Genetic evaluation studies in Tabapuã cattle breed using random regression models-Review

[Estudos de avaliações genéticas em bovinos da raça Tabapuã utilizando modelos de regressão aleatória-Revisão]

“Revisão/Review”

Wéverton José Lima Fonseca¹, Amauri Felipe Evangelista², Laylson da Silva Borges³, Gleissa Mayone Silva Vogado⁴, Thales de Lima Silva¹, Leandro de Oliveira Guerra⁵, João Marcos Monteiro Batista⁶, Kássio Alexandre Marques⁵, Severino Cavalcante de Sousa Júnior⁶

¹Programa de Pós-Graduação em Zootecnia, Universidade Estadual do Sudoeste da Bahia (UESB), Itapetinga-BA, Brasil.
²Programa de Pós-Graduação em Zootecnia, Universidade Federal do Paraná (UFPR), Curitiba-PR, Brasil.
³Programa de Pós-Graduação em Ciência Animal, Universidade Federal do Piauí (UFPI), Teresina-PI, Brasil.
⁴Universidade Federal do Piauí (UFPI), Bom Jesus-PI, Brasil.
⁵Programa de Pós-Graduação em Zootecnia, Universidade Federal do Paraíba (UFPB), Areia-PB, Brasil.
⁶Universidade Federal do Piauí (UFPI), Parnaíba-PI, Brasil.

*Autor para correspondência/Corresponding author: E-mail: fonsecawjl@gmail.com

Abstract

The objective of this research is to estimate the genetic parameters and covariance components for growth characteristics in cattle from Tabapuã breed through random regression models (RRM). The estimates of covariance functions are considered nowadays, interesting alternatives to work with longitudinal data that allow the description of covariances changes due to time and prediction of variances and covariances along the growth curve. The RRM has been used to estimate the variance components and genetic parameters of milk production on the control day of the lactation curve and meat production in many species, including bodyweight of beef cattle and among these the cattle from Tabapuã breed, enabling us to work with features of growth curves that are measured repeatedly during the animal's life, known as longitudinal data. These models used in genetic evaluation in cattle are considered special cases of covariance functions that allow estimation of coefficients of genetic covariance functions directly through the method of restricted maximum likelihood.

Keywords: longitudinal data; heritability; animal model.

Resumo

Objetivou-se com esta pesquisa, estimar os parâmetros genéticos e componentes de covariância para características de crescimento em bovinos da raça Tabapuã por meio de modelos de regressão aleatória (MRA). As estimativas de funções de covariância são consideradas, até hoje, alternativas interessantes para se trabalhar com dados longitudinais que permitem a descrição de mudanças das covariâncias em função do tempo e predição de variâncias e covariâncias ao longo da curva de crescimento. Os MRA vêm sendo utilizados para estimar os componentes de variância e parâmetros genéticos de produção de leite no dia do controle de curva de lactação e carne de diversas espécies, entre elas os pesos de bovinos de corte e dentre estes os de bovinos da raça Tabapuã, possibilitando trabalhar com características de curvas crescimento que são medidas repetidamente na vida do animal, denominados dados longitudinais. Estes modelos utilizados na avaliação genética em bovinos, são considerados casos especiais de funções de covariância que permitem estimar diretamente os coeficientes das funções de covariância genética através do método da máxima verossimilhança restrita.

Keywords: dados longitudinais; herdabilidade; modelo animal.
Introduction

One of the main objectives of beef cattle breeding is to obtain precocity in weight gain and finishing of the animals. Under this perspective, it is important to adopt animal breeding tools, especially selection of genetically superior animals for economic relevant characteristics. However, growth characteristics, weaning weight, yearling and weight gain at different ages stand out as selection criterion, presenting means to high magnitude heritability and providing higher genetic gains per generation at more advanced ages (Laureano et al., 2011).

Studies of genetic evaluations of the growth curve of beef cattle as body weight and measures in the early stages of the animal development, are important features in determining the economic efficiency and used in the genetic evaluation of breeding programs. The genetic evaluation of beef cattle depends essentially on the availability of estimation of genetic parameters for features of economic interest (Sousa Júnior et al., 2010).

The conventional analysis to estimate (co)variance components and genetic parameters of animals of the population as growth characteristics are performed by means of finite-dimensional models. However, it is necessary to estimate covariance components to obtain the parameters and genetic values, and the choice of methods used in estimation of covariance components and genetic parameters are fundamental in genetic evaluations (Lacerda et al., 2014).

The random regression models (RRM), called as infinite-dimensional models, are being used to estimate the variance components and genetic parameters of weight in beef cattle. The RRM have become an alternative standard for genetic analysis of longitudinal data, where one of the barriers of these models is related to availability of memory and computational time for conducting large-scale genetic evaluations (Sousa Júnior et al., 2014).

The features that are interesting to the genetic improvement, measured several times during the life of the animal, are defined as longitudinal data. The simplest model for this analysis is the repeatability model, assuming that all measures acquired overtime represent the same feature. These models assume unrealistic premises, since they consider that the same individual measures represent a single feature, and that the correlations between the various observations obtained from the same animal are equal to one (Morris and Amyes, 2012).

Generally, random regression models by Legendre polynomial functions up to cubic order have been used mostly for modeling of average growth trajectories of the animals. However, the differences in order for adjustment of average trajectories have influenced the variance functions, but little attention has been given to adjusting this effect (Scalez et al., 2014). Thereby, the RRM allows adjusting the random growth curve for each animal and obtaining the variance and covariance structures to the interested random effects. In addition, the use of these models allows modeling of residual variance under different structures (Sarmento et al., 2010).

Thus, the aim of this research was to present the studies that increased variance and genetic evaluation components using random regression models by Legendre orthogonal polynomials to growth characteristics, and stimulate more study initiatives on Tabapuã cattle, due to the great need to acknowledge those animals productive capacity.

Genetic parameters of growth

Aiming genetic parameter estimation for animal growth characteristics, Albuquerque and Meyer (2001), Nobre et al. (2003) and Dias et al. (2005) reported that when the heritability for birth weight (0.26 to 0.33) is higher, the estimates of time of weaning (0.11 to 0.16) decrease and these estimates increase according to age (0.21 and 0.25 to weight to 570 days old and 0.20 for weight to 683 days old). Research on adult weight of *Bos indicus* suggested the existence of genetic variation sufficient to obtain genetic change in animal breeding programs aiming to decrease, increase or keep it constant (Talhari et al., 2003; Mercadante et al., 2004).

According to Dias et al. (2005), several works on genetic parameter estimation for the Zebu breeds, until the nineties, were conducted with small number of records of the animals and the genetic parameters were estimated at predetermined ages, besides there was no separation of direct and maternal additive genetic effects. Genetic evaluations of components of (co)variance for growth characteristics, are obtained considering the weights at standard ages (birth weight, weaning, yearling and final) or the weight gains between the two ages, but estimation of parameters, including the maternal effect, for
weight at a young age range are scarce in the literature (Albuquerque and Meyer, 2001a).

For these parameters to be trusted, genetic evaluation methodologies must be applied impeccably so that predicted genetic values are accurately representing an effective instrument in the pursuit of genetic progress (Lopes et al., 2011). The genetic values for growth traits in beef cattle are predicted and the components of variance are estimated considering the weight to standard ages such as birth, weaning, annually and at the final age, using solo or multi-characteristics analysis (Albuquerque and El Faro, 2008).

The increase in recent discussions among geneticists of beef cattle is the better way to improve the accuracy of the existing genetic evaluations (Spiedel et al., 2010). Although most of this discussion is focused on the incorporation of information from DNA markers for current genetic evaluation schemes, increasing the number of usable records would also add accuracy to the expected differences in progeny (EDP) (Schaeffer, 2006; Goddard and Hayes, 2007). The use of random regression techniques can allow the incorporation of observations taken at any time through out the animal’s life, which by increasing the accuracy of the DEP’s, could result in an increase in the rate of genetic gain (Williams et al., 2009).

The increase in the production chain of beef cattle through animal genetics improvement program, has been achieved by selection, crossings and the combination of both (Bueno et al., 2012). The genetic evaluation in populations consisting of different breeds, combines the information from all animals belonging to the population under study, such as pure and mixed-race, and contemplates the analysis of direct racial and maternal effects, besides heterosis effects (Williams et al., 2010). The accuracy of prediction of the animal genetic value, will depend on the accurate estimation of the averages for specific racial composition of deviations of animals relating to the environment and the (co)variances among pure or mixed-race relatives (Cardoso and Tempelman, 2004).

Random regression models

The random regression models (RRM) make it possible to estimate coefficients of covariance functions using the restricted maximum likelihood method. These are suitable for analysis of longitudinal data, because they overcome the deficiency of conventional methods of quantitative genetic analyses, which intrinsically consider the phenotypic values (continuous processes) as discrete processes. The random regression coefficients related to time are obtained from each animal. The RRM allows adjusting the growth curves for each individual and obtaining variances and (co)variances structures between the different measures through the functions of (co)variances for the random effects of interest (Sarmento et al., 2010).

In recent years the random regression models (RRM) have been adopted as a standard technique for studies referred to the analysis of longitudinal data in genetic improvement. Currently this methodology has been applied in many species of domestic animals such as beef cattle (Dias et al., 2006) and milk cattle (Sousa Junior et al., 2014). In the RRM, the variances of heterogeneous residuals can be used to improve the partition of total variation, improving parameter estimation in the process of maximizing the restricted likelihood function (Neto et al., 2011). The study of longitudinal data such as growth feature, can be achieved by using random regression models that estimate genetic parameters at any age within each range considered (Sousa Júnior et al., 2010).

Another advantage of these models is to estimate with greater accuracy the genetic and phenotypic covariance components, as they consider measurement over the time, using information from all the observations, therefore obtaining estimates for the temporary environment variances or measurement errors (Kirkpatrick et al., 1990). The random regression coefficients can be adjusted for each source of variation of the model. The covariance functions (CF) are considered an alternative to work with longitudinal data by allowing the description of the gradual change of the (co)variances over time and the prediction of variances and (co)variances for points along the growth curve (Dias et al., 2006).

In Brazil, researches using random regression models have been conducted with the aim of studying growth characteristics of beef cattle (Sousa Júnior et al., 2012). According to Sakaguti et al. (2003) while studying the Tabapuã breed, the RRM was applied to estimate functions of (co)variance for weights between 365 and 650 days old, however, due to the computational limitations, linear and quadratic models were adjusted only to the effects of the animal, not considering the maternal effects. The authors
concluded that the random regression models allowed to estimate variance components at any age, as well as provided additional useful parameters to genetic evaluations of breed cattle from Tabapuã breed.

According to Albuquerque and Meyer (2001) works carried out in herds of Nelore cattle, estimated (co)variance genetic functions using these random regression models (RRM) for records of birth weights to 630 days old and concluded that the random regressions described properly the (co)variances changes with age. However, Nobre et al. (2003) analyzed the growth curve of the Nelore breed to (683 days old at birth), applying multi-characteristics models and random regression models, they observed that the random regression models were more sensitive to sampling problems compared with the multi-characteristics models.

**Longitudinal data**

Among the features that are interesting to the animal breeding measured several times during the lifetime of the animal are regarded as longitudinal data. Thus, the simplest model for the analysis of longitudinal data consists basically in the repeatability model, assuming that all measures acquired over time represent the same feature. It is possible to analyze characteristics of interest to animal breeding, measured repeatedly throughout the animal's life, such as monthly production of milk and weight gain, being these, called characteristics repeated in time or longitudinal (Sousa Júnior et al., 2010).

An alternative consists in considering each measure as a different characteristic and conduct a joint analysis on a multi-characteristics model, which is a common approach in analysis for several features of interest (Faria et al., 2007). Traditionally, these characteristics are evaluated through repeatability models. These models assume unrealistic premises, since they consider that the measurements on the same individual represent a single feature, and that the correlations between the various observations obtained from the same animal are equal to one (Morris and Amyes, 2012).

The characteristics can be analyzed by means of multi-characteristics models, assuming that in each adjusted weight measure obtained for standard ages in beef cattle are a different feature. Genetic multi-characteristics evaluations predict genetic values for different ages by the incorporation of genetic and residual (co)variances between the ages (Spiedel et al., 2010). Another important point in this type of analysis is that the number of parameters to be estimated grows significantly with the increasing number of features, offering an alternative in the genetic evaluation of longitudinal data, as in the case of body weights at different ages, because their resulting functions of (co)variance are the most suitable for the analysis of these models.

The random regression models have become standard on analysis of data from animal production systems, especially in the last decade (Henderson, 1982; Laird and Ware, 1982). This model presupposes that the genetic variances and permanent environment do not vary over time where the genetic and permanent environmental correlations between repeated measures are equal to the unit. According to Silva et al. (2008), it is possible for an animal A to be genetically superior to an animal B in certain age, but not in another, regarding an specific evaluated characteristic.

By describing studies in the area of animal breeding, it can be verified that the application of this methodology is, mostly, in modeling of waste correlated in studies evaluation of on animal non-linear growth curves (Pala et al., 2005). In case of estimating (co) variance components, most researches limited its use only for the purpose of permanent environment of the animal (Carvalheira et al., 1998; Albuquerque and Meyer, 2001). Therefore, a small number of works in which it was used the parameterization to estimate genetic components, used in many of them the functions of structured ante premises (Jaffrézic et al., 2004).

**Covariance functions**

The covariance functions (CF) can be defined as a continuous function that provides the (co)variances of characteristics measured in different points of a trajectory such as lactation, and to describe the (co)variances between the measures taken under certain ages as days in lactation (Van Der Werf and Schaeffer, 1997). According to Meyer and Hill (1997) it was demonstrated that the coefficients of the (co)variance function can be estimated from the random regression models. Thus, different functions can be applied to adjust the trajectory over time. Among the parametric functions it stands out the exponential function of Wilmink (Wilmink, 1987) and the logarithmic function of Ali and Schaeffer (Ali and Schaeffer, 1987).
In the search for more parsimonious random regression models (RRM), many alternatives have been suggested, among them, the use of structures of different residual variances, by grouping classes with similar residual variations or by functions of residual variances using orthogonal polynomials (Pelicioni et al., 2009). The modeling for weight adjustments for animals that have not been evaluated exactly in the standard age and elimination of observations collected at ages far from the considered as standard, are necessary procedures, because the model considers that the observations of the different animals were collected exactly at the same age and are subject to the same covariance parameters (Valente et al., 2008).

Estimates of covariance functions (CF) are considered, until now, interesting alternatives for working with analysis of longitudinal data, and thus allows the description of change of (co)variances according to the time and the prediction of variances and (co)variances for points along the curve of animal growth. The (CFs) are equivalent to “finite dimension” of the matrix of (co)variances in multivariate analysis of “finite dimension” and may be obtained by the variance and covariance matrices of these models, or from the covariance matrices between the regression coefficients estimated by random regression models (Meyer, 1998).

In studies conducted by Kirkpatrick and Heckman (1989) they presented advantages of functions of (co)variance models over the traditional multi-characteristics models. The first advantage is that the covariance functions produce a description for every point along a continuous scale of measures allowing the (co)variances between ages not measured be easily obtained by interpolation. Another advantage is that every covariance function has a set of self values and self functions that provide information about the direction in which the average curve (growth, lactation, etc.) has greater chance to be modified by selection, because of it higher genetic variance.

The use of random regression models are a special case of covariance functions, and allows us to directly estimate the coefficients of covariance functions by the method of restricted maximum likelihood (Dias et al., 2006). In studies of experiment to growth selection, used RRM on genetic parameter estimation for days to birth at Nellore females, indicating that its application to the records of days to the birth provided detailed analysis of the behavior of genetic (co)variances and the genetic value of the reproductive performance of females during their lifetime, which might be appropriate in many studies (Mercadante et al., 2002).

Conclusion

The cattle from Tabapuã breed are still slightly studied in Brazil, and its potential may be underestimated due to lack of proper studies and research about this breed, overseeing their most important production characteristics. Therefore, more studies and efforts should be made to increase the knowledge about cattle from Tabapuã breed capabilities.

References


Carvalheira, J.G.V.; Blake, R.W.; Pollak, E.J.; Quaas, R.L.; Duran-Castro, C.V. Application of an autoregressive process to estimate genetic parameters and breeding values for daily milk yield in a tropical herd of Lucena cattle and in...


