Random Regression Models for Estimation of Covariance Functions, Genetic Parameters and Prediction of Breeding Values for Rib Eye Area in a Colombian \textit{Bos indicus-Bos taurus} Multibreed Cattle Population

Abstract

In this paper we present an application of random regression models (RRM) to obtain restricted maximum likelihood estimates of covariance functions and predictions of breeding values for longitudinal records of rib eye area measured by ultrasound (REA) in a Colombian multibreed cattle population. The dataset contained 708 records from 340 calves progeny of 37 sires from nine breeds mated to Gray Brahman Cows. The mixed model was a RRM that used Legendre polynomials (LP) of order 1 to 3. Fixed effects were age of animal, dam parity, contemporary group (herd*year*season*sex), breed additive genetic and heterosis, whereas direct and maternal additive genetic and maternal permanent environment were random effects. Residual
variances were modeled either as constant or changing across the growth trajectory. Models were compared with two Information Criteria, the corrected Akaike’s and the Schwartz’s Bayesian. According to these criteria the best model was the one with first order LP and constant residual variance. Given that with this model estimated maternal additive genetic and permanent environment covariance functions showed that these effects were not accurately disentangled, a parsimonious model without maternal additive genetic effects was used to obtain genetic parameters and breeding values. Direct additive genetic variance decreased until 150 days and then increased. Maternal permanent environment variance increased with age. Direct heritability estimates for REA at 4 months, weaning, 12 and 15 months (considered as target ages), were 0.003, 0.007, 0.034 and 0.058, respectively. Direct additive correlations ranged from \(-0.7\) to 1. Maternal permanent environmental correlations were close to unity across the entire range of ages. Estimates of (co)variance components showed the need to validate results with larger multigenerational multibreed populations before implement RRM in regional or national genetic evaluation procedures in Colombia.

Key words: Animal population, Covariance functions, Mixed model.

Resumen

En este trabajo presentamos una aplicación de modelos de regresión aleatoria (RRM) para obtener estimadores de máxima verosimilitud restringida de funciones de covarianza y predicciones del valor genético para datos longitudinales de área de ojo del lomo medidos por ultrasonido (REA) en una población bovina multirracial en Colombia. El conjunto de datos contenía 708 registros de 340 animales descendientes de 37 toros de 9 razas apareados con hembras Brahman Gris. Los modelos mixtos empleados fueron RRM que usaron polinomios de Legendre (LP) de orden 1 a 3. Los efectos fijos fueron edad del animal, número de partos de la madre, grupo contemporáneo (hacienda*ano*época*sexo), efectos genéticos aditivos de raza y heterosis, mientras que los efectos genéticos aditivos directos y maternos y de ambiente permanente materno fueron aleatorios. Las varianzas residuales se modelaron como constantes o cambiantes a través de la trayectoria de crecimiento. Los modelos fueron comparados mediante el criterio de información de Akaike corregido y el de información bayesiana de Schwartz. Según esos criterios, el mejor modelo fue aquel con LP de orden 1 y varianza residual constante. Dado que con este modelo las estimaciones de las funciones de covarianza genética aditiva materna y de ambiente permanente materno indicaron que estos dos efectos no se separaron adecuadamente, un modelo más parsimonioso sin los efectos genéticos aditivos maternos fue empleado para obtener parámetros y valores genéticos. La varianza genética aditiva directa decreció hasta 150 días y luego aumentó. La varianza de ambiente permanente materno aumentó con la edad. Las estimaciones de heredabilidad directa para REA a los 4 meses, destete, 12 y 15 meses (considerados como edades de referencia) fueron 0.003, 0.007, 0.034 y 0.058, respectivamente. Las correlaciones aditivas directas varieron de \(-0.7\) a 1. Las correlaciones de ambiente permanente materno fueron cercanas a la unidad a través de todo el rango de edades. Las estimaciones de componentes de (co)varianza mostraron la
1. Introduction

Modeling of longitudinal records with Legendre polynomials (LP) was proposed by Kirkpatrick, Lofsvold & Bulmer (1990) to describe direct additive genetic covariances among records at any pair of ages in a continuous form. The LP are solutions to the Legendre's differential equation and they are orthogonal. This property allows describing patterns of genetic variation through a growth trajectory. Continuous functions representing covariances among records are called covariance functions (Kirkpatrick et al. 1990). Meyer (1998) suggested that coefficients of covariance functions could be estimated as covariances among random regression coefficients by fitting linear mixed models. Advantages of random regression over multiple trait models (MTM) involve the inclusion of all available data without pre-adjustment to particular ages, no lose of records taken outside certain age ranges, and reduction in the number of parameters to be estimated by fitting parsimonious models (Kirkpatrick et al. 1990, Meyer & Hill 1997). Until today, these models have not been implemented for genetic analysis in Colombia. Carcass quality is important in the current beef market. Thus, there exists great interest in carcass traits measured by ultrasound like the rib eye area (REA), because they are closely related to the true carcass values and meat yields (Hougton & Turlington 1992). Genetic evaluation of carcass traits has been implemented in animal breeding programs in different countries and species (Wilson 1992, Hassen, Wilson & Rouse 2003, Fischer, van der Werf, Banks, Ball & Gilmour 2006, Choy, Lee, Kim, Choi, Choi & Hwang 2008). However, few genetic studies have considered ultrasound carcass traits in a longitudinal manner either in purebred or crossbred cattle (Fischer et al. 2006, Speidel, Enns, Brigham & Keeman 2007, Mercadante, El Faro, Pinheiro, Cyrillo, Bonilha & Branco 2010). Jiménez, Manrique & Martínez (2010) conducted the only study in Colombia on ultrasound carcass traits in cattle under pasture conditions using purebred Brahman. In low tropical areas of Colombia, there are limiting environmental conditions for livestock production. Consequently, crossbreeding between native Creole or European (Bos taurus) with Zebu (Bos indicus) breeds is frequently used as a strategy to increase beef production while maintaining adaptability (FEDEGAN 2006). This mating strategy has created a need to establish genetic evaluation programs involving animals from temperate and tropically adapted breeds for carcass traits. These programs must take into consideration that 72% of the Colombia’s cattle population is Zebu (mainly Brahman) (FEDEGAN 2006). Thus, the objective of this research was to show how to apply the RRM to obtain restricted maximum likelihood estimates of covariance functions and predictions of breeding values for longitudinal records of rib eye area measured by ultrasound (REA) in a Colombian multibreed cattle population.
2. Materials and Methods

All of the practices involving manipulation of animals that were performed to obtain records in this research were approved by the Animal Bio-ethics Committee of the National University of Colombia (Approval letter number: CBE-FMVZ-012, July, 2010).

2.1. Breeds, Matings and Animal’s Management

To construct the multibreed population, 37 bulls from 9 breeds were mated to third-parity Gray Brahman (GB) cows and heifers. Sire breeds were Gray Brahman (GB; n = 12), Red Brahman (RB; n = 4), Guzerat (GUZ; n = 3), Romosinuano (ROM; n = 3), Blanco Orejinegro (BON; n = 3), Simmental (SIM; n = 3), Braunvieh (BVH; n = 3), Normand (NOR; n = 3) and Limousin (LIM; n = 3). These Bos taurus breeds (Creole and temperate) were chosen because they are frequently used for crossbreeding programs with zebu cattle in Colombia’s low tropical beef production systems. Brahman was included because it has the largest cattle population in the country (Jiménez et al. 2010), and GUZ is a Bos indicus breed with increasingly higher representation in Colombia that has not been studied as a single breed or in crosses with Brahman. Females were chosen on the basis of a normal reproductive cycle and a healthy reproductive system. Subsequently, cows and heifers were randomly allocated to males, and artificially inseminated using a fixed-time protocol. Firstly, females received a progesterone implant (CIDR, Pfizer, NY, USA) and 2 mg of estradiol benzoate. Eight days later, the CIDR implants were removed, and 1 cm$^3$ of F2 α prostaglandin (Estrumate, Schering Plough S.A., Kenilworth, NJ, USA) was applied, followed by an injection of 1 mg of estradiol benzoate 24 hours later. Females were artificially inseminated 54 hours after progesterone implant removal. Calves were born in 2008 and 2009. Table 1 shows the number of sires per breed and the number of calves per breed group by year and total.

Table 1: Number of sires per breed and number of calves per breed group by year of birth.

<table>
<thead>
<tr>
<th>Sire breed</th>
<th>Number of sires</th>
<th>Calf breed group</th>
<th>Number of calves</th>
</tr>
</thead>
<tbody>
<tr>
<td>BON</td>
<td>3</td>
<td>BON X GB</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>33</td>
</tr>
<tr>
<td>BVH</td>
<td>3</td>
<td>BVH X GB</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>GB</td>
<td>12</td>
<td>BG X GB</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>97</td>
</tr>
<tr>
<td>GUZ</td>
<td>3</td>
<td>GUZ X GB</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>27</td>
</tr>
<tr>
<td>LIM</td>
<td>3</td>
<td>LIM X GB</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>33</td>
</tr>
<tr>
<td>NOR</td>
<td>3</td>
<td>NOR X GB</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>RB</td>
<td>4</td>
<td>BR X GB</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>34</td>
</tr>
<tr>
<td>ROM</td>
<td>3</td>
<td>ROM X GB</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>28</td>
</tr>
<tr>
<td>SIM</td>
<td>3</td>
<td>SIM X GB</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>31</td>
</tr>
</tbody>
</table>

Total 37 222 118 340

BON = Blanco Orejinegro; BVH = Braunvieh; GB = Gray Brahman; GUZ = Guzerat; LIM = Limousin; NOR = Normand; RB = Red Brahman; ROM = Romosinuano; SIM = Simmental.
Animals were kept in two herds located in Southern Cesar, municipality of Aguachica, Colombia. The ecosystem in this micro region is a very dry tropical forest. This region has a mean annual temperature of 28 °C, a height above sea level of 50 m, a relative humidity of 80% and sandy-loam soils. Because of its environmental conditions, Southern Cesar is considered to be better suited for beef cattle production than other regions in Colombia. The feeding system was based on pastures. Grass species were Brachipará (*Brachiaria plantaginea*), Guinea (*Panicum maximum*) and Angleton (*Dichantium aristatum*). Pastures were not fertilized. Animals were provided with an 8% phosphorus mineral supplement (GANASAL®, Colombia). Mineral supplement consumption was *ad libitum*. The grazing system was rotational with a rotation period of 60 days. All calves were weaned between 7 and 8 months of age and males were castrated at 12 months of age.

2.2. Records

The REA records were taken by a certified technician of the Colombian Zebu Cattle Breeders Association (ASOCEBU, Bogotá D.C., Colombia) using an Aquila Esaote model device (Pie Medical Equipment B.V., Maastricht, Limburg, The Netherlands). Once ultrasound images were collected, they were analyzed to check quality and to obtain the REA values (cm²) using the Echo Image Viewer software of Pie Medical (Pie Medical Equipment B.V., Maastricht, Limburg, The Netherlands). The total number of REA records was 708. Age of animals ranged from 70 to 492 days. Records were intended to be taken approximately at four, eight (weaning), twelve and fifteen months. Mean ages at each of these data collection points were: 120, 233, 332 and 445 days. At 4 months of age, calves are more dependent on the cow’s milk production that at weaning. This is due to the fact that at this stage the calf has not finished its transition from pre-ruminant to ruminant (Van Soest 1994). Thus, REA measurements taken at this age are useful to evaluate maternal effects (both genetic and non genetic).

2.3. Genetic Analysis

Mixed models procedures were carried out to obtain restricted maximum likelihood (REML) estimates of covariance components and best linear unbiased predictors (BLUP) of animal breeding values (BV). The following effects were assumed to be fixed in the mixed model: Contemporary group (herd*year*season*sex subclass), breed group additive effects, non additive effects (individual heterosis), dam parity (heifer or third parity cow) and age of the animal (linear and quadratic effects). In a first approach, the random effects were: Direct additive genetic, maternal additive genetic, maternal permanent environment, and residual. Seasons within years were defined as rainy or dry. The first season was a rainy season from mid April to mid August of 2009, the second was a dry season from mid August to mid December of 2009, the third was a dry season from mid December of 2009 to mid April of 2010, and the fourth was a rainy season from mid April to mid August of 2010. The GB and RB bulls were grouped as a single breed (BR). Thus, there were 8 breed groups for calves: BR x GB, BON X GB, BVH X GB, GUZ X GB,
LIM X GB, NOR X GB, ROM X GB and SIM X GB. Breed group effects were modeled as a continuous function of breeds over time. This function was a linear LP. Additive genetic breed group effects were modeled in such a way because individual random deviations and breed group solutions are required to obtain BV at a particular age in a multibreed population (Elzo & Wakeman 1998). In addition, because of the orthogonality of LP, the block of the mixed model equations corresponding to breed group effects was an identity matrix, thus, multicollinearity and confounding problems that are commonly present among genetic fixed effects in multibreed populations (Elzo & Famula 1985) could be alleviated at least partially.

To estimate covariance functions (CF) for the following effects: Direct additive genetic (DAGCF), maternal additive genetic (MAGCF) and maternal permanent environment (MPECF) and to compute BV, the regression variables used were normalized LP (LP with norm 1), evaluated at age of animal when records were collected. Orders of LP ranged from 1 to 3. The following combinations of LP to describe direct additive, maternal additive and maternal permanent environment CF were used: one (LP1), 2(LP2) and 3(LP3) for the 3 covariance components, and 3 for direct additive genetic covariances and 2 for maternal additive genetic and permanent environment covariances (LP32). The orders of LP were defined taking into account data set size and literature reports (Fischer et al. 2006, Mercadante et al. 2010). The residual variance was modeled in two ways. The first one assumed that the residual variance was the same along the entire growth trajectory (LP1HOM, LP2HOM, LP3HOM, LP32HOM), and the second one assumed a step function (LP1HET, LP2HET, LP3HET, LP32HET) across 3 age intervals (70 ≤ age ≤ 230 days, 230 < age ≤ 365 days, and 365 < age ≤ 492 days). Residuals were assumed to be independent and normally distributed. Thus, there were a total of 8 random regression models to compare: LP1HET, LP2HET, LP3HET, LP32HET, LP1HOM, LP2HOM, LP3HOM, and LP32HOM. Models comparison was made through the Schwartz’s Bayesian Information Criterion (BIC) and the Corrected Akaike’s Information Criterion (AICC):

\[
BIC = -2 \log L + K \log(N - r)
\]
\[
AICC = AIC + \frac{(2(K + 1)(K + 2))}{(N - K - 2)}
\]

Where AIC is the Akaike’s information criterion, K is the number of parameters, N is the number of records, logL is the natural logarithm of the likelihood function and r is the rank of the fixed part of the model, that is, the rank of the incidence matrix for all fixed effects in the model. The AICC was preferred over the AIC in our study because of the small data set size, which is suggested by Littell, Milliken, Stroup, Wolfinger & Schabenberger (2006). However, estimated covariance functions showed a strong negative correlation among maternal additive genetic and maternal environmental effects, which indicated that these effects were not accurately separated. Thus, a parsimonious version of the model selected in the first approach (LP1HOM) considering only maternal permanent environmental effects and denoted as LP1HOMS was used to compute variance-covariance components, genetic parameters and BV. The number of variance-covariance parameters ranged from 7 for the most parsimonious model (LP1HOMS) to 33 for
model LP4HET (Table 2). In matrix notation the RRM used was as follows:

\[
y = X\beta + Q_{ga}g_a + Q_hh + \Phi_a\alpha + \Phi_p\psi + e
\]

\[
Var \begin{bmatrix}
a \\
p \\
e
\end{bmatrix} = A \otimes K_a I \otimes K_p R
\]

\[
E[y] = X\beta + Q_{ga}g_a + Q_hh
\]

\[
Var(y) = \Phi_a(A \otimes K_a)\Phi_a' + \Phi_p(I \otimes K_p)\Phi_p' + R
\]

<table>
<thead>
<tr>
<th>Model</th>
<th>AICC</th>
<th>BIC</th>
<th>Number of variance covariance parameters</th>
<th>Log L¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP1HET</td>
<td>3394.74</td>
<td>3448.66</td>
<td>12</td>
<td>−1685.15</td>
</tr>
<tr>
<td>LP2HET</td>
<td>3412.64</td>
<td>3506.42</td>
<td>21</td>
<td>−1684.65</td>
</tr>
<tr>
<td>LP3HET</td>
<td>3435.54</td>
<td>3581.68</td>
<td>33</td>
<td>−1683.10</td>
</tr>
<tr>
<td>LP32HET</td>
<td>3426.62</td>
<td>3555.42</td>
<td>29</td>
<td>−1683.03</td>
</tr>
<tr>
<td>LP1HOM</td>
<td>3392.82</td>
<td>3437.8</td>
<td>10</td>
<td>−1686.25</td>
</tr>
<tr>
<td>LP2HOM</td>
<td>3410.66</td>
<td>3495.62</td>
<td>19</td>
<td>−1685.78</td>
</tr>
<tr>
<td>LP3HOM</td>
<td>3508.34</td>
<td>3645.82</td>
<td>31</td>
<td>−1721.70</td>
</tr>
<tr>
<td>LP32HOM</td>
<td>3426.66</td>
<td>3546.72</td>
<td>27</td>
<td>−1685.22</td>
</tr>
<tr>
<td>LP1HOMS</td>
<td>3386.66</td>
<td>3418.21</td>
<td>7</td>
<td>−1686.25</td>
</tr>
</tbody>
</table>

¹Natural logarithm of the restricted likelihood function.

Where \( y \) = vector containing the REA records, \( \beta \) = vector of unknown fixed effects of contemporary group, dam parity and age of animal, \( g_a \) = vector of fixed additive genetic group effects (modeled as a continuous function of time) which correspond to the mean effects of genes from a given breed (Elzo 2010), \( h \) = vector of fixed non additive genetic effects (individual heterosis) these are the effects due to the presence of alleles from different breeds in one locus (Elzo 2010), \( a \) = vector of random regression coefficients for direct additive genetic effects, which are the sum of effects of individual genes affecting REA (Kempthorne 1957, Lynch & Walsh 1998), \( p \) = vector containing random regression coefficients for maternal permanent environmental effects, which correspond to those effects explained by the environment proportioned to the calf by its dam, maternal effects are genetic to the dam and environmental to the calf, \( e \) = random vector of residuals, \( X, Q_{ga}, Q_h, \Phi_a, \Phi_p \) were known incidence matrices respectively relating vectors \( \beta, g_a, h, a, p \) to REA records and super index “′” denotes transposition. Columns in \( X \) relating records to fixed effects of age contained second order LP evaluated at each age; columns for the other fixed effects contained zeroes and ones. Matrix \( Q_{ga} \) contained linear LP evaluated at the expected fraction of each breed in an animal times the age of the animal, and matrix \( Q_h \) contained probabilities of alleles of different breeds occurring at one locus in an animal (Elzo & Famula 1985) and it was calculated as: \( H_l = 1 - \sum_{b=1}^{B}(R_p \times R_m)_i \), where \( R_p \) and \( R_m \) are the expected fractions of each breed in sire and dam of the animal and \( b \) is the number of breeds, matrices \( \Phi_a, \Phi_m, \Phi_p \) contained LP evaluated at the ages of the animals when records were taken (Meyer 1998); matrices \( K_a \) and \( K_p \) contained the coefficients for additive genetic, and maternal permanent environmental covariance.
functions, $A$ was the additive relationship matrix, $\otimes$ represents the Kronecker product, and $R$ was the residual covariance matrix which had the form $R = I \sigma^2_e$.

The mixed models analyses were performed with software WOMBAT (Meyer 2007) using an average information (AI) algorithm. Different starting values were used to ensure that estimates corresponded to global maximums. Convergence was declared when change of value of the natural logarithm of the restricted likelihood function in two consecutive iterations was lower than $5 \times 10^{-4}$. Model effects were estimated by solving the mixed model equations:

$$
\begin{bmatrix}
X'R^{-1}X & X'R^{-1}Q_a & X'R^{-1}Q_p & X'R^{-1}k_a & X'R^{-1}k_p \\
Q_a'R^{-1}Q_a & Q_a'R^{-1}Q_p & Q_a'R^{-1}k_a & Q_a'R^{-1}k_p \\
Q_p'R^{-1}Q_p & Q_p'R^{-1}k_p & Q_p'R^{-1}k_p & Q_p'R^{-1}k_p \\
\end{bmatrix}
\begin{bmatrix}
\beta \\
k_a \\
k_p \\
\end{bmatrix}
= 
\begin{bmatrix}
X'y' \\
Q_a'y' \\
Q_p'y' \\
\end{bmatrix}
$$

The eigenfunctions (EF) of a CF are continuous smooth functions representing a possible deformation in the mean growth trajectory (Kirkpatrick et al. 1990). Thus, the EF were calculated to study variation patterns throughout the REA growth curve. Each EF has a correspondent eigenvalue. Only EF whose eigenvalues together explained at least 80% of the respective variance component were computed. The EF were computed for direct additive genetic CF from eigenvectors of $K_a$ matrix as:

$$
\psi_i(t) = < c_{\psi_i}, \phi_t^* >
$$

where $c_{\psi_i}$ is the $i^{th}$ eigenvector of the matrix $K_a$ and $\phi_t^*$ is a vector with LP evaluated at $t^*$ (age $t$ standardized to the real interval $[-1, 1]$) and the operator $< \cdot, \cdot >$ represents the internal or dot product between vectors. The age $t$ was standardized to the real interval $[-1, 1]$ by using the following expression (Kirkpatrick et al. 1990):

$$
t^* = \frac{2(t - t_{\min})}{t_{\max} - t_{\min}} - 1
$$

where $t_{\min}$ and $t_{\max}$ are the minimum and maximum ages at which records were taken. Matrices of covariance components for additive direct genetic effects and maternal permanent environmental effects as well as BV for REA at 4 target ages were obtained using the REML estimates of covariance matrices among random regression coefficients obtained at convergence which are equal to the coefficient matrices of corresponding CF (Meyer 1998). Target ages were 120, 230, 365 and 450 days, and the corresponding REA values were denoted as REA4, REAW, REAY and REAF.

Covariance matrices for REA at target ages were computed using the CF which were obtained as the product of a matrix containing LP evaluated at those ages ($\Phi$), the correspondent coefficients matrix ($K_a$ for direct additive covariance, and $K_p$ for maternal permanent environmental covariance) and the transpose of matrix $\Phi$ (Kirkpatrick et al. 1990, Meyer 1998):

$$
cov_j = \Phi K_j \Phi'
$$
Random regression models for genetic longitudinal data

where, \( Cov_j \) is the covariance matrix for the \( j^{th} \) covariance component (additive genetic or maternal permanent environment). The matrix \( \Phi \) was obtained as the product of two matrices. The first is matrix \( M = (m_{ij})_{d \times k} = t_i^{j-1} \), where \( t_i^{*} \) is the \( i^{th} \) age standardized to the real interval \([-1, 1]\), \( d \) is the number of ages considered (4 in this case) and \( k - 1 \) is the order of the LP. The second matrix was \( \Lambda_{k \times k} \), which contained the coefficients of the LP. Thus, \( \Phi = MA \) (Kirkpatrick et al. 1990). Consequently,

\[
Cov_j = \Phi K_j \Phi' = MA \Lambda_j \Lambda' M' = MC_j M', \quad \text{where} \quad C_j = \Lambda K_j \Lambda'
\]

By using matrix \( C_j \) instead of matrix \( K_j \) for representing the \( j^{th} \) CF, \( cov_j \) is calculated directly as a function of the age standardized to the interval \([-1, 1]\) (i.e., \( t^{*} \)). This equivalent form was used to compute critical points of CF. The extremes of the CF were also assessed in order to detect the global maximum and minimum values of each CF.

The BV were computed for REA4, REAW, REAY and REAF for all individuals in the population (sires, dams, and offspring). The additive breeding value for animal \( i \) at age \( t \) (\( BV_{it} \)) was computed by adding two terms. The first term was a weighted sum of probabilities of alleles of breed \( b \) in animal \( i \) times the generalized least squares estimate of breed \( b \) (deviated from BR) at time \( t \), \( b = 1, 2, \ldots, 7 \). The second term was the BLUP of the random solution for each individual. This value was computed as the internal (or dot) product between a vector containing LP evaluated at age \( t \) and a vector whose entries were the BLUP for random regression coefficients of animal \( i \). Thus, \( BV_{it} \) was computed as:

\[
BV_{it} = <\phi_{bt}, \hat{g}_b> + <\phi_t, \hat{a}_i>
\]

where \( \phi_{bt} \) is a vector of LP evaluated at the product of the fraction of breed \( b \) \((b = 1, 2, \ldots, 7)\) in animal \( i \) times calf age \( t \) standardized to real interval \([-1, 1]\), \( \hat{g}_b \) is the generalized least squares solution of the fixed coefficient for breed additive genetic effects, \( \phi_t \) is a vector of LP evaluated at calf age \( t \) standardized to real interval \([-1, 1]\), and \( \hat{a}_i \) is the BLUP vector of the random coefficients for animal \( i \).

3. Results

3.1. Model Selection

As stated before, estimated covariance functions, covariance components, genetic parameters and breeding values were computed using model LP1HOMS. Although this model was selected given the evidence of correlation among maternal additive genetic and environmental effects, according to AICC and BIC values, this was the best model since it had the smallest AICC and BIC values (Table 2).
3.2. REML Estimates of Covariance Functions and Covariance Components

Direct additive genetic (DAGC) and maternal permanent environment (MPEC) covariances between pairs of ages $t_1$ and $t_2$ such that $70 \leq t_1, t_2 \leq 492$, were described by the following CF (DAGCF, and MPECF, respectively) obtained with model LP1HOMS using $\text{cov}_j = \Phi K_j \Phi'$:

$$
\begin{align*}
DAGC(t_1, t_2) &= \begin{bmatrix} \phi_0(t_1^*), & \phi_1(t_1^*) \end{bmatrix} \begin{bmatrix} 1.5900 & 1.2435 \\ 1.2435 & 1.1589 \end{bmatrix} \begin{bmatrix} \phi_0(t_2^*) \\ \phi_1(t_2^*) \end{bmatrix} \\
MPEC(t_1, t_2) &= \begin{bmatrix} \phi_0(t_1^*), & \phi_1(t_1^*) \end{bmatrix} \begin{bmatrix} 53.482 & 4.5003 \\ 4.5003 & 0.3787 \end{bmatrix} \begin{bmatrix} \phi_0(t_2^*) \\ \phi_1(t_2^*) \end{bmatrix}
\end{align*}
$$

where $t_i^*$ is the $i^{th}$ age standardized in the real interval $[-1, 1]$, and $\phi_j(t_i^*)$, $j = 0, 1$, is the $j^{th}$ LP evaluated at $i^{th}$ age. The equivalent forms of these 2 CF, using $\text{cov}_j = MC_jM'$, were as follows:

$$
\begin{align*}
DAGC(t_1, t_2) &= \begin{bmatrix} 1 \\ t_1^* \end{bmatrix} \begin{bmatrix} 0.7950 & 1.0769 \\ 1.0769 & 1.7382 \end{bmatrix} \begin{bmatrix} 1 \\ t_2^* \end{bmatrix} \\
MPEC(t_1, t_2) &= \begin{bmatrix} 1 \\ t_1^* \end{bmatrix} \begin{bmatrix} 26.7405 & 3.8972 \\ 3.8972 & 0.5680 \end{bmatrix} \begin{bmatrix} 1 \\ t_2^* \end{bmatrix}
\end{align*}
$$

These functions are defined (domain) for the following set: $D = [70, 492] \times [70, 492]$. The partial derivatives were:

$$
\begin{align*}
\frac{\partial CF_j}{\partial t_1^*} &= c_{12} + c_{22}t_1^*; & \frac{\partial CF_j}{\partial t_2^*} &= c_{12} + c_{22}t_2^*
\end{align*}
$$

where $c_{ij}$ is the $(i-j)^{th}$ entry of the matrix $C$ and $CF_j$ is the $j^{th}$ CF ($j = \text{DAGCF or MPECF}$), and $t_i^*$ are standardized calf ages at $[-1, 1]$. By equating these expressions to zero yielded that the critical arguments of the CF were $\frac{-c_{12}}{c_{22}}$ for both $t_1^*$ and $t_2^*$ (because the 2 CF were symmetric).

To determine if the critical points obtained from the last expression were maximums, minimums or saddle points the determinant of the Hessian matrix was computed. Because these functions are polynomials, the Clairaut’s theorem (Stewart 2008) applies making the Hessian matrix to be symmetric. This matrix was:

$$
H = \begin{bmatrix} \frac{\partial^2 CF_j}{\partial t_1^{2*}} & \frac{\partial^2 CF_j}{\partial t_1^* \partial t_2^*} \\ \frac{\partial^2 CF_j}{\partial t_2^* \partial t_1^*} & \frac{\partial^2 CF_j}{\partial t_2^{2*}} \end{bmatrix} = \begin{bmatrix} 0 & c_{22} \\ c_{22} & 0 \end{bmatrix}
$$

Thus: $|H| = -(c_{22}^2)$, and the critical point is a saddle point. Variance functions (VF) are special cases of CF when $t_1^* = t_2^*$. Because there is a single age, VF are univariate. Critical points computed for CF and VF could be outside the range of calf ages (i.e., outside their domain). If this happens, these critical points should be ignored because in regression analysis values outside the domain (range of calf ages) would have no valid interpretation (Draper & Smith 1981).
Random regression models for genetic longitudinal data

The DAGCF had a saddle point located at 150 days. Thus, covariances before 150 days tended to decrease with age. After 150 days, the pattern was more complex. Covariances among ages lower than 150 days and ages greater than 150 days tended to decrease as distance among them increased (Figure 1). On the other hand, covariances among ages greater than 150 days tended to increase as the animals grew older. The MPECF was positive throughout the entire domain (Figure 1). The minimum value of MPECF (19.51 cm$^4$) was located at coordinates (in days): (70, 70) while the maximum (35.10 cm$^4$) was located at (492, 492). The analysis of derivatives showed that MPECF had a critical point outside the range of calf ages in this study. As indicated before, VF are special cases of CF, because by definition they are the covariance of a random variable with itself. Consequently, the diagonals of the CF correspond to VF. According to the analysis of first and second derivatives of the direct additive genetic variance function, direct additive genetic variance (DAGV) had a global minimum located at 150 days (0.13 cm$^4$). The largest value of DAGV was 4.69 cm$^4$ at 492 days. As shown in Table 3 for the target ages, DAV was 0.16 for REA4, 0.38 for REAW, 1.93 for REAY and 3.64 cm$^4$ for REAF. The DAGC were negative for REA4-REAY ($-0.13$ cm$^4$) and REA4-REAF ($-0.23$ cm$^4$) and the biggest covariance value was among REAY and REAF (2.64 cm$^4$). For the target ages, maternal permanent environment variance (MPEV) ranged from 21.12 (REA4) to 33.35 cm$^4$ (REAF). MPEC had its lowest value (22.93 cm$^4$) among REA4 and REAW and the largest (31.60 cm$^4$) for REAY-REAF (Table 3). Considering the entire range of ages, MPEV had its maximum value at 492 days (35.10 cm$^4$) and the minimum (19.51 cm$^4$) at 70 days.

![Figure 1: Plots of direct additive genetic (DAGC (cm$^4$); left), and maternal permanent environment (MPEC (cm$^4$); right) covariances.](image)

REML estimate of residual variance was 25.55 cm$^4$. Because phenotypic variance (PhV) is the sum of genetic and environmental variance components, it also increased as animals grew older. Its minimum value was 45.45 cm$^4$ at 70 days and its maximum was 65.35 cm$^4$ at 492 days. Plots of DAGV and MPEV are shown in Figure 2.
3.3. Heritability and Ratio of MPEV to PhV

The direct heritability (the ratio of DAGV to PhV) estimates (Dh), were low at the entire trajectory. The Dh reached a global minimum at 150 days (0.003) and its maximum at 492 days (0.072). The estimate of Dh at 70 days was 0.008. The Dh estimates at the 4 target age points were 0.003 (REA4), 0.007 (REAW), 0.034 (REAY) and 0.058 (REAF) (Table 3). The trend of Dh across the range of calf ages is shown in Figure 3. The ratio of MPEV to phenotypic variance (MPr) ranged from 0.43 at 70 days to 0.54 at 492 days. The MPr had an upward trend.
through the REA trajectory (Figure 3). The MP\(r\) estimates for the target ages were 0.45 for REA4, 0.49 for REAW, 0.52 for REAY and 0.53 for REAF (Table 3).

**3.4. Correlations**

The estimates of direct additive genetic (DAGR), maternal permanent environment (MPER) and Phenotypic (PhR) correlations at the 4 target ages are shown in Table 3. Estimates of DAGR formed a plateau close to unity approximately after 240 days. The DAGR between REA at 70 days and REA at other ages were negative after 193 days and had its lowest value at 492 days (−0.71). For target ages, DAGR estimates ranged from −0.29 among REA4 and REAF to 0.99 among REAY and REAF (Table 3). The MPER estimates were close to unity throughout the entire range of ages considered. The PhR estimates were always positive and ranged from moderate to high. For the 4 selected age points, PhR values ranged from 0.47 (REA4-REAW) to 0.57 (REAY-REAF).

**3.5. Eigenfunctions**

The first eigenvalue for DAGCF was 2.64 and it accounted for 95.9% of total DAGV. Thus, for DAGCF only the first EF (DAGEF1) was computed. The first eigenvector of the coefficient matrix associated with DAGCF was \((0.7651 \quad 0.6439)\)\(^T\), and the DAGEF1 was:

\[
DAGEF1 = 0.5358 + 0.7991t\
\]

Figure 4 shows a graph of this function across the entire range of calf ages. The DAGEF1 was an increasing function, but it was not positive at the entire range trajectory. The point where this function crossed the age axis was 136 days.

---

**Figure 3**: Graphics of continuous functions describing direct heritability (left), and ratio of maternal permanent environmental variance (MPEV) to phenotypic variance (PhV) (right).
behavior of the EF was a consequence of the estimates obtained here for DAGR. As described previously, there were negative DAGR between early and late calf ages.

![Figure 4](image.png)

**Figure 4:** First eigenfunction of the direct additive genetic (DAGEF1) covariance function.

### 3.6. Breeding values

Descriptive statistics for BLUP of BV in general and discriminated by sire breed are shown in Table 4. Values for sire breeds were obtained using information from the bulls and the overall values were obtained from BV of all animals. Overall mean BV were 0.41 for REA4, 0.72 for REAW, 1.26 for REAY and 1.55 cm² for REAF. Values presented in Table 4 indicate that on average LIM bulls had the highest BV for REA. Sires of BVH and NOR breeds had the smallest BV at the 4 target age points. In the Creole cattle group, ROM sires had greater mean BV than BON sires. Finally, for the Bos indicus breeds, GUZ bulls had the greatest mean BV.

### 4. Discussion

#### 4.1. Model Selection

Selection of the most parsimonious model (LP1HOMS) as the best model by BIC and AICC implies that the larger log likelihood values obtained with other models was insufficient to counterbalance BIC and AICC penalties due to the higher number of required parameters. Consequently, the BIC and AICC values of those other models were larger than the values for model LP1HOMS (Table 2).
The use of heterogeneous error structures was reported for Nellore cattle in tropical conditions (Mercadante et al. 2010), for crossbred Australian cattle under pasture and feedlot conditions (Mirzaei, Verbyla & Pitchford 2011), and for lambs (Fischer et al. 2006). However, heterogeneous error structure models in these studies were not compared with models fitting a homogeneous residual variance structure. For Colombian Buffaloes, it was found that a model fitting within animal homogeneous variance structure described better REA data (Bolívar, Cerón-Muñoz, Elzo, Ramírez & Agudelo 2011). Meyer (2000), suggested that seasonal variations could be responsible for the heterogeneity in the measurement error. Given that the heterogeneous error variance approach did not show a better fit here, it indicates that environmental factors such as weaning and castration of bulls were not important sources of environmental variation in this multibreed population.

The order of LP used to estimate DAGCF was in agreement with the results found by Mercadante et al. (2010) who compared orders 1, 2 and 3 using AIC.
Carlos Alberto Martínez, et al.

and BIC as model selection criteria. However, they did not consider LP of order 1 to model random non genetic effects. In that study, orders of LP to model those effects were either 2 or 3. Mercadante et al. (2010) found that the model considering the lower orders of fit for both direct additive genetic and permanent environmental effects was the best 1. The LP of order one were also reported to be sufficient to explain direct additive genetic effects for weight data in crossbred cattle cows (Arango, Cundiff & Van Vleck 2004).

Considering the small size of the dataset in this study and that a model with only 7 parameters that permitted the use of all records was selected, RRM seem to be a good option to model longitudinal ultrasound data. If a four-trait model assuming zero covariance between direct and maternal additive effects had been fitted here, the number of parameters needed would have been $4 \times (4 \times (4 + 1)/2) = 40$, which is more than 4 times greater than the number of parameters estimated with the LP1HOMS model. Even if two-trait models had been utilized, a total of 6 two-trait analysis would have had to be performed to estimate the full covariance matrix for REA at the 4 target ages. In addition, because each analysis would be performed separately, there would have been no certainty for the estimated six-trait covariance matrix to be positive definite.

4.2. REML Estimates of Covariance Functions and Covariance Components

The direct additive genetic variance function corresponding to the DAGCF when $t_1 = t_2$ (Figure 2) was concave up with a global minimum at 150 days of age. Thus, the increase in the magnitude of the variance after the minimum point was always positive and greater as the animals grew older. Among the few literature reports using RRM to model ultrasound longitudinal data, a smoother pattern for DAGV (in the age interval 60 to 360 days) was reported for eye muscle depth (a ultrasonic measure at the same point where REA is taken, but measuring depth not area) in lambs (Fischer et al. 2006). Although they found that additive genetic variance did not have great changes, it had a concave up shape. A Nellore cattle study under pasture and feedlot conditions in a tropical region was conducted by Mercadante et al. (2010) in Brazil. However they did not discuss the covariance tendencies. The very low values of DAGV around 150 days here may have been due to computing artifacts rather than biology. Numerical problems have been reported for RRM using LP as base functions (Nobre, Misztal, Tsuruta, Bertrand, Silva & Lopes 2003, Bohmanova, Misztal & Bertrand 2005, Bertrand, Misztal, Robins, Bohmanova & Tsuruta 2006).

The DAGV did not decrease after weaning but it increased with the calf’s age. Maternal effects have been found to be important for REA and other ultrasound traits (Speidel et al. 2007). These results suggested that maternal effects would need to be considered in models for genetic analysis of postweaning growth traits. No other literature reports were found for longitudinal REA data considering maternal effects in cattle.
4.3. Heritability and Ratio of MPEV to PhV

The Dh values followed the same trajectory as DAGV. Low values of Dh (particularly at 150 days) could be due to numerical problems related to the population structure and small size of dataset. The only literature report found for Dh of REA in cattle using RRM showed higher values than those reported in the current study. That study considered a range of ages from 323 to 773 days in a Brazilian Nellore cattle population and Dh estimates ranged from 0.31 to 0.42 (Mercadante et al. 2010). The Dh for REA at slaughter for Australian crossbred cattle in pasture conditions until 18 months of age and then placed in feedlot conditions was estimated to be 0.40 (Mirzaei et al. 2011). In a Colombian purebred Brahman population under similar management conditions (pastures and mineral supplementation) to those in this study, Dh for REAF was 0.37 (Jiménez et al. 2010). For Red Angus animals of ages between 300 and 480 days and with a single ultrasonic REA measurement, Speidel et al. (2007) found a Dh estimate of 0.35. Crews & Kemp (1999) suggested that maternal effects were unimportant for the genetic evaluation of carcass traits (including REA) in a multibreed population. However, they did not use RRM because they considered REA data only at slaughter. Thus, differences in the data structure (longitudinal vs. simple), the model used, and the fact that presumably maternal effects have a small effect on traits measured at slaughter could explain the different results. In agreement with results here, for Red Angus cattle, Speidel et al. (2007) concluded (based on a likelihood ratio test) that inclusion of maternal effects improved the ability of genetic models to account for variability on carcass traits. The MPr estimates increased smoothly with age. The MPr had medium to high values across all ages and had a total (maximum value - minimum value) change of 10.8 percentage units. For live weight, under similar conditions and for a Bos indicus (Nellore) beef cattle population, Albuquerque & Meyer (2001) found a similar pattern for MPr. No research including maternal permanent environmental effects for REA data in cattle was found in the literature. The MPr values did not decrease after weaning, thus, the permanent maternal environmental effects were important for post weaning development phases. This suggests that remnants of pre-weaning permanent environmental cow effects continued to influence calf REA until 492 days of age. Maternal effects are mainly explained for cow’s milk production (genetic to the dam and environmental to the calf). Considering the values of MPr (0.43 to 0.54), it seems that a key point to obtain animals with greater REA, which are expected to have a greater meat production, would be to implement an adequate selection program that includes both direct growth and maternal milk production. It has to be taken into account that although maternal additive genetic effects were not included in the model due to estimation problems, they are still present. On the other hand, the unique maternal effect term in the model is possibly accounting for both: Additive genetic and permanent environment maternal effects.
4.4. Correlations

As the DAGR formed a plateau after approximately 240 days, for genetic evaluation purposes, when considering REA data with ages greater than 240 days (for example, from weaning to greater ages), it will be possible to use a repeatability model. The simplicity of this model will make it desirable, especially for small data sets as present one. For live weight records, a similar conclusion was found by Arango et al. (2004) for crossbred beef cows in a temperate region.

The negative DAGR between ages at the beginning of the trajectory and final ages indicated that those genes controlling REA at ages near to 70 days are antagonist to genes controlling this trait at ages near to 492 days. Taking into account that what matters is REA at ages near slaughter, animals could be selected for REA at ages after 240 days (because of the plateau formed by DAGR occurred after that point). Because MPER values were medium to high across calf ages, it appears that maternal permanent environmental effects exerted a positive effect on REA preweaning, and this effect persisted until 492 days of age. As a general observation taking into account, MPr and MPER values for this population, maternal effects appeared to be important to obtain greater REA.

4.5. Eigenfunctions

The proportion of DAGV explained by the first eigenvalue (95.9%) was in the range of proportions found by Mercadante et al. (2010). Such range was 84% to 99% depending on the model used. A similar proportion (90%) was described for Longissimus muscle depth at the same point where REA was taken in lambs (Fischer et al. 2006). As the DAGEF1 crossed the age axis at 136 days, this is a critical age because selection for greater REA values before this trajectory point will tend to negatively deform the mean population REA growth curve for later ages. Considering only ages after that point, selection for direct additive genetic effects will increase REA mean population growth curve. Thus, selection for REA could be performed after 136 days, i.e., roughly 4 months of age under field conditions. However, considering the high DAGR between 136 days and 240 days of age, a practical age to perform selection for REA would be at weaning.

4.6. Breeding Values

Given the small number of sires considered in the current study (especially for Bos taurus breeds) results should be viewed with caution. As expected, all genetic additive direct breed effects were estimable. Thus, the use of orthogonal functions to describe fixed genetic effects when modeling longitudinal data could be useful in order to prevent estimability problems. No research that considered breed effects as a continuous function of age of calf was found in the literature.

Range of BV for REAF of BR sires (Table 4) was smaller than the range reported by Jiménez et al. (2010) for purebred Brahman cattle under pasture conditions in Colombia. They reported EPD values ranging from $-2.84$ to $3.47$ cm$^2$. 

thus, the BV (twice the EPD) ranged from −5.68 to 6.94 cm$^2$. As in the current study, BV were deviated from BR. The range of BV for purebred BR animals (non parents; −0.82 to 1.12 cm$^2$) was smaller than those reported by Jiménez et al. (2010) suggesting that the amount of genetic variability in the dataset here was smaller than in the Brahman population analyzed by these authors.

The BLUP of BV suggested that among the tested sires and under the conditions of the study LIM bulls had the greatest mean genetic merit for REA at all target ages (Table 4). When all of the sires were ranked according to individual BV, LIM sires were always those with the greatest values. Consequently, the LIM breed would have to be considered for crossbreeding programs with Brahman cows under pasture conditions in the Southern Cesar region of Colombia. The LIM breed had been reported to have greater additive genetic effects for REA at different ages when compared to Bos indicus and Bos taurus breeds in temperate areas under feedlot or high supplement conditions (Ríos-Utrera, Cundiff, Gregory, Koch, Dikeman, Koolmaraie & Van Vleck 2006, Williams, Aguilar, Rekaya & Bertrand 2010). According to the results of this research, in tropical regions and under pasture conditions, LIM animals also showed a good performance for this trait.

5. Final Remarks

It should be mentioned that genetic parameters and breeding values were estimated with limited accuracy due to the structure and small size of the available multibreed population. Estimates of (co)variance components showed that it is necessary to validate the results of this research with substantially larger multigenerational populations before implement RRM in regional or national genetic evaluation procedures. Thus, there is a need to continue obtaining longitudinal ultrasound information from different beef cattle herds where the breeds studied here are represented. Results suggested that maternal effects were important, both preweaning and postweaning. Thus, maternal effects (genetic and non-genetic) appeared to be relevant effects to be included in models for genetic evaluation of REA pre and postweaning under pasture conditions in Colombia.

Acknowledgments

We sincerely thank two referees for helpful comments and suggestions which led to improve this paper.
References


Random regression models for genetic longitudinal data


Kempthorne, O. (1957), An Introduction to Genetic Statistics, John Wiley.


Stewart, J. (2008), Cálculo en varias variables. Trascendentes tempranas, 6 edn, Cengage Learning, México DF.

