

ORIGINAL ARTICLE OPEN ACCESS

Genomic Prediction Ability Using Linear and Threshold Approaches for Different Stayability Definitions in Nellore Females

Leticia Silva Pereira¹  | Larissa Bordin Temp²  | Eduarda da Silva Oliveira²  | Jorge Hidalgo³  | Cláudio Ulhoa Magnabosco⁴ | Fernando Baldi²

¹National Association of Breeders and Researchers (ANCP), Ribeirão Preto, São Paulo, Brazil | ²School of Agriculture and Veterinary Science, São Paulo State University, Jaboticabal, São Paulo, Brazil | ³Department of Animal and Dairy Science, University of Georgia, Athens, Georgia, USA | ⁴Embrapa Cerrados, Brasília, Federal District, Brazil

Correspondence: Leticia Silva Pereira (leticiapsvet@gmail.com)

Received: 12 September 2025 | **Revised:** 3 November 2025 | **Accepted:** 15 November 2025

Keywords: fertility | genomic selection | liability scale | longevity

ABSTRACT

Enhancing female longevity through regular calvings improves herd replacement rates and reduces the costs associated with replacing low reproductive efficiency animals. Stayability (STAY), defined as a cow's ability to remain productive in the herd, is a categorical trait that challenges traditional genetic evaluation due to its non-normal distribution. This study aimed to estimate genomic predictions for different STAY definitions—based on the number of calvings at specific ages in Nellore females—and to compare the predictive ability of linear and threshold models using the linear regression (LR) method. Phenotypic and genotypic data from 187 herds provided by the Nellore Brazil breeding program (ANCP, Ribeirão Preto, Brazil) were used. Four STAY definitions (STAY48-2, STAY48-3, STAY54-2, STAY54-3, STAY72-3) were evaluated. Genomic estimated breeding values (GEBV) were obtained using univariate linear and threshold models implemented in the BLUPF90 software family. Variance components were transformed from liability to observed scale. Heritability estimates ranged from 0.16 to 0.22 on the liability scale and 0.07 to 0.09 on the observed scale. Threshold models showed superior predictive ability compared to linear models, with higher accuracies (0.531 to 0.698 vs. 0.451 to 0.532), lower bias (−0.0004 to 0.008 vs. 0.027 to 0.096) and dispersion values closer to the ideal (0.932 to 1.000 vs. 0.811 to 0.848). Among the definitions, STAY48 with at least two or three calvings demonstrated the most consistent performance, representing a promising criterion for genetic evaluation in Nellore cattle.

1 | Introduction

An efficient reproductive lifespan of cows is essential for the success of beef cattle production systems. Regardless of the production system adopted, reproductive traits are paramount in defining selection criteria (Phocas et al. 1998; Brumatti et al. 2011). Stayability (STAY) is a trait that measures a cow's ability to remain in the herd as a function of its productivity, therefore, it is an economically relevant trait in livestock

production systems (Portes et al. 2021). Increased female longevity, with regular calvings, can optimise herd replacement rates, which may help reduce costs associated with replacing animals with low reproductive efficiency (Santana et al. 2013). In general, STAY is measured late in the animal's productive life, only in females and exhibits low heritability estimates. These factors can hinder direct selection for this trait, as they tend to increase the generation interval and reduce genetic gain per year (Teixeira et al. 2017; Ramos et al. 2021; Silva et al. 2024).

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Journal of Animal Breeding and Genetics* published by John Wiley & Sons Ltd.

In selection herds, the implementation of reproductive technologies and genomics has reduced the age at first calving and increased the voluntary culling rate, thereby decreasing the average age of the herd, with a consequent reduction in generation interval (GI). In *Zebu* breeds like Nellore, females are raised in extensive systems in the tropics, where they are subjected to an environment with seasonal feed resources, with STAY being a measure of productivity over several reproductive cycles. Therefore, it is a trait that reflects not only reproductive efficiency but also the ability to adapt to environmental and nutritional conditions, as well as maternal ability.

With the intense selection for sexual precocity, especially in the last 14 years in the Nellore breed, that is, with females challenged between 10 and 14 months of age (Pereira et al. 2025), those that manage to conceive, calving, wean a calf and conceive again at regular reproductive intervals make their longevity an even more effective measure of the functional reproductive longevity of these females in real production environments. These changes in herd structure have necessitated a reevaluation of the selection criteria for productive longevity in females challenged at an early age in selection herds, such as those involved in employment. Earlier definitions of the trait (such as at 48 or 54 months of age) are being replaced by traditional definitions later, like those based on permanence up to 72 months.

As a categorical trait (success or failure), STAY does not follow a normal distribution, which increases the complexity of genetic evaluation models (Cavani et al. 2015; Costa et al. 2019). In this context, genetic values are predicted using threshold models, which assume the existence of an underlying standard variable (liability) that is presumed to have a joint distribution with the observed binary phenotypes (Hidalgo, Tsuruta, et al. 2024). Threshold models require a nonlinear system of equations imposing challenges for convergence and computing time. These challenges are lessened in linear models, which are favoured as well because in practice the advantages provided by threshold models over linear models are minimal for raking purposes. Cappelloni et al. (2022), estimated variance components and genetic values for traits in different pig breeds, comparing threshold and linear models. Their findings indicated that both models provided similar estimates, but the linear model was recommended due to its greater practicality and computational efficiency.

Misztal et al. (1989) identified significant differences in computational cost between threshold and linear models, estimating an increase of approximately five times in computational time for the threshold models. Similarly, Weller et al. (1988) and Clutter et al. (1989) reported an increase of about three times the time required to run threshold models compared to linear models. Supporting these findings, Hidalgo, Misztal, et al. (2024) also highlighted that the convergence time for threshold models was approximately five times longer than that of linear models. The linear models offer a robust approximation, and previous studies in the literature (Weller et al. 1988; Heringstad et al. 2003; Koeck et al. 2010) have reported highly correlated estimated breeding values (EBVs) between linear and threshold models (Hidalgo, Misztal, et al. 2024).

Model validation for categorical traits presents challenges, especially in the context of genetic prediction. The linear regression (LR) method has been proposed as a valuable tool for evaluating the performance of these models and estimating potential improvements in predictive accuracy (Bermann et al. 2021). Given this context, in the present study we aimed to estimate genomic predictions for different STAY definitions (considering varying calvings in Nellore females) and to compare, using the LR method, the predictive ability achieved using linear and threshold models. Considering that STAY is a categorical trait, exploring different statistical approaches becomes essential to ensure more reliable estimates.

2 | Material and Methods

2.1 | Ethics Statement

This study was exempt from evaluation by the Animal Ethics Committee (CEUA), as established by Law No. 11,794 of 08/10/2008 and Normative Resolution No. 51 of 05/19/2021 from the National Council for Animal Experimentation Control (CONCEA) because the data was obtained from an existing database.

2.2 | Phenotypes and Quality Control

The phenotypic and genotypic information was provided by the Nellore Brazil breeding program, coordinated by the National Association of Breeders and Researchers (ANCP, Ribeirão Preto, Brazil). The animals belonged to 187 farms in Brazil's central-west, southeast, northeast and northern regions. The heifers were exposed to reproduction at 10–14 months of age, and the mating season lasted 3–4 months.

2.2.1 | Longevity

Stayability (STAY, %) was considered in this study. To calculate traditional stayability, all females in the database were used, regardless of whether they had been challenged between 10 and 14 months of age or not. To calculate the new stayability approaches in early females, only challenged females were used. For STAY, females born between 2000 and 2019 were classified according to the number of calvings at different ages, as described below (Pereira et al. 2025):

- STAY48-2: Challenged females that produced two calves by 48 months of age received a score of 2 (success), while those that did not meet the previous criterion received a score of 1 (failure).
- STAY48-3: Challenged females that produced three calves by 48 months of age received a score of 2 (success), while those that did not meet the previous criterion received a score of 1 (failure).
- STAY54-2: Challenged females that produced two calves by 54 months of age received a score of 2 (success), while those that did not meet the previous criterion received a score of 1 (failure).

- STAY54-3: Challenged females that produced 3 calves by 54 months of age received a score of 2 (success), while those that did not meet the previous criterion received a score of 1 (failure).
- STAY72-3: Challenged females that produced 3 calves by 72 months of age were given a score of 2 (success), while those that did not meet the above criterion were given a score of 1 (failure).

The relationship matrix contained 720,239 animals born from 1990 to 2019 with an average inbreeding coefficient of 0.01% and the percentage of inbred individuals being 0.02%. Contemporary groups (CG) were defined as the combination of farm, management group, year and birth season (dry season from April to September and the rainy season from October to March). The data structure and descriptive statistics for different STAY definitions are presented in Table 1.

2.3 | Genotypes and Quality Control

A total of 402,130 animals were genotyped with the low-density panel (Clarifide Nellore 3.0) containing more than 54,000 markers. The genotype quality control was performed using the PreGSF90 program (Aguilar et al. 2014). Data quality control removed: samples and markers with a call rate <0.90, markers located on sex and mitochondrial chromosomes, markers with mendelian conflicts >1%, minor allele frequency <0.05, markers with redundant position and markers with deviation from Hardy Weinberg Equilibrium ($p < 10^{-10}$). After quality control 42,349 SNPs were retained in the database for analysis.

2.4 | Statistical Model and Genetic Analyses

Univariate statistical analyses were performed to predict breeding values by fitting threshold and linear models, using single-step genomic best linear unbiased prediction (ssGBLUP). Computations were done using CBLUP90IOD2OMP1 (v. 3.38) for threshold models and BLUP90IOD3 (v. 3.138) for linear models (Misztal et al. 2014). Variance components for STAY were transformed from liability to observed scale to obtain the genomic estimated breeding values (GEBV) using linear models. This conversion process relied on the formula originally developed

by Robertson (refer to Dempster and Lerner (1950)), which was later expanded to multiple-categorical traits by Gianola (1979).

The linear model, under matrix notation, can be written as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e},$$

where: $\mathbf{y}$ is the vector of phenotypic observations; $\boldsymbol{\beta}$ is the vector of fixed effects (CG); $\mathbf{u}$ is the vector of random additive animal effects; $\mathbf{e}$ is the vector of random residuals; $\mathbf{X}$ is the incidence matrix relating the effects in $\boldsymbol{\beta}$ to observations in $\mathbf{y}$; and $\mathbf{Z}$ is the incidence matrix relating the levels in $\mathbf{u}$ to observations in $\mathbf{y}$.

Under multivariate normality assumption, the (co)variance matrix for the random effects is given by:

$$\text{Var} \begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{H}\sigma_u^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_e^2 \end{bmatrix},$$

where: $\mathbf{H}$ is a joint covariance matrix for non-genotyped and genotyped animals (Legarra et al. 2009); $\mathbf{I}$ is an identity matrix of proper order; σ_u^2 is the additive genetic variance for the trait being analysed; and σ_e^2 is the residual variance.

The inverse of $\mathbf{H}$ was computed as in Aguilar et al. (2010):

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

where $\mathbf{A}^{-1}$ is the inverse of pedigree-based relationship matrix for all animals included in the pedigree, $\mathbf{A}_{22}^{-1}$ is the inverse of the pedigree-based relationship matrix for the genotyped animals. The genomic relationship matrix ($\mathbf{G}$) was constructed based on the first method proposed by VanRaden (2008) as $\mathbf{G} = \mathbf{Z}\mathbf{Z}'$, in which $\mathbf{Z} = (\mathbf{M} - \mathbf{P}) / \left[2 \sum_{j=1}^n p_j(1 - p_j) \right]$ with $\mathbf{M}$ being a matrix with genotypes coded as {0, 1, 2, 5 (missing)}, number of rows equal to the number of genotyped animals and number of columns equal to the number of SNP. $\mathbf{P}$ is a matrix with twice the observed frequency of the second allele p at locus j (p_j) in the entire genotype file. It incorporates a blending approach in which 5% of the pedigree-based relationship matrix for genotyped animals ($\mathbf{A}_{22}$) is included to mitigate singularity issues.

TABLE 1 | Number of records (N) and contemporary groups (CG), mean (SD) and percentage (%) of animals within each category of several stayability definitions in Nellore cattle.

Trait	N	CG	Mean (SD)	Success (2) %	Failure (1) %
STAY48-2	182,064	2741	1.14 (0.35)	26,152 (14.36%)	155,912 (85.24%)
STAY48-3	182,064	2741	1.13 (0.33)	23,685 (13.01%)	158,379 (86.99%)
STAY54-2	181,613	2717	1.11 (0.31)	20,865 (11.49%)	160,748 (88.51%)
STAY54-3	181,613	2717	1.11 (0.31)	20,040 (11.03%)	161,573 (88.97%)
STAY72	646,918	15,967	1.38 (0.48)	246,079 (38.03%)	400,839 (61.97%)

Note: STAY48-2: female's ability to remain in the herd up to 48 months of age, calving to at least two times; STAY48-3: female's ability to remain in the herd up to 48 months of age, calving to at least three times; STAY54-2: female's ability to remain in the herd up to 54 months of age, calving to at least two times; STAY54-3: female's ability to remain in the herd up to 54 months of age, calving to at least three times; STAY72-3: female's ability to remain in the herd up to 72 months of age, calving to at least three times.

Additionally, $\mathbf{G}$ was rescaled to align with $\mathbf{A}_{22}$ by adjusting its diagonal and off-diagonal elements (VanRaden 2008). The GEBV was calculated using the algorithm for proven and young (APY; Misztal et al. 2014) to ensure computational feasibility and efficiency. The APY method relies on a recursive approach applied to a subset of animals for the inversion of the $\mathbf{G}$ matrix, which is divided into core (c: 20,105) and non-core (n: 382,025) animals. The number of core animals is selected to approximate the dimensionality of $\mathbf{G}$ or the number of effective single nucleotide polymorphisms (SNPs) (Misztal et al. 2014).

A threshold model was implemented via Bayesian inference, considering CG as a systematic effect and animal additive as a random effect. Under matrix notation, the model for a latent variable (liability) is: $\mathbf{l} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$, where: $\mathbf{l}$ is the vector of liabilities; $\boldsymbol{\beta}$, $\mathbf{u}$, $\mathbf{e}$, $\mathbf{X}$ and $\mathbf{Z}$, were defined previously. Given $\boldsymbol{\beta}$ and $\mathbf{u}$, the elements of $\mathbf{l}$ are conditionally independent and distributed as: $\mathbf{l}|\boldsymbol{\beta}, \mathbf{u}, \sigma_u^2, \sigma_e^2 \sim \text{MVN}\{\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}, \mathbf{I}\sigma_e^2\}$. It was also assumed that: $\boldsymbol{\beta}|\mathbf{I}, \sigma_\beta^2 \sim \text{MVN}\{\mathbf{0}, \mathbf{I}\sigma_\beta^2\}$; $\mathbf{u}|\mathbf{H}, \sigma_u^2 \sim \text{MVN}\{\mathbf{0}, \mathbf{H}\sigma_u^2\}$, where $\mathbf{H}$ is, the relationship matrix; and $\mathbf{e}|\mathbf{I}, \sigma_e^2 \sim \text{MVN}\{\mathbf{0}, \mathbf{I}\sigma_e^2\}$, where $\mathbf{I}$ is an identity matrix of proper order; $\sigma_u^2|\nu_u \sim \chi^{-2}(\nu_u, S_u^2)$ and $\sigma_e^2|\nu_e \sim \chi^{-2}(\nu_e, S_e^2)$, being χ^{-2} an inverted chi-square distribution with ν degrees of belief and S scale matrix.

2.5 | Variance Components and Heritability Estimates

The variance components were estimated via Gibbs sampling implemented in the GIBBSf90+ software (Misztal et al. 2014) using the default prior for random and systematic effects (i.e., Gaussian prior for random effects and flat for systematic). In the Gibbs Sampling implementation, 300,000 iterations were employed, with an initial burn-in of 100,000 iterations and a sampling interval of 100 iterations (thin). The POSTGIBBSf90 software (Misztal et al. 2002) was used for posterior inference. Convergence was assessed through visual inspection of the generated Markov chains and by applying convergence diagnostics as follows: (1) Geweke's test (Geweke 1992), which compares the initial and final segments of the chain to detect convergence failures.

Rejection of the null hypothesis in this test indicates convergence, with probabilities below 0.05 providing evidence against chain convergence (Buzanskas et al. 2010); (2) Heidelberg and Welch test (Welch and Heidelberger 1983) was utilised, employing Cramer-von Mises statistics to evaluate the null hypothesis of stationarity in the generated samples and (3) Raftery and Lewis test (Raftery and Lewis 1992). All methods are implemented in the 'boa' R package (Smith 2007).

Variance components were estimated on the liability scale and then transformed to the observed scale to implement a linear-linear model for predicting breeding values. The transformation was based on the following formula (Dempster and Lerner 1950), derived for binary traits and described by Hidalgo, Tsuruta, et al. (2024): $h_o^2 = z^2 h_l^2 / \alpha (1 - \alpha)$, where h_o^2 is the additive heritability on the observed scale, z^2 is the height of the ordinate of the standard normal probability density function at

the point corresponding to the threshold between categories, estimated from the incidence of success (α) of the trait, h_l^2 is the heritability on the liability scale. Also, the formula allowed to obtain the additive ($\sigma_u^2 = z^2 h_l^2$) and phenotypic ($\sigma_p^2 = \alpha (1 - \alpha)$) variances on the observed scale. The residual variance in the observed scale was obtained by $\sigma_e^2 = \sigma_p^2 - (\sigma_u^2 + \sigma_{pe}^2 + \sigma_{cg}^2)$, where σ_{pe}^2 and σ_{cg}^2 are the random effects of the permanent environment and the contemporary group of the cow, in our case zero.

2.6 | Evaluation of Model Performance: Prediction Accuracy, Bias, Dispersion and Correlation

Scenarios were compared using the LR (Legarra and Reverter 2018) method. The whole dataset (represented by subscript w) included all sources of information for all animals (Table 2). The partial dataset was created by removing the last 2 years of data. This was done to ensure an enough number of focal animals for validation. Focal animals were young, genotyped females without their own (or progeny) records in the partial dataset.

The validation statistics were accuracy (acc), bias (δ), dispersion (b_1) and correlation (corr), computed as follows:

$$\begin{aligned} \text{acc} &= \sqrt{\frac{\text{cov}(\hat{\mathbf{u}}_w, \hat{\mathbf{u}}_p)}{(1 - \bar{F})\hat{\sigma}_u^2}}, \\ \delta &= \frac{\bar{\hat{\mathbf{u}}}_p - \bar{\hat{\mathbf{u}}}_w}{\hat{\sigma}_u}, \\ b_1 &= \frac{\text{cov}(\hat{\mathbf{u}}_w, \hat{\mathbf{u}}_p)}{\text{var}(\hat{\mathbf{u}}_p)}, \\ \text{corr} &= \frac{\text{cov}(\hat{\mathbf{u}}_w, \hat{\mathbf{u}}_p)}{\sqrt{\text{var}(\hat{\mathbf{u}}_w)\text{var}(\hat{\mathbf{u}}_p)}}, \end{aligned}$$

where: $\text{cov}(\cdot)$ is the sample covariance; $\hat{\mathbf{u}}_w$ is the vector of GEBV from the whole dataset; $\hat{\mathbf{u}}_p$ is the vector of GEBV from the partial dataset; $\bar{F}$ is the average pedigree-inbreeding coefficient among focal animals; $\hat{\sigma}_u^2$ is the estimated additive genetic variance; $\bar{\hat{\mathbf{u}}}_p$ is the average of $\hat{\mathbf{u}}_p$ (likewise for $\bar{\hat{\mathbf{u}}}_w$); and $\text{var}(\cdot)$ is the sample variance.

TABLE 2 | Number of records in the whole and partial datasets and validation animals used in the linear regression method for the different scenarios for the longevity trait in the Nellore breed.

	STAY 48	STAY 54	STAY72
Whole	182,064	181,613	646,918
Partial	172,852	172,852	610,623
Validation ^a	9212	8761	10,704

^aYoung animals with genotypes and phenotypes.

3 | Results and Discussion

The heritability estimates for the different categories of the STAY trait were low (Table 3), ranging from 0.16 to 0.22 on the liability scale. When transformed to the observed scale, the heritability estimates ranged from 0.07 to 0.09 and were lower than those obtained on the liability scale.

The genomic prediction parameters for the validation animals for each trait are presented in Table 4. Ramirez-Valverde et al. (2001) and Koeck et al. (2010) reported a high correlation (≥ 0.96) between EBVs of linear and threshold models. Varona et al. (1999) found no advantage of the threshold model over the linear model for calving ease. Similar results between the two models were also observed by Weller et al. (1988) in evaluations of calf mortality and dystocia in Holstein cattle, as well as by Matos et al. (1997) in reproductive traits of sheep.

The accuracy values obtained for the different STAY definitions were higher in the threshold model, with values ranging from 0.531 to 0.698, compared to the linear model, whose values ranged from 0.451 to 0.532. Ramirez-Valverde et al. (2001), when analysing the trait of calving difficulty in beef cattle, reported that the accuracy of the threshold model EBVs was 10% higher than that of linear model EBVs, in line with our results as

we found improvements in accuracy from 10% (0.503 vs. 0.558; STAY48-2) and up to 35% (0.451 vs. 0.698; STAY72-3). Our results indicate that the model based on the observable scale led to a reduction in accuracy, while increasing bias compared to models operating on the liability scale (Table 4). As an illustration, for STAY72-3 the bias using a threshold model was -0.017 , whereas it was 0.096 when using a linear model.

Considering the categorical nature of STAY, the application of the threshold model is theoretically the most appropriate choice, as proposed by Gianola (1982). However, from a practical standpoint, the linear model offers greater applicability (Varona et al. 1999; Silvestre et al. 2019). Another important factor to consider is the increased computational effort required by the threshold model. Although advantages of the threshold model over linear models have been reported in simulated data (Meijering and Gianola 1985), studies using real data have yielded variable results (Silvestre et al. 2019; Matos et al. 1997).

Conceptually, the greater the data availability, the more robust the approximation of linear models can become, even with deviations from normality (Matos et al. 1997). However, when considering the evaluation of reproduction traits, such as fertility expressed as a binary trait, litter size with four categories and ovulation rate with five categories in a small

TABLE 3 | Posterior means and high probability density (HPD^a) for direct additive genetic (σ_a^2), residual (σ_e^2) variances and heritability (SD, standard deviation) for different categories of stayability in Nellore cattle.

Trait	σ_a^2	σ_a^{2b}	σ_e^2	σ_e^{2b}	h^2	h^{2b}	$\pm$ SD	HPD
STAY48-2	0.24	0.010	1.01	0.113	0.20	0.08	0.01	0.16 to 0.22
STAY48-3	0.26	0.009	1.00	0.104	0.20	0.08	0.01	0.17 to 0.23
STAY54-2	0.27	0.008	1.00	0.094	0.22	0.07	0.02	0.18 to 0.24
STAY54-3	0.28	0.008	1.00	0.090	0.22	0.07	0.02	0.19 to 0.25
STAY72-3	0.19	0.023	1.00	0.213	0.16	0.09	0.01	0.14 to 0.18

Note: STAY48-2: female's ability to remain in the herd up to 48 months of age, calving to at least two times; STAY48-3: female's ability to remain in the herd up to 48 months of age, calving to at least three times; STAY54-2: female's ability to remain in the herd up to 54 months of age, calving to at least two times; STAY54-3: female's ability to remain in the herd up to 54 months of age, calving to at least three times; and STAY72-3: female's ability to remain in the herd up to 72 months of age, calving to at least three times.

^aCredibility intervals at 95%.

^bObservable scale.

TABLE 4 | Genomic prediction accuracy, bias, dispersion and correlation of genomic predictions for different stayability definitions in Nellore cattle using threshold (Th) and linear models (Li).

Trait	Accuracy		Bias		Dispersion		Correlation	
	Th	Li	Th	Li	Th	Li	Th	Li
STAY48-2	0.558	0.503	0.002	0.027	0.997	0.842	0.997	0.986
STAY48-3	0.531	0.532	0.008	0.028	0.932	0.847	0.955	0.986
STAY54-2	0.551	0.510	0.135	0.031	0.974	0.848	0.892	0.988
STAY54-3	0.567	0.514	-0.0004	0.031	1.000	0.848	1.000	0.988
STAY72-3	0.698	0.451	-0.017	0.096	0.977	0.811	0.995	0.959

Note: STAY48-2: female's ability to remain in the herd up to 48 months of age, calving to at least two times; STAY48-3: female's ability to remain in the herd up to 48 months of age, calving to at least three times. STAY54-2: female's ability to remain in the herd up to 54 months of age, calving to at least two times; STAY54-3: female's ability to remain in the herd up to 54 months of age, calving to at least three times and STAY72-3: female's ability to remain in the herd up to 72 months of age, calving to at least three times.

population of sheep, Matos et al. (1997) did not identify superiority of the threshold model in relation to the linear model, since the predictive abilities were similar: 0.21 for the linear model and 0.20 for the threshold model. Casellas et al. (2007) also reported minimal differences in the predictive ability of linear and threshold models in a similar study for the genetic evaluation of litter size and days until parturition in a sheep population. Carlén et al. (2006) reported in a simulation study that there were minimal gains in replacing the linear model with a threshold model or with survival analysis in the genetic evaluation of mastitis in dairy cattle.

Regarding bias, the lowest values were generally observed for the threshold model, while the linear model yielded slightly higher bias. The results indicated that the bias estimates obtained with the threshold model were close to zero (ranging from -0.0004 to 0.008), except for STAY54-2, which presented a higher bias value (0.135). For the linear model, bias values were slightly higher, ranging from 0.031 to 0.096 , except for STAY54-2, where the bias (0.031) was lower than that observed with the threshold model. Generally, more biased predictions were associated with lower accuracy.

The results indicated that the dispersion values for the linear model differed from the expected value of 1 (0.811 to 0.827). In contrast, dispersion values observed in the threshold model were closer to or at the ideal value (0.932 to 1.00). The STAY is defined as the probability of a female remaining in the herd until a certain age, reflecting her capacity for reproductive efficiency and is relevant for the new classifications proposed for those early females, which, according to the descriptive table, can observe a similar success rate. This pattern is due to the low reconception rate in the second reproductive season, as STAY is a trait with a high influence from the environmental component (i.e., low heritability) and high culling of animals for different reasons, such as being empty at the end of the reproductive season.

For example, for females that had two and three calvings by 48 months of age, the success rate was 14.36% and 13.01% , respectively. The success rate was even lower for females that had two and three calvings by 54 months of age, 11.49% and 11.03% , respectively. The latter can be a possible explanation for STAY54-2 presenting higher biases and lower genetic correlations of genetic values in both models; SATY72-3, for example, showed a success rate of 38.03% and better LR statistics, that is, higher accuracy and correlations, and lower bias and dispersion.

The imbalance in class frequencies can directly influence the predictive ability of the models, especially when using the linear model. In low success rate scenarios such as the 48- and 54-month classifications, where there are high discarding rates, linear models tend to underestimate the breeding values, as demonstrated in the results found in the bias and dispersion values. Even though linear models do not consider the binary nature of binary traits, the correlation of genetic values between the linear and threshold models followed the same trend, approximately close to 1, except for STAY54-2, where the correlation values of the threshold and linear models were 0.892 and 0.988 , respectively. In practical terms, the reclassification of animals in both models would be minimal.

According to Kizilkaya et al. (2014), in certain situations, linear models can present similar efficiency to threshold models. This behaviour is expected as the number of categories increases and approaches normality. The least favourable situation is for binary data, as predictive capacity declines as the number of categories, as reported by the same authors. In their results, these authors also reported that the efficiency of the analysis of binomial data transformed into continuous data also tends to reduce as the observed frequency moves away from 0.5 (i.e., with more imbalance between categories).

Although the threshold model is theoretically the most appropriate approach for analysing categorical traits, practical results do not always conform to this expectation. Working with longevity in the Italian Jersey breed, Fabris et al. (2025) demonstrated that EBVs can be estimated using a linear model with nearly equivalent performance, but with less computational effort and greater routine feasibility. Previous studies have also reported contrasting outcomes, showing cases in which threshold models performed better (Colonia et al. 2022; Camargo Júnior et al. 2025), worse (Carlén et al. 2006) or comparably to linear approaches (Campos et al. 2018).

4 | Conclusion

Overall, converting the trait to a linear model resulted in increased bias and reduced accuracy of genomic predictions compared to the threshold model. The threshold model demonstrated superior predictive ability (in terms of bias, dispersion and accuracy), particularly for the earlier definitions (STAY48 and STAY54), and is therefore the most suitable approach for genomic prediction of stayability in this dataset.

Among the different STAY definitions evaluated for early-challenged females, the STAY48 definition, requiring at least two or three calvings, stood out for its consistent and satisfactory performance in both models, representing a promising alternative for use in genetic evaluations. On the other hand, although the prediction accuracies of the earlier definitions are slightly lower, they represent more efficient traits for guiding culling and replacement decisions in breeding programs, as they facilitate informed decision-making, allow for greater selection intensity and consequently accelerate genetic and productive progress in the herd.

It is important to acknowledge the trade-off between predictive accuracy, heritability and the age at which the phenotype is measured for selection purposes. Although the latter definition, STAY72-3 (measured at 6 years of age), presented slightly higher accuracy values and lower bias in the threshold model, this category showed lower heritability (0.16) compared to the earlier definitions, whose heritabilities ranged between 0.20 and 0.22 . Thus, its use as a selection criterion tends to reduce the expected response to selection per unit of time, since the decision is postponed to later stages of the productive life. Therefore, the use of the definitions measured at 48 or 54 months is more appropriate, depending on the specific breeding objective of the production system.

Acknowledgements

The Article Processing Charge for the publication of this research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) (ROR identifier: 00x0ma614).

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

- Bermann, M., A. Legarra, M. K. Hollifield, Y. Masuda, D. Lourenco, and I. Misztal. 2021. "Validation of Single-Step GBLUP Genomic Predictions From Threshold Models Using the Linear Regression Method: An Application in Chicken Mortality." *Journal of Animal Breeding and Genetics* 138, no. 1: 4–13. <https://doi.org/10.1111/jbg.12507>.
- Brumatti, R. C., J. B. S. Ferraz, J. P. Eler, and E. I. B. Formigoni. 2011. "Desenvolvimento de Índice de Seleção em Gado Corte Sob o Enfoque de um Modelo Bioeconômico." *Archivos de Zootecnia* 60, no. 230: 205–213. <https://doi.org/10.4321/S0004-05922011000200005>.
- Buzanskas, M. E., D. A. Grossi, F. Baldi, et al. 2010. "Genetic Associations Between Stayability and Reproductive and Growth Traits in Canchim Beef Cattle." *Livestock Science* 132, no. 1–3: 107–112. <https://doi.org/10.1016/j.livsci.2010.05.008>.
- Camargo Júnior, R. N. C., C. V. D. Araújo, M. Gomes, et al. 2025. "Comparison Between Linear Mixed Model and Threshold Model in the Estimation of Variance Components in Age at First Calving and Milk Production in Buffaloes." *Frontiers in Veterinary Science* 12: 1649690. <https://doi.org/10.3389/fvets.2025.1649690>.
- Campos, G. S., F. A. Reimann, P. I. Schimdt, et al. 2018. "Threshold and Linear Models for Genetic Evaluation of Visual Scores in Hereford and Braford Cattle." *Animal Production Science* 59, no. 4: 619–627. <https://doi.org/10.1071/AN17436>.
- Cappelloni, M., M. Gallo, and A. Cesarani. 2022. "Use of Threshold and Linear Models to Estimate Variance Components and Breeding Values for Disease Resistance in Italian Heavy Pigs." *Italian Journal of Animal Science* 21, no. 1: 488–492. <https://doi.org/10.1080/1828051X.2022.2050471>.
- Carlén, E., U. Emanuelson, and E. Strandberg. 2006. "Genetic Evaluation of Mastitis in Dairy Cattle Using Linear Models, Threshold Models, and Survival Analysis: A Simulation Study." *Journal of Dairy Science* 89, no. 10: 4049–4057. [https://doi.org/10.3168/jds.S0022-0302\(06\)72448-1](https://doi.org/10.3168/jds.S0022-0302(06)72448-1).
- Cavani, L., D. A. Garcia, L. O. D. Carreño, et al. 2015. "Estimates of Genetic Parameters for Reproductive Traits in Brahman Cattle Breed." *Journal of Animal Science* 93, no. 7: 3287–3291. <https://doi.org/10.2527/jas.2015-8970>.
- Clutter, A. C., P. J. Berger, and J. M. Mattison. 1989. "Threshold Model Analysis of Dystocia in Dairy Cattle When Progeny Information Is Limited." *Journal of Dairy Science* 72, no. 12: 3264–3272. [https://doi.org/10.3168/jds.S0022-0302\(89\)79487-X](https://doi.org/10.3168/jds.S0022-0302(89)79487-X).
- Colonia, S. R. R., A. D. C. Oliveira, F. Pilonetto, et al. 2022. "Genetic Parameters for Milk Yield, Casein Percentage, Subclinical Mastitis Incidence and Sexual Precocity Using Bayesian Linear and Threshold Models." *Animal Production Science* 62, no. 8: 792–801. <https://doi.org/10.1071/AN20313>.
- Costa, E. V., H. T. Ventura, R. Veroneze, F. F. Silva, M. A. Pereira, and P. S. Lopes. 2019. "Bayesian Linear-Threshold Censored Models for Genetic Evaluation of Age at First Calving and Stayability in Nellore Cattle." *Livestock Science* 230: 103833. <https://doi.org/10.1016/j.livsci.2019.103833>.
- Fabris, A., F. Tiezzi Mazzoni Della Stella Maestri, R. Finocchiario, M. Marusi, and M. Cassandro. 2025. "Genetic Evaluation of Longevity in Italian Jersey." *Italian Journal of Animal Science* 24, no. 1: 68–77. <https://doi.org/10.1080/1828051X.2024.2440124>.
- Geweke, J. 1992. "Evaluating the Accuracy of Sampling Based Approaches to the Calculation of Posterior Moments." *Bayesian Statistics* 4: 169–193.
- Gianola, D. 1982. "Theory and Analysis of Threshold Characters." *Journal of Animal Science* 54, no. 5: 1079–1096. <https://doi.org/10.2527/jas1982.5451079x>.
- Heringstad, B., R. Rekaya, D. Gianola, G. Klemetsdal, and K. A. Weigel. 2003. "Genetic Change for Clinical Mastitis in Norwegian Cattle: A Threshold Model Analysis." *Journal of Dairy Science* 86, no. 1: 369–375. [https://doi.org/10.3168/jds.S0022-0302\(03\)73615-7](https://doi.org/10.3168/jds.S0022-0302(03)73615-7).
- Hidalgo, J., I. Misztal, S. Tsuruta, et al. 2024. "Transforming Estimated Breeding Values From Observed to Probability Scale: How to Make Categorical Data Analyses More Efficient." *Journal of Animal Science* 102: skae307. <https://doi.org/10.1093/jas/skae307>.
- Hidalgo, J., S. Tsuruta, D. Gonzalez, et al. 2024. "Converting Estimated Breeding Values From the Observed to Probability Scale for Health Traits." *Journal of Dairy Science* 107, no. 11: 9628–9637. <https://doi.org/10.3168/jds.2024-24767>.
- Kizilkaya, K., R. L. Fernando, and D. J. Garrick. 2014. "Reduction in Accuracy of Genomic Prediction for Ordered Categorical Data Compared to Continuous Observations." *Genetics Selection Evolution* 46: 1–13. <https://doi.org/10.1186/1297-9686-46-37>.
- Koeck, A., B. Heringstad, C. Egger-Danner, C. Fuerst, and B. Fuerst-Waltl. 2010. "Comparison of Different Models for Genetic Analysis of Clinical Mastitis in Austrian Fleckvieh Dual-Purpose Cows." *Journal of Dairy Science* 93, no. 9: 4351–4358. <https://doi.org/10.3168/jds.2009-2955>.
- Matos, C. A. P., D. L. Thomas, D. Gianola, M. Perez-Enciso, and L. D. Young. 1997. "Genetic Analysis of Discrete Reproductive Traits in Sheep Using Linear and Nonlinear Models: II. Goodness of Fit and Predictive Ability." *Journal of Animal Science* 75, no. 1: 88–94. <https://doi.org/10.2527/1997.75188x>.
- Meijering, A., and D. Gianola. 1985. "Linear Versus Nonlinear Methods of Sire Evaluation for Categorical Traits: A Simulation Study." *Genetics Selection Evolution* 17, no. 1: 115–132.
- Misztal, I., D. Gianola, and J. L. Foulley. 1989. "Computing Aspects of a Nonlinear Method of Sire Evaluation for Categorical Data." *Journal of Dairy Science* 72, no. 6: 1557–1568. [https://doi.org/10.3168/jds.S0022-0302\(89\)79267-5](https://doi.org/10.3168/jds.S0022-0302(89)79267-5).
- Misztal, I., A. Legarra, and I. Aguilar. 2014. "Using Recursion to Compute the Inverse of the Genomic Relationship Matrix." *Journal of Dairy Science* 97: 3943–3952. <https://doi.org/10.3168/jds.2013-7752>.
- Misztal, I., S. Tsuruta, T. Strabel, B. Auvray, T. Druet, and D. Lee. 2002. "BLUPF90 and Related Programs (BGF90)." *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production* 28, no. 7: 743.
- Pereira, L. S., L. B. Temp, G. Gubiani, M. Teodoro, C. U. Magnabosco, and F. Baldi. 2025. "Stayability in the Era of Early-Challenged Females: Genetic Parameters and Correlations With Economically Relevant Traits." *Journal of Animal Breeding and Genetics*. <https://doi.org/10.1111/jbg.70026>.

- Phocas, F., C. Bloch, P. Chapelle, F. Bécherel, G. Renand, and F. Ménéssier. 1998. "Developing a Breeding Objective for a French Purebred Beef Cattle Selection Programme." *Livestock Production Science* 57, no. 1: 49–65. [https://doi.org/10.1016/S0301-6226\(98\)00157-2](https://doi.org/10.1016/S0301-6226(98)00157-2).
- Portes, J. V., G. R. O. Menezes, L. O. C. Silva, et al. 2021. "Selection Indexes for Nellore Production System in the Brazilian Pantanal." *Revista Brasileira de Zootecnia* 50: e20200264. <https://doi.org/10.37496/rbz5020200264>.
- Raftery, A. E., and S. Lewis. 1992. "How Many Iterations in the Gibbs Sampler." *Bayesian Statistics* 4, no. 2: 763–773.
- Ramos, P. V. B., L. O. C. Silva, B. C. Perez, et al. 2021. "Stayability and Consecutive Rebreding Ability Associated to Carcass and Growth Traits in Brazilian Nellore Cattle: A Bayesian Framework." *Livestock Science* 245: 104416. <https://doi.org/10.1016/j.livsci.2021.104416>.
- Santana, M. L., J. P. Eler, A. B. Bignardi, and J. B. S. Ferraz. 2013. "Genetic Associations Among Average Annual Productivity, Growth Traits, and Stayability: A Parallel Between Nelore and Composite Beef Cattle." *Journal of Animal Science* 91, no. 6: 2566–2574. <https://doi.org/10.2527/jas.2012-5856>.
- Silva, D. O., G. A. Fernandes Júnior, L. F. S. Fonseca, et al. 2024. "Genome-Wide Association Study for Stayability at Different Calvings in Nellore Beef Cattle." *BMC Genomics* 25, no. 1: 1–14. <https://doi.org/10.1186/s12864-024-10020-y>.
- Silvestre, A., Â. Martins, V. Santos, and J. Colaço. 2019. "Genetic Parameters of Calving Ease in Dairy Cattle Using Threshold and Linear Models." *Italian Journal of Animal Science* 18, no. 1: 80–87. <https://doi.org/10.1080/1828051X.2018.1482801>.
- Smith, B. J. 2007. "BOA: An R Package for MCMC Output Convergence Assessment and Posterior Inference." *Journal of Statistical Software* 21: 1–37.
- Teixeira, D. B. A., G. A. Fernandes Júnior, D. B. S. Silva, et al. 2017. "Genomic Analysis of Stayability in Nellore Cattle." *PLoS One* 12, no. 6: e0179076. <https://doi.org/10.1371/journal.pone.0179076>.
- Varona, L., I. Misztal, and J. K. Bertrand. 1999. "Threshold-Linear Versus Linear-Linear Analysis of Birth Weight and Calving Ease Using an Animal Model: II. Comparison of Models." *Journal of Animal Science* 77, no. 8: 2003–2007. <https://doi.org/10.2527/1999.7782003x>.
- Welch, P. D., and P. Heidelberger. 1983. "Simulation Run Length Control in Presence of an Initial Transient." *Operations Research* 31: 1109–1144.
- Weller, J. I., I. Misztal, and D. Gianola. 1988. "Genetic Analysis of Dystocia and Calf Mortality in Israeli-Holsteins by Threshold and Linear Models." *Journal of Dairy Science* 71, no. 9: 2491–2501. [https://doi.org/10.3168/jds.S0022-0302\(88\)79836-7](https://doi.org/10.3168/jds.S0022-0302(88)79836-7).
- Aguilar, I., I. Misztal, S. Tsuruta, A. Legarra, and H. Wang. 2014. "PREGSF90–POSTGSF90: Computational Tools for the Implementation of Single-Step Genomic Selection and Genome-Wide Association With Ungenotyped Individuals in BLUPF90 Programs." In *Proceedings of the 10th World Congress of Genetics Applied to Livestock Production*, vol. 3. WCGALP, Vancouver, BC, Canada.
- Gianola, D. 1979. "Heritability of Polychotomous Characters." *Genetics* 93, no. 4: 1051–1055. <https://doi.org/10.1093/genetics/93.4.1051>.
- Dempster, E. R., and I. M. Lerner. 1950. "Heritability of Threshold Characters." *Genetics* 35, no. 2: 212–236. <https://doi.org/10.1093/genet/ics/35.2.212>.
- Legarra, A., I. Aguilar, and I. Misztal. 2009. "A Relationship Matrix Including Full Pedigree and Genomic Information." *Journal of Dairy Science* 92, no. 9: 4656–4663. <https://doi.org/10.3168/jds.2009-2061>.
- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. "Hot Topic: A Unified Approach to Utilize Phenotypic, Full Pedigree, and Genomic Information for Genetic Evaluation of Holstein Final Score." *Journal of Dairy Science* 93, no. 2: 743–752. <https://doi.org/10.3168/jds.2009-2730>.
- VanRaden, P. M. 2008. "Efficient Methods to Compute Genomic Predictions." *Journal of Dairy Science* 91, no. 11: 4414–4423. <https://doi.org/10.3168/jds.2007-0980>.
- Legarra, A., and A. Reverter. 2018. "Semi-Parametric Estimates of Population Accuracy and Bias of Predictions of Breeding Values and Future Phenotypes Using the LR Method." *Genetics Selection Evolution* 50, no. 1. <https://doi.org/10.1186/s12711-018-0426-6>.
- Ramirez-Valverde, R., I. Misztal, and J. K. Bertrand. 2001. "Comparison of Threshold vs Linear and Animal vs Sire Models for Predicting Direct and Maternal Genetic Effects on Calving Difficulty in Beef Cattle." *Journal of Animal Science* 79, no. 2: 333. <https://doi.org/10.2527/2001.792333x>.