1978;6(2):461–464.
24. Azevêdo DMMR, Alves AA. Bioclimatology applied to dairy cattle production in the tropics. Teresina: Embrapa Mid-North, 2009.
25. Carvalho JH. Economic potential of the hard-footed bovine. *Embrapa Meio-Norte*. Infoteca-e. 2002.
26. Carvalho GMC, Almeida MDO, Azevêdo DMMR, et al. Origin, formation and conservation of the Pé-Duro cattle, the bovine of the Brazilian Northeast. *Embrapa Meio-Norte*. Infoteca-e; 2010.

Citation: Soares KS, Tomazella V, Júnior S, et al. Longevity evaluation of cattle Curraleiro Pé-Duro breed using the inverse Gaussian frailty model. *Biom Biostat Int J*. 2022;11(3):111‒117. DOI: 10.15406/bbij.2022.11.00363