Livestock Science 154 (2013) 28–33

Contents lists available at SciVerse ScienceDirect

Livestock Science

journal homepage: www.elsevier.com/locate/livsci

Age at first calving of Nellore cattle in the semi-arid region of northeastern Brazil using linear, threshold, censored and penalty models

Carlos Henrique Mendes Malhado [a,e,*], Ana Claudia Mendes Malhado [b], Raimundo Martins Filho [c], Paulo Luiz Souza Carneiro [a], Akin Pala [d], José Adrián Carrillo [e]

[a] Universidade Estadual do Sudoeste da Bahia, Av: José Moreira Sobrinho, Jequié, BA 45200-000, Brazil
[b] Universidade Federal do Alagoas, Av. Lourival Melo Mota, s/n, Tabuleiro do Martins, Maceió, AL 57072-900, Brazil
[c] Universidade Federal do Ceará, Campus do Cariri, Avenida Tenente Raimundo Rocha, Cidade Universitária, Juazeiro do Norte, CE 63000-000, Brazil
[d] Canakkale Onsekiz Mart University, Ziraat Fak., ofis 222, Canakkale, Turkey
[e] Department of Animal and Avian Sciences, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Article history:
Received 3 July 2012
Received in revised form
23 January 2013
Accepted 27 February 2013

Keywords:
Data splitting
Genetic parameter
Missing method
Heritability

ABSTRACT

Data are typically discarded when there are inconsistencies in a database, or when information is discrepant or out of range for the biological characteristics of the species being analyzed. However, such loss of information could have considerable implications for genetic evaluation of cattle. Here, we use different models to estimate genetic parameters for age at first calving in order to prevent data elimination. We used five approaches based on trait distribution to define the limits of censure/disposal: linear model, censorship, penalty and missing methods, and threshold model (binary and polychotomous). Data splitting and Pearson correlation were used to evaluate fitting and comparison of models. The lowest heritabilities were estimated for the missing method and the binary model. Exclusion of outliers from the data considerably affects the estimation of genetic parameters and the ranking of sires. Moreover, models that suffered from data elimination generated the worst classifications in terms of the comparison of models. Researchers should be extremely careful when deciding to discard data. For example, an age at first calving up to 72 months could be considered an outlier for other countries or even other areas in Brazil. Although this value is very high, it could be realistic under arid or semi-arid conditions. In these situations, the penalty or censure models are the most appropriate methods of analyses.

1. Introduction

The breeding of beef cattle in Brazil is generally well established and adheres to modern principles to ensure high productivity. However, in the northeastern region (particularly in the semi-arid zone) livestock rearing still follows traditional practices with low levels of contemporary practices innovation.

Growth related traits are the traditional targets for selection programs in beef cattle and, unsurprisingly, the majority of research studies in the northeastern region of

* Corresponding author at: Universidade Estadual do Sudoeste da Bahia Av: José Moreira Sobrinho 45200-000 Jequié, BA, Brazil. Tel.: +55 7391919131.
E-mail addresses: carlosmalhado@gmail.com (C.H.M. Malhado), anaclaudiamalhado@gmail.com (A.C.M. Malhado), rmartinsfilho@yahoo.com.br (R. Martins Filho), plscarneiro@gmail.com (P.L.S. Carneiro), dr.akin.pala@gmail.com (A. Pala), carrillo@umd.edu (J. Adrián Carrillo).

1871-1413/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.livsci.2013.02.021

Brazil have focused on these characteristics (e.g. Malhado et al., 2008). However, reproductive traits appear to be the most economically important characteristic in many production systems and breeding operations, strongly affecting the efficiency of cattle rearing. Herds in semi-arid tropical regions typically show low reproductive efficiency and breeding strategies are often severely compromised by the inaccuracy of the available reproductive records.

Selection based on reproductive information provides several challenges (Urioste et al., 2007a). Most notably, reproductive performance is a complex trait made up of many components, and no-completely satisfactory measure for it has been established (Rust and Groeneveld, 2001).

Survival analysis and adapted linear models (based on census data) have been recently proposed to estimate sires breeding values for days open in dairy cattle (Gonzalez-Recio et al., 2006). Both analyses gave similar sires ranking, but a linear model adapted for census data was easier to implement—considering the magnitude and scope of a national genetic evaluation program for days open in dairy cattle (Gonzalez-Recio et al., 2006). In a censored modeling approach, missing data are replaced using imputation (Urioste et al., 2007a) or augmentation methods (Gonzalez-Recio et al., 2006; Urioste et al., 2007a).

One of the reproductive traits that is most easily obtained and routinely registered in most breeding females is age at first calving (AFC) (Malhado et al., 2010; Boligon, Albuquerque, 2011). Moreover, AFC could potentially be analyzed employing a threshold approach. Usually the trait is analyzed by defining two categories, known as heifer pregnancy (Eler et al., 2004; Boligon, Albuquerque, 2011). However, a study with Nellore cattle argued that the use of polychotomous categories has a higher ability to detect genetic variation than the traditional binary approach (Marcondes et al., 2005).

Data sets on reproductive traits of Nelore cattle in Northeastern Brazil include missing data due to pasture mating—although the overall number of records collected per year tends to be high. Furthermore, a considerable proportion of the AFC records are extremely high and deviate from expected levels. These characteristics make complicate data handling and may produce results that are difficult to interpret. The reported values for AFC could include annotation errors, typing errors, or missing data due to non-reported numbers or incorrect reports.

In addition, the outliers could be caused by animals that had problems during development. The major reason for the delayed onset of puberty in the northeastern region is probably the poor body condition of the females due to the strong seasonality in forage production. This is often augmented by a lack of supplemental feeding during the dry season after weaning, when heifers are still growing (Azevêdo et al., 2006). Moreover, some breeders adopt reproductive management programs that only consider the weight of the heifers rather than their age, which also contributes to the high mean for the age at first calving. Understandably, researchers have difficulties distinguishing between the origins of these values and usually adopt a cut-off point that determines which data is discarded and which enters the database.

The consequences of this information disposal are largely unknown, but could have a detrimental impact on genetic evaluations. Here, we use different models to prevent such data elimination. The specific objective of this study being to estimate genetic parameters for age at first calving (AFC) in Nellore cattle from northeastern Brazil using five approaches: linear, censored, threshold, penalty and missing method.

2. Material and methods

2.1. Data

The Brazilian Association of Zebu Breeders (ABCZ) provided the data that included 4959 dams with recorded age at first calving. The animals were born between 1970 and 2010 in the semi-arid ecosystem in the center of the Northeast region of Brazil. All cows in the database had known sire and dam records. The mean age at first calving was 1418.46 ± 518.83 days (46.5 months).

2.2. Statistical models

We used the trait distribution to define the limits of censure or disposal (Fig. 1). Three different points were used: the third quartile (Q3) with a value of 59.11 months (59), the ninth decile (D9) with 75.44 (75) months and the ninth point seventy-fifth decile (D9.75) with 93.24 (93) months.

Five approaches were employed to evaluate AFC:

(1) Linear model (LIN): all animals were used in the analysis, without any restrictions.
(2) Censored traits (CEN): created by "integrating out" the conditional distribution of the censored data. If a censored phenotype y^* (age at first calving) was observed, and the effect affecting the phenotype (contemporary group) were known, the real, unobserved phenotype y^r follows a truncated normal distribution:
$f(yr|y^*;\ \text{contemporary group}) \sim N(\mu_{cg}, \sigma^2)\ I\ [a,b]$, being μ_{cg} the phenotypic mean of contemporary group, and

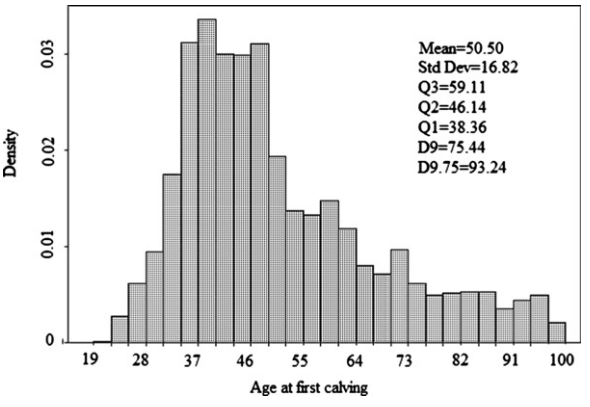

Fig. 1. Histogram and descriptive statistics for age at first calving.