PAULO ROBERTO COSTA NOBRE

ANALYSES OF SEQUENTIAL WEIGHTS OF NELLORE CATTLE USING MULTIPLE TRAIT AND RANDOM REGRESSION MODELS

VIÇOSA
MINAS GERAIS – BRAZIL
2001

PAULO ROBERTO COSTA NOBRE

ANALYSES OF SEQUENTIAL WEIGHTS OF NELLORE CATTLE USING MULTIPLE TRAIT AND RANDOM REGRESSION MODELS

A Dissertation submitted to the Genetics and Breeding Graduate Program of the Federal University of Viçosa in partial fulfillment of the requirements of the degree Doctor Scientiae.

VIÇOSA
MINAS GERAIS – BRAZIL
2001

PAULO ROBERTO COSTA NOBRE

ANALYSES OF SEQUENTIAL WEIGHTS OF NELLORE CATTLE USING MULTIPLE TRAIT AND RANDOM REGRESSION MODELS

A Dissertation submitted to the Genetics and Breeding Graduate Program of the Federal University of Viçosa in partial fulfillment of the requirements of the degree Doctor Scientiae.

APPROVED: November 13th, 2001

Prof. Robledo de Almeida Torres (Committee Member)	Prof. Ignacy Misztal

Prof. Roberto Augusto de Almeida Torres Júnior

Pesq. Dr. Luiz Otávio Campos da Silva

Prof. Paulo Sávio Lopes
(Adviser)

ACKNOWLEDGEMENTS

I wish to express my appreciation to the Federal University of Viçosa and to The University of Georgia, for the opportunity of working in the program and their support.

Appreciation is also extended to the Foundation for the Coordination of Higher Education and Graduate Training – CAPES Foundation, to National Research Council for Development – CNPq, and the Minas Gerais State Foundation for Research Support – FAPEMIG for the sponsorship of my graduate studies.

I would also like to thank the to Embrapa Beef Cattle – Brazilian Agricultural Research Corporation for providing the data for this study.

I would like to express my gratitude to my advisers, Prof. Paulo Sávio Lopes and Prof. Ignacy Misztal for their guidance and supervision during the course of the graduate program. It was a privilege and a pleasure to be their student.

I am very thankful for the opportunity to learn from Dr. Kepler Euclides Filho and Dr. Luiz Otávio Campos da Silva in a special friendly team environment.

I wish to express my appreciation to Prof. Joseph Keith Bertrand for his friendship and support throughout last year.

My special appreciation is also extended to my committee members, Prof. Robledo de Almeida Torres and Prof. Adair José Regazzi for their assistance and contributions.

I would like to express my appreciation to Prof. Roberto Augusto de Almeida Torres Júnior for his contribution.

Appreciation is expressed to Dr. Deukhwan Lee and Dr. Shogo Tsuruta for their ongoing advice and support.

My special thanks are also for my fellow graduate students in the Federal University of Viçosa and in The University of Georgia for their friendship.

A very special thank you to Katherine and Robyn for some entertainment and also for making me happy and hopeful as they are when I needed.

The friendly environment provided by Cláudio Araújo, Miguel Canseco and Ronaldo Encarnação was greatly appreciated, thanks for making life easier.

Finally, a very special thank you to my family for their love, encouragement and care. My father, my mother, my sisters, and my brothers, the completion of this work has only been possible because of your never ending support and understanding.

TABLE OF CONTENTS

ABSTRACT

NOBRE, Paulo Roberto Costa, D.S., Federal University of Viçosa, November 2001. **Analyses of sequential weights of Nellore cattle using multiple trait and random regression models.** Adviser: Paulo Sávio Lopes. Committee Members: Robledo de Almeida Torres and Adair José Regazzi.

The objective of the first study was to obtain genetic parameters for sequential weights of beef cattle using RRM on data sets with missing and no missing traits, and to compare these estimates with those obtained by MTM. Growth curves of Nellore cattle were analyzed using body weights measured at ages ranging from 1 day (birth weight) to 733 days. Two data samples were created: one with 71,867 records from herds with missing traits and the other with 74,601 records from herds with no missing traits. Records preadjusted to a fixed age were analyzed by a multiple trait model (MTM), which included the effects of contemporary group, age of dam class, additive direct, additive maternal, and maternal permanent environment. Analyses were by restricted maximum likelihood (REML) with 5 traits at a time. The random regression model (RRM) included the effects of age of animal, contemporary group, age of dam class, additive direct, additive maternal, permanent environment, and maternal permanent environment. Legendre cubic polynomials were used to describe the random effects. Estimates of covariances by MTM were similar for both data sets, although those from the missing data set showed more variability from age to age. The estimates from RRM were similar to those from MTM only for the complete-trait case and showed large artifacts for the case of missing traits.

Estimates of additive direct-maternal correlations under RRM for some ages approached -1.0, and most likely contained artifacts. If many traits are missing, the best approach to obtaining parameters for RRM would be conversion from smoothed MTM estimates. The purpose of the second study was estimation of parameters of models and data sets as in the first study by a Bayesian methodology – Gibbs sampling, and to make comparisons with their estimates by REML. Analyses were by a Bayesian method for all 9 traits. MTM estimated covariance components and genetic parameters for birth weight and sequential weights and RRM for all ages. Estimates of additive direct variance from herds with missing traits increased from birth weight through weight at 551 to 651 days with MTM. However, this component also increased for the sample with no missing traits after this age. Additive direct and residual estimated variance with RRM increased over all ages for both samples. For MTM, additive direct and maternal heritabilities were greater from the sample with herds with missing traits than those values from herds with no missing traits. The estimates from RRM were slightly lower than those from MTM for the sample with no missing traits; however, additive maternal heritabilities from MTM were greater than those using RRM. The estimated additive direct genetic correlations for each pair of traits were slightly higher for the first age (birth weight) using MTM than RRM. The range of additive maternal genetic correlations was lower than that for additive direct genetic correlations with MTM and RRM. Due to the fact that covariance components based on RRM were inflated for herds with missing traits, MTM should be used and converted to covariance functions. As well, for analyses with standard models where inferences on shapes of parameters are not important, analyses by REML may be more robust. The first goal of the third study was to implement the genetic evaluation of weights for a large population of beef cattle using the random regression model. The second goal was to compare these evaluations with those obtained from a multitrait evaluation. Expected progeny differences (EPD) were computed by two methods: a finite method using sparse factorization (SF) and interating (IT) by preconditioned conjugate gradient (PCG). The correlations between EPDs from MTM and RRM by SF and IT were ≤ .43 until the random regressions were orthogonalized. After orthogonalization high computing requirements of RRM were reduced by removing regressions corresponding to very low eigenvalues and by replacing the random error effects with weights. Correlations between EPDs from MTM and RRM for the additive direct effect were .87, .89, .89, .87, and .86 for W1 (weight at 60 days), W2 (weight at 252 days), W3 (weight at 243 days), W5 (weight at 426 days), and W7 (weight at 601 days), respectively. The corresponding correlations for the additive maternal effect were .85, .86, .88, .85 and .84, respectively. These low correlations were mostly due to differences in variances between the models and, to a lesser degree, due to better accounting for environmental effects and more data by RRM. The RRM applied to beef weights may be poorly conditioned numerically.

RESUMO

NOBRE, Paulo Roberto Costa, D.S., Universidade Federal de Viçosa, novembro de 2001. **Análises de pesos seqüenciais de gado Nelore usando modelos de características múltiplas e regressões aleatórias.** Orientador: Paulo Sávio Lopes. Conselheiros: Robledo de Almeida Torres e Adair José Regazzi.

O objetivo do primeiro estudo foi estimar parâmetros para pesos seqüenciais de gado de corte, por meio de modelos de regressão aleatória (RRM), em características com informações perdidas e completas. Analisaram-se curvas de crescimento de gado Nelore mediante o uso de pesos corporais coletados, do nascer aos 733 dias de idade. Duas amostras foram geradas; a primeira era constituída de 71.867 medidas provenientes de rebanhos com informações perdidas, e a segunda, de 74.601 medidas oriundas de rebanhos com informações completas. Os pesos pré-ajustados a idades fixas foram analisados por meio de um modelo de características múltiplas (MTM), cinco características por vez, no qual foram incluídos efeitos de grupo contemporâneo, classe de idade da vaca, aditivo direto, aditivo materno e ambiente materno permanente. No modelo de regressão aleatória (RRM) foram incluídos efeitos de idade do animal, grupo contemporâneo, classe de idade da vaca, aditivo direto, ambiente permanente, aditivo materno e ambiente materno permanente. Polinômios cúbicos de Legendre foram utilizados na descrição dos efeitos aleatórios. Estimativas de covariâncias por meio de MTM foram similares em ambas as amostras, apesar de as obtidas da amostra com informações perdidas terem apresentado maior variabilidade entre as idades. As estimativas obtidas pelo RRM foram similares às obtidas pelo MTM somente para o caso de características completas e mostraram grande variabilidade para o caso de características com informações perdidas. Estimativas de correlações entre os efeitos aditivos direto e materno, por meio de RRM, foram iguais a -1.0, em algumas idades. Se várias informações forem perdidas, a melhor aproximação para obter parâmetros por meio de RRM seria a conversão das estimativas obtidas por meio de MTM. O segundo estudo objetivou estimar parâmetros por meio de modelos e características com informações perdidas e completas, à semelhança do primeiro estudo, mediante metodologia Bayesiana – Gibbs sampling, e efetivar comparações com as estimativas obtidas por meio da metodologia REML. As análises por meio do MTM foram para nove características. Estimaram-se componentes de covariâncias e parâmetros genéticos para específicos pontos seqüenciais, por meio do MTM; entretanto, por meio do RRM, tais estimativas foram obtidas para todas as idades. Estimativas de variâncias aditivas diretas para a amostra com informações perdidas aumentaram, do nascer à idade de 551 a 651 dias, pelo MTM, e em todas as idades, na amostra com informações completas. Estimativas de variâncias aditiva direta e residual, mediante RRM, aumentaram ao longo de todas as idades, em ambas as amostras. Pelo MTM, heritabilidades aditivas direta e materna foram maiores na amostra de rebanhos com informações perdidas do que na de rebanhos com informações completas. As estimativas obtidas pelo RRM foram ligeiramente menores do que aquelas obtidas pelo MTM na amostra com informações completas. Heritabilidades aditivas maternas pelo MTM foram maiores do que aquelas obtidas pelo RRM. As estimativas de correlações genéticas aditivas diretas foram levemente maiores para peso ao nascer, quando se utilizou MTM do que quando se empregou RRM. A amplitude das correlações genéticas aditivas maternas foi menor do que a do efeito genético aditivo direto, pelo MTM e pelo RRM. Tendo em vista que os componentes de covariância baseados em RRM são influenciados por informações perdidas, recomendam-se o MTM e a conversão destes componentes em funções de covariância. Além disso, nas análises com modelos-padrão em que inferências dos parâmetros não são importantes, o REML deve ser escolhido. Um terceiro trabalho objetivou a implementação de avaliação genética em bovinos de corte, utilizando modelo de regressão aleatória. Além disso, as avaliações foram comparadas com aquelas estimadas por meio de um modelo de características múltiplas. Dois métodos foram considerados nas análises: um método finito, FSPAKF90 (*Factorization sparse matrix package*), e o de iteração nos dados, PCG (*Preconditioned conjugate gradient*). As correlações entre as diferenças esperadas nas progênies (DEP), estimadas pelo MTM e pelo RRM, foram muito baixas antes de se terem as regressões aleatórias ortogonais. Grande demanda computacional dos RRM foi reduzida pela remoção das regressões correspondentes a pequenas variâncias e também pela substituição dos efeitos aleatórios do erro por específica ponderação. Correlações entre DEPs, estimadas pelo MTM e pelo RRM para efeito aditivo direto, foram .87, .89, .89, .87 e .86 para W1 (peso aos 60 dias), W2 (peso aos 152 dias), W3 (peso aos 243 dias), W5 (peso aos 426 dias) e W7 (peso aos 601 dias), respectivamente. As correlações correspondentes, para efeito aditivo materno, foram .85, .86, .88, .85 e .84, respectivamente. Estimativas obtidas pelos RRM em informações ponderais de gado de corte podem não ser adequadas, em virtude das propriedades numéricas desses modelos. Em geral, baixas correlações são devidas a diferenças em variâncias entre modelos, número insuficiente de graus de liberdade para estimar os efeitos de ambiente e informações perdidas nos RRM.

CHAPTER I – INTRODUCTION

In domestic animals, traits of economic importance are usually dependent on many pairs of genes. Environmental factors interact with gene effects in an additive or interactive ways, promoting variation in traits. Therefore, the genetic progress sought by the breeder through the application of selective methods of breeding will depend on his capacity to use, for reproduction, animals of superior breeding value for the characteristics desired. In this context, progress in obtaining the characteristics will depend on, among other factors, identifying superior animals and determining whether this superiority is due to the genes the animal possesses or to the environment in which it lives.

Breeding value is a measure used in the process of sire evaluation, and it corresponds to a prediction of the value of the genes of individual that can be transmitted to the following generation, if the animal is selected for breeding. In the selection of sires in beef cattle, it is important to provide breeding values that are useful to the commercial industry. Thus, breeders would be providing breeding stock that should transmit half of its sample of germplasm to facilitate an additional gain in meat or of live weight in the progeny. On the other hand, the commercial producers sell kilograms of meat, not the genetic value. However, they need to acquire sires that possess desirable genetic values in order to obtain advantages for all traits, especially reproductive traits. In this way, breeders can select genetically superior animals, and commercial producers can benefit from knowing the prediction of sire genetic values.

Geneticists and breeders use genetic selection as a way of increasing the efficiency of beef cattle production. Thus, they need methodologies, that predict the individuals' breeding values. These values are dependent on the data and the (co)variance components of all traits.

Several methodologies have been presented for the estimation of these components. The Restricted Maximum Likelihood (REML) method uses all available information, has stable convergence, and takes selection into account. Because of these properties, it is a standard method to estimate covariance components for use in animal genetic evaluation. Bayesian methods have been used to solve many problems related to parameter estimation in animal populations as well.

The Best Linear Unbiased Prediction (BLUP) is a standard methodology for beef cattle, and the evaluations use models that include the effects of direct genetic, maternal genetic, and maternal permanent environmental factors.

Most beef evaluations have been more concerned with multiple trait selection. However, multiple trait models provide predictions for only point estimates of birth, weaning, yearling, and post-yearling weights. In contrast, the random regression model for longitudinal data, which was outlined by LAIRD and WARE (1982), describes the traits at all points rather than at finite points. The development of more efficient algorithms and the addition of hardware processing capacity makes the genetic analyses of longitudinal data by random regression models possible.

Although random regression models have desirable features, the estimates of parameters by these models may be poor (MISZTAL, 2000). This could be due to large number of parameters, scarcity of data at extreme points, and inadequate properties of the estimation methodologies. Therefore, parameters estimated by random regression models need to be validated.

The first objective of this study was to estimate the covariance components of sequential traits in beef cattle by REML and Gibbs sampling. The second objective was to validate parameters obtained from random regression models by multiple trait methodology.

CHAPTER II – REVIEW OF LITERATURE

This literature review summarizes studies of sequential weights in beef cattle on models, methodologies, and genetic parameters.

Data analyses and models

In single trait analysis the basic assumption is that a single measurement arises from a single unit. In multitrait analysis, a number of different traits are measured from each experimental unit; for example, body weight, and these measurements may be correlated. When the same physical quantity is measured sequentially over time on each experimental unit, they are described as repeated measurements, which can be a special form of a multivariate case.

Growth traits such as birth weight, weaning weight (155 to 255 days), yearling weight (305 to 365 days), and post yearling weight (500 to 600 days) are recorded and routinely analyzed in genetic evaluations in beef cattle (ABCZ, 2001 and CNPGC, 2001). Therefore, most beef genetic evaluations provide predictions for only point estimates of these weights. However, a more comprehensive description of individual animal growth can be estimated using these weights and others measured later in life (KAPS et al., 2000).

For traits that are measured over time, a simple repeatability model that assumes constant variances and correlation does not hold. In standard analyses, this has been taken into account by treating records in different age intervals as different traits (MEYER, 2000).

The beef cattle industry is interested in the growth of the animal throughout its lifetime because of genetic and phenotypic relationships among the growth traits. Important traits, such as relative and absolute growth rates, maturing rate, and mature size can be derived from growth functions, such as those described by BRODY (1945) and RICHARDS (1959).

DeNISE and BRINKS (1985) showed that Brody's curve was especially dependent on the input data. The curve fit actual weights poorly when expected to project beyond the range of information; for example, when birth and mature weights were absent. Richard's function appeared to be less dependent on the range of data.

BELTRAN et al. (1992) reported that Brody's model gave adequate predictions of weights from 18 months to maturity, but early weights were slightly overestimated. KAPS et al. (2000) reported that beef breed genetic evaluations are performed only for points along the growth curve and concluded that those data could be used also to evaluate animals for developmental and maturing patterns. Growth and maturing traits derived from a growth curve and their relationships with direct and maternal weaning weight could be considered as additional information for use in genetic evaluation programs.

Random regression models have recently been recognized as ideally suited for the analysis of longitudinal data in animal breeding (MEYER, 2000). LAIRD and WARE (1982) outlined a random regression mixed model for longitudinal data with both growth curve and repeatability models as special cases. HENDERSON JR. (1982) pointed out that the mixed model equations could be used to analyze the random-coefficient model.

There are two advantages of analyses with repeated measurements: possibility of analyzing continuous changes, and increased statistical power. In animals random regression models have first been proposed to model test day production records of dairy cows (SCHAEFFER and DEKKERS, 1994). Applications in this area were reported by JAMROZIK and SCHAEFFER (1997), JAMROZIK et al. (1997), VAN DER WERF et al. (1998), VEERKAMP and THOMPSON (1999), and STRABEL and MISZTAL (1999).

Recent conceptual and computational advancements have made the genetic analysis of continuous traits possible. Three methods have been described in the literature. First, random regression models have been widely used for the analysis of longitudinal data in the traditional statistical literature (DIGGLE et al., 1994) and recently have been applied in the breeding context (JAMROZIK et al., 1997). Second, the use of orthogonal polynomials to approximate covariance matrices was initially suggested by KIRKPATRICK and HECKMAN (1989) and is closely related to the random regression models (MEYER and HILL, 1997 and MEYER, 1998). Third, the character process model was recently proposed by PLETCHER and GEYER (1999) and is based on theories of stochastic processes. JAFFREZIC and PLETCHER (2000) compared and contrasted random regression, orthogonal polynomials, and character process models. The authors concluded that random regression and orthogonal polynomial methods work quite well when the correlation structure remains high over time. The advantage of the character process appears to be the ability to model the variance and correlation separately.

MEYER (1998 and 2000) and SAKAGUTI (2000) reported applications of random regression models in beef cattle. MEYER (1998) reported that random regression models provide a valuable tool to model repeated records in animal breeding, especially if the traits measured change gradually. Random regression models allow covariance functions to be formulated to describe genetic and environmental covariances among records over time. Moreover, random regression models impose a structure on covariance matrices. By regressing on random orthogonal polynomials of the continuous scale variable, the coefficients of covariance functions can be estimated as covariances among the regression coefficients.

MEYER (2000) concluded that random regression models are well suited for analysis of growth data, even if they are subject to seasonal variation, which cannot be separated. Regression on orthogonal polynomials as general purpose functions is recommended as well, especially if higher orders of fitting are desirable. In the same paper, MEYER (2000) reported that random regression models accommodate repeated records for traits, which change gradually and continually over time, and which do not require stringent assumptions about constancy of variances and covariances.

SAKAGUTI (2000) reported that the random regression models allowed a better goodness of fitting to the weight data, estimated breeding values and variance components for weights at any pair of ages, and provided additional and very useful parameters for beef cattle genetic evaluation.

VARONA et al. (1997) and VILLALBA et al. (2000) studied growth curves in beef cattle. VARONA et al. (1997) reported that the application of this procedure to complex biological production curves might allow altering the shape of curves of production through genetic selection on its underlying variables. This indicates that in beef production it would be possible to alter maturing rate without changes in adult weight.

VILLALBA et al. (2000) estimated the preweaning growth curves of calves for several environmental factors. Estimates of individual variability of the parameters describing the growth curve were also obtained. The authors concluded that the use of random regression coefficients improved the adjustment for weaning weight.

A procedure for obtaining approximate reliabilities of estimated breeding values under a random regression model was presented by JAMROZIK et al. (2000). The accuracy of the proposed approximation was tested using a multiple trait random regression test day for dairy production traits. The approximation was shown to be relatively unbiased for both bulls and cows.

Growth curves for the Nellore breed were estimated using the Von Bertalanffy nonlinear model and the Gauss Newton model by FREITAS et al. (2000). The authors concluded that the animal's performance, from birth to 750 days, was adequately estimated because the coefficients of determination were greater than 98.0%.

Random regression models with parameters estimated by REML and Gibbs sampling are powerful tools in animal breeding and genetics because of the use of the hierarchical models that account for several levels of identifiable biological variability.

Variance Components – Methodologies

Models used in animal breeding require knowledge of parameters, which usually are variance components. These components are useful in animal breeding for a number of reasons. Firstly, functions of variances provide heritabilities, repeatabilities, and prediction error variances of genetic evaluations. Secondly, they help when designing experiments to determine the necessary sample size to detect significant differences. They are also useful in predicting expected genetic values.

Different methodologies have been presented for the estimation of variance components. The analysis of variance (ANOVA) was one of the first methods to estimate these components. However, it required that the analyzed data were balanced. In general, animal field data is unbalanced. Therefore, there is no one unique solution to the sum of squares.

Henderson's Methods

Modifications in the ANOVA method were presented by HENDERSON (1953) for unbalanced data, becoming known as Henderson's methods I, II and III. The last method was used for many years.

Henderson's method I is only applicable to random models in populations with complete random mating and no selection. This method is unbiased and translation invariant. Henderson's method II is an unbiased, translation invariant procedure and can be used with mixed models provided that interactions between fixed and random factors, or nesting of random factors within fixed factors, do not exist in the model. Henderson's method III is capable of handling general mixed models and has been called the fitting constants method because it uses reductions in sums of squares to fitting submodels of the full model. This method is unbiased and translation invariant, but is not uniquely defined in the sense that more reductions can be computed than are necessary to estimate the variance components (SCHAEFFER, 1998)

Method III has been shown in some situations to provide more accurate estimates of covariance components than methods I and II. The reason for greater accuracy is due to fitting submodels with two or more random factors simultaneously rather than separately, as would be the case in methods I and II (SCHAEFFER, 1998). Disadvantages of the method III are the possibility of negative variances estimates and the lack of uniqueness due to the fact that there is no unique set of sum of squares (SEARLE et al. 1992)

The Maximum Likelihood Method

The Maximum Likelihood Method (ML) of estimating covariance components was proposed by HARTLEY and RAO (1967). The ML methodology consists of maximizing the function of likelihood of observations in relation to the fixed effects and to the covariance components of random effects of the model.

Assuming that the mixed model for covariance component estimation is:

$$y = X\beta + Zu + e$$

where:

y is a vector of observations,
β is a vector of fixed effects,
u is a vector of random effects,
X and Z are known incident matrices, and
e is a vector of random residual effects.

The first and second moments are as follows:

$$E[y] = X\beta \qquad E[u] = 0 \qquad E[e] = 0,$$

and

$$Var\begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}$$

$$Var[y] = V = ZGZ' + R$$

Under the multivariate normality assumption of u and e, $y \sim MVN\{X\beta, V\}$, the likelihood to be maximized is:

$$L = L(\hat{a}, V|y) = (2\eth)^{-\frac{N}{2}}|V|^{-\frac{1}{2}} exp\{-\tfrac{1}{2}(y - X\hat{a})' V^{-1}(y - X\hat{a})\}$$

The logarithm of the likelihood function is given by:

$$l = ln(L) = -\tfrac{N}{2} ln(2\eth) - \tfrac{1}{2} ln|V| - \tfrac{1}{2}(y - X\hat{a})' V^{-1}(y - X\hat{a})$$

To obtain the estimates of β***,*** *G* and *R****,*** the logarithm of likelihood function needs to be maximized, for example, by equating derivatives of *L* with respect to β*, G* and *R* to zero.

The ML is translation invariant, but biased in that it does not take into account the degrees of freedom used for estimating the fixed effects.

The Restricted Maximum Likelihood Method (REML) proposed by PATERSON and THOMPSON (1971) eliminates the bias present in the ML method. REML, however, is derived from linear combinations of *y* so that these combinations do not contain any fixed effects. The likelihood function for REML is based on:

$$K'y \sim MVN\{0, K'VK\},$$

where:

K is a matrix of linear combinations of *y* such that: $K'X\beta = 0$ for all β*.* Then, the likelihood function *(L)* is:

$$L = L(\hat{a}, V|K'y) = (2ð)^{-\frac{1}{2}(N\text{-}rank(X))}|K'VK|^{-\frac{1}{2}}exp\{-\tfrac{1}{2}y'K(K'VK)^{-1}K'y\}$$

and its logarithm is:

$$ln(L) = -\frac{1}{2}(N - rank(\mathbf{X}))ln(2ð) - \frac{1}{2}ln|\mathbf{K'VK}| - \frac{1}{2}\mathbf{y'K(K'VK)^{-1}K'y}$$

HARVILLE (1977) showed that:

$$ln|K'VK| = ln|V| + ln|X'V^{-1}X|,$$

and

$$y'K(K'VK)^{-1}K'y = y'Py,$$

where:

$$P = V^{-1} - V^{-1}X(X'V^{-1}X)^{-}X'V^{-1},$$

Hence, *ln(L)* can be written as

$$ln(L) = -\tfrac{1}{2}(N - rank(X))ln(2ð) - \tfrac{1}{2}ln|V| - \tfrac{1}{2}ln|X'V^{-1}X| - \tfrac{1}{2}y'Py$$

In practice, the alternative form for the logarithm of likelihood function is based on the following equations:

$$ln|V| = ln|R| + ln|G| + ln|Z'R^{-1}Z + G^{-1}|\ ,$$

and

$$ln|X'V^{-1}X| = ln|C| - ln|Z'R^{-1}Z + G^{-1}|$$

where:

C is a coefficient matrix of the mixed model equation.

Notice that $-\tfrac{1}{2}(N - rank(X))ln(2ð)$ is a constant (*K*), therefore, the logarithm of likelihood can be written as:

$$ln(L) = -\tfrac{1}{2}\{\, k + ln|R| + ln|G| + ln|C| + y'Py\,\},$$

In addition, SEARLE et al. (1992) showed that:

$$y'Py = (y - X\,\hat{a}^0\,)'V^{-1}(y - X\,\hat{a}^0\,) = y'R^{-1}(y - X\,\hat{a}^0 - Z\hat{u})$$

Combining these results one has:

$$-2ln(L) = k + ln|R| + ln|G| + ln|C| + y'R^{-1}(\,y - X\,\hat{a}^0 - Z\hat{u}\,)$$

The formulas for estimating the variance components are derived by maximizing this function directly or equating its derivative to zero.

Considering that a mixed model with several uncorrelated random effects is:

$$y = X\hat{a} + \sum_{i}^{r} Z_i u_i + e$$

where:

y is a vector of observations,

β is a vector of fixed effects,

u_i is a vector of random effects ***i***,

X and Z_i are known incident matrices,

e is a vector of random residual effects, and

r is the number of random effects.

The first and second moments are:

$E[y] = X\beta \qquad E[u] = 0 \qquad E[e] = 0,$

and

$Var[u_i] = A_i\sigma_i^2$

$Var[e] = R = I\sigma_e^2$

$Var[y] = \sum_i^r Z'G_iZ + R$

where:

σ_i^2 is variance for random effect ***i***,

σ_e^2 is variance for residual effect, and

A_i is the known incident matrix for random effect *i*.

The derivatives of the logarithm of likelihood function with respect to particular components lead to the following:

Estimation of σ_i^2

$$ln(L) = -\tfrac{1}{2}(N - rank(\mathbf{X}))ln(2\eth) - \tfrac{1}{2}ln|V| - \tfrac{1}{2}ln|X'V^{-1}X| - \tfrac{1}{2}y'Py$$

$$\frac{\partial ln(L)}{\partial\sigma_i^2} = -\tfrac{1}{2}tr(V^{-1}\frac{\partial V^{-1}}{\partial\sigma_i^2}) + \tfrac{1}{2}tr\{(X'V^{-1}X)^{-}X'V^{-1}\frac{\partial V^{-1}}{\partial\sigma_i^2}V^{-1}X\} + \tfrac{1}{2}y'P\frac{\partial V^{-1}}{\partial\sigma_i^2}Py$$

$$\frac{\partial ln(L)}{\partial\sigma_i^2} = -\tfrac{1}{2}tr\{[V^{-1} - V^{-1}X(X'V^{-1}X)^{-}X'V^{-1}]\frac{\partial V^{-1}}{\partial\sigma_i^2}\} + \tfrac{1}{2}y'P\frac{\partial V^{-1}}{\partial\sigma_i^2}Py$$

$$\frac{\partial ln(L)}{\partial \sigma_i^2} = -\tfrac{1}{2} tr(P\frac{\partial V^{-1}}{\partial \sigma_i^2}) + \tfrac{1}{2} y'P\frac{\partial V^{-1}}{\partial \sigma_i^2}Py$$

$$\frac{\partial ln(L)}{\partial \sigma_i^2} = -\tfrac{1}{2} tr(PZ_iZ_i') + \tfrac{1}{2} y'PZ_iZ_i'Py$$

$$\frac{\partial ln(L)}{\partial \sigma_i^2} = -\tfrac{1}{2}\{\frac{q_i}{\sigma_i^2} - \frac{\sigma_e^2}{\sigma_i^4} tr(A_i^{-1}C^{ii})\} + \tfrac{1}{2}\frac{1}{\sigma_i^4}(\hat{u}_i'A_i^{-1}\hat{u}_i)$$

After equating the derivatives to zero and multiplying equation by $\hat{\sigma}_i^4$, the equation for $\hat{\sigma}_i^2$ becomes:

$$\hat{\sigma}_i^2 = \frac{\hat{u}_i'A_i^{-1}\hat{u}_i + \hat{\sigma}_e^2 tr(A_i^{-1}C^{ii})}{q_i}$$

where:

C^{ii} is a diagonal block of the inverse of C corresponding to the random effect *i,* and q_i is the dimension of the vector u_i.

Estimation of σ_e^2

$$ln(L) = -\tfrac{1}{2}(N - rank(X))ln(2\eth) - \tfrac{1}{2}ln|V| - \tfrac{1}{2}ln|X'V^{-1}X| - \tfrac{1}{2}y'Py$$

$$\frac{\partial ln(L)}{\partial \sigma_e^2} = -\tfrac{1}{2} tr(V^{-1}\frac{\partial V^{-1}}{\partial \sigma_e^2}) + \tfrac{1}{2} tr\{(X'V^{-1}X)^{-} X'V^{-1}\frac{\partial V^{-1}}{\partial \sigma_e^2}V^{-1}X\} + \tfrac{1}{2} y'P\frac{\partial V^{-1}}{\partial \sigma_e^2}Py$$

$$\frac{\partial ln(L)}{\partial \sigma_e^2} = -\tfrac{1}{2} tr\{[V^{-1} - V^{-1}X(X'V^{-1}X)^{-} X'V^{-1}]\frac{\partial V^{-1}}{\partial \sigma_e^2}\} + \tfrac{1}{2} y'P\frac{\partial V^{-1}}{\partial \sigma_e^2}Py$$

$$\frac{\partial ln(L)}{\partial \sigma_e^2} = -\tfrac{1}{2} tr(P\frac{\partial V^{-1}}{\partial \sigma_e^2}) + \tfrac{1}{2} y'P\frac{\partial V^{-1}}{\partial \sigma_e^2}Py$$

$$\frac{\partial ln(L)}{\partial \sigma_e^2} = -\tfrac{1}{2} tr(P) + \tfrac{1}{2} y'PPy$$

$$\frac{\partial ln(L)}{\partial \sigma_e^2} = -\tfrac{1}{2}[\{N - rank(X)\} - \frac{1}{\sigma_i^2}\hat{u}_i' A_i^{-1} \hat{u}_i\ ] + \tfrac{1}{2}[\frac{1}{\sigma_e^2}(y'Py - \hat{u}_i' A_i^{-1} \hat{u}_i \alpha_i)]$$

Equating the derivatives to zero leads to:

$$\{N - rank(X)\} - \frac{1}{\hat{\sigma}_i^2}\hat{u}_i' A_i^{-1} \hat{u}_i \ = \frac{1}{\hat{\sigma}_e^2}(y'Py - \hat{u}_i' A_i^{-1} \hat{u}_i \alpha_i)$$

Multiplying both sides with $\hat{\sigma}_e^2$:

$$\{N - rank(X)\}\hat{\sigma}_e^2 - \alpha_i \hat{u}_i' A_i^{-1} \hat{u}_i \ = y'Py - \hat{u}_i' A_i^{-1} \hat{u}_i \alpha_i$$

and, the estimate for $\hat{\sigma}_e^2$ becomes:

$$\hat{\sigma}_e^2 = \frac{y'Py}{N - rank(X)}$$

or

$$\hat{\sigma}_e^2 = \frac{y'(y - X\hat{a}^0 - \sum_{i=1}^{r} Z_i \hat{u}_i)}{N - rank(X)}$$

These formulas are equivalent and can be derived differently as Expectation Maximization (EM) REML algorithm (DEMPSTER et al., 1977).

Following HENDERSON (1984), a faster converging version of the same formulas are:

$$\hat{\sigma}_i^2 = \frac{\hat{u}_i' A_i^{-1} \hat{u}_i}{n_i - tr(A_i^{-1} C^{ii})\dfrac{\hat{\sigma}_e^2}{\hat{\sigma}_i^2}}$$

where:

n_i is the number of levels in the random effect *i*.

The residual estimated is:

$$\hat{\sigma}_e^2 = \frac{y'(y - X\hat{a}^0 - \sum_{i=1}^{r} Z_i \hat{u}_i)}{n_r - r(X'X)}$$

where:

n_r is the number of records, and $r(X'X)$ is the rank of the fixed-effect part of the mixed model equations.

For multiple traits, following MISZTAL (2000) let us consider G as:

$$G = \begin{bmatrix} \cdots & \cdots & \cdots \\ \cdots & \sigma_{ijkl}^2 A_{ijkl} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

and the system of equations as:

$$(W'R^{-1}W + G^{-1})\,\hat{\theta} = W'R^{-1}y$$

$$W = [X, Z]$$

$$\hat{\theta} = \begin{bmatrix} \hat{a}^0 \\ \hat{u} \end{bmatrix}$$

where:

A_{ijkl} can either be a numerator relationship matrix or an identity matrix for a block corresponding to effects *i* and *j* and traits *k* and *l*.

Denote $A^{ijkl} = A_{ijkl}^{-1}$

The first-derivative REML in multiple traits is similar to single traits for random effect variances:

$$\hat{\sigma}_{ijkl}^2 = \frac{\hat{u}_{ik}' A^{ijkl} \hat{u}_{jl} + tr(A^{ijkl} C^{ijkl})}{n_{ik}}$$

where:

n_{ik} is the number of levels for effect i and k ($n_{ik} = n_{jl}$).

The formula for the residual variances when all traits are recorded is:

$$\hat{\sigma}^2_{e_{ij}} = \frac{\hat{e}'_i \hat{e}_j + tr(C^{ij} W'_i W_j)}{n_r}$$

where:

$$\hat{e}_i = y_i - X\hat{a}_i^0 - Z_i \hat{u}_i$$

and,

n_r is the number of records.

One interesting property of the REML method is that, for balanced data, the solutions for the system of equations in REML produce identical results to the ANOVA method estimator. In the same way as that of the ML method, its limitation is high computing cost.

Both ML and REML equations are nonlinear. Therefore, solutions for variance estimation are obtained by numerical iteration. Several iterative algorithms have been proposed. Newton-Raphson is an iterative procedure for maximizing or minimizing nonlinear functions based on second derivatives. But, the computation of second derivatives of the log likelihood function is quite expensive. PATTERSON and THOMPSON (1971) suggested the Fisher scoring (FS) algorithm, which is a procedure characterized by replacing the computation of second-order partial derivatives by their expected values.

Various algorithms have been presented to overcome these problems to deal with large data sets. The derivative free (DF) algorithm, which does not require derivatives, was described by GRASER et al. (1987). The DF algorithm has, at most, half the numerical accuracy of the derivative algorithm, and its computational cost is proportional to t^5, where t is the number of traits.

MISZTAL (1997) reported that the numerical precision was a problem with DF, and the computational requirement was still a significant problem in multiple trait analysis.

The maximum point can be obtained by searching the surface of the likelihood function by Simplex or other methods. This method is efficient and accurate with a small number of variance components (SEARLE et al., 1992).

The Expectation Maximization algorithm is an efficient procedure for obtaining REML estimates of variance components because it guarantees convergence within the parameter space. The E-step involves obtaining the expectation of the sufficient statistics conditional on the incomplete data, $E(u'u \mid y)$. The M-step is to maximize the likelihood of the complete data. However, MISZTAL (1990) reported that the EM algorithm is especially slow with some data structures such as highly unbalanced data or with lowly heritable traits.

The average information (AI) algorithm avoids multiplication of large matrices. Instead of using the information matrix, the average of information matrix and the expectation of information matrix are used in the equation, resulting in simplification of formulas. JOHNSON and THOMPSON (1995) showed that the AI algorithm performed well when compared to the DF and EM algorithms. As reported by GILMOUR et al. (1995) the computational requirements of the AI algorithm for each round are similar to those of the EM algorithm, but the number of iterations is usually much lower.

The method of REML has been preferable for estimating variance components in animal breeding because the estimates are within the parameter space, all pedigree information is used, and the estimates are unbiased under selection (THOMPSON, 1973; SORENSEN and KENNEDY, 1984). Its disadvantages include high cost for large models and relatively difficult programming, especially for nonstandard models, although various implementations of REML differ widely in their computing properties (MISZTAL, 2000).

Bayesian Methods and Gibbs Sampling

The application of Markov Chain Monte Carlo (MCMC) methods, of which Gibbs sampling is a special case, can be used in a way to obtain a Bayesian inference about the parameter of interest. Bayesian statistics differ from traditional statistics in a number of ways. Even so, many procedures in statistics can have both a Bayesian and a traditional interpretation. Bayesian methods consider prior information with a particular probability and combine those with information contained in the sample of data .

Following SCHAEFFER (1998) a prior density is the information about the unknown parameters (*b* and σ_i^2) before the sample *(y)* is observed. Then,

$$(b, \sigma_i^2, i = 0,1,..,s).$$

The posterior density represents knowledge of the parameters after the sample is observed. The parameters are as follows:

$$(b, \sigma_i^2, i = 0,1,..,s \mid y).$$

Bayesian inference involves parameters of the posterior distribution, such as the mean, variance, mode and median. Bayesian inference has been used in many animal breeding applications to report variance components.

Gibbs sampling is a numerical integration method useful for integrating out all other parameters from the joint distribution. This methodology was applied to estimate variance components in a general mixed model by WANG et al. (1993), in models including maternal effects by JENSEN et al. (1994), and in multiple trait models by VAN TASSEL and VAN VLECK (1996). Gibbs sampling generates random samples from the marginal posterior distributions iteratively by sampling from the conditional distributions. Marginal distributions are often needed to make appropriate inferences. However, the complexity of the joint posterior distributions usually make analytical derivation of marginal distributions impossible.

Assuming that the mixed model for covariance component estimation is:

$$y = X\beta + Zu + e$$

where:

y is a vector of observations,

β is a vector of fixed effects,

u is a vector of random effects,

X and Z are known incident matrices, and

e is a vector of random residual effects.

The first and second moments are as follows:

$$E[y] = X\beta \qquad E[u] = 0 \qquad E[e] = 0,$$

and

$$Var\begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}$$

$$Var[y] = V = ZGZ' + R$$

The Bayesian approach is to derive the joint posterior distribution by application of the Bayes theorem.

Following GIANOLA and FERNANDO (1986) let us consider θ be as a vector of parameters and y a data vector, then:

$$p(\theta\,, y) = p(\theta)\, p(y \mid \theta) = p(y)\, p(\theta \mid y),$$

or

$$p(\theta \mid y) = p(\theta)\frac{p(y \mid \theta)}{p(y)}$$

where:

$p(\theta)$ is a prior probability for θ, and

$p(\theta \mid y)$ is a posterior probability function of θ.

Since $p(y)$ does not involve parameter (θ), in practice, the Bayesian inference can be written as:

$$p(\theta \mid y) \propto p(\theta)\, p(y \mid \theta).$$

For a simple animal model, θ includes β, u, σ^2_a, σ^2_e. The conditional distribution that generates y given θ is:

$$y \mid \beta, u, \sigma^2_a, \sigma^2_e \sim N\{X\beta + Zu,\ I\sigma^2_e\}$$

$$p(y \mid \beta, u, \sigma^2_a, \sigma^2_e) \propto \frac{1}{(\sigma^2_e)^{n/2}} exp\{-\frac{1}{2\sigma^2_e}[y - X\beta - Zu]' [y - X\beta - Zu]\}$$

Prior Distributions

To complete the Bayesian inference, following SORENSEN (1988), all unknown parameters in the model will be assigned to prior distributions. Generally, the flat prior is assigned to β, so:

$p(\beta) \propto$ constant

It is assumed that u_i is mutually multivariate normal and independent of β.

$$[u \setminus G, \sigma^2_a] \sim N\{0, A\sigma^2_a\}$$

Scaled inverted χ^2 distributions can be applied as priors for variance components:

$$p(\sigma^2_e | \nu_e, s^2_e) \propto (\sigma^2_e)^{-\frac{\nu_e}{2}-1} exp\{-1/2\nu_e\ s^2_e/\sigma^2_e\}$$

$$p(\sigma^2_a | \nu_u, s^2_a) \propto (\sigma^2_a)^{-\frac{\nu_u}{2}-1} exp\{-1/2\nu_u\ s^2_a/\sigma^2_a\}$$

where:

s^2_e and s^2_a are arbitrary prior values for a particular variance, and ν is the degree of belief.

Joint Posterior Distributions

The joint posterior distribution of the unknown parameters β, u, σ_a^2, σ_e^2) is proportional to the product of likelihood function and the joint prior distribution. As described by GIANOLA and FERNANDO (1986), the joint posterior distribution becomes a normal gamma form as follows:

$$p(\beta, u, \sigma_a^2, \sigma_e^2 \mid y) \propto p(\beta)p(u \mid \sigma_a^2)p(\sigma_a^2)p(\sigma_e^2)p(y \mid \beta, u, \sigma_a^2, \sigma_e^2)$$

Then,

$$p(\beta, u, \sigma_a^2, \sigma_e^2 \mid y) \propto \frac{1}{(\sigma_e^2)^{n/2+1}} exp\{-\frac{1}{2\sigma_e^2}[y - X\beta - Zu]' \; [y - X\beta - Zu]\}$$

$$x \; \frac{1}{(\sigma_u^2)^{q/2+1}} \; exp\{-\frac{1}{2\sigma_u^2} u'Au\}]$$

Full Conditional Posterior Distributions

In order to implement Gibbs sampling, the full conditional density of particular unknown parameters must be obtained by regarding all other parameters from the joint posterior distribution as known. The conditional distribution for β and u are:

$$\beta_i \mid \beta_{-i}, u, \sigma_a^2, \sigma_e^2, y \sim N\{\hat{a}_i^0, (x_i'x_i)^{-1}\sigma_e^2\},$$

$$u_i \mid \beta, u_{-i}, \sigma_a^2, \sigma_e^2, y \sim N\{\hat{u}_i, (z_i'z_i + A_{i,i}^{-1}\alpha)^{-1}\sigma_e^2\},$$

where:

$\beta_i^0 = (x_i'x_i)^{-1}x_i' \; (y - X_{-i}\beta_{-i} - Zu),$

$\hat{u}_i = (z_i'z_i + A_{i,i}^{-1}\alpha)^{-1}z_i' \; (y - X\beta - Z_{-i}u_{-i} - A_{i,-i}^{-1}\alpha \, u_{-i}),$

β_i is an element i of vector β,

β_{-i} is a vector β with the element i excluded,

u_i is an element i of vector u,

u_{-i} is a vector u with the element i excluded,

A^{-1} is the inverse of additive genetic relationship matrix,

$A^{-1}_{i,i}$ is an element *i* of matrix A^{-1},

$A^{-1}_{i,-i}$ is a row *i* of matrix A^{-1} with the element *i* excluded,

X_{-i} is the matrix *X* with column *i* deleted,

x_i is a vector of column *i* from matrix *X*,

Z_{-i} is the matrix *Z* with column *i* deleted,

z_i is a vector of column *i* from matrix *Z*, and α is $\frac{\sigma_e^2}{\sigma_a^2}$.

The conditional distribution for σ_a^2 is scaled inverted χ^2:

$$\sigma_a^2 \mid \beta, u, \sigma_e^2, y \sim \tilde{\nu}_u \tilde{s}_a^2 \chi^{-2}_{\tilde{\nu}_u}$$

where:

$\tilde{s}_a^2 = u' Gu \,/\, \tilde{\nu}_u$, and $\tilde{\nu}_u = \mathrm{q_i} + \nu_u$

The conditional distribution for σ_e^2 is scaled inverted χ^2:

$$\sigma_e^2 \mid \beta, u, \sigma_a^2, y \sim \tilde{\nu}_e \tilde{s}_e^2 \chi^{-2}_{\tilde{\nu}_e}$$

where:

$\tilde{s}_e^2 = (y - X\beta - Zu)' (y - X\beta - Zu) / \tilde{\nu}_e$, and $\tilde{\nu}_e = \mathrm{n} + \nu_e$

In multiple traits, following MISZTAL (2000), the marginal distributions for u's and β's can be obtained as in single traits, by using multidimensional analyses of univariate distribution. For variances, a univariate inverted Chi-square distribution becomes a multivariate inverted Wishart.

Consider a *t* trait model with a single random effect. Denote u_i and e_i as i^{th} trait random samples for the random effect and the residual, respectively.

The distribution for the variance components of the random effect is:

$$G_o \sim IW\,[n_g + v_g\,,\ (v_g\,S_g + P_a)^{-1}\,]$$

where:

n_g is the number of levels in the random effect,

v_g is degrees of belief (equal to -t-1 for flat priors), and

$$P_a = \begin{bmatrix} u'_1 G^{11} u_1 & u'_1 G^{12} u_2 & \dots \\ u'_2 G^{21} u_1 & u'_2 G^{22} u_2 & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

The distribution for the residual variance components is:

$$R_o \sim IW\,[n + v_e\,,\ (v_e\,S_e + P_e)^{-1}\,]$$

where:

n is the maximum number of records (including missing),

v_e is degrees of belief (equal to -t-1 for flat priors), and

$$P_e = \begin{bmatrix} e'_1\,e_1 & e'_1\,e_2 & \dots \\ e'_2\,e_1 & e'_2\,e_2 & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

If some records are missing, they are predicted from the distribution with known records. Assuming that subscript m denote quantities associated with missing traits, and p associated with present traits. The missing records can be predicted as:

$$y_m \sim MVN\,[E(y_m|\dots),\ Var(y_m|\dots)]$$

where:

$$E(y_m|\dots) = X_m\beta + Z_m u + R_{mp}\,R_{pp}^{-1}\,(y_p - X_p\beta + Z_p u)$$

$$Var(y_m|\dots) = N(\,0\,,\ R_{mm} - R_{mp}^{-1} R_{pp}^{-1} R_{pm})$$

Forming Samples

Gibbs sampling can be used in various ways. SORENSEN (1998) reported that these differing methods affect the degree of correlation between samples, estimates of the Monte Carlo variance, and rates of convergence. GEYER (1992) proposed Markov Chain Monte Carlo inference based on a single long chain. However, GELFAND and SMITH (1990) suggested generating m independent Gibbs chains of the length k, and keeping the last sample of each chain. GELMAN and RUBIN (1992) suggested using the average of the samples in each chain instead of only the last sample, and described a way of choosing the starting values for the chains. The sub-sampling of the chain strategy is to use a single long chain every n^{th} sample (CASELLA and GEORGE ,1992).

WANG et al. (1993) reported that samples from Gibbs sampling will show an initial period of burn-in, followed by a repetitive pattern. The burn-in period usually lasts between 500 and 1000 samples, although for slowly converging sampler that number could be considerably higher. The subsequent samples are correlated, and in order to make inferences about parameters of the model, it makes sense to analyze only every n sample, where n can range from 10 to 1000.

The Gibbs sampling convergence is tied to the convergence of the method of Gauss-Seidel, and to priors. When the corresponding convergence of Gauss-Seidel is slow, the convergence of Gibbs sampling will also be slow. One way to increase the convergence speed in multiple traits is to use block iteration, where samples are generated by block of traits (JAMROZIK and SCHAEFFER, 2000).

Estimation of Marginal Posterior Densities

Various methods can be used to obtain the marginal posterior densities of interest parameters.

The first method, proposed by TANNER and WONG (1987), relies on the ergodic theorem and states that the average of a function of interest provides a consistent estimator of its expectation. If the chain is running to equilibrium and *n* is large, the empirical mean and variance of the marginal posterior density can be estimated as:

$$\hat{E}(X) = \hat{\mu}_x = \frac{1}{n}\sum_{i=1}^{n} x_i$$

and

$$V\hat{a}r(X) = \frac{1}{n}\sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2$$

where:

x_i is the Gibbs sample corresponding to a parameter of interest at cycle *i*,

n is the number of cycle.

GELFAND and SMITH (1990) proposed a second method which is based on the average of conditional densities. Then,

$$\hat{p}(x) = \frac{1}{n}\sum_{i=1}^{n} p(x \mid y_i)$$

where:

$\hat{p}(x)$ is the marginal density of parameter of interest,

y_i is the Gibbs sample for conditional variables drawn from the marginal distribution of *p(y)* at cycle *i.*

Thus, the mean of the marginal density is given by:

$$\hat{E}(X) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} x_j p(x_j \mid y_i)$$

where the summation over *i* is over *n* samples of the Gibbs chain, and the summation over *j* is over the *m* chosen values of *X*. This estimator is called a Rao-Blackwellized estimator or a mixture estimator.

A third method was proposed by SILVERMAN (1986). It is based on the normal kernel densities. Therefore,

$$\hat{p}(x) = \frac{1}{nh}\sum_{i=1}^{n} \frac{1}{\sqrt{2\pi}} exp[-\frac{1}{2}(\frac{x - x_i}{h})^2]$$

where:

x_i is the Gibbs sample for parameter of interest at iteration *i*,

h is a "window width" constant given by the user which determines the smoothness of the estimated density.

SORENSEN (1998) reported that the ergodic estimator is simple to obtain but generally has larger variance than other estimators. According to the same author, the disadvantage of the mixture estimator is that the closed form of conditional density is required and extra information for conditional variables needs to be saved.

Monte Carlo Variance and Effective Chain Size

SORENSEN (1998) suggested the Monte Carlo variance for assessing the accuracy of the estimators. It can be estimated from the mean of the marginal density from independent chains. Two methods are commonly used; one is based on times series, and the other is known as the method of batching.

Consider the sequence of Gibbs sample $x_1, x_2, ..., x_n$ where *n* is number of cycles and x_i is the Gibbs sample corresponding to the parameter of interest at cycle *i*. The lag time auto-covariance of the sequence is estimated as:

$$\hat{\gamma}(t) = \frac{1}{n}\sum_{i=1}^{n-t}(x_i - \hat{\mu})(x_{i+t} - \hat{\mu})$$

where:

$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample mean of the chain and *t* is the lag.

The lag(t) auto-correlation is estimated as:

$$\hat{\rho}(t) = \frac{\hat{\gamma}(t)}{\hat{\gamma}(0)}$$

where:

$\hat{\gamma}(0)$ is the variance of the chain,

$\hat{\gamma}(t)$ is the covariance between samples x_i and x_{i+t}.

If n is large, the variance of the mean of the chain is given by (SORENSEN, 1998):

$$V\hat{a}r(\hat{\mu}) = \hat{\gamma}(0)(1 + 2\sum_{t=1}^{\infty} \hat{\rho}(t)) = \hat{\gamma}(0) + 2\sum_{t=1}^{\infty} \hat{\gamma}(t)$$

The effective chain size or the effective number of independent samples can be computed as described by SORENSEN et al. (1995):

$$\hat{\O} = \frac{\hat{\gamma}(0)}{V\hat{a}r(\hat{\mu})}$$

Consider that the chain of total length n is divided into b batches, with each batch having m Gibbs samples ($n = bm$). The variance using the method of batching can be estimated from:

$$V\hat{a}r(\hat{\mu}) = \frac{1}{b(b-1)} \sum_{i=1}^{b} (B_i - \hat{\mu})^2$$

where:

B_i is the sample mean for batch i.

Genetic parameters

The genetic parameters, heritabilities, and correlations are of paramount importance in all beef breeding programs. A large number of studies have investigated the estimates of various genetic parameters. These values can be estimated not only based on multiple traits, but also on random regression models (MEYER and HILL, 1997). The heritability of a metric character is one of its most important properties, and while only the phenotypic values of individuals can be directly measured, it is the breeding value that determines their influence on the next generation (FALCONER and MACKAY, 1989).

Estimates of direct and maternal heritabilities for growth traits in beef cattle were reported by KOCH and CLARK (1955), NELSEN and KRESS (1967), SCHAEFFER and WILTON (1981), ALENDA and MARTIN (1987), JOHNSTON et al. (1992), MEYER (1993), ELER et al. (1995), ROBINSON (1996a), SOUZA et al. (1997), DODENHOFF et al. (1998), and MATTOS et al. (2000). The values obtained by these authors indicate that progress in the field can be made through selection for growth characters of cattle and that maternal effects decreased in importance after weaning.

Correlated traits are of interest for three reasons: pleiotropic action of genes, changes brought by selection and because characters are correlated with natural selection (FALCONER and MACKAY, 1989). Genetic correlations were calculated for several economic characteristics by GARRICK et al. (1989), WALDRON et al. (1993), MEYER (1995), ELER et al. (1995), ROBINSON (1996a), VARONA et al. (1997), and FERREIRA et al. (1999). The genetic correlations suggest that correlated responses in other growth characters can be expected. Most of the studies presented moderate and positive estimates; however, a small number of studies reported that correlations between direct genetic effects for growth were close to 1.

Growth during the suckling period is affected both by the offspring, whose growth is measured, and by his dam, who furnishes the developmental environment. The trait, as measured, is the phenotypic value of the offspring, but it is composed of at least two components, offspring growth and maternal effect contributed by a related individual, the dam (WILLHAM, 1972). The maternal effect is strictly environmental relative to the offspring, but phenotypic differences among dams for the maternal effect are expressed in the phenotypic values of their offspring.

Negative covariances between direct and maternal effects were detected by VESELY and ROBINSON (1971), BERTRAND and BENYSHEK (1987), TRUS and WILTON (1988), ROBINSON (1996b), LEE and POLLAK (1997a), LEE and POLLAK (1997b), MEYER (1997), DODENHOFF et al. (1998), SCARPATI et al. (1998), and MATTOS et al. (2000). However, KOCH

(1972) and SKAAR (1985) reported positive correlations between direct and maternal effects. Because of the apparent negative correlation between direct and maternal effects, selection for both the direct and maternal components of maternal traits is advised.

Accurate estimates of genetic parameters are essential for any efficient selection program. Most estimates of genetic parameters have been based on records collected at experiment stations under research conditions. Records collected under farm conditions may yield different results. In this way, research is required to confirm the advantages of random regression models over multitrait models using field data.

LITERATURE CITED

ABCZ – Associação Brasileira dos Criadores de Zebu. CPD – Controle de Desenvolvimento Ponderal – Edição 2000. Disponível em: <http://www.abcz.org.br>. Acesso em: 10 jan. 2001.

ALENDA, R.; MARTIN, T.G. Genetic parameters and consequences of selection for growth traits in a beef herd selected for yearling weight. **Journal of Animal Science**, v.64, n.2, p.366-372, 1987.

BELTRAN, J.J. et al. Growth pattern of two lines of Angus cattle selected using predicted growth parameters. **Journal of Animal Science**, v.70, n. 3, p.734-741, 1992.

BERTRAND, J.K.; BENYSHEK, L.L. Variance and covariance estimates for maternally influenced beef growth traits. **Journal of Animal Science**, v.64, n.3, p.728-734, 1987.

BRODY, S. Time relations of growth of individuals and populations. In: __________ **Bioenergetics and growth**. New York: Reinhold, 1945. p.484-663.

CASELLA, G.; GEORGE, E.I. Explaining the Gibbs sampler. **Journal of Statistical Association**, v.46, p.167-174, 1992.

CNPGC. MA/ABCZ/EMBRAPA. Sumário nacional de touros das raças zebuínas: Edição 2000. Disponível em: <http://www.cnpgc.embrapa.br>. Acesso em: 10 jan. 2001.

DEMPSTER, A.P.; LAIR, N.M.; RUBIN, D.B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of Royal Statistical Society: Series B**, v.39, p.1-38, 1977.

DENISE, R.S.; BRINKS, J.S. Genetic and environmental aspects of the growth curve in beef cows. **Journal of Animal Science**, v.61, n.6, p.1431-1440, 1985.

DIGGLE, P.J.; LITANG, K.Y.; ZEGER, S.L. **Analysis of longitudinal data**. Oxford: Oxford University Press, Oxford. 1994.

DODENHOFF, J. et al. Parameter estimates for direct, maternal, and grandmaternal genetic effects for birth weight and weaning weight in Hereford cattle. **Journal of Animal Science**, v.76, n.10, p.2521-2527,1998.

ELER, J.P. et al. Estimation of variances due to direct and maternal effects for growth traits of Nelore cattle. **Journal of Animal Science**, v.73, n.11, p.3253-3258, 1995.

FALCONER, D.S., MACKAY, T.F.C. **Introduction to quantitative genetics**. Edinburgh: Longman. 1989. 464p.

FERREIRA, G.B.; MacNEIL, M.D.; VAN VLECK, L.D. Variance components and breeding values for growth traits from different statistical models. **Journal of Animal Science**, v.77, n.10, p.2641-2650, 1999.

FREITAS, A.R. et al. Curvas de crescimento na raça Nelore. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 37, 2000. Viçosa, MG. **Anais**… Viçosa, MG: SBZ, 2000, p. 209.

GARRICK, D.J. et al. Variance heterogeneity in direct and maternal weight traits by sex and percent purebred for Simmental-sired calves. **Journal of Animal Science**, v.67, n.10, p.2515-2518, 1989.

GELFAND, A.E.; SMITH, A.F.M. Sampling-based approaches to calculating marginal densities. **Journal of American Statistical Association**, v.85, n.410, p.398-409, 1990.

GELMAN, A., RUBIN, D.B. Inference from iterative simulation using multiple sequences. **Statistical Science**, v.7, n.4, p.457-511, 1992.

GEYER, C.J. Practical Markov chain Monte Carlo. **Statistical Science**, v.7, n.4, p.473-511, 1992.

GIANOLA, D.; FERNANDO, R.L. Bayesian methods in animal breeding theory. **Journal of Animal Science**, v.63, n.1, p.217-244, 1986.

GILMOUR, A.R.; THOMPSON, R.; CULLIS, B.R. Average information REML: an efficient algorithm for variance parameter estimation in linear mixed models. **Biometrics,** v.51, n.4, p.1440-1450, 1995.

GRASER, H.U.; SMITH ,S.P.; TIER, B. A derivative free approach for estimating variance components in animal models by restricted maximum likelihood. **Journal of Animal Science**, v.64, n.5, p.1362-1370, 1987.

HARTLEY, H.O.; RAO, J.N.K. Maximum likelihood estimation for the mixed model analysis of variance model. **Biometrika,** v.54, n.1, p.93-108, 1967.

HARVILLE, D.A. Maximum likelihood approaches to variance component estimations and to related problems. **Journal of American Statistical Association**, v.72, p.320-338, 1977.

HENDERSON JUNIOR, C.R. Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random regressions. **Biometrics**, v.38, n.3, p.623-640, 1982.

HENDERSON, C.R. **Application of linear models in animal models**. Ithaca: Cornell University, 1984. 462 p.

HENDERSON, C.R. Estimation of variance and covariance components. **Biometrics**, v.9, n.1, p.226-252, 1953.

JAFFREZIC, F.; PLETCHER, S.D. Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. **Genetics**, v.156, n.2, p. 913-922, 2000.

JAMROZIK, J.; SCHAEFFER, L.R. Comparison of two computing algorithms mixed model equations for multiple trait random regression test day models. **Livestock Production Science**, v.67, n.4, p.143-153, 2000.

JAMROZIK, J.; SCHAEFFER, L.R. Estimates of genetic parameters for a test day model with random regressions for production of first lactation Holsteins. **Journal of Dairy Science**, v.80, n.1-2, p.762-770, 1997.

JAMROZIK, J.; SCHAEFFER, L.R.; DEKKERS, J.C.M. Genetic evaluation of dairy cattle using test day yields and a random regression model. **Journal of Dairy Science**, v.80, n.6, p.1217-1226, 1997.

JAMROZIK, J.; SCHAEFFER, L.R.; JANSEN, G.B. Approximate accuracies of prediction from random regressions models. **Livestock Production Science**, v.66, n.1-2, p.85-92, 2000.

JENSEN, J. et al. Bayesian inference on variance and covariance components for traits influenced by maternal and direct genetic effects using the Gibbs sampler. **Acta Agriculturae Scandinavica**, v, 44, p. 193-201, 1994.

JOHNSON, D.L.; THOMPSON, R. Restricted maximum likelihood estimation of variance components for univariate animal models using sparse matrix techniques and a average information. **Journal of Dairy Science**, v. 78, n.2, p. 449-456, 1995.

JOHNSTON, D.J. et al. Estimates of genetic parameters for growth and carcass traits in Charolais cattle. **Canadian Journal Animal Science**, v.72, n.3, p.493-499, 1992.

KAPS, M.; HERRING, W.O.; LAMBERSON, W.R. Genetic and environmental parameters for traits derived from the Brody growth curve and their relationship with weaning weight in Angus cattle. **Journal of Animal Science**, v.78, n.6, p.1436-1442, 2000.

KIRKPATRIC, M.; HECKMAN, N. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. **Journal Mathematical Biology**, v.27, p.429-450, 1989.

KOCH, R.M. The role of maternal effects in animal breeding: VI. Maternal effects in beef cattle. **Journal of Animal Science**, v.35, n.6, p.1316-1323, 1972.

KOCH, R.M.; CLARK, R.T. Genetic and environmental relationships among economic characters in beef cattle. I. Correlation among paternal and maternal half-sibs. **Journal of Animal Science**, v.14, n.3, p.775-785, 1955.

LAIRD, N.M.; WARE, J.H. Random effects models for longitudinal data. **Biometrics**, v.38, n.3, p.963-974, 1982.

LEE, C.; POLLAK, E.J. Influence of sire misidentification on sire x interaction variance and direct-maternal genetic covariance for weaning weights in beef cattle. **Journal of Animal Science**, v.75, n.1, p.2858-2863, 1997b.

LEE, C.; POLLAK, E.J. Relationship between sire x year interactions and direct-maternal genetic correlation for weaning weight of Simmental cattle. **Journal of Animal Science**, v.75, n.11, p.68-75, 1997a.

MATTOS, D.; BERTRAND, J.K.; MISZTAL, I. Investigation of genotype by environment interactions for weaning weight for Herefords in three countries. **Journal of Animal Science**, v.78, n.8, p.2121-2126, 2000.

MEYER, K. Estimates of covariance components for growth traits of Australian Charolais cattle. **Australian Journal of Agricultural Research**, v.44, n.7, p.1501-1508, 1993.

MEYER, K. Estimates of genetic parameters and breeding values for New Zealand and Australian Angus cattle. **Australian Journal of Agricultural Research**, v.46, n.6, p.1219-1229, 1995.

MEYER, K. Estimates of genetic parameters of weaning weight of beef cattle accounting for direct-maternal environmental covariances. **Livestock Production Science**, v.52, n.3, p.187-199, 1997.

MEYER, K. Estimating covariance functions for longitudinal data using a random regression model. **Genetic Selection Evolution**, v.30, p.221-240, 1998.

MEYER, K. Random regressions to model phenotypic variation in monthly weights of Australian beef cows. **Livestock Production Science**, v.65, n.1, p.19-38, 2000.

MEYER, K.; HILL, W.G. Estimation of genetic and phenotypic covariance functions for longitudinal or ‘repeated’ records by restricted maximum likelihood. **Livestock Production Science**, v.47, n.3, p.185-200, 1997.

MISZTAL, I. **Computational techniques in animal breeding**. Athens, GA: University of Georgia, 2000. 176p.

MISZTAL, I. Estimation of variance components with large scale dominance models. **Journal of Dairy Science**, v. 78, n.4, p.965-974, 1977.

MISZTAL, I. et al. Strategies for estimating the parameters needed for different test-day models. **Journal of Dairy Science**, v. 83, n.4, p.1125-1134, 2000.

MISZTAL, I. Restricted maximum likelihood estimation of variance components of animal model using sparse matrix inversion and a super computer. **Journal of Dairy Science**, v.73, n.1, p.163-172, 1990.

NELSEN, T.C.; KRESS, D.D. Estimates of heritabilities and correlation for production characters of Angus and Hereford calves. **Journal of Animal Science**, v.48, n.2, p.286-292, 1979.

PATTERSON, H.D.; THOMPSON, R. Recovery of inter-block information when block sizes are unequal. **Biometrika**, v. 58, n.3, p. 545-554, 1971.

PLETCHER, S.D.; GEYER, C.J. The genetic analysis of age dependent on traits: modeling a character process. **Genetics**, v. 153, n.2, p. 825-833, 1999.

RICHARDS, J. F. A flexible growth function for empirical use. **Journal of Experimental Botany**, v.10, n.29, p.290-300, 1959.

ROBINSON, D.L. Estimation and interpretation of direct and maternal genetic parameters for weights of Australian Angus cattle. **Livestock Production Science**, v.45, n.1, p.1-11, 1996b.

ROBINSON, D.L. Models which might explain negative correlations between direct and maternal genetic effects. **Livestock Production Science**, v.45, n.2-3, p.111-122, 1996a.

SAKAGUTI, E.S. **Funções de covariâncias e modelos de regressão aleatória na avaliação genética do crescimento de bovinos jovens da raça Tabapuã**. 2000, 81f. Dissertação (Doutorado em Genética e Melhoramento) – Departamento de Zootecnia. Universidade Federal de Viçosa, Viçosa, MG. 2000.

SCARPATI, M.T.V. et al., Modelos animais alternativos para estimação de componentes de (co) variância e de parâmetros genéticos e fenotípicos do peso ao nascer na raça Nelore. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 36, 1998, Botucatu, SP. **Anais**… Botucatu, SP: SBZ, 1998, v.3, p. 512.

SCHAEFFER, L.R. **Variance component estimation course notes**. Armidale, NSW. University of New England, 1998. 155 p.

SCHAEFFER, L.R.; DEKKERS, J.C.M. Random regressions in animal models for test-day production in dairy cattle. In: WORLD CONGRESS GENETIC APPLIED LIVESTOCK PRODUCTION, 5, 1994, Guelph: **Proceedings...** Guelph. 1994. p.443-446.

SCHAEFFER, L.R.; WILTON, J.W. Estimation of variances and covariances for use in multiple trait beef sire evaluation model. **Canadian Journal of Animal Science**, v.61, n.3, p.531-538, 1981.

SEARLE, S.R.; CASELLA, G.; McCULLOCH, C.E. **Variance components**. New York: John Wiley & Sons, 1992. 501p.

SILVERMAN, B.W. **Density estimation for statistics and data analysis**. London: Chapman and Hall, 1986.

SKAAR, B.R. **Direct genetic and maternal variances and covariance component estimates from Angus and Hereford field data**. 1985, 116f. Dissertation (PhD in Animal breeding) – Iowa State University, Ames. 1985.

SORENSEN, D.A. et al. Bayesian analysis in threshold models using Gibbs sampling. **Genetic Selection Evolution,** v. 27, p. 229-249, 1995.

SORENSEN, D.A. **Gibbs sampling in quantitative genetics**. Tjele: Danish Institute of Human Science, 1988. 191p. (Internal Report, 82).

SORENSEN, D.A.; KENNEDY, B.W. Estimation of response to selection using least-squares and mixed model methodology. **Journal of Animal Science**, v.58, n.5, p.1097-1106. 1984.

SOUZA, J.C. et al. Estimativa de parâmetros genéticos para o peso ao desmame de animais da raça Nelore no Brasil. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34, 1997, Juiz de Fora, MG. **Anais**… Juiz de Fora: SBZ, 1997, v.3, p.251-253.

STRABEL, T.; MISZTAL, I. Genetic parameters for first and second lactation milk yields of Polish Black and White cattle with random regression test-day models. **Journal of Dairy Science**, v. 82, n.12, p. 2805-2810, 1999.

TANNER, M.A.; WONG, W.A. The calculation of posterior distributions by data augmentation. **Journal of the American Statistical Association,** v. 82, n.398, p. 528-540, 1987.

THOMPSON, R. The estimation of variance and covariance components with an application when records are subject to culling. **Biometrics**, v. 29, n.3, p. 527-550, 1973.

TRUS, D.; WILTON, J.W. Genetic parameters for maternal traits in beef cattle. **Canadian Journal of Animal Science**, v.68, n.1, p.119-128, 1988.

VAN DER WERF, J.H.J.; GODDARD, M.E.; MEYER, K. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. **Journal of Dairy Science**, v.81, n.12, p.3300-3308, 1998.

VAN TASSELL, C.P.; VAN VLECK, L.D. Multiple trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based (co)variance components inference. **Journal of Animal Science**, v.74. n.11, p.2586-2597, 1996.

VARONA, L. et al. Multiple trait genetic analysis of underlying biological variables of production functions. **Livestock Production Science**, v.47, n.3, p.201-209, 1997.

VEERKAMP, R.F.; THOMPSON, R. A covariance function for feed intake, live weight, and milk yield estimated using a random regression model. **Journal of Dairy Science**, v.82, n.7, p.1565-1573, 1999.

VESELY, J.A.; ROBINSON, O.W. Genetic and maternal effect on preweaning growth and type score in beef calves. **Journal of Animal Science**, v.32, n.5, p.825-831, 1971.

VILLALBA, D. et al. Preweaning growth curves in Brown Swiss and Pirenaica calves with emphasis on individual variability. **Journal of Animal Science**, v.78, n.5, p.1132-1140, 2000.

WALDRON, D.F. et al. Maternal effects for growth traits in beef cattle. **Livestock Production Science** , v.34, n.1-2, p.57-70, 1993.

WANG, C.S.; RUTLEDGE, J.J.; GIANOLA, D. Marginal inferences about variance components in a mixed linear model using Gibbs sampling. **Genetic Selection Evolution**, v.25, p.41-62, 1993.

WILLHAM, R.L. The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. **Journal of Animal Science**, v.35, n.6, p.1288-1293, 1972.

CHAPTER III – ANALYSES OF GROWTH CURVES OF NELLORE CATTLE BY RRM AND REML

ABSTRACT

Growth curves of Nellore cattle were analyzed using body weights measured at ages ranging from 1 day (birth weight) to 733 days. Two data samples were created: one with 71,867 records from herds with missing traits and the other with 74,601 records from herds with no missing traits. Records preadjusted to a fixed age were analyzed by a multiple trait model (MTM), which included the effects of contemporary group, age of dam class, additive direct, additive maternal, and maternal permanent environment. Analyses were by restricted maximum likelihood (REML) with 5 traits at a time. The random regression model (RRM) included the effects of age of animal, contemporary group, age of dam class, additive direct, additive maternal, permanent environment, and maternal permanent environment. Legendre cubic polynomials were used to describe the random effects. Estimates of covariances by MTM were similar for both data sets, although those from the missing data set showed more variability from age to age. The estimates from RRM were similar to those from MTM only for the complete-trait case and showed large artifacts for the case of missing traits. Estimates of additive direct-maternal correlations under RRM for some ages approached -1.0, and most likely contained artifacts. If many traits are missing, the best approach for obtaining parameters for RRM would be conversion from smoothed MTM estimates.

Key Words: beef cattle, multiple trait, random regression.

RESUMO

Curvas de crescimento de gado Nelore foram analisadas com base nos pesos corporais coletados, do nascer aos 733 dias de idade. Duas amostras foram geradas; a primeira era constituída de 71.867 medidas provenientes de rebanhos com informações perdidas, e a segunda, de 74.601 medidas oriundas de rebanhos com informações completas. Os pesos pré-ajustados a idades fixas foram analisados por meio de um modelo de características múltiplas (MTM), cinco características por vez, no qual foram incluídos efeitos de grupo contemporâneo, classe de idade da vaca, aditivo direto, aditivo materno e ambiente materno permanente. No modelo de regressão aleatória (RRM) incluíram-se efeitos de idade do animal, grupo contemporâneo, classe de idade da vaca direto, aditivo, ambiente permanente, aditivo materno e ambiente materno permanente. Polinômios cúbicos de Legendre foram utilizados na descrição dos efeitos aleatórios. Estimativas de covariâncias por meio de MTM foram similares em ambas as amostras, apesar de as obtidas da amostra com informações perdidas terem apresentado maior variabilidade entre as idades. As estimativas obtidas pelo RRM foram similares às obtidas pelo MTM somente para o caso da amostra com informações completas e mostraram grande variabilidade para o caso da amostra com informações perdidas. Estimativas de correlações entre os efeitos aditivos direto e materno, por meio de RRM, foram iguais a -1.0, em algumas idades. Se o número de informações perdidas for elevado, a melhor aproximação para obter parâmetros para RRM seria a conversão das estimativas obtidas por meio de MTM.

Palavras-chave: bovinos de corte, características múltiplas, regressão aleatória.

INTRODUCTION

The current genetic evaluation of growth for beef cattle uses multiple trait methodology, where animals are evaluated for weights at several points of age (ABCZ, 2001; BIF, 1996 and CNPGC, 2001). Because the actual weights are recorded at different ages, actual records made within specific intervals are preadjusted to fixed points, and records outside the intervals are not utilized. Both preadjustment and removing out of age range records lowers the accuracy of the evaluation.

The application of random regression models (Villaba, 2000) allows use of all available records, therefore, preadjustment to constant ages is not needed. However, such models may be complicated and hard to compute as they require choices on type and degree of random regression functions. Meyer (2000) evaluated the use of up to 9th degree Legendre polynomial in beef cattle analysis. Lower degree polynomials did not approximate the variance well, while the higher degree polynomials resulted in modeling artifacts. Also, computing with models using high-degree polynomials is very expensive.

In parameter estimation with random regression models, parameters corresponding to the extreme of trajectories or where the data is sparse may be poor (Meyer, 1999). In beef cattle with sequential recording, later observations may be increasingly missing, resulting in poor estimates of parameters for later ages. If the parameters of the random regression models (RRM) are poor, the evaluation with this model may be less accurate than with multiple trait models (MTM). One methodology for assessing the quality of parameters in RRM is to compare estimates obtained by RRM with estimates from multiple trait models. Although the MTM estimates may be biased or less accurate compared to the underlying model due to preadjustments, these estimates are less likely to be affected by extremes of trajectories.

The purpose of this study was to obtain genetic parameters for sequential growth of beef cattle using RRM with data sets with missing and no missing observations and to compare these estimates with those obtained by MTM.

MATERIALS AND METHODS

Data

Data were collected by the Brazilian Zebu Breeders Association (ABCZ) and provided by the Brazilian Agricultural Research Corporation (EMBRAPA). The data consisted of records on 619,989 Nellore animals; progeny of 11,847 sires, and 273,263 dams raised under Brazilian pasture conditions. The records were collected from 1975 to 1999.

Traits considered were birth weight (BW), 10 to 110 days weight (W1 or weight at 60 days), 102 to 202 days weight (W2 or weight at 152 days), 193 to 293 days weight (W3 or weight at 243 days), 283 to 383 days weight (W4 or weight at 333 days), 376 to 476 days weight (W5 or weight at 426 days), 467 to 567 days weight (W6 or weight at 517 days), 551 to 651 days weight (W7 or weight at 601 days), and 633 to 733 days weight (W8 or weight at 683 days).

Edits included eliminating records of animals outside the range of three standard deviations from the overall mean for each weight, and eliminating records outside of the range in age classes provided above. Table 1 presents characteristics of the data.

TABLE 1 – Characteristics of the data set by traits

Traits [1]	N [2]	Sires [3]	Dams [3]	Herds [3]	Mean (kg)	SD (kg)
BW (1)	608,870	11,847	273,263	4,747	29.81	2.65
W1 (60)	519,664	10,862	245,078	4,547	74.09	22.21
W2 (152)	451,273	10,191	224,496	4,329	133.22	28.64
W3 (243)	421,347	10,033	215,785	4,153	180.18	33.43
W4 (333)	312,869	8,916	173,148	3,702	206.11	37.96
W5 (426)	258,460	8,014	148,972	3,424	234.64	44.57
W6 (517)	216,486	7,280	129,260	3,166	270.30	51.88
W7 (601)	159,853	6,250	102,544	2,829	298.58	56.50
W8 (683)	37,565	3.505	32,813	1,642	328.73	61.37
Records	2,986,387					

[1] Numbers within parenthesis refer to the mean age (days); [2] Number of records; [3] With progeny in the data set.

Two sample data sets were formed by randomly sampling: one from all herds and one from herds with no missing traits. The number of animals in the pedigree file was 21,055 and 16,161 for the samples, respectively. Both samples are described in Table 2. Dams were 1.5 through 20 years of age at calving. Classes of age of dam were defined every year but the first class was for dams whose age at calving was less than 3 years, and the last class was for dams whose age at calving was greater than 15 years. The season of measurement was defined every three months, i.e., October to December; January to March; April to June; and July to September, respectively.

The sample of herds with missing traits was obtained from herds with more than 500 birth weight records, and an average contemporary group size greater than 5 within each herd. The sample of herds with no missing traits was obtained from herds with more than 50 birth weight records, and an average contemporary group size greater than 5 within each herd. Single record contemporary groups were eliminated from both samples, and then 5% of the herds that remained were sampled.

TABLE 2 – Characteristics of the samples by traits

	A (with missing traits)				B (with no missing traits)			
Traits	N[1]	CG[2]	Mean (Kg)	SD (Kg)	N[1]	CG[2]	Mean (Kg)	SD (Kg)
BW	14,752	636	29.85	2.13	8,229	658	29.99	2.83
W1	12,014	601	75.92	20.32	8,229	619	71.81	19.69
W2	10,488	579	134.29	26.30	8,229	616	131.20	24.85
W3	9,640	556	181.10	29.61	8,229	616	177.61	28.58
W4	7,563	492	206.08	34.21	8,229	617	200.23	32.42
W5	6,756	439	239.14	40.37	8,229	615	230.66	37.62
W6	5,882	410	279.85	46.71	8,229	614	267.57	42.71
W7	4,397	347	312.06	49.82	8,229	609	299.65	47.19
W8	555	112	333.98	56.77	8,229	611	327.38	50.74

[1] Number of observations; [2] Contemporary group (herd, year, season of the measurement, and sex of the calf).

Models

Two models were used for analyses: a multiple trait model (MTM) and a random regression model (RRM). The MTM was:

$$y = X\beta + Z_1 d + Z_2 m + Z_3 mp + e$$

where:

y was a vector of records preadjusted to fixed age; β was a vector of fixed effects (contemporary group and age of dam class); d was a vector of additive direct genetic random effects of the animal; m was a vector of additive maternal genetic random effects; mp was a vector of random effects of maternal permanent environment; X was the incidence matrix for fixed effects; Z_1, Z_2, and Z_3 were incidence matrices for animal, maternal, and maternal permanent environmental effects, respectively; and e was the vector of residual random effects.

The variances and covariances were defined as follows:

$$V=Var\begin{bmatrix} d \\ m \\ mp \\ e \end{bmatrix}=\begin{bmatrix} G\otimes A & & 0 & 0 \\ & & 0 & 0 \\ 0 & 0 & MP\otimes I_c & 0 \\ 0 & 0 & 0 & R\otimes I_n \end{bmatrix}$$

where:

G was a 10 x 10 covariance matrix of random genetic effects; MP was a 5 x 5 covariance matrix of random maternal permanent environmental effects; R was a 5 x 5 covariance matrix of random residual effects; A was the additive genetic relationship matrix; I_c was an identity matrix whose order was the number of dams; I_n was an identity matrix whose order was the number of animals; and $\otimes$ was the direct product operator.

(Co)variance components were estimated for five traits at a time by REMLF90 program (Misztal, 2001). All possible five traits combinations of the possible nine traits were fit and the parameters presented were based on averages from analyses of models that contained that particular parameter.

The random regression model (RRM) was defined as follows:

$$y_{ijklm} = \sum_{d=1}^{3} \hat{a}_d z_d + \sum_{d=0}^{3} cg_{di} z_{di} + \sum_{d=0}^{3} cad_{dj} z_{dj} + \sum_{d=0}^{3} d_{dk} z_{dk} + \sum_{d=0}^{3} p_{dk} z_{dk}$$

$$+ \sum_{d=0}^{3} m_{dl} z_{dl} + \sum_{d=0}^{3} mp_{dl} z_{dl} + \sum_{d=0}^{3} r_{dm} z_{dm} + \varepsilon_{ijklm}$$

where:

y_{ijklm} was the observation on contemporary group i, age of dam class j, animal k, dam l, and record m; β_d was the fixed regression coefficient d for age of animal; cg_{di} was the fixed regression coefficient d for contemporary group i; cad_{dj} was the fixed regression coefficient d for age of dam class j; d_{dk} and p_{dk} were random regression coefficients d for additive direct and permanent environmental effects of animal k; m_{dl} and mp_{dl} were random regression coefficients d for additive maternal and maternal permanent environmental effects of dam l; r_{dm} was the random regression coefficient d for residual effect of record m; z_d , z_{di} , z_{dj} , z_{dk} , z_{dl} , and z_{dm} were Legendre polynomials; and ε_{ijklm} were residual effects.

The mixed model equation with random regressions in matrix notation was:

$$y = X\beta + Z_1 d + Z_2 p + Z_3 m + Z_4 mp + Z_5 r + e$$

where:

y was the vector of records; β was the vector of fixed regressions; d, p, m, mp, and r were vectors for additive direct genetic, permanent environment, additive maternal genetic, maternal permanent environmental and residual effects, respectively; X was the incidence matrix for fixed effects; and Z_1, Z_2, Z_3, Z_4, and Z_5 were incidence covariate matrices for additive direct genetic, permanent environment, additive maternal genetic, maternal permanent environment and residual random effects, respectively; and *e* was residual random effects.

The variances and covariances were defined as follows:

$$V = Var\begin{bmatrix} d \\ m \\ p \\ mp \\ r \\ e \end{bmatrix} = \begin{bmatrix} \multicolumn{2}{c}{G\otimes A} & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ 0 & 0 & P\otimes I_k & 0 & 0 & 0 \\ 0 & 0 & 0 & MP\otimes I_l & 0 & 0 \\ 0 & 0 & 0 & 0 & R\otimes I_m & 0 \\ 0 & 0 & 0 & 0 & 0 & I_n\sigma_e^2 \end{bmatrix}$$

where:

G was an 8 x 8 covariance matrix of random regression for genetic effects; *P, MP* and *R* were 4 x 4 covariance matrices of random regression for permanent environment, maternal permanent environment and residual effects, respectively; σ_e^2 was assumed constant residual variance; *A* was additive genetic relationship matrix; I_k was an identity matrix whose order was the number of animals; I_l was an identity matrix whose order was the number of dams; I_m was an identity matrix whose order was the number of records; *In* was the number of records; and ⊗ was the direct product operator.

(Co)variance components for the RRM were estimated by REMLF90 program (Misztal, 2001).

RESULTS AND DISCUSSION

An important problem of modeling a trajectory with polynomials is what order of polynomials to use; too low order results in insufficient modeling capability and too high order results in excessive computation and modeling artifacts. Preliminary analyses were carried out to characterize the pattern of variation in the data. The age of the animal was fit as a cubic covariable; then, in this study, the order of polynomials was cubic.

Table 3 presents eigenvalues obtained from (co)variance matrices.

TABLE 3 – Eigenvalues for genetic (G), permanent environment (Pe), maternal permanent environment (MPe), and residual (R) effects with random regression model

Effects	Eigenvalues							
G	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	.99787	15.23651	22.55033	36.18708	160.652681
Pe	$< 10^{-4}$	6.92246	39.47664	240.07977	---	---	---	---
MPe	$< 10^{-4}$	1.79353	3.65663	63.686211	---	---	---	---
R	$< 10^{-2}$	13.17468	40.65412	72.11852	---	---	---	---

Eigenvalues showed in Table 3 were calculated from (co)variance matrices based on RRM. The values close to 0 indicates that the order of fitting was sufficient.

Variance Components

Table 4 shows covariance components averaged over 5 traits with MTM analyses.

TABLE 4 – Mean covariance components for additive direct (D), additive maternal (M), direct and maternal additive (AM), maternal permanent environment (MPe), and residual (R) effects with multiple trait model

				Mean age			Sample A (with missing traits)				Sample B (with no missing traits
ys)	D	M	AM	Mpe	R	D	M	AM	MPe	R	
1	0.67	0.43	-0.35	0.02	1.27	1.07	0.30	-0.28	0.16	3.79	
0	17.28	7.63	-5.47	14.62	68.20	8.70	7.86	-4.21	4.23	61.70	
52	50.26	28.15	-11.85	49.60	170.10	20.61	33.75	-18.34	18.78	130.80	
43	78.33	39.41	-16.36	84.31	295.30	41.61	44.03	-29.40	21.96	233.83	
33	86.95	25.04	-5.17	75.34	417.90	52.86	31.99	-30.48	15.73	312.13	
26	127.50	16.85	-5.99	75.81	473.60	74.16	24.78	-31.54	14.64	356.60	
17	161.35	15.10	-6.62	65.58	583.40	127.60	21.64	-35.17	14.91	384.43	
01	267.67	28.00	-23.00	61.97	618.90	160.13	21.04	-37.74	13.77	443.33	
83	153.26	26.99	-25.56	35.19	585.15	181.05	21.52	-40.14	14.20	512.10	

The estimates of additive direct variance for weight at different ages were dependent on the age of the animals, and the values increased over all ages. Additive direct variance for the sample from herds with no missing traits showed greater consistency than those for herds with missing traits.

Additive maternal variance components decreased after the weight at 243 days for both samples. Also, after 243 days the estimates for the sample from herds with no missing traits showed more consistent trend than those from herds with missing traits. The estimates of the covariance components between the direct and maternal additive effects were negative for both samples, and these values were weaker for the sample from herds with missing traits after weight at 60 days.

The maternal permanent environment variance was higher for the sample from herds with missing traits. For both samples these values decreased after the age at 243 days. However, residual variance components increased over ages. The residual variances for the sample from herds with no missing traits were lower than those for the sample with all herds. In addition, the residual variance components for herds with no missing traits showed more consistent trend than those from herds with missing traits.

The mean covariance components estimated at different ages from the samples with RRM are presented in Table 5.

TABLE 5 – Mean covariance components for additive direct (D), additive maternal (M), direct and maternal additive (AM), permanent environment (Pe), maternal permanent environment (MPe), and residual (R) effects with random regression model

Mean	Sample A						Sample B					
Age	(with missing traits)						(with no missing traits)					
(days)	D	M	AM	Pe	MPe	R	D	M	AM	Pe	MPe	R
1	0.24	0.25	-0.03	0.06	0.04	1.72	0.85	0.26	-0.15	0.28	0.13	4.40
60	26.70	11.61	-11.29	24.45	15.82	59.98	12.75	10.91	-10.85	26.35	7.65	46.28
152	70.98	33.47	-26.05	86.75	51.06	113.39	36.90	32.09	-24.87	87.99	21.16	91.68
243	96.68	43.56	-28.10	158.44	76.74	101.10	54.98	45.35	-26.82	149.57	26.83	91.90
333	129.96	42.24	-31.94	242.21	89.16	101.33	73.68	50.20	-30.46	211.95	26.02	79.57
426	188.63	34.88	-45.73	325.31	88.64	134.94	102.60	49.97	-43.56	272.97	23.60	84.28
517	253.18	31.46	-53.95	386.02	80.53	115.19	139.19	49.02	-51.37	321.11	23.29	93.21
601	332.50	30.64	-47.65	459.59	74.30	54.98	176.35	50.16	-45.39	363.84	24.10	80.36
683	558.62	37.89	-69.92	684.12	89.85	357.92	216.36	59.80	-66.54	433.51	31.54	86.31

Estimates of the variance components for weights with the RRM (Table 5) were slightly greater than those calculated with the MTM (Table 4). After birth weight (1 day) estimates of the additive direct variance for all ages with RRM were greater than those with MTM. Also, additive maternal variance decreased after the weight at 243 days for both samples. The estimates of the covariance components between the direct and maternal additive effects were slightly weaker with RRM than those with MTM. On the contrary, the additive maternal covariance components were slightly higher with RRM than those with MTM. On the other hand, the maternal permanent environment variance estimated with both models presented the same slope; i.e., decreased after the weight at 243 days; even so, the values with RRM were slightly higher than those with MTM.

The residual (permanent environment plus error) components with the RRM were similar to those with the MTM with the sample with no missing traits. On the contrary, these components with the RRM were higher than those with MTM from birth weight (1 day) through the weight at 152 days and the weight at 683 with the sample with missing traits. The estimates increased after the weight at 683 days using the RRM for the sample from herds with missing traits. This may be due to the number of records at this age. The number of records for weight at 633 to 733 days was only 3.8% (Table 2) of the number of records at birth weight because farmers cull a number of calves at weaning and at yearling. Therefore, RRM estimates might be inflated at the extremes of the trajectory. This was less important for MTM than RRM because MTM provides point estimates.

Figures 1 and 2 show all curves for covariances components for the specific points (ages) with MTM from both samples, respectively.

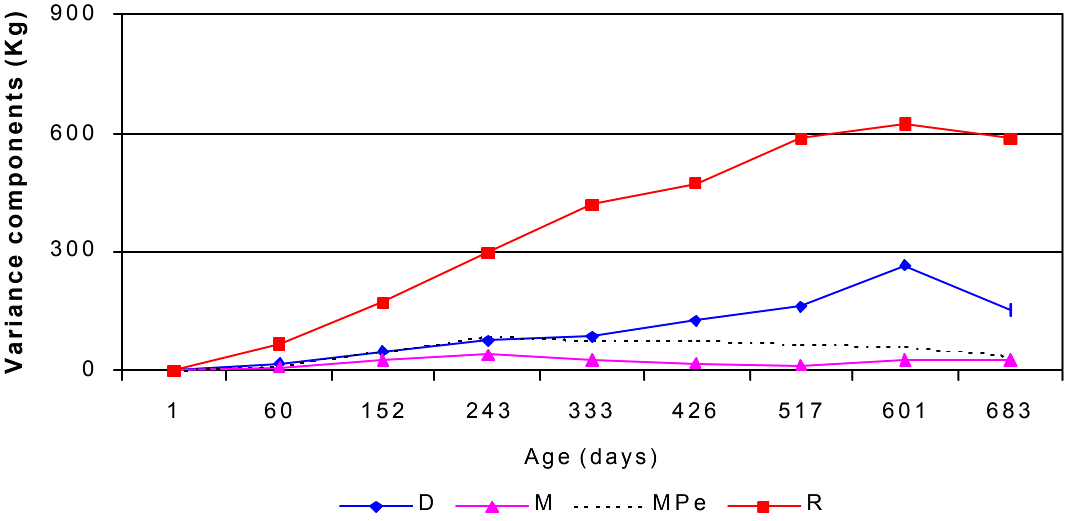

FIGURE 1 – Variance components additive direct (D), additive maternal (M), maternal permanent environment (MPe), and residual (R) at different ages with multiple trait model for herds with missing traits.

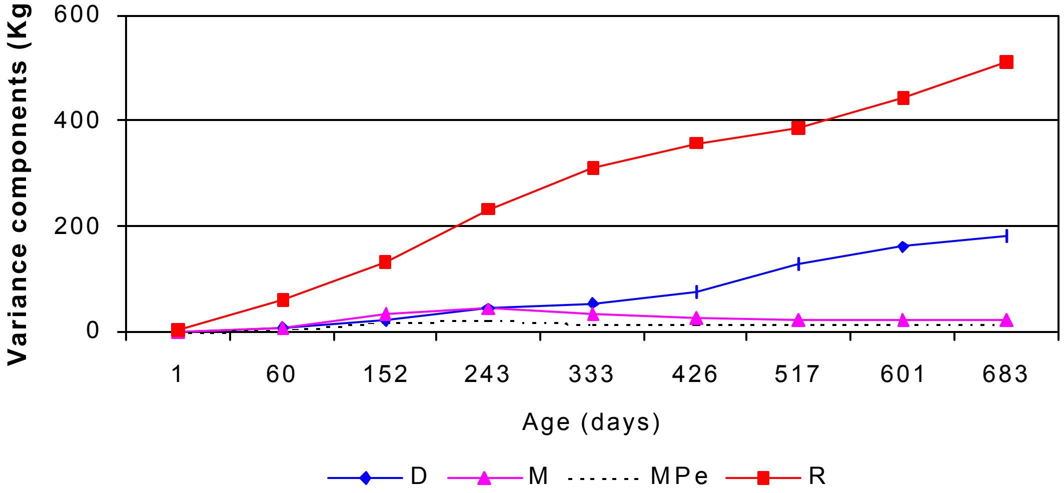

FIGURE 2 – Variance components additive direct (D), additive maternal (M), maternal permanent environment (MPe), and residual (R) at different ages with multiple trait model for herds with no missing traits.

Curves fit based on MTM show that residual and additive direct variance components increased over ages; however, these values decreased at later ages for the sample from herds with missing traits. Additive maternal and maternal permanent environment variances differed over the samples; both estimates were higher for the sample from herds with missing traits.

Figures 3 and 4 show all curves with RRM from the samples, respectively.

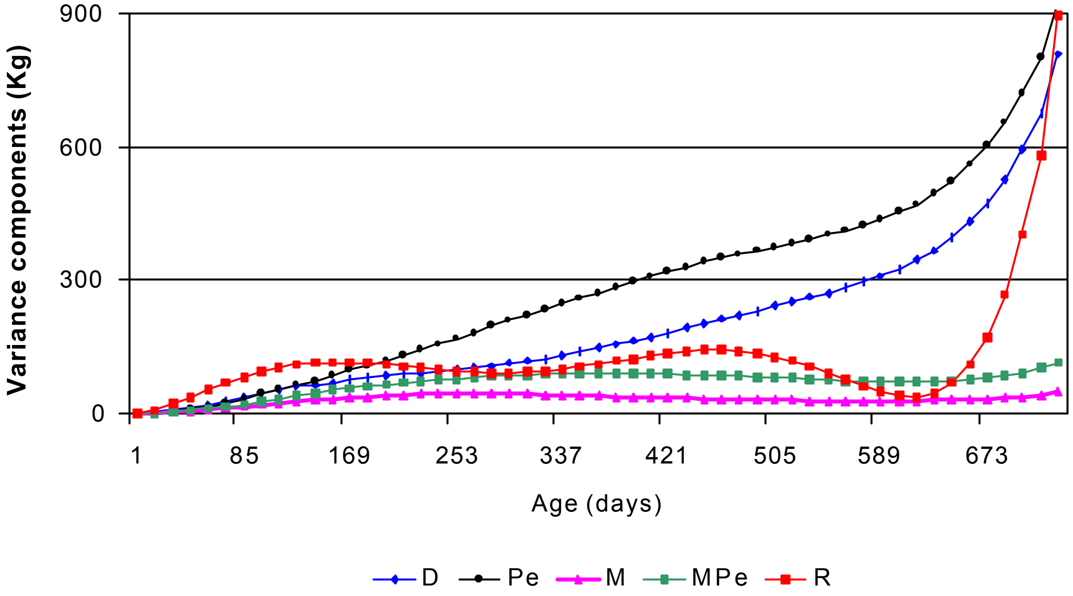

Figure 3 – Variance components additive direct (D), permanent environment (Pe), additive maternal (M), maternal permanent environment (MPe), and residual (R) at different ages with random regression model for herds with missing traits.

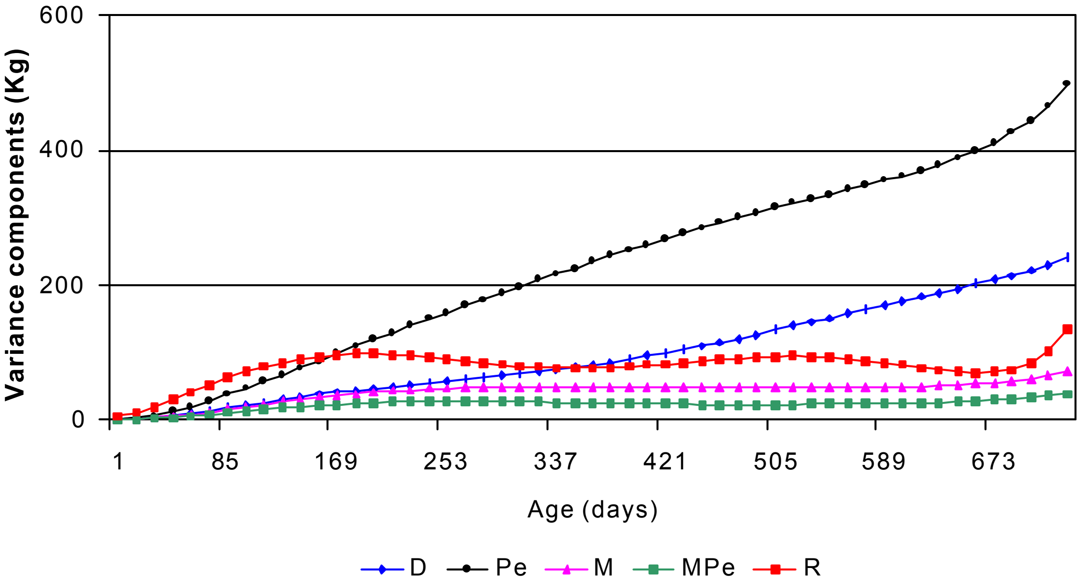

FIGURE 4 – Variance components additive direct (D), permanent environment (Pe), additive maternal (M), maternal permanent environment (MPe), and residual (R) at different ages with random regression model for herds with no missing traits.

The curves (Figure 4) fit using RRM for herds with no missing traits showed less increase over the growth of the animals than those fit for the sample from herds with missing traits (Figure 3). This might be a result of the missing records over all traits; however, all curves presented the same trajectory. According to Meyer (1999) this could be explained by small numbers of observations and sampling variation in the partitioning of total variation.

Results show that RRM were capable of describing the growth curve over ages as reported by Meyer (2000). In addition, as reported by Gringnola et al. (1998), genetic changes over time for body weight were observed in this study. The general pattern of the shape of the curves with MTM was similar to those with RRM from birth weight (1 day) through the weight at age at 601 days.

Additive Direct and Additive Maternal Correlations

Table 6 shows the additive direct and additive maternal genetic correlations with MTM from both samples.

TABLE 6 – Mean estimated additive (r_d) and maternal (r_m) genetic correlations with multiple trait model[1]

Mean age (days)	1	60	152	243	333	426	517	601	683
1	--	.82(.48)	.55(.32)	.46(.37)	.42(.32)	.51(.38)	.49(.32)	. 39(.32)	.17(.32)
60	.47(.17)	--	.86(.85)	.74(.81)	.47(.63)	.56(.50)	.48(.42)	.53(.43)	.31(.44)
152	.33(.14)	.77(.86)	--	.99(.92)	.62(.78)	.67(.66)	.43(.61)	.50(.60)	.37(.60)
243	.30(.12)	.66(.81)	.95(.90)	--	.81(.85)	.79(.71)	.54(.61)	.62(.65)	.41(.58)
333	.31(.10)	.38(.73)	.64(.86)	.79(.92)	--	.87(.92)	.78(.84)	.65(.83)	.43(.77)
426	.32(.13)	.57(.64)	.75(.76)	.84(.85)	.65(.92)	--	.88(.98)	.87(.97)	.54(.94)
517	.17(.12)	.47(.54)	.55(.70)	.22(.82)	.37(.82)	.75(.70)	--	.90(.96)	.38(.97)
601	.36(.11)	.43(.51)	.44(.66)	.48(.70)	.39(.70)	.62(.73)	.69(.86)	--	.51(.98)
683	.20(.14)	.29(.51)	.42(.60)	.44(.67)	.68(.67)	.48(.88)	.58(.78)	.44(.90)	--

[1] Numbers within parentheses refer to estimates from herds with no missing traits.

r_d - above diagonal; r_m - below diagonal.

All the additive direct genetic correlations were positive for both samples. The additive direct genetic correlations between birth weight (1 day) and all other weights were higher for the sample from herds with missing traits than those from herds with no missing traits, except for the weight at 683 days. However, after the weight at 60 days, most of the additive direct genetic correlations for the sample from herds with no missing traits were higher than those from herds with missing traits.

Analogous estimates were reported by Meyer (1993), who reported moderate additive direct genetic correlations between birth weight and the other weights, ranging from .44 to .67, while estimates between weaning, yearling and final weights were close to 1.0. The results of this study, based on estimates from herds with missing traits, were similar to those reported by

Robinson (1996), whose correlations between birth weight and later weights varied from .52 to .59. Figures 5 and 6 show the additive direct genetic correlations for specific points with MTM from the samples, respectively.

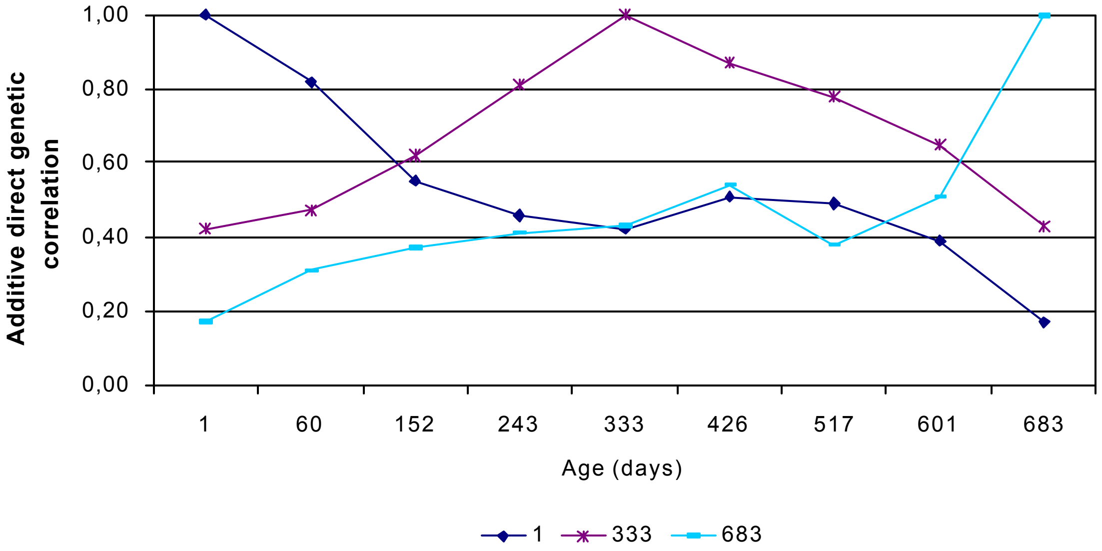

FIGURE 5 – Additive direct genetic correlations at different ages with multiple trait model for herds with missing traits.

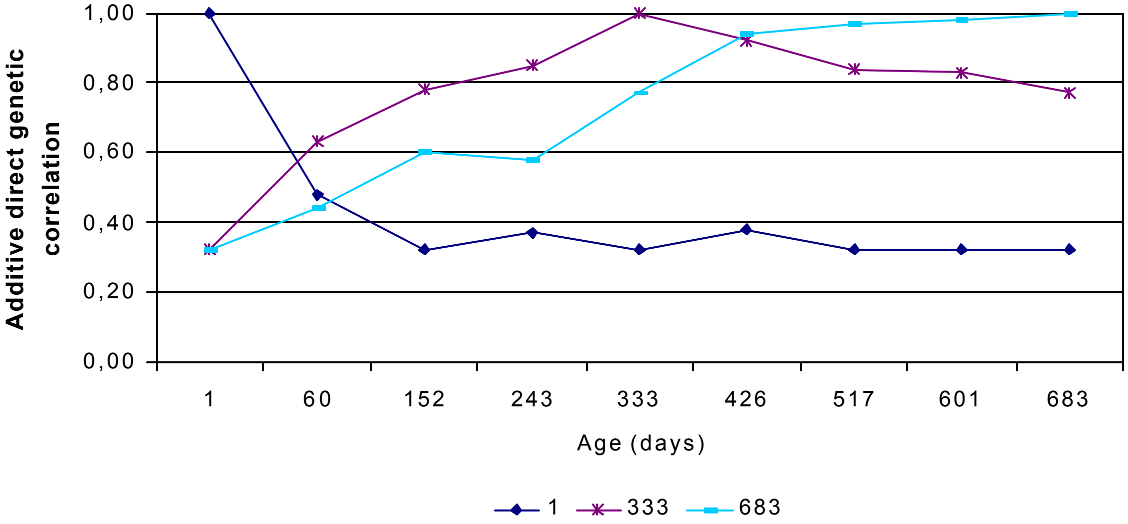

FIGURE 6 – Additive direct genetic correlations at different ages with multiple trait model for herds with no missing traits.

Although the additive maternal genetic correlations were positive for both samples (Table 6), the additive maternal genetic correlations between birth weight (1 day) and all the other weights were small for herds with no missing traits. After this age all estimates for herds with no missing traits were more consistent than those for herds with missing traits.

Additive maternal genetic correlations estimated with MTM (Table 6) were higher for weights from birth weight through the weight at 426 days than those after this age. The estimates of additive maternal genetic correlation were within the range of previous reports (Waldron et al., 1993; Eler et al., 1995). However, these estimates were lower than those reported by Robinson (1996).

For both samples the additive maternal genetic correlations with MTM are illustrated in Figures 7 and 8, respectively.

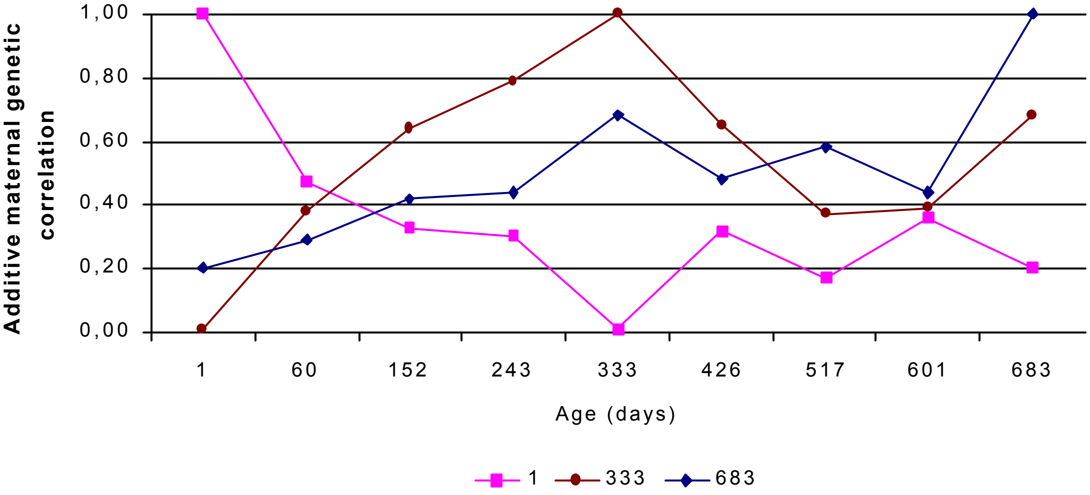

FIGURE 7 – Additive maternal genetic correlations at different ages with multiple trait model for herds with missing traits.

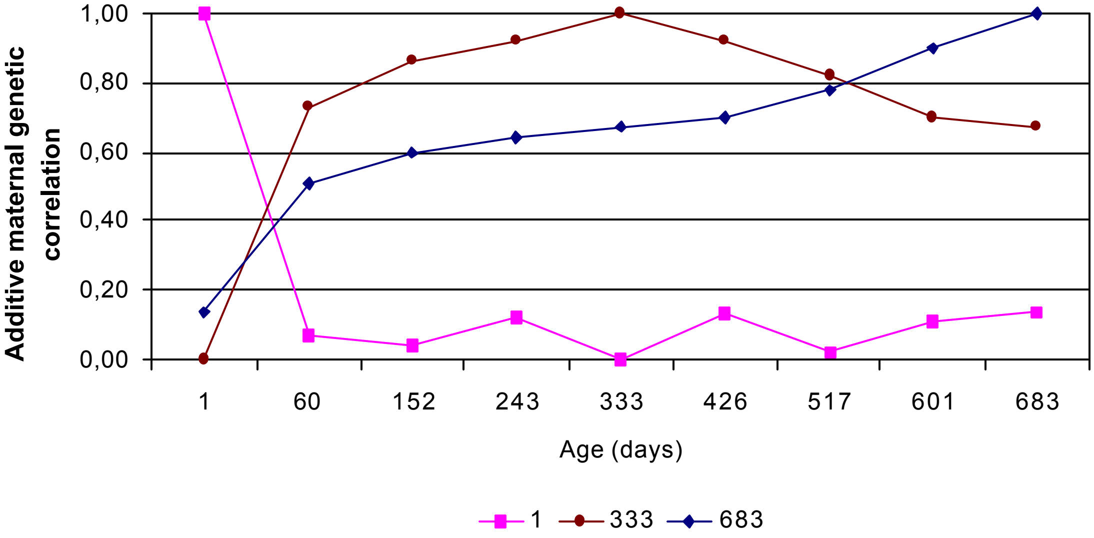

FIGURE 8 – Additive maternal genetic correlations at different ages with multiple trait model for herds with no missing traits.

The additive direct and additive maternal correlations with RRM for both samples are presented in Table 7.

TABLE 7 – Mean estimated additive direct (r_d) and additive maternal (r_m) genetic correlations with random regression model[1]

Mean age (days)	1	85	169	253	337	421	505	589	673
1	--	.53(.38)	.55(.35)	.55(.36)	.49(.37)	.41(.37)	.37(.38)	.36.(39)	.36(.43)
85	.25(.41)	--	.97(.98)	.84(.91)	.60(.77)	.38(.59)	.26(.44)	.29(.36)	.41(.37)
169	.19(.34)	.99(.99)	--	.94(.97)	.77(.87)	.58(.72)	.46(.58)	.45(.50)	.49(.48)
253	.16(.30)	.97(.96)	.99(.99)	--	.94(.96)	.81(.86)	.71(.76)	.64(.68)	.57(.65)
337	.13(.26)	.90(.91)	.94(.95)	.98(.99)	--	.96(.97)	.89(.90)	.80(.84)	.63(.80)
421	.10(.22)	.78(.83)	.84(.89)	.91(.94)	.97(.98)	--	.98(.98)	.90(.95)	.68(.90)
505	.08(.18)	.61(.73)	.69(.80)	.78(.88)	.88(.94)	.97(.99)	--	.96(.99)	.77(.95)
589	.10(.16)	.48(.67)	.55(.74)	.65(.82)	.76(.89)	.88(.95)	.97(.99)	--	.91(.98)
673	.17(.17)	.46(.67)	.50(.72)	.57(.78)	.65(.83)	.75(.88)	.85(.92)	.94(.97)	--

[1] Numbers within parentheses refer to estimates from herds with no missing traits.

r_d - above diagonal; r_m - below diagonal.

Figures 9 and 10 show the additive direct genetic correlation on a continuous scale with RRM, respectively, for both samples.

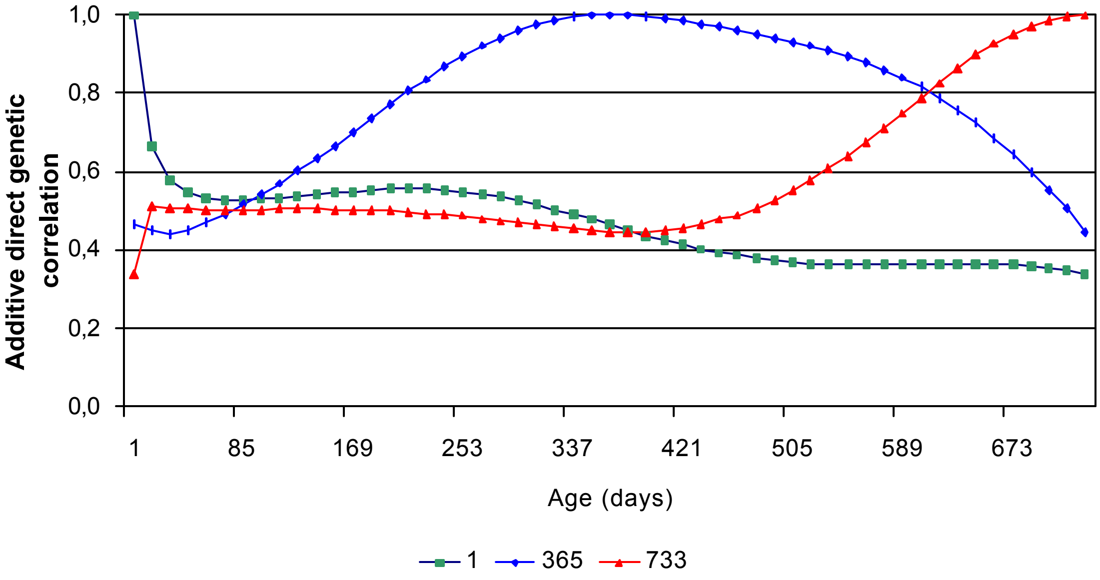

FIGURE 9 – Additive direct genetic correlations at different ages with random regression model for herds with missing traits.

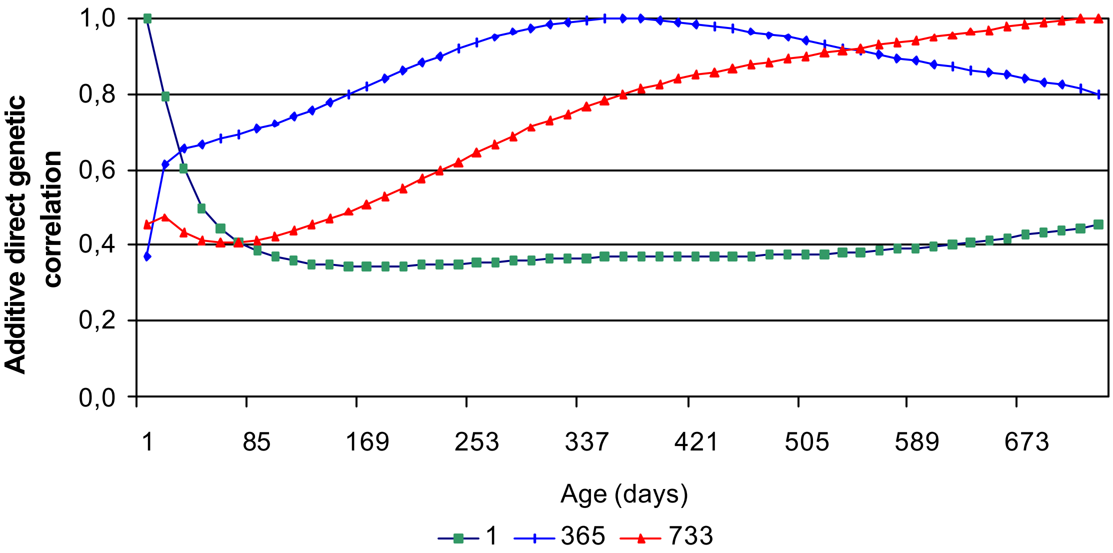

FIGURE 10 – Additive direct genetic correlations at different ages with random regression model for herds with no missing traits.

The additive direct genetic correlations with RRM (Table 7 and Figure 9) were all positive, and the highest correlations were estimated between the weight at 365 days (Figure 9) and the weights at other ages. The additive direct genetic correlations based on weight at later ages (mean age 673 days) were slightly higher than those based on birth weight. The estimates with MTM were, in general, lower than those with RRM for weights from birth weight to the weight at 243 days. However, additive direct genetic correlations with RRM were higher after the age at 243 days. In addition, the estimates from herds with no missing traits were higher than those from herds with missing traits in higher ages for both models (Tables 6 and 7). With RRM (Table 6) the additive maternal genetic correlations were less variable than those with MTM (Table 5). After the age at 243 days the additive maternal genetic correlations were, in general, higher than those with MTM.

The curves for additive maternal genetic correlations over all ages with RRM are presented in Figures 11 and 12. The estimates between birth weight (1 day) and the weights at other ages decreased as age increased. In contrast, the estimates calculated based on the weight at 733 days increased from birth weight through the weight at 426 days at constant rates, but the rate of increase in the correlation was greater after this age. The values obtained from weight at 365 days were all positive and higher than those estimated based on the birth weight and the weight at 733 days. The additive maternal genetic correlations estimated for herds with missing traits were slightly lower than those estimated for herds with no missing traits; however, the smoothness of the curves were similar (Figures11 and 12).

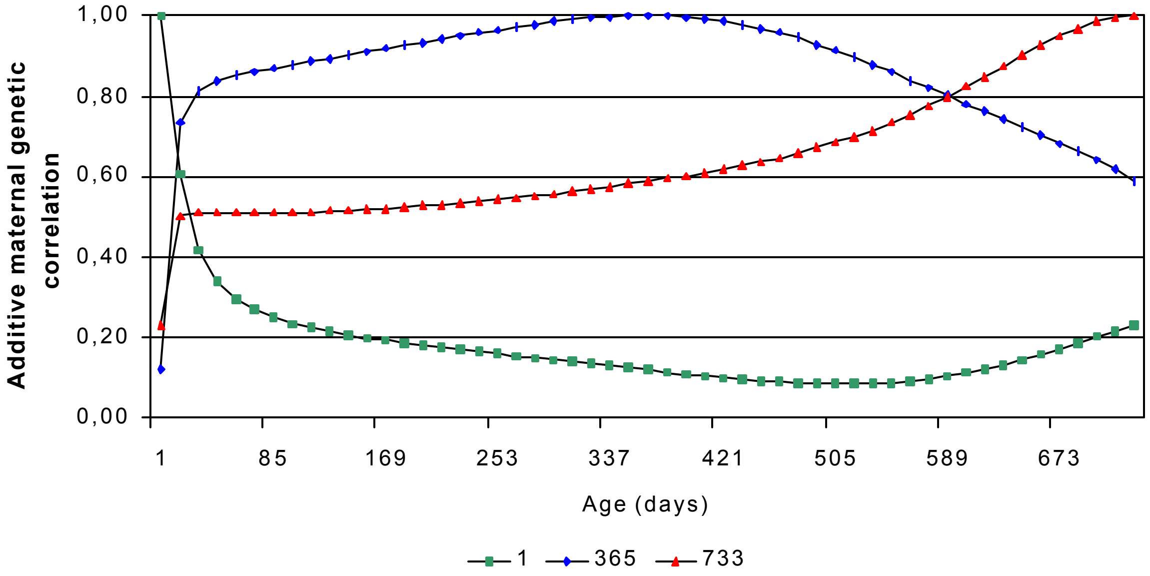

FIGURE 11 – Additive maternal genetic correlations at different ages with random regression model for herds with missing traits.

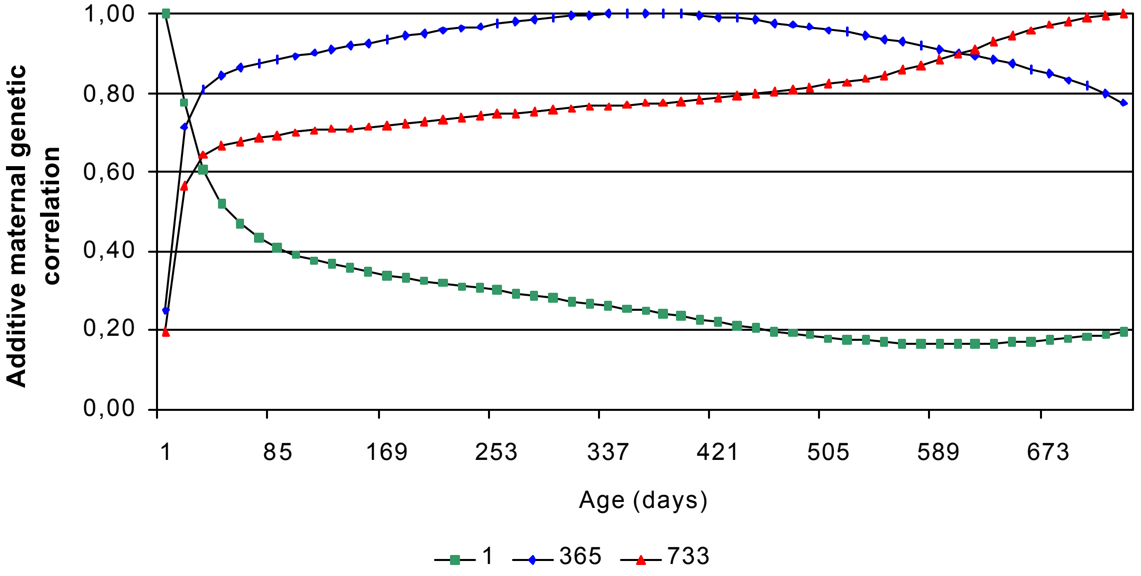

FIGURE 12 – Additive maternal genetic correlations at different ages with random regression model for herds with no missing traits.

Heritabilities

The mean additive direct and additive maternal heritabilities at different ages based on MTM for both samples are presented in Table 8.

TABLE 8 – Mean additive direct heritabilities (h^2_d), additive maternal heritabilities (h^2_m) at different ages with multiple trait model

Mean age (days)	Herds with missing traits		Herds with no missing traits	
	h^2_d	h^2_m	h^2_d	h^2_m
1	33	21	21	6
60	17	7	11	10
152	18	10	11	18
243	16	8	13	14
333	14	4	14	8
426	18	2	17	6
517	20	2	25	4
601	28	3	27	4
683	20	3	26	3

The highest additive direct heritability with MTM was at birth weight for herds with missing traits, and at the weight at 601 for herds with no missing traits (Table 8). The highest additive maternal heritability with MTM was at birth weight for herds with missing traits. In contrast, for herds with no missing traits the highest value was at the weight at 152 days. With MTM, the additive maternal heritability average was 6.7 and 8.1 from the samples, with missing and no missing traits, respectively.

Figures 13 and 14 show the curves for additive direct and additive maternal heritabilities with MTM for herds with missing and no missing traits, respectively.

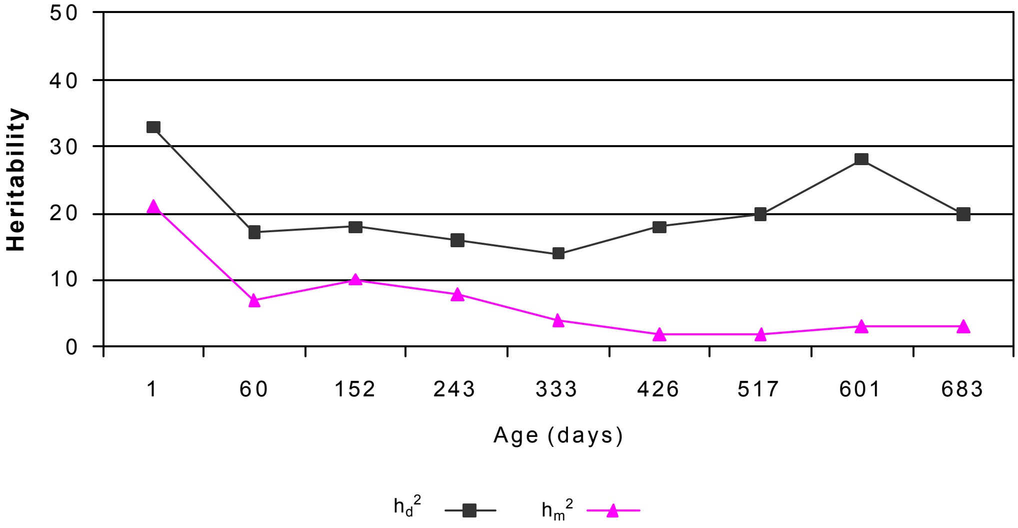

FIGURE 13 – Additive direct (h_d^2), and additive maternal (h_m^2) heritabilities at different ages with multiple trait model for herds with missing traits.

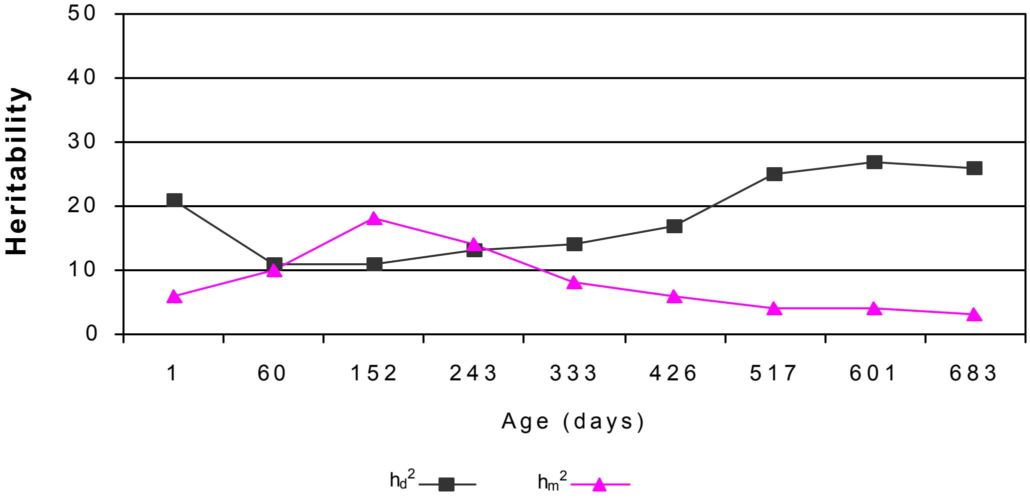

Figure 14 – Additive direct (h_d^2), and additive maternal (h_m^2) heritabilities at different ages with multiple trait model for herds with no missing traits.

Both additive direct and additive maternal heritability estimates for birth weight with MTM may be biased because when birth weight is not recorded, producers submit a breed average estimate for each sex.

Table 9 shows the mean estimates of the heritabilities, additive direct and additive maternal, at different ages with RRM from both samples.

TABLE 9 – Mean additive direct heritabilities (h^2_d), additive maternal heritabilities (h^2_m) at different ages with random regression model

Mean age (days)	Herds with missing traits		Herds with no missing traits	
	h^2_d	h^2_m	h^2_d	h^2_m
1	10	11	14	8
60	18	8	12	10
152	20	9	14	10
243	20	9	15	12
333	21	7	17	11
426	24	5	19	9
517	29	4	22	8
601	35	3	25	7
683	33	2	26	7

The highest additive direct heritability with RRM was for weights at later ages in both samples. The average additive direct heritability estimate with RRM for herds with missing traits was 23.3 and for herds with no missing traits was 18.2. The highest additive maternal heritability was at birth weight with the sample with missing traits; however, the highest additive maternal heritability was at 243 days with the sample with no missing traits. The average for herds with missing traits was 6.4 and for herds with no missing traits was 9.1.

The estimates of heritability for additive direct genetic effects were slightly higher with RRM than those with MTM after birth weight for herds with missing traits. However, these estimates with RRM were similar to those with MTM for herds with no missing traits. The estimates of heritability for additive maternal genetic effect showed less consistent trend. In addition, a small amount of variation was observed for additive maternal heritability over all ages for both samples under RRM and also with MTM.

The curves for additive direct and additive maternal heritabilities with RRM are presented in Figures 15 and 16 for both samples, respectively.

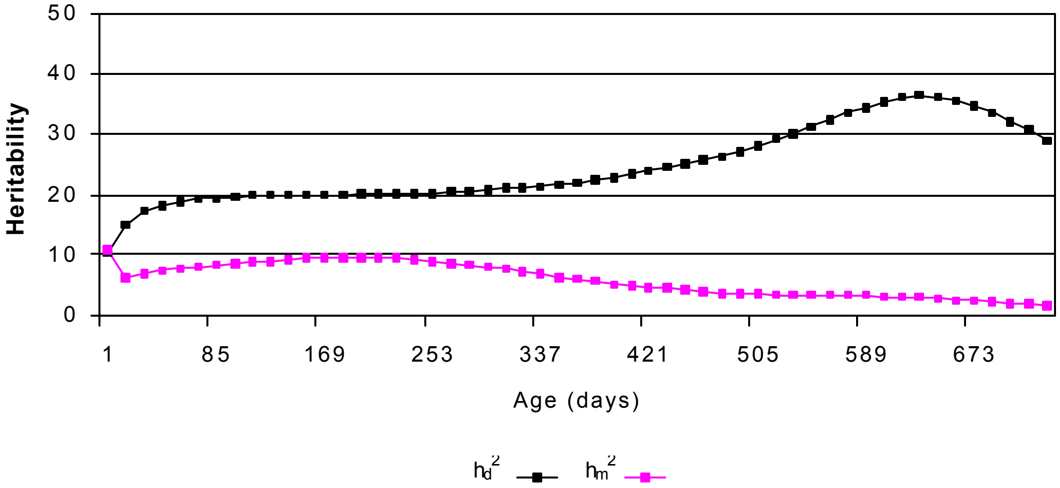

FIGURE 15 – Additive direct (h_d^2), and additive maternal (h_m^2) heritabilities at different ages with random regression model for herds with missing traits.

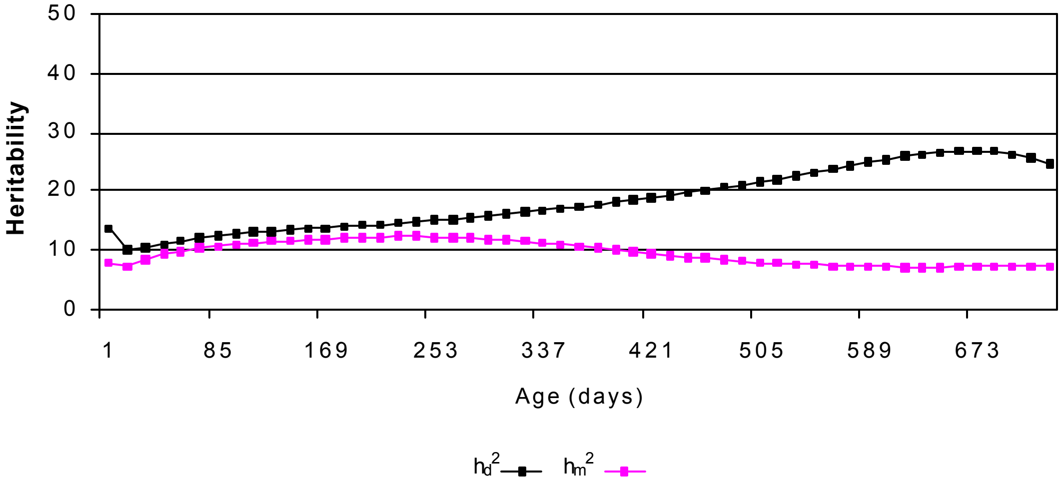

FIGURE 16 – Additive direct (h_d^2), and additive maternal (h_m^2) heritabilities at different ages with random regression model for herds with no missing traits.

The estimates of heritability for additive direct effects with both models and samples were slightly smaller than previously reported in the literature. According to Meyer (1992), one reason could be that models not accounting for additive maternal effects may yield substantially higher estimates of additive genetic variance, and, as a result, higher estimates of additive heritability. The estimates of heritability for additive maternal effects decreased in importance with time after weaning. Similar results were reported by Waldron et al. (1993) and Eler et al. (1995).

Additive Direct Maternal Correlations

Table 10 reports the mean estimates of the correlation between direct and maternal additive effects with MTM from the samples.

TABLE 10 – Mean correlations between direct and maternal additive effects at different ages with multiple trait model

Mean age (days)	Herds with missing traits	Herds with no missing traits
1	-.64	-.51
60	-.48	-.51
152	-.34	-.70
243	-.29	-.69
333	-.11	-.74
426	-.13	-.74
517	-.13	-.67
601	-.27	-.65
683	-.40	-.64

The correlation between direct and maternal additive effects was negative for all ages from both samples. The estimates for the sample with no missing traits were stronger than those for the sample with missing traits after the weight at 1 day (birth weight). These results were in agreement with the majority of reports in the literature (Bertrand and Benyshek, 1987; Garrick et al., 1989; Eler et al., 1995, and Mattos et al., 2000).

Figures 17 and 18 show the curves for the mean estimates of the correlation between direct and maternal additive effects with MTM from the samples.

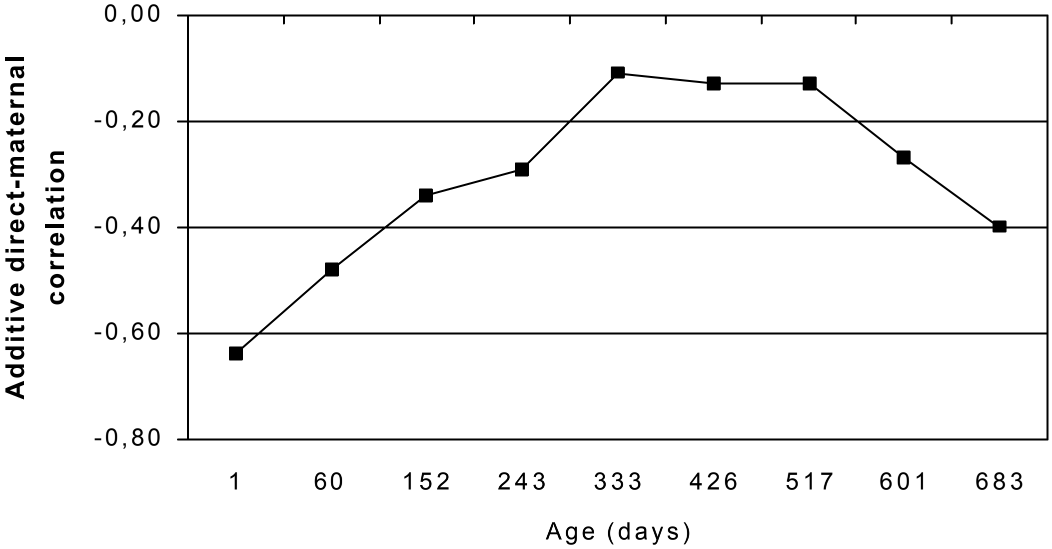

FIGURE 17 – Additive direct-maternal correlations at different ages with multiple trait model for herds with missing traits.

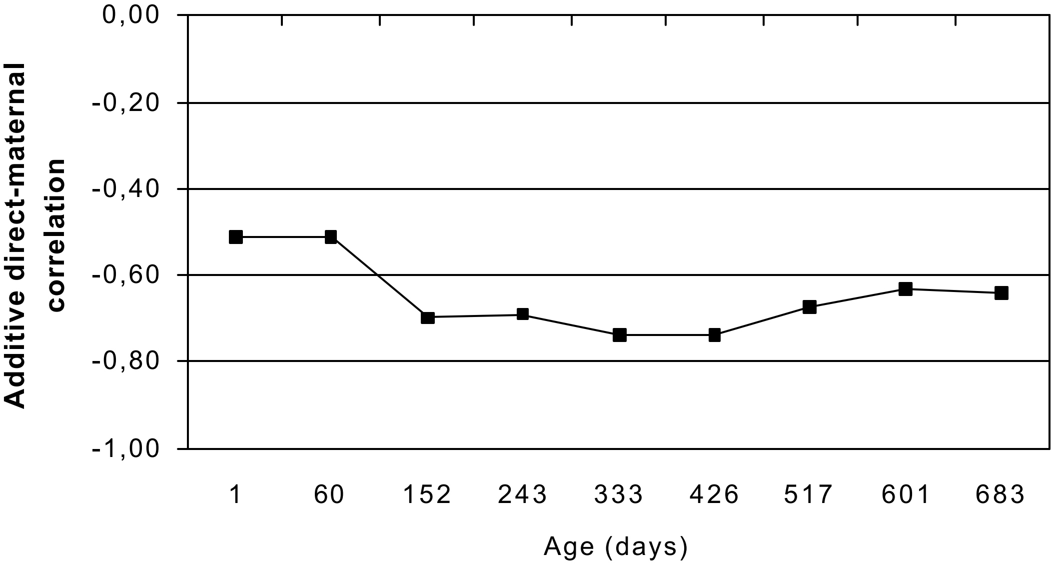

FIGURE 18 – Additive direct-maternal correlations at different ages with multiple trait model for herds with no missing traits.

The mean estimates of the correlation between direct and maternal additive effects with RRM from both samples are presented in Table 11.

TABLE 11 – Mean correlations between direct and maternal additive effects at different ages with random regression model

Mean age (days)	Herds with missing traits	Herds with no missing traits
1	-.13	-.23
60	-.66	-.91
152	-.54	-.65
243	-.43	-.54
333	-.43	-.50
426	-.55	-.61
517	-.61	-.62
601	-.48	-.48
683	-.47	-.57

Using RRM, the estimates for herds with missing traits were weaker than MTM after birth weight; however, for the herds with no missing traits, the estimates with RRM were stronger than those with MTM after weight at 152 days.

Figures 19 and 20 show the curves for these values for all ages from herds with missing and no missing traits with RRM.

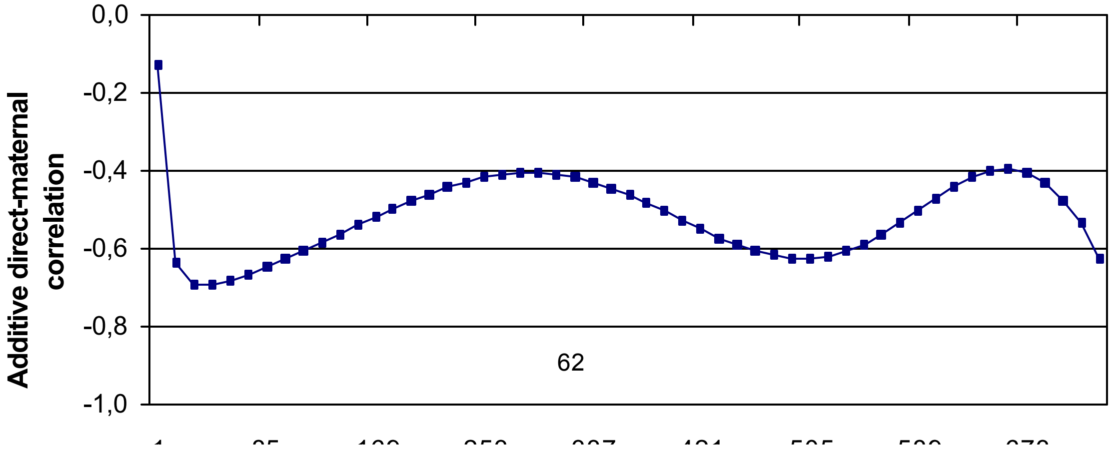

FIGURE 19 – Additive direct-maternal correlations at different ages with random regression model for herds with missing traits.

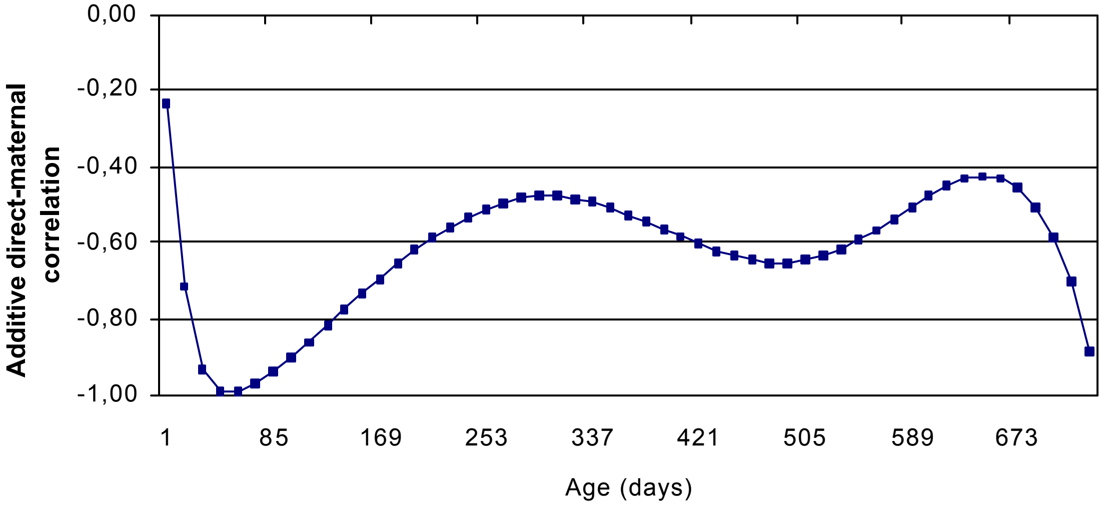

FIGURE 20 – Additive direct-maternal correlations at different ages with random regression model for herds with no missing traits.

The magnitude of the correlations between direct and additive maternal genetic effects for the sample with no missing traits seems to be inflated. One reason could be that the additive maternal information was too limited. This sample was based on herds with more than 50 birth weight records; consequently, the number of herds sampled was high (81) and less than 10% of the bulls were sires and maternal grandsires at the same time.

IMPLICATIONS

Many studies have considered genetic aspects of growth by first estimating parameters for growth curves, and subsequently estimating variance components for the growth curve parameters. Such analysis could be improved upon by the use of random regression models. Main differences between these approaches are that the first may be less able to estimate curves for animals with missing data, and more generally, does not use information from relatives. The value of information from relatives is well known to animal breeders, not only in improving accuracy, but also in accounting for directional selection.

Growth curves in this study by RRM obtained with missing trait data showed large increases of variance at later ages and also the direct maternal correlation approached the unrealistic value of -1.0. It is likely that real data sets have many missing traits, and that data without missing traits is obtained by elimination of incomplete records, possibly resulting in selection bias. If an evaluation by RRM is desired, estimates of parameters by RRM may not be satisfactory. A practical argument against analyzing large animal breeding data sets is that more random regression coefficients rapidly increase computing demands, and also estimates of parameters by RRM with small data sets may not be accurately fit.

Several strategies may be used for obtaining "better" estimates. One could be to provide priors or provide constraints for parameters to limit artifacts. It is well know that Legendre polynomials are orthogonal polynomials; therefore, equal weight was given to measurements earlier and later in life of the animals.

However, later records are less representative of the animal, being subject to accumulated selection and environmental effects. Consequently, it might be preferable to reduce emphasis for records at extremes of the growth curves. It is possible to calculate orthogonal polynomials, which decrease the spacing of later records. The other one would be to accept MT parameters and, possibly after "smoothing", convert them to RRM scale. Finally, the parameters by RRM could be corrected and smoothed based on MT and literature information. With MT methodology, parameters estimated for birth weight (BW) are optimal because that trait is obtained at the same age and subsequently no preadjustment is necessary. To ensure that evaluation of BW by RRM are optimal, it may be necessary to "adjust" RRM parameters so they would equal those of MT for birth weight.

The results of this study indicate that progress in the field can be made through selection for growth traits of beef cattle. Due to the high contribution of additive maternal effects to the phenotypic variance of the growth curves, those effects should be taken into account in genetic evaluations of growth for beef cattle. The genetic correlations suggest that correlated responses in other growth traits can be expected. In addition, methods of selection accounting for both direct and maternal additive effects would result in greater economic response to long-term selection than selection based only on additive direct genetic effects because the genetic correlations between direct and maternal additive effects were negative for all growth traits analyzed. In summary, future selection emphasis could be placed for three weights. The first should be before weaning (around 90 days), the second one at weaning (around 240 days), and the last after weaning (around 420 days). Finally breeding goals for any efficient selection program should take not only growth traits but also those of the reproductive complex and maturing traits.

CONCLUSIONS

The growth curves by MTM for samples with missing and no missing data were similar, which indicates that the MT methodology is resistant to missing traits if the missing pattern is sequential.

Growth curves by RRM were different; however, the RRM curves with no missing traits were similar to those by MTM.

Most artifacts in RRM are at extremes of growth and in genetic correlations.

LITERATURE CITED

ABCZ. 2001. Associação Brasileira dos Criadores de Zebu. CPD – Controle de Desenvolvimento Ponderal. Available at: http://www.abcz.org.br. Accessed Jan. 10, 01.

Bertrand, J.K. and L.L. Benyshek. 1987. Variance and covariance estimates for maternally influenced beef growth traits. J. Anim. Sci. 64:728-734.

BIF. 1996. Guidelines for uniform beef improvement programs. Beef Improvement Federation.

CNPGC. 2001. MA/ABCZ/EMBRAPA. Sumário das raças zebuínas de corte – 2000. Available at: http://www.cnpgc.embrapa.br. Accessed Jan. 10, 01.

Eler, J.P., L.D. Van Vleck, J.B.S. Ferraz and R.B. Lôbo. 1995. Estimation of variances due to direct and maternal effects for growth traits of Nelore cattle. J. Anim. Sci. 73:3253-3258.

Garrick, D.J., E.J. Pollack, R.L. Quaas and L.D. Van Vleck. 1989. Variance heterogeneity in direct and maternal weight traits by sex and percent purebred for Simmental-sired calves. J. Anim. Sci. 67:2515-2518.

Grignola, F.E., J. Jamrozik, L. Varona, I. Misztal and K. Bertrand. 1998. A random regression approach to estimate variance components for longitudinal records in beef cattle via Gibbs sampling. J. Anim. Sci. 76 (Suppl. 1): 57 (Abstr.).

Mattos, D., J.K. Bertrand, and I. Misztal. 2000. Investigation of genotype by environment interactions for weaning weight for Herefords in three countries. J. Anim. Sci. 78:2121-2126.

Meyer, K. 1992. Variance components due to direct and maternal effects for growth traits of Australian beef cattle. Livest. Prod. Sci. 31:179-204.

Meyer, K. 1993. Estimates of covariance components for growth traits of Australian Charolais cattle. Aust. J. Agric. Res. 44:1501-1508.

Meyer, K. 1999. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. J. Anim. Breed. Genet. 116:181-205.

Meyer, K. 2000. Random regression to model phenotypic variation in monthly weights of Australian beef cows. Livest. Prod. Sci. 65:19-38.

Misztal, I. 2001. REMLF90 manual: Available at: ftp://nce.ads.uga.edu/pub/ignacy/blupf90/. Accessed Mar 22, 01.

Robinson, D.L. 1996. Estimation and interpretation of direct and maternal genetic parameters for weights of Australian Angus cattle. Livest. Prod. Sci. 45:1-11.

Villalba, D., I. Casasús, A. Sanz, J. Estany and R. Revilla. 2000. Preweaning growth curves in Brown Swiss and Pirenaica calves with emphasis on individual variability. J. Anim.Sci. 78:1132-1140.

Waldron, D.F., C.A. Morris, R.L. Baker and D.L. Robinson. 1993. Maternal effects for growth traits in beef cattle. Livest. Prod. Sci. 34:57-70.

CHAPTER IV – ANALYSES OF GROWTH CURVES OF NELLORE CATTLE BY BAYESIAN METHOD VIA GIBBS SAMPLING

ABSTRACT

Growth curves of Nellore cattle were analyzed using body weights measured at ages ranging from 1 day (birth weight) to 733 days. Two data samples were created: one with 79,849 records from herds that had missing traits and another with 74,601 from herds with no missing traits. Records preadjusted to a fixed age were analyzed by a multiple trait model (MTM), which included the effects of contemporary group, age of dam class, additive direct, additive maternal, and maternal permanent environment. Analyses were by a Bayesian method for all 9 traits. The random regression model (RRM) included the effects of age of animal, contemporary group, age of dam class, additive direct, permanent environment, additive maternal, and maternal permanent environment. Legendre cubic polynomials were used to describe random effects. MTM estimated covariance components and genetic parameters for birth weight and sequential weights and RRM for all ages.Estimates of additive direct variance from herds with missing traits increased from birth weight through weight at 551 to 651 days with MTM. However, this component also increased for the sample with no missing traits after this age. Additive direct and residual estimated variance with RRM increased over all ages for both samples. For MTM, additive direct and maternal heritabilities were greater from the sample with herds with missing traits than those values from herds with no missing traits. The estimates from RRM were slightly lower than those from MTM for the sample with no missing traits; however, additive maternal heritabilities from MTM were greater than those using RRM. The estimated additive direct genetic correlations for each pair of traits were slightly higher for the first age (birth weight) using MTM than RRM. The range of additive maternal genetic correlations was lower than that for additive direct genetic correlations with MTM and RRM. Due to the fact that covariance components based on RRM were inflated for herds with missing traits, MTM should be used and converted to covariance functions.

Key Words: beef cattle, multiple trait, random regression.

RESUMO

Curvas de crescimento de gado Nelore foram analisadas com base nos pesos corporais coletados, do nascer aos 733 dias de idade. Duas amostras foram geradas; a primeira era constituída de 79.849 medidas provenientes de rebanhos com informações perdidas, e a segunda, de 74.601 medidas oriundas de rebanhos com informações completas. Os pesos pré-ajustados a idades fixas foram analisados por meio de um modelo de características múltiplas (MTM), nove características, no qual se incluíram efeitos de grupo contemporâneo, classe de idade da vaca, aditivo direto, aditivo materno e ambiente materno permanente. Polinômios cúbicos de Legendre foram utilizados em um modelo de regressão aleatória (RRM), no qual foram incluídos efeitos de idade do animal, grupo contemporâneo, classe de idade da vaca, aditivo direto, ambiente permanente, aditivo materno e ambiente materno permanente. Componentes de covariâncias e parâmetros genéticos foram estimados, para específicos pontos seqüenciais, por meio do MTM. Essas estimativas foram calculadas por meio do RRM, para todas as idades. Estimativas de variâncias aditivas diretas para a amostra com informações perdidas aumentaram, do nascer à idade de 551 a 651 dias, pelo MTM, e em todas as idades, na amostra com informações completas. Estimativas de variâncias aditiva direta e residual pelo RRM aumentaram ao longo de todas as idades, em ambas as amostras. Pelo MTM, heritabilidades aditivas direta e materna foram maiores na amostra de rebanhos com informações perdidas do que nas de rebanhos com informações completas. As estimativas pelo RRM foram ligeiramente menores do que aquelas obtidas pelo MTM, na amostra com informações completas. Heritabilidades aditivas maternas pelo MTM foram maiores do que aquelas calculadas pelo RRM. As estimativas de correlações genéticas aditivas diretas foram levemente maiores, para peso ao nascer, pelo MTM do que pelo RRM. A amplitude das correlações genéticas aditivas maternas foi menor do que aquela obtida para efeito genético aditivo direto, por meio de ambos os modelos. Tendo em vista que os componentes de covariância baseados em RRM são influenciados por informações perdidas, recomendam-se o MTM e a conversão destes componentes em funções de covariância.

Palavras-chave: bovinos de corte, características múltiplas, regressão aleatória.

INTRODUCTION

One of the most popular topics in beef breeding currently is modeling continuous growth via longitudinal models (Meyer, 2000). In such models, weights at any age can be accommodated, and predictions can be obtained for any day of life. The application of such models requires knowledge of variance components. Additionally, decisions need to be made regarding the complexity of the model as growth can be approximated by many different functions.

Traditionally, variance components were estimated by Restricted Maximum Likelihood – REML (Thompson, 1973; Sorensen and Kennedy, 1984). This method has good properties under selection and software packages are available. Unfortunately, computing with REML is very intensive, prohibiting use with more than a few traits at a time and with larger data sets. Also, REML formulas are very complicated to develop for less traditional models. Parameters estimated by REML are modes of joint restricted likelihood and asymptotic standard deviations.

The emergence of Gibbs sampling (GS) led to the growing popularity of Bayesian methodology. With Bayesian methodology (BM), one can obtain functions of marginal distributions for any parameter; such functions include means, modes, standard deviations, and high posterior density intervals. Formulas to implement GS are relatively simple even for complicated models. With optimization, BM via GS can be used to analyze much larger data sets and with more traits than REML.

With proper implementation, convergence of REML formulas is easily established. With BM via GS, iterations should continue beyond the burn-in period and until enough effective samples of required parameters are obtained. However, formulas to establish the burn-in and the number of effective samples do not seem perfect (Schaeffer, 1998). In practice, the convergence of GS depends on the data structure. While the convergence is very fast in simulated models, that convergence may be drastically reduced with real data sets, where traits are missing and pedigrees are incomplete.

Despite theoretically better properties of BM-GS, REML and BM-GS usually result in very similar parameter estimates. Differences are usually treated with suspicion and are generally believed to be due to incomplete convergence or incorrect implementation of BM-GS.

In Chapter III, nine sequential weights of Nellore cattle were analyzed by REML using random regression and multiple trait models. In the last model, computing costs limited estimation to the most five traits at a time.

Two data sets were used: one with all traits present and one with traits successively missing; the ninth trait had only 3.8% of records present. These models and data sets are ideal to compare computing properties of implementations of REML and BM-GS. The purpose of this study was the estimation of parameters of models and data sets as in the previous Chapter by BM-GS, and to make comparisons with their estimates by REML.

MATERIALS AND METHODS

Data

Data were collected by the Brazilian Zebu Breeders Association (ABCZ) and provided by the Brazilian Agricultural Research Corporation (EMBRAPA). The data consisted of records on 619,989 Nellore animals; progeny of 11,847 sires, and 273, 263 dams raised under Brazilian pasture conditions. The records were collected from 1975 to 1999.

Traits considered were birth weight (BW), 10 to 110 days weight (W1 or weight at 60 days), 102 to 202 days weight (W2 or weight at 152 days), 193 to 293 days weight (W3 or weight at 243 days), 283 to 383 days weight (W4 or weight at 333 days), 376 to 476 days weight (W5 or weight at 426 days), 467 to 567 days weight (W6 or weight at 517 days), 551 to 651 days weight (W7 or weight at 601 days), and 633 to 733 days weight (W8 or weight at 683 days).

Edits included eliminating records of animals outside the range of three standard deviations from the overall mean for each weight, and eliminating records outside of the range in age classes provided above. Table 1 summarizes characteristics of the data.

TABLE 1 – Characteristics of the data set by traits

Traits [1]	N [2]	Sires [3]	Dams [3]	Herds [3]	Mean (kg)	SD (kg)
BW (1)	608,870	11,847	273,263	4,747	29.81	2.65
W1 (60)	519,664	10,862	245,078	4,547	74.09	22.21
W2 (152)	451,273	10,191	224,496	4,329	133.22	28.64
W3 (243)	421,347	10,033	215,785	4,153	180.18	33.43
W4 (333)	312,869	8,916	173,148	3,702	206.11	37.96
W5 (426)	258,460	8,014	148,972	3,424	234.64	44.57
W6 (517)	216,486	7,280	129,260	3,166	270.30	51.88
W7 (601)	159,853	6,250	102,544	2,829	298.58	56.50
W8 (683)	37,565	3,505	32,813	1,642	328.73	61.37
Records	2,986,387					

[1] Numbers within parenthesis refer to the mean age (days); [2] Number of records; [3] With progeny in the data set.

Two sample data sets were formed by randomly sampling herds; one from all herds and one from herds with no missing traits. The number of animals in the pedigree file was 20,186 and 16,161 for the samples, respectively. Both samples are described in Table 2. Dams were 1.5 through 20 years of age at calving. Classes of age of dam were defined every year; the first class was for dams whose age at calving was less than 3 years, and the last class was for dams whose age at calving was greater than 15 years. The season of measurement was defined every three months, i.e., January to March; April to June; July to September; and October to December, respectively.

The sample of herds with missing traits was obtained from herds with more than 500 birth weight records, and an average contemporary size group greater than 25 within each herd. The sample of herds with no missing traits was obtained from herds with more than 50 birth weight records, and an average contemporary group size greater than 5 within each herd. Single record contemporary groups were eliminated, and then 7.5% of the herds that remained were sampled.

TABLE 2 – Characteristics of the samples by traits

	A (with missing traits)				B (with no missing traits)			
Traits	N[1]	CG[2]	Mean (Kg)	SD (Kg)	N[1]	CG[2]	Mean (Kg)	SD (Kg)
BW	14,952	257	29.21	2.33	8,229	658	29.99	2.83
W1	13,515	295	72.72	21.69	8,229	619	71.81	19.69
W2	12,180	289	132.27	27.26	8,229	616	131.20	24.85
W3	11,064	285	181.27	32.09	8,229	616	177.61	28.58
W4	8,244	254	204.06	36.74	8,229	617	200.23	32.42
W5	7,526	239	232.22	43.19	8,229	615	230.66	37.62
W6	6,824	228	269.52	49.20	8,229	614	267.57	42.71
W7	4,580	195	299.59	54.98	8,229	609	299.65	47.19
W8	964	76	335.17	61.39	8,229	611	327.38	50.74

[1] Number of records; [2] Contemporary group (herd, year, season of the measurement, and sex of the calf).

Models

Two models were used for analyses: a multiple trait model (MTM) and a random regression model (RRM). The MTM was:

$$y = X\beta + Z_1 d + Z_2 m + Z_3 mp + e$$

where:

y was a vector of records preadjusted to fixed age; β was a vector of fixed effects (contemporary group and age of dam class); d was a vector of additive direct genetic random effects of the animal; m was a vector of additive maternal genetic random effects; mp was a vector of random effects of maternal permanent environment; X was the incidence matrix for fixed effects; Z_1 , Z_2 , and Z_3 were incidence matrices for animal, maternal, and maternal permanent environmental effects, respectively; and e was the vector of residual random effects.

The variances and covariances were defined as follows:

$$V = Var \begin{bmatrix} d \\ m \\ mp \\ e \end{bmatrix} = \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{G \otimes A}} & 0 & 0 \\ & & 0 & 0 \\ 0 & 0 & MP \otimes I_c & 0 \\ 0 & 0 & 0 & R \otimes I_n \end{bmatrix}$$

where:

G was an 18 x 18 covariance matrix of random genetic effects; MP was a 9 x 9 covariance matrix of random maternal permanent environmental effects; R was a 9 x 9 covariance matrix of random residual effects; A was the additive genetic relationship matrix; I_c was an identity matrix whose order was the number of dams; I_n was an identity matrix whose order was the number of animals; and $\otimes$ was the direct product operator.

(Co)variance components were estimated for nine traits simultaneously by the GIBBS2F90 program (Misztal, 2001). This program stores only single trait matrices and manipulates them for block-diagonal sampling with joint sampling for correlated effects (Misztal, 2001). The POSTGIBBS2F90 program (Tsuruta, 2001) was used for post gibbs analyses. The chain length was 100,000 and the burn-in period was 40,000 for both samples.

The mixed model with random regression (RRM) was defined as follows:

$$y_{ijklm} = \sum_{d=1}^{3} \hat{a}_d z_d + \sum_{d=0}^{3} cg_{di} z_{di} + \sum_{d=0}^{3} cad_{dj} z_{dj} + \sum_{d=0}^{3} d_{dk} z_{dk} + \sum_{d=0}^{3} p_{dk} z_{dk}$$

$$+ \sum_{d=0}^{3} m_{dl} z_{dl} + \sum_{d=0}^{3} mp_{dl} z_{dl} + \sum_{d=0}^{3} r_{dm} z_{dm} + \varepsilon_{ijklm}$$

where:

y_{ijklm} was the observation on contemporary group i, age of dam class j, animal k, dam l, and record m; β_d was the fixed regression coefficient d for age of animal; cg_{di} was the fixed regression coefficient d for contemporary group i; cad_{dj} was the fixed regression coefficient d for age of dam class j; d_{dk} and p_{dk} were random regression coefficients d for additive direct and permanent environmental effects of animal k; m_{dl} and mp_{dl} were random regression coefficients d for additive maternal and maternal permanent environmental effects of dam l; r_{dm} was the random regression coefficient d for residual effect of record m; z_d, z_{di}, z_{dj}, z_{dk}, z_{dl}, and z_{dm} were Legendre polynomials; and ε_{ijklm} were residual effects.

The mixed model equation with random regressions in matrix notation was:

$$y = X\beta + Z_1 d + Z_2 p + Z_3 m + Z_4 mp + Z_5 r + e$$

where:

y was the vector of records; β was the vector of fixed regressions; d, p, m, mp, and r were vectors for additive direct genetic, permanent environment, additive maternal genetic, maternal permanent environment and residual effects, respectively; X was the incidence matrix for fixed effects; and Z_1, Z_2, Z_3, Z_4, and Z_5 were incidence covariate matrices for additive direct genetic, permanent environment, additive maternal genetic, maternal permanent environment and residual random effects, respectively; and e was residual random effects.

The variances and covariances were defined as follows:

$$V = Var\begin{bmatrix} d \\ m \\ p \\ mp \\ r \\ e \end{bmatrix} = \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{G \otimes A}} & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ 0 & 0 & P \otimes I_k & 0 & 0 & 0 \\ 0 & 0 & 0 & MP \otimes I_l & 0 & 0 \\ 0 & 0 & 0 & 0 & R \otimes I_m & 0 \\ 0 & 0 & 0 & 0 & 0 & I_n \sigma_e^2 \end{bmatrix}$$

where:

G was an 8 x 8 covariance matrix of random regression for genetic effects; *P, MP* and *R* were 4 x 4 covariance matrices of random regression for permanent environment, maternal permanent environment and residual effects, respectively; σ_e^2 was assumed constant residual variance; *A* was additive genetic relationship matrix; I_k was an identity matrix whose order is the number of animals; I_l was an identity matrix whose order was the number of dams; I_m was an identity matrix whose order was the number of records; I_n was the number of records; and $\otimes$ was the direct product operator.

(Co)variance components for the RRM were estimated by the GIBBS2F90 program (Misztal, 2001). The chain length was 50,000 and the burn-in period was 10,000 for both samples.

RESULTS AND DISCUSSION

Variance components

The analyses by MT model via Gibbs sampling took up to 21M of memory and an average of 32 days of computing. The same analyses for only 5 traits at a time with REML, as reported in Chapter III, took up to 485M of memory and an average of 12 days. The analyses by RRM took up to 496M of memory and an average of 22 days of computing. The same analyses by REML, as reported in Chapter III, took up to 1.2G of memory and an average of 30 days.

Table 3 shows a summary of the mean covariance components estimated at different ages from the samples with MTM analyses via Gibbs sampling.

TABLE 3 – Mean covariance components for additive direct (D), additive maternal (M), direct and maternal additive (AM), maternal permanent environment (MPe), and residual (R) effects with multiple trait model

Mean age	Sample A (with missing traits)					Sample B (with no missing traits)				
(days)	D	M	AM	MPe	R	D	M	AM	MPe	R
1	10.6	0.68	-0.66	0.02	4.04	1.32	0.95	-0.73	0.21	3.38
60	28.46	22.84	-16.28	8.38	62.30	10.49	17.63	-2.34	5.91	56.54
152	90.02	63.18	-46.86	39.89	151.50	28.50	57.52	-0.29	20.41	121.20
243	121.60	106.60	-69.10	65.31	275.50	56.60	86.40	-0.98	22.01	215.60
333	138.90	108.80	-62.34	62.66	384.00	71.70	82.51	-9.87	28.93	281.90
426	172.20	114.70	-73.65	73.24	456.10	83.35	65.70	-17.36	27.40	321.40
517	279.80	143.40	-124.70	67.54	531.60	136.90	81.25	-7.98	38.76	350.00
601	433.70	158.30	-169.00	82.98	540.70	188.10	107.80	-30.31	37.08	394.30
683	373.20	182.20	-236.40	109.10	687.80	198.20	83.65	-3.06	49.24	462.30

The estimates of the additive direct variance for weight at different ages increased over all ages. However, when calculated for the sample from herds with missing traits the values were greater than when estimated for the sample for herds with no missing traits. Also, additive maternal variance components increased across all ages for the sample from herds with missing traits. The estimates for both samples were consistent and presented a small increment after the age at 243 days. The estimates of direct and maternal additive covariance components decreased over time with MTM for herds with missing traits and was variable for the sample with no missing traits. The maternal permanent environment variance was higher for the sample from herds with missing traits. The values increased over all ages, as did the additive maternal variance components for this sample.

Residual variance components increased over ages for the two samples. The values for the sample from herds with missing traits were higher than those for the sample with herds with no missing traits. The residual variance components for herds with no missing traits were more consistent than those for herds with missing traits as well.

The mean covariance components estimated at different ages from the samples with RRM are presented in Table 4.

TABLE 4 – Mean covariance components for additive direct (D), additive maternal (M), direct and maternal additive (AM), permanent environment (Pe), maternal permanent environment (MPe), and residual (R) effects with random regression model

Mean Age (days)	Sample A (with missing traits)						Sample B (with no missing traits)					
	D	M	AM	Pe	MPe	R	D	M	AM	Pe	MPe	R
1	66.04	21.50	-30.45	31.74	0.33	141.67	23.80	27.12	-15.11	31.16	2.78	172.61
60	45.05	26.46	-14.62	38.38	10.71	96.76	25.56	26.58	-20.98	91.39	14.22	61.52
152	71.31	42.26	-22.41	104.32	14.67	125.22	49.26	41.21	-29.78	191.57	39.15	73.43
243	105.76	52.70	-28.04	186.27	15.74	125.13	70.86	53.45	-29.52	274.89	55.59	88.40
333	135.77	56.36	-30.56	261.43	15.71	111.60	89.47	58.11	-33.43	346.90	63.07	71.10
426	171.50	58.08	-37.36	324.30	14.06	122.46	113.91	57.26	-48.28	405.67	64.14	60.78
517	216.28	62.10	-43.86	370.30	13.06	140.32	145.32	56.22	-57.05	444.00	60.90	54.55
601	264.44	69.18	-41.65	414.10	15.21	131.11	178.01	60.63	-50.61	471.37	56.01	24.45
683	322.95	79.45	-43.62	505.93	16.85	159.49	209.75	82.66	-72.62	521.21	61.52	57.64

Estimates of the phenotypic variance for weights with RRM were greater than those calculated with the MTM from herds with no missing traits for all ages, and only for the first ages (before 243 days) from herds with missing traits. The additive direct variance at birth weight was higher when estimated with RRM than when estimated with MTM, not only for the sample with herds with missing traits, but also for the sample with no missing traits. This might be an indication that birth weight is a different trait. After this age (1 day) all estimates increased over all ages; however, additive maternal variance components were higher with MTM than with RRM when estimated for the sample with missing traits. In addition, the values obtained with MTM increased across all ages, while with RRM they stabilized. The estimates of direct and maternal additive covariance with RRM for both samples decreased as age increased, and the values presented a slight variation among ages.

The maternal permanent environment variance showed a similar trend with both models and samples; small variation was observed after 243 days of age. On the other hand, the residual (permanent environment plus error) components with the RRM were not consistent over all ages for herds with missing traits. However, the residual (permanent environment plus error) components with RRM were higher than those with MTM for herds with no missing traits. The estimates with RRM for the first weight (birth weight) and for the last weight (633 to 733 days) might be overestimated. The reason may be chain length insufficient, and also at 683 days the number of records may have reduced the degree of freedom.

Figures 1 and 2 show all curves for covariance components for the specific points (ages) with MTM from both samples, respectively.

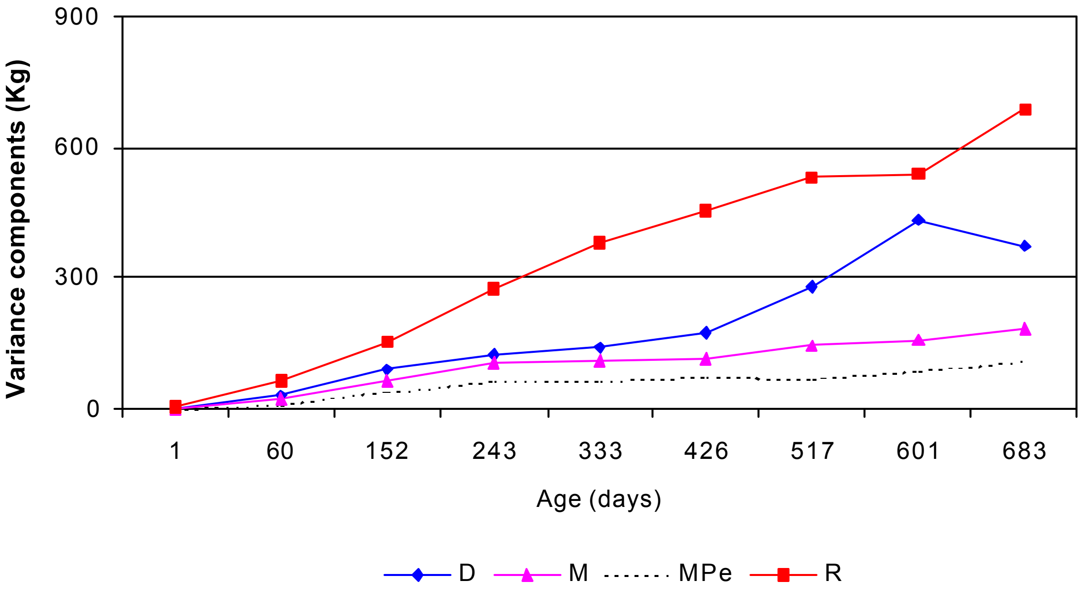

FIGURE 1 – Variance components for additive direct (D), additive maternal (M), maternal permanent environment (MPe), and residual (R) at different ages with multiple trait model for herds with missing traits.

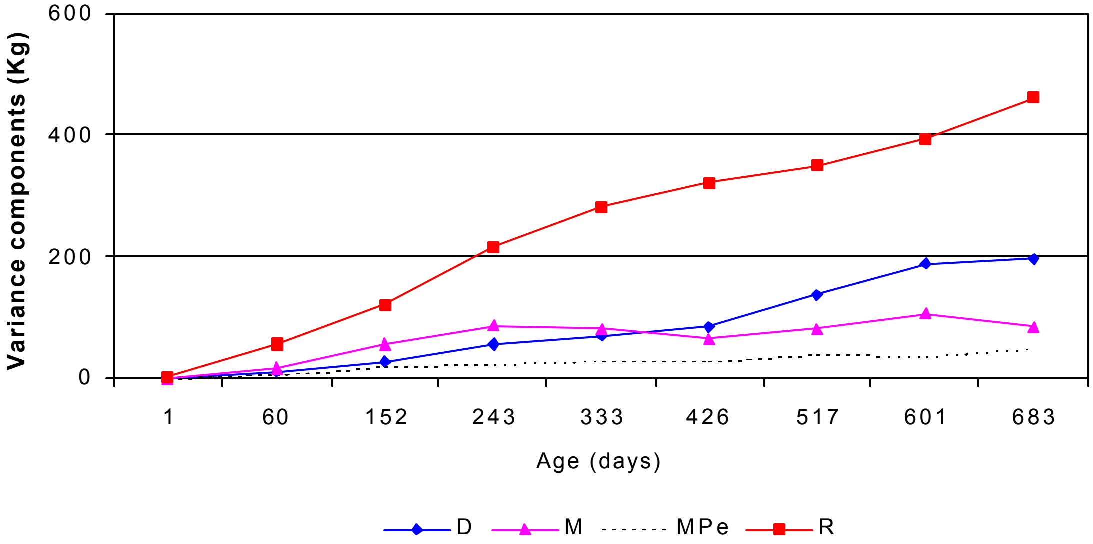

FIGURE 2 – Variance components for additive direct (D), additive maternal (M), maternal permanent environment (MPe), and residual (R) at different ages with multiple trait model for herds with no missing traits.

Curves fit with MTM show that residual and additive direct variance components increased over ages. However, additive maternal and maternal permanent environment variances showed little change over the growth of the animals, not only for the sample from herds with missing traits, but also for the sample from herds with no missing traits (Figures 1 and 2).

Figures 3 and 4 show all curves with RRM from both samples, respectively.

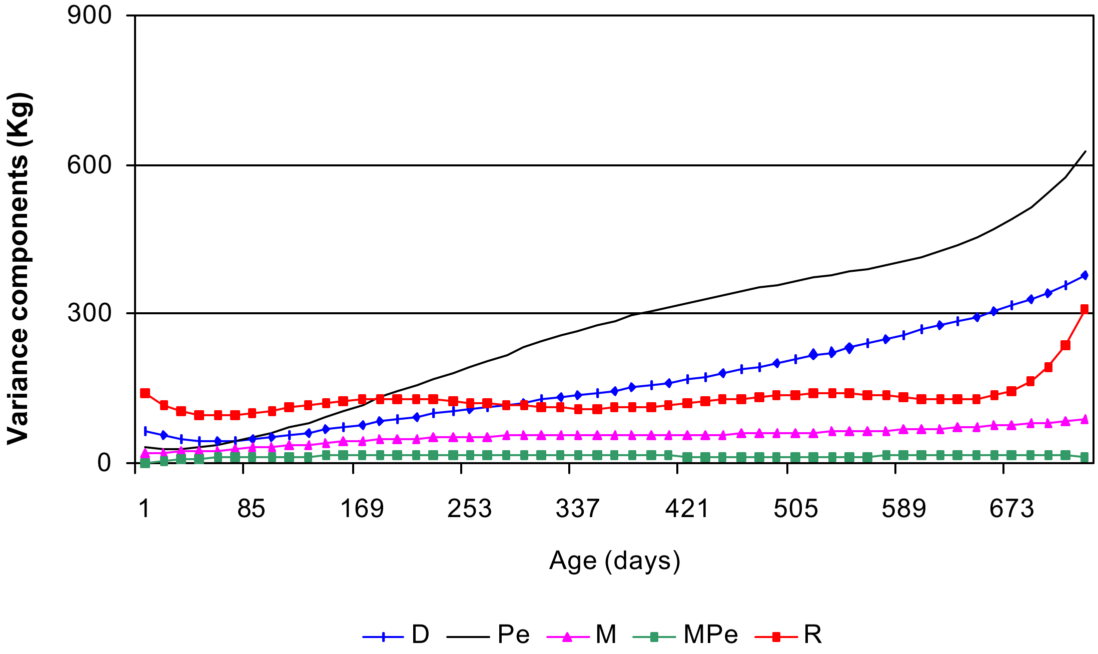

FIGURE 3 – Variance components for additive direct (D), permanent environment (Pe), additive maternal (M), maternal permanent environment (MPe), and residual (R) at different ages with random regression model for herds with missing traits.

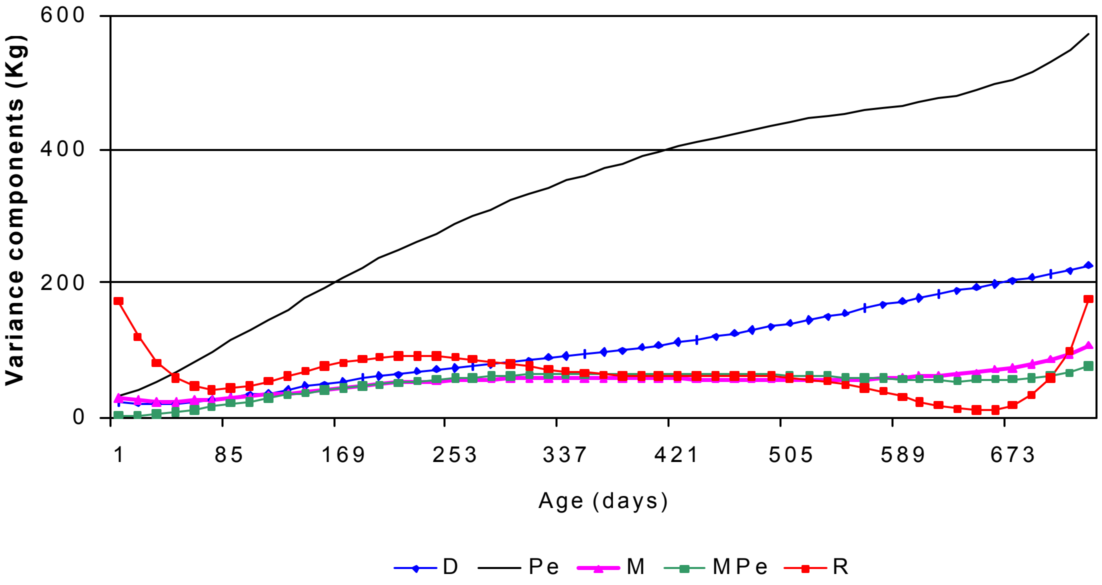

FIGURE 4 – Variance components for additive direct (D), permanent environment (Pe), additive maternal (M), maternal permanent environment (MPe), and residual (R) at different ages with random regression model for herds with no missing traits.

The results with RRM are in partial agreement with those reported by Grignola et al. (1998) and Meyer (1999). Values of variance or covariance components implausible at the extremes of the ages in the data were also reported by Meyer (1999). According to the author, this could be in part explained by small numbers of observations and sampling variation in the partitioning of total variation. In the same study, Meyer (1999) suggested that this is likely to be caused by the large effect that values furthest from the mean have in a regression analysis, and then reported that a transformation of ages to logarithmic scale reduced the influence of the later ages in the data.

The curves fit with MTM (Figures 1 and 2) were similar to those fit with RRM (Figures 3 and 4). There was little difference in estimates of covariance components for additive direct and additive maternal between the models. Similarly, there was good agreement between maternal permanent environment and residual covariance components estimated at different ages with the two models.

The estimates for the sample from herds with no missing traits were similar to those reported in Chapter III. The mean covariance components for additive maternal, direct and maternal additive, and maternal permanent environment were slightly lower than those by REML. This most likely was caused by insufficient chain length.

Additive Direct and Additive Maternal Correlations

Table 5 presents the additive direct and additive maternal genetic correlations with MTM from both samples, respectively.

TABLE 5 – Mean estimated additive direct (r_d) and additive maternal (r_m) genetic correlations with multiple trait model[1]

Mean age (days)	1	60	152	243	333	426	517	601	683
1	--	.64(.33)	.46(.12)	.41(.12)	.39(.14)	.28(.21)	.22(.09)	. 23(.13)	.53(.14)
60	.41(.25)	--	.78(.62)	.66(.58)	.48(.35)	.45(.18)	.34(.11)	.41(.16)	.54(.15)
152	.36(.14)	.79(.68)	--	.91(.80)	.65(.67)	.56(.52)	.31(.48)	.38(.48)	.64(.45)
243	.33(.23)	.71(.59)	.90(.82)	--	.82(.83)	.69(.68)	.43(.53)	.40(.59)	.65(.54)
333	.25(.16)	.58(.47)	.73(.73)	.89(.86)	--	.83(.88)	.60(.77)	.43(.74)	.70(.65)
426	.17(.18)	.47(.26)	.63(.54)	.73(.69)	.81(.83)	--	.82(.94)	.68(.92)	.62(.87)
517	.09(.24)	.37(.20)	.44(.45)	.55(.54)	.65(.75)	.78(.81)	--	.87(.92)	.71(.92)
601	.11(.10)	.37(.14)	.52(.43)	.52(.44)	.51(.60)	.64(.67)	.80(.80)	--	.57(.93)
683	.31(.06)	.38(.14)	.59(.45)	.67(.37)	.75(.51)	.76(.61)	.79(.74)	.77(.86)	--

[1] Numbers within parentheses refer to estimates from herds with no missing traits; r_d - above diagonal; r_m - below diagonal.

The additive direct genetic correlations were positive for both samples. Additive direct correlations between birth weight and all other weights and between the weight at 60 days and all other weights were higher for the sample from herds with missing traits than those from herds with no missing traits. On the contrary, after the weight at 152 days, most of the additive direct genetic correlations for the sample from herds with no missing traits were higher than those from herds with missing traits.

Alenda and Martin (1987) and Garrick et al. (1989) reported estimates with similar magnitude to those reported in this study. On the other hand, Meyer (1995) reported correlations between additive direct genetic effects for growth close to 1. Figures 5 and 6 show the additive direct genetic correlations for specific points (age) with MTM from both samples, respectively.

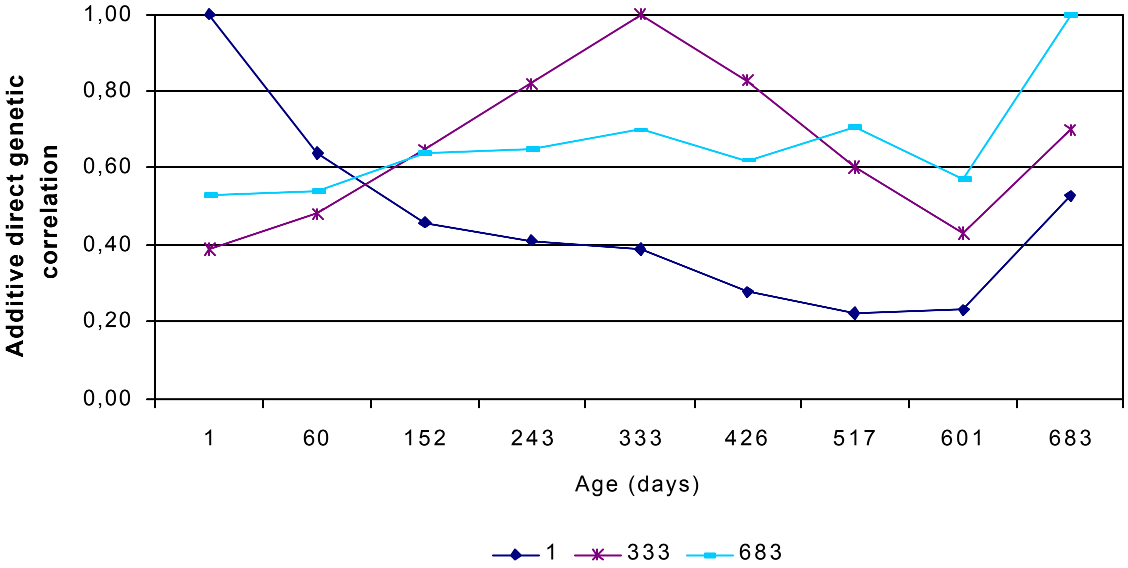

FIGURE 5 – Additive direct genetic correlations at different ages with multiple trait model for herds with missing traits.

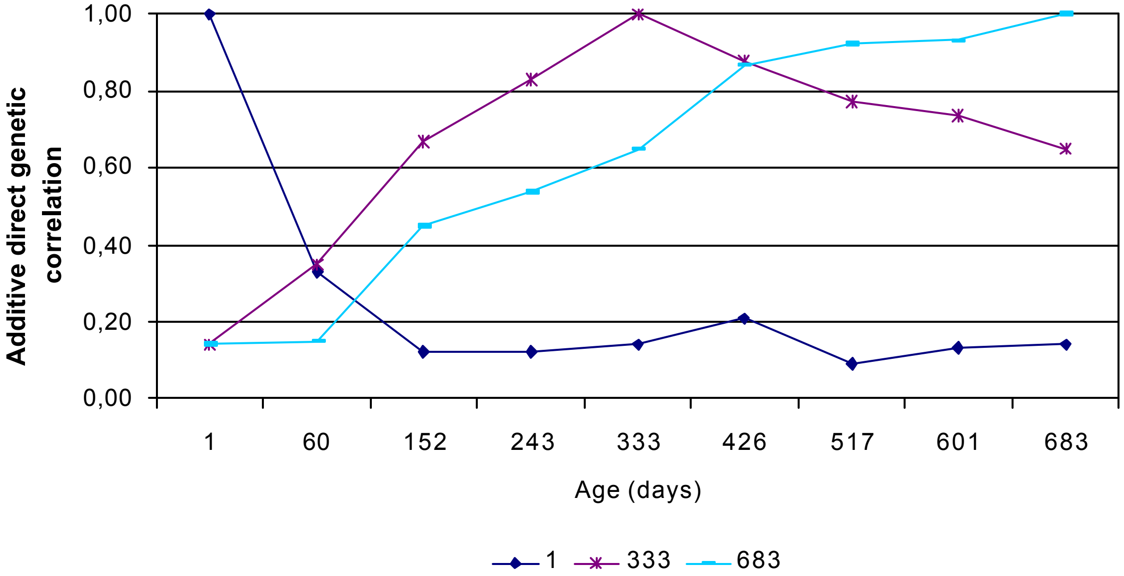

FIGURE 6 – Additive direct genetic correlations at different ages with multiple trait model for herds with no missing traits.

The estimates of the additive maternal genetic correlations estimated with MTM are presented in Table 5. All the values were positive, but the magnitudes were moderate. Additive maternal genetic correlations between birth weight (1 day) and all the other weights were smaller than those between all the other weights for both samples. In addition, values estimated for the sample from herds with no missing traits were lower than those for the sample from herds with missing traits. In general the estimates decreased after the weight at 243 days. However, these values are in agreement with values of other reports such as those obtained by Garrick et al. (1989), Waldron et al. (1993) and Eler et al. (1995). Additive maternal genetic effects that are not important for birth weight and small for weaning weight were reported by Meyer (1995).

The maternal genetic correlations with MTM from both samples are shown in Figures 7 and 8, respectively.

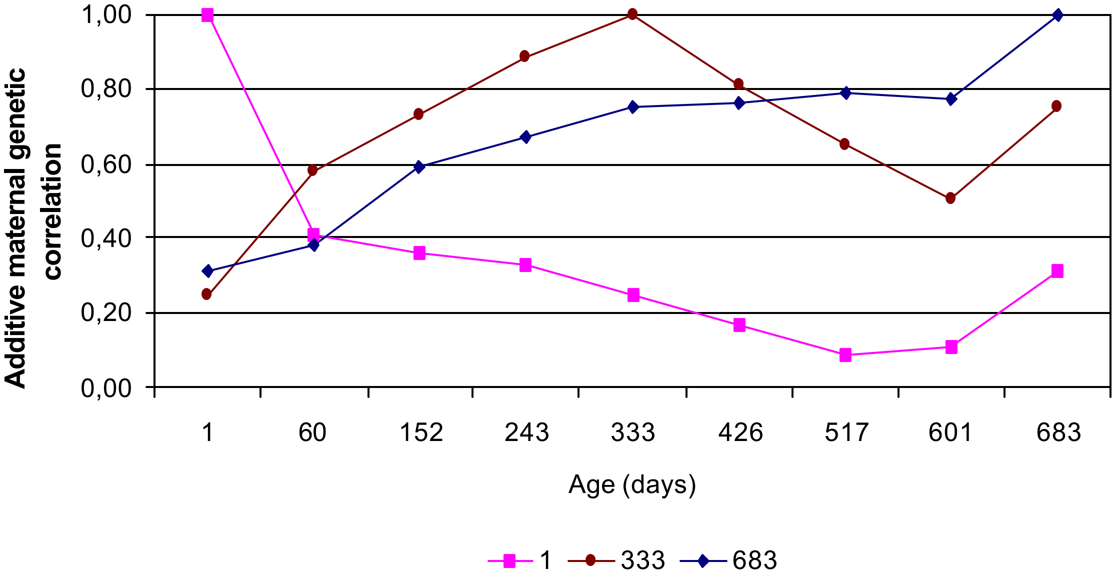

FIGURE 7 – Additive maternal genetic correlations at different ages with multiple trait model for herds with missing traits.

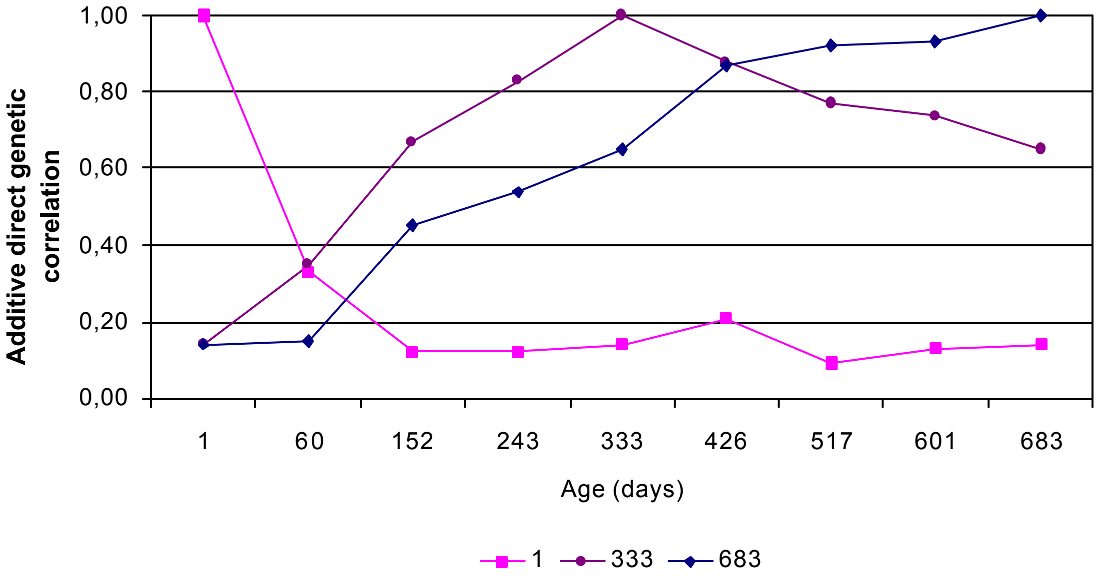

FIGURE 8 – Additive maternal genetic correlations at different ages with multiple trait model for herds with no missing traits.

The additive direct and additive maternal correlations with RRM for both samples are presented in Table 6.

TABLE 6 – Mean estimated additive direct (r_d) and additive maternal (r_m) genetic correlations with random regression model[1]

Mean age (days)	1	85	169	253	337	421	505	589	673
1	--	.65(.56)	.30(.32)	.20(.31)	.22(.39)	.29(.47)	.37(.52)	.43.(53)	.46(.50)
85	.77(.65)	--	.91(.96)	.83(.93)	.78(.89)	.73(.82)	.68(.73)	.66(.66)	.67(.63)
169	.54(.31)	.95(.92)	--	.98(.99)	.93(.94)	.85(.85)	.76(.74)	.70(.66)	.68(.64)
253	.40(.14)	.87(.82)	.98(.98)	--	.98(.98)	.92(.91)	.84(.82)	.78(.75)	.73(.73)
337	.29(.05)	.79(.73)	.93(.92)	.98(.98)	--	.98(.98)	.93(.92)	.87(.87)	.81(.84)
421	.22(.01)	.69(.63)	.85(.84)	.93(.93)	.98(.98)	--	.98(.98)	.94(.95)	.89(.93)
505	.20(.01)	.60(.54)	.76(.73)	.85(.83)	.93(.92)	.98(.98)	--	.98(.99)	.94(.97)
589	.25(.04)	.57(.47)	.69(.62)	.79(.72)	.88(.81)	.95(.89)	.99(.97)	--	.98(.99)
673	.36(.10)	.61(.44)	.70(.53)	.78(.59)	.86(.66)	.93(.74)	.98(.84)	.99(.94)	--

[1] Numbers within parentheses refer to estimates from herds with no missing traits; r_d - above diagonal; r_m - below diagonal.

The additive direct genetic correlations with RRM were all positive. The estimates of the additive direct genetic correlations with RRM were not consistent for early ages. After the weight at 152 days these estimates were higher than those with MTM for the sample with missing traits. However, the estimates of the additive direct genetic correlations with RRM and the sample with no missing traits were higher than those with MTM for all ages. The estimates of the additive maternal genetic correlations with RRM were not consistent at birth weight with herds with no missing traits. After this age (1 day) all additive maternal genetic correlations were higher than those with MTM. On the other hand, additive maternal genetic correlations with RRM for herds with missing traits were higher than those with MTM over the traits (Tables 5 and 6). In addition, for the same age the correlations were higher for early ages and lower for later ages. The highest estimates were obtained after the weight at 243 days with both models.

Figures 9 and 10 show the additive direct genetic correlation on a continuous scale with RRM, respectively, from both samples.

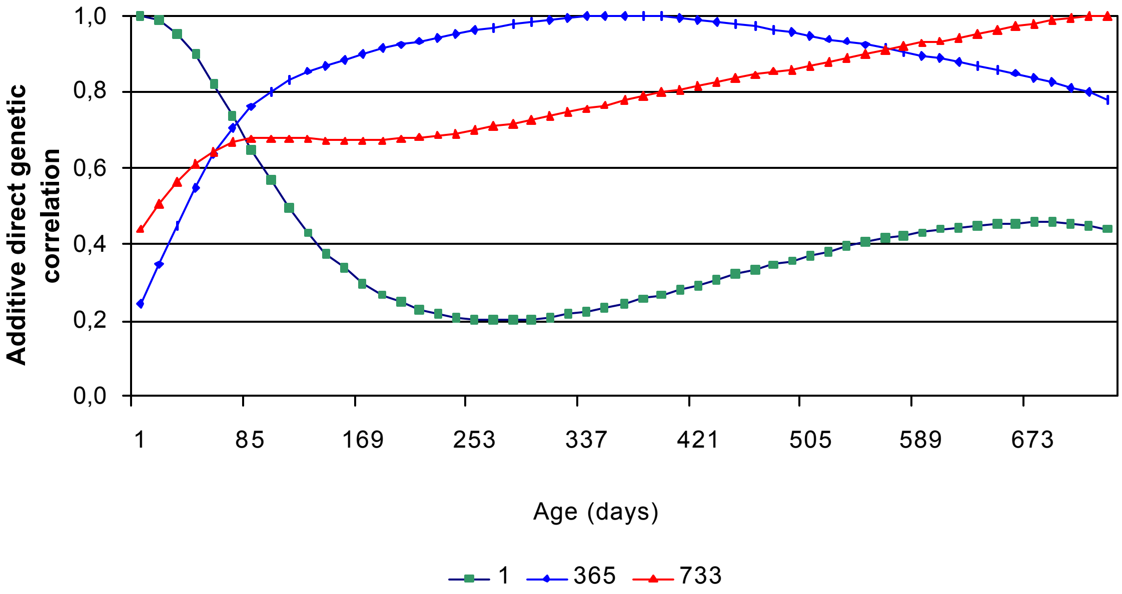

FIGURE 9 – Additive direct genetic correlations at different ages with random regression model for herds with missing traits.

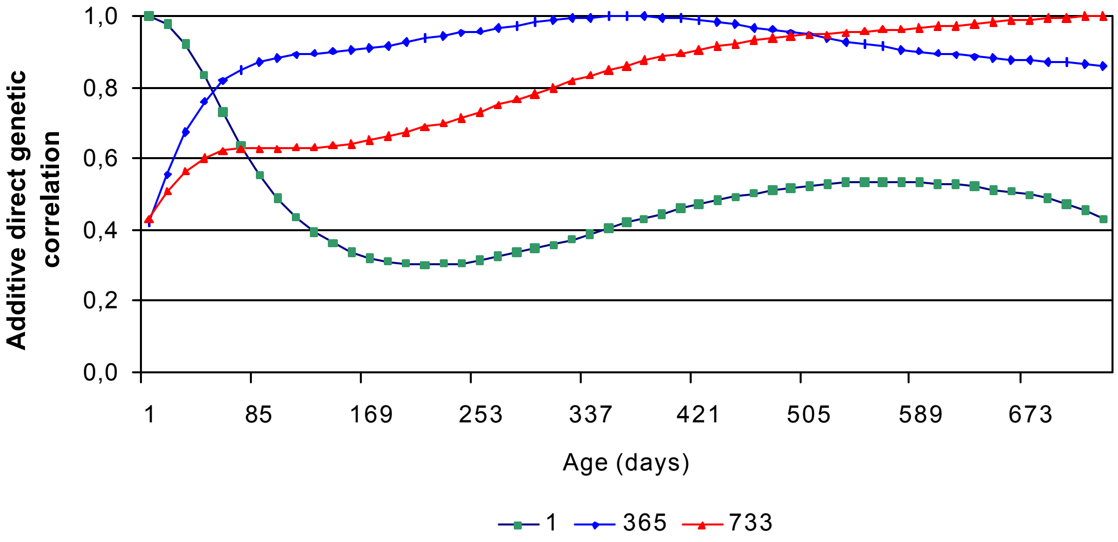

FIGURE 10 – Additive direct genetic correlations at different ages with random regression model for herds with no missing traits.

The shape of the curves with RRM for additive direct genetic correlations for herds with missing traits was similar to those for herds with no missing traits. The estimates for the sample from herds with no missing traits were slightly higher than those from herds with missing traits. In addition, lower correlations were estimated with RRM for both samples based on birth weight. This weight showed low correlation with other weights after the weight at 243 days. The correlations based on weight at 365 days were slightly higher when estimated for the sample with no missing traits. The values based on the weight at 733 days presented less variation when estimated for herds with no missing traits (Figures 9 and 10).

Figures 11 and 12 present the additive maternal genetic correlations with RRM at all the ages considered in this study. The estimates for the sample from herds with missing traits were slightly higher than those from herds with no missing traits. The correlations between birth weight and other weights were lower than those based on the weight at 365 days and 733 days. Higher values were estimated with weight at 365 days, and those estimated with weight at 733 were moderate in magnitude.

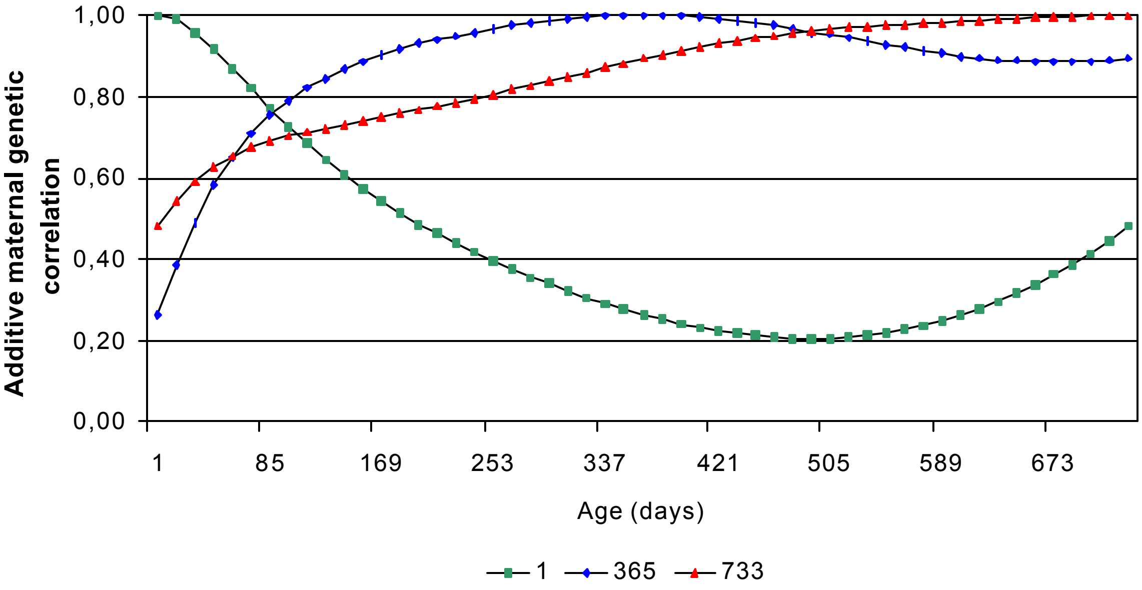

FIGURE 11 – Additive maternal genetic correlations at different ages with random regression model for herds with missing traits.

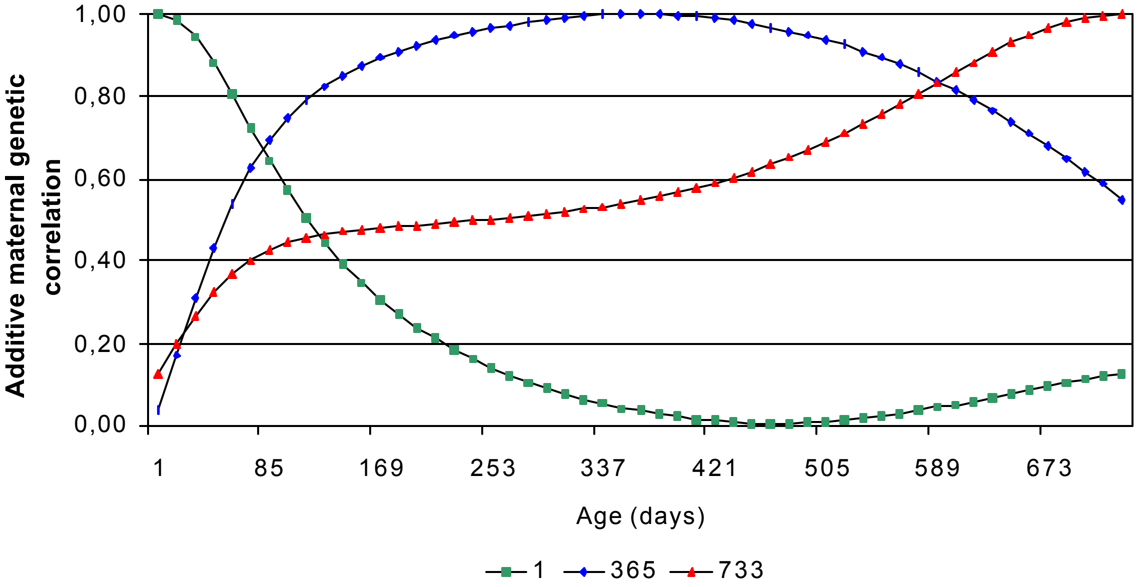

FIGURE 12 – Additive maternal genetic corrrelations at different ages with random regression model for herds with no missing traits .

Heritability

The mean additive direct and additive maternal heritabilities at different ages based on MTM from both samples for the sequential weights are presented in Table 7.

TABLE 7 – Mean additive direct heritabilities (h^2_d), additive maternal heritabilities (h^2_m) at different ages with multiple trait model

Mean age (days)	Herds with missing traits		Herds with no missing traits	
	h^2_d	h^2_m	h^2_d	h^2_m
1	21	13	26	18
60	27	22	12	20
152	30	21	13	25
243	24	21	15	23
333	22	17	16	18
426	23	15	17	14
517	31	16	23	14
601	41	15	27	15
683	33	16	25	11

The highest additive direct heritability with MTM was at the last weights, not only for the sample from herds with missing traits, but also for the sample from herds with no missing traits. The average estimate with MTM from herds with missing traits was 28.0 and from herds with no missing traits was 19.3.

The highest additive maternal heritability estimated from herds with missing traits was at early weights, and the average heritability for this sample 17.3; for the sample with herds with no missing traits the highest additive maternal heritability was at the weight at 152 days. The average for the additive maternal heritability based on this sample was 17.6.

Figures 13 and 14 show the curves for additive drect and additive maternal heritabilities with MTM from herds with missing and no missing traits, respectively.

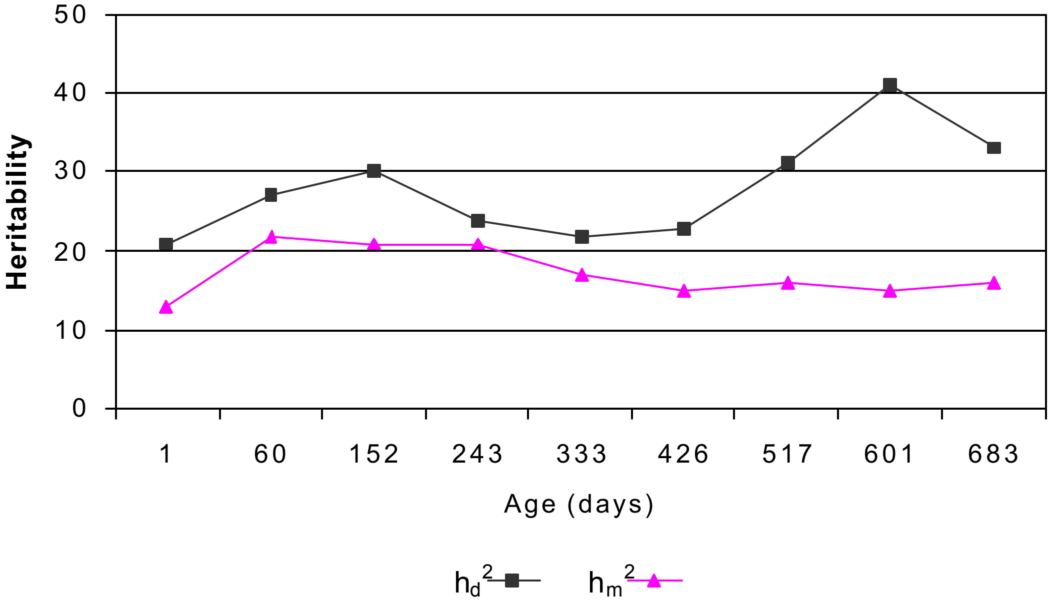

FIGURE 13 – Additive direct (h_d^2), and additive maternal (h_m^2) heritabilities at different ages with multiple trait model for herds with missing traits.

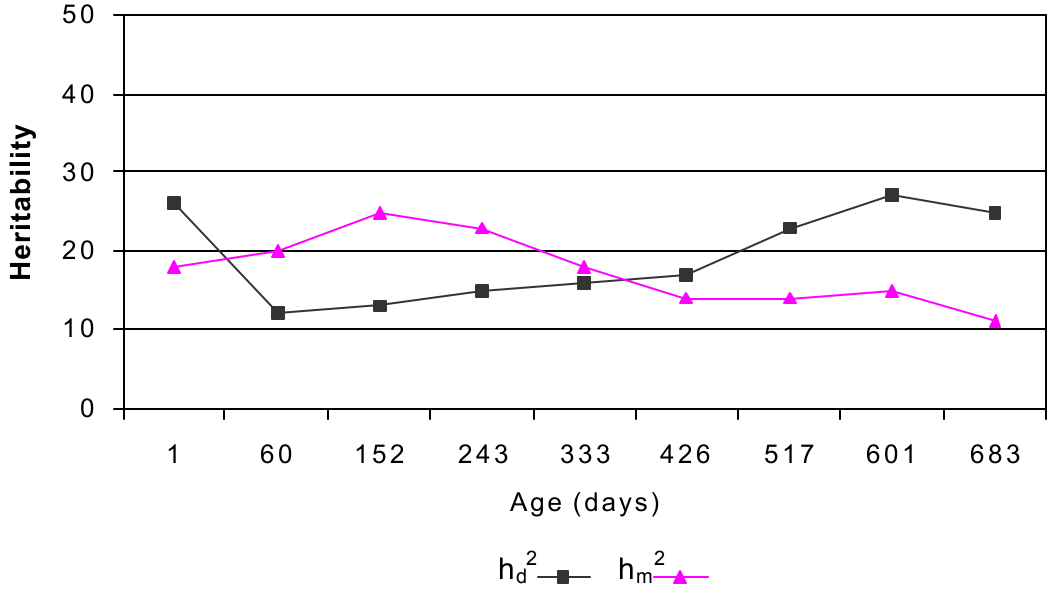

FIGURE 14 – Additive direct (h_d^2), and additive maternal (h_m^2) heritabilities at different ages with multiple trait model for herds with no missing traits.

Heritability estimates additive direct and additive maternal, for birth weight from both samples with MTM may be biased because when this weight is not recorded producers can submit a breed average estimate for each sex. In addition, the estimates for additive maternal effect were slightly higher than previous results reported by Waldron et al. (1993) and Eler et al. (1995) when estimated with multiple trait models.

Table 8 shows the mean estimates of the heritabilities, additive direct and additive maternal, at different ages with RRM from both samples.

TABLE 8 – Mean additive direct heritabilities (h^2_d), additive maternal heritabilities (h^2_m) at different ages with random regression model

Mean age (days)	Herds with missing traits		Herds with no missing traits	
	h^2_d	h^2_m	h^2_d	h^2_m
1	25	8	9	11
60	21	12	12	12
152	20	12	12	11
243	22	11	13	10
333	23	10	14	9
426	25	8	16	8
517	27	8	19	7
601	30	8	23	8
683	30	7	23	9

The additive direct heritabilities estimated with RRM (Table 8) were similar to those estimated with MTM (Table 7). On the other hand, additive maternal heritability obtained using RRM showed a small variation, ranging from 7 to 12 for both samples. The estimates of heritability for additive direct and additive maternal effects were slightly higher for both samples with MTM (Table 7).

The curves for additive direct and additive maternal heritabilities with RRM are presented in Figures 15 and 16 for both samples, respectively.

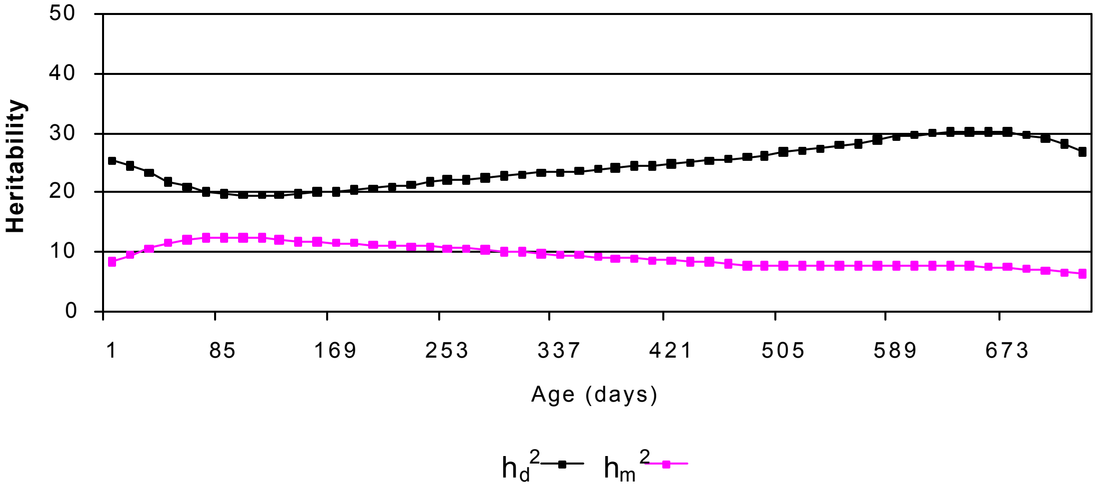

FIGURE 15 – Additive direct (h_d^2), and additive maternal (h_m^2) heritabilities at different ages with random regression model for herds with missing traits.

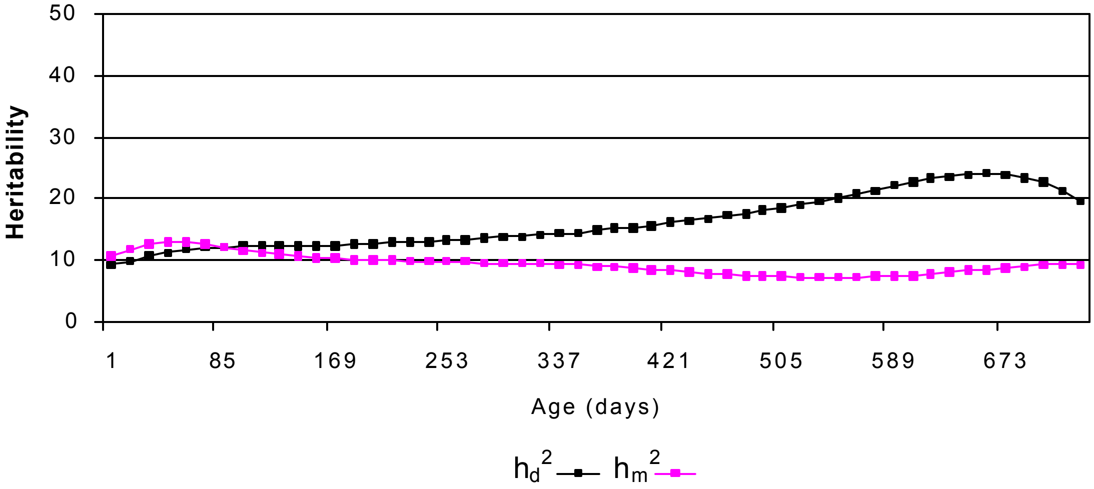

FIGURE 16 – Additive direct (h_d^2), and additive maternal (h_m^2) heritabilities at different ages with random regression model for herds with no missing traits.

Using RRM, the estimates of heritability for additive maternal effects decreased in importance after the weight at 243 days. This result is in agreement with the majority of reports in the literature (Garrick et al., 1989; Waldron at al., 1993 and Robinson, 1996a).

Additive Direct-Maternal Correlations

Table 9 reports the mean estimates of the correlations between direct and maternal additive effects from MTM for the two samples.

TABLE 9 – Mean correlations between direct and maternal additive effects at different ages with multiple trait model

Mean age (days)	Herds with missing traits	Herds with no missing traits
1	-.78	-.65
60	-.64	-.17
152	-.62	-.07
243	-.61	-.01
333	-.51	-.13
426	-.52	-.23
517	-.62	-.08
601	-.64	-.21
683	-.91	-.02

The magnitude of the correlation between direct and maternal additive genetic effects was negative across models and ages with both samples. Similarly, with MTM, Vesely and Robinson (1971), Bertrand and Benyshek (1987), Trus and Wilton (1988), Garrick et al. (1989), Eler et al. (1995), Robinson (1996a) and mattos et al. (2000) reported negative correlations between direct and maternal additive effects.

Figures 17 and 18 show the curves for the mean estimates of the correlation between direct and maternal additive effects with MTM from the samples.

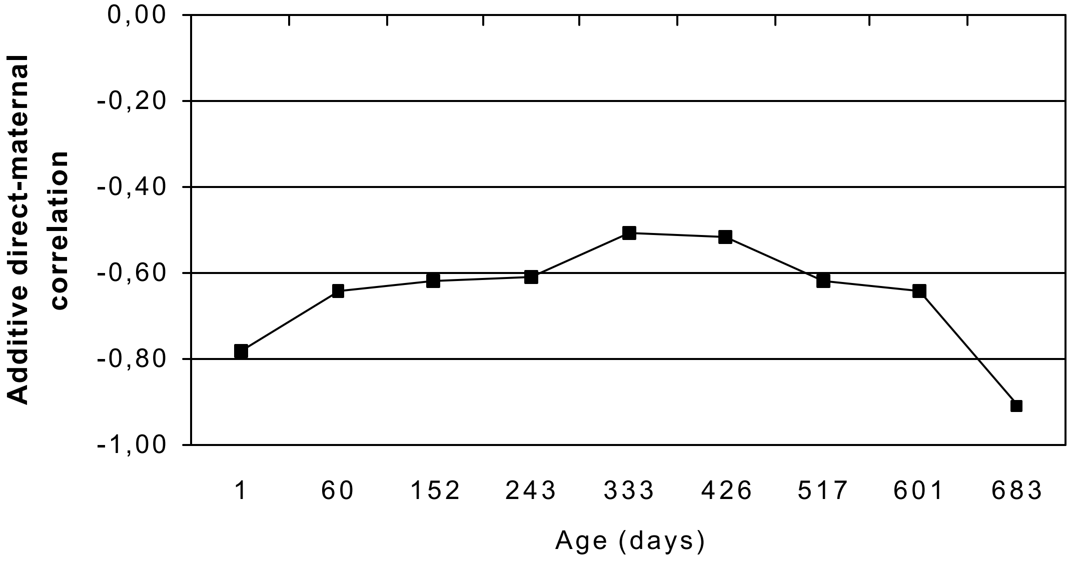

FIGURE 17 – Additive direct-maternal genetic correlation at different ages with multiple trait model for herds with missing traits.

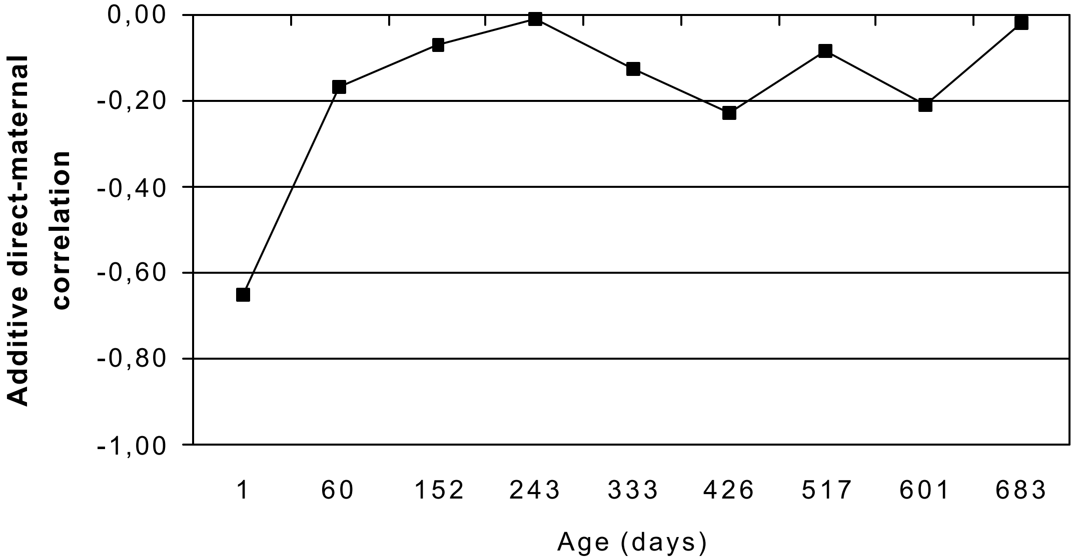

FIGURE 18 – Additive direct-maternal genetic correlation at different ages with random regression model for herds with no missing traits.

The mean estimates of the correlation between direct and maternal additive effects with RRM from both samples are presented in Table 10.

TABLE 10 – Mean correlations between direct and maternal additive effects at different ages with random regression model

Mean age (days)	Herds with missing traits	Herds with no missing traits
1	-.81	-.59
60	-.42	-.86
152	-.41	-.68
243	-.38	-.47
333	-.35	-.45
426	-.37	-.60
517	-.38	-.65
601	-.31	-.49
683	-.27	-.47

The correlations estimated using MTM from the sample with missing traits were stronger than those from the sample with no missing traits. These values presented a small variation when based on RRM with both samples.

The sign and magnitude of the genetic correlation between direct and maternal additive effects of growth traits is critical. According to Willham (1980) the estimation of additive maternal effects and their covariance components is inherently problematic since additive direct and additive maternal effects are generally confounded. Also, the expression of the additive maternal effects is sex limited, and occurs late in life of the female and lags by one generation.

Robinson (1996b) reported that estimates of additive maternal correlations may be negative not only because of genetic antagonisms, but also because of additional sire or sire x year variation or negative dam-offspring covariances. The author concluded that these causes will often be difficult to distinguish from true negative genetic correlations. The best method of detection is to fit the alternative models and assess their significance using likelihood ratio tests.

Moderate to large estimates of additive maternal genetic correlation in this study indicate that these may have been inflated by such environmental covariance, i.e.; that the models of analyses considered may not have been comprehensive enough (Meyer, 1992).

Figures 19 and 20 show the plots for the mean estimates of the correlations between direct and maternal additive effects with RRM from both samples.

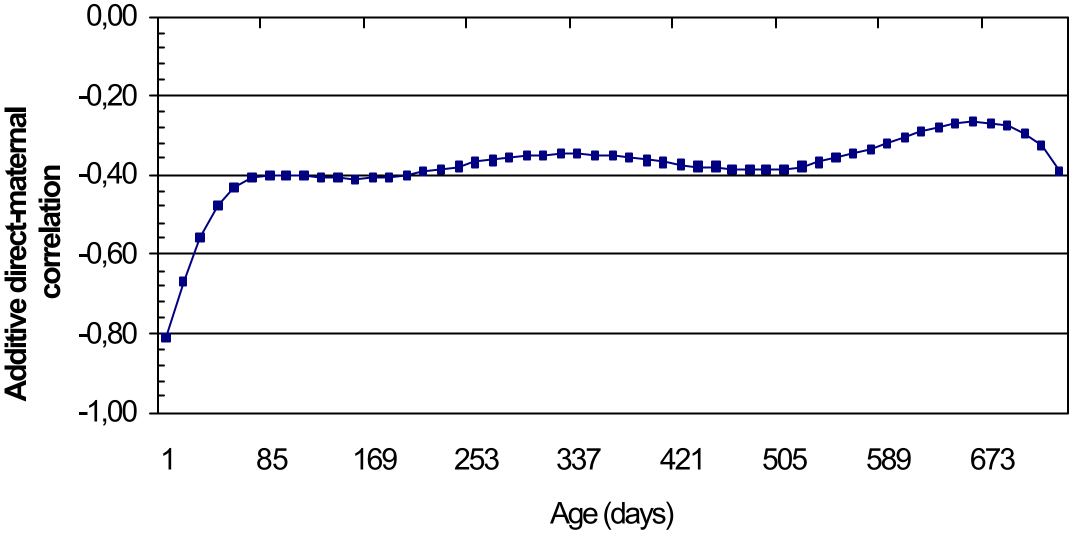

FIGURE 19 – Additive direct-maternal genetic correlation at different ages with random regression model for herds with missing traits.

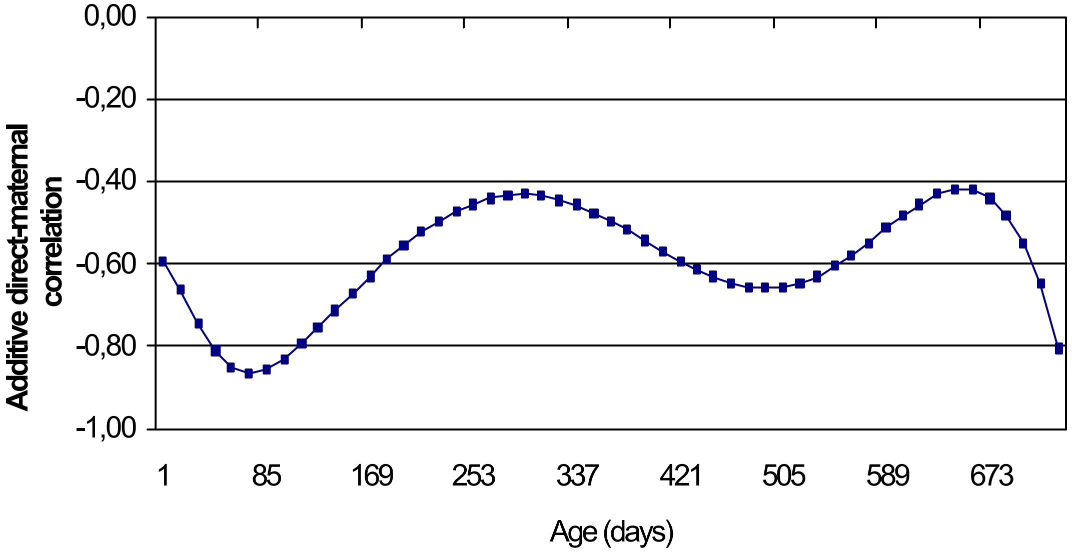

FIGURE 20 – Additive direct-maternal genetic correlation at different ages with random regression model for herds with no missing traits.

Postgibbs Analyses

Tables 11 and 12 show a summary including the mean, standard deviation, high posterior density interval, effective number size, auto correlations (lag 1, 10, 50) and independent chain size for additive direct, additive maternal, direct and maternal additive, maternal permanent environment and residual components for all traits from the two samples, respectively. The chain length was 100,000 rounds. Burn-in period was evaluated based on plots of samples from the Gibbs chain. It appeared that 40,000 rounds of sampling were sufficient for the burn-in.

The range of the mean covariance components for the same trait from the two samples data sets differed considerably, which must be due to the initial fluctuation in sample means. All high posterior density (HPD) intervals were narrow indicating sharp posterior distributions, which were expected given the large amount of information on the data. The posteriors means were approximately at the center of the HPD intervals reflecting symmetric posteriors distributions.

Due to the fact that independent samples are needed for the posterior estimates and the auto correlations for all components from the two samples were high (Tables 11 and 12), many rounds of iteration would be required for each sample used in the density estimates.

The effective number samples should be larger for parameters from the data set with no missing traits, especially for parameters corresponding to later weights. For many parameters, the opposite was found. This means that formulas used to calculate the effective sample size were inaccurate, and despite appearance of burn in, probably many more rounds were in fact required for convergence.

Figures 21 to 29 show the plots for additive, maternal and additive maternal components for all nine traits: birth weight, W1, W2, W3, W4, W5, W6, W7 and W8, for the sample with missing traits. Figures 30 to 38 show the plots for the sample with no missing traits.

Even though these posterior density estimates could be improved, some interesting information is available in the Figures. For example, the range of the mean covariance additive maternal effect is the same as of additive direct effect (Figure 24) for the sample with missing traits. Figures 31, 32 and 33 show the mean covariance component of maternal effect higher than additive effect for the sample with no missing traits. The mean covariance component of direct and maternal effect appears regularly in both samples for all traits as well; thus, strong evidence to support a negative value.

TABLE 11 – Mean, standard deviation (SD), high posterior density interval, effective number size, auto correlations and independent chain size with multiple trait model with missing traits

Mean age (days)	VC¹	Mean	SD	High posterior density interval (95%)		Effective number size	Auto correlations lag			Independent chain size
							1	10	50	
1	A	1.060	0.028	1.006	1.114	17.7	.96	.80	.48	448
	M	0.677	0.021	0.637	0.717	16.3	.97	.85	.61	450
	AM	-0.661	0.024	-0.708	-0.613	15.1	.97	.85	.60	448
	MPe	0.021	0.001	0.019	0.023	30.3	.98	.87	.50	310
	R	4.046	0.016	4.015	4.077	17.4	.92	.78	.49	456
60	A	28.46	0.493	27.490	29.430	54.3	.96	.76	.24	130
	M	22.84	1.094	20.700	24.990	8.0	.96	.76	.41	1314
	AM	-16.28	0.396	-17.060	-15.510	53.6	.97	.76	.28	124
	MPe	8.382	0.257	7.878	8.886	18.6	.96	.76	.40	452
	R	62.30	0.276	61.750	62.840	69.9	.76	.57	.20	124
152	A	90.02	3.216	83.720	96.330	17.4	.98	.86	.61	502
	M	63.18	1.086	61.050	65.310	42.8	.95	.74	.32	156
	AM	-46.86	1.470	-49.740	-43.970	32.9	.97	.83	.50	150
	MPe	39.89	1.121	37.690	42.090	21.7	.95	.74	.37	418
	R	151.5	1.619	148.400	154.700	22.0	.84	.72	.46	338
243	A	121.6	7.627	106.700	136.600	5.6	.97	.85	.63	1374
	M	106.6	1.916	102.800	110.300	36.1	.95	.71	.22	238
	AM	-69.10	2.944	-74.870	-63.330	14.4	.96	.76	.41	852
	MPe	65.31	1.939	61.510	69.110	16.5	.95	.77	.47	808
	R	275.5	3.411	268.900	282.200	10.6	.77	.64	.42	932
333	A	138.9	3.517	132.000	145.800	20.3	.95	.73	.40	462
	M	108.8	3.522	101.900	115.800	14.6	.96	.78	.41	688
	AM	-62.34	3.042	-68.300	-56.370	18.7	.97	.80	.52	378
	MPe	62.66	2.372	58.010	67.310	17.3	.97	.81	.53	718
	R	384.0	1.990	380.100	387.900	41.4	.61	.46	.26	440
426	A	172.2	7.872	156.700	187.600	11.9	.97	.81	.51	858
	M	114.7	5.407	104.100	125.300	11.9	.97	.82	.43	778
	AM	-73.65	5.629	-84.680	-62.620	12.4	.98	.84	.59	866
	MPe	73.24	2.298	68.740	77.740	24.3	.96	.78	.37	358
	R	456.1	3.887	448.500	463.700	23.6	.71	.55	.29	696
517	A	279.8	8.842	262.500	297.100	24.9	.97	.81	.42	310
	M	143.4	4.998	133.600	153.200	23.8	.97	.82	.32	378
	AM	-124.7	7.057	-138.500	-110.800	19.7	.98	.85	.47	364

	MPe	67.54	2.790	62.070	73.010	21.9	.98	.82	.41	350
	R	531.6	4.653	522.400	540.700	36.0	.80	.64	.33	260
601	A	433.7	9.337	415.300	452.100	32.2	.96	.75	.34	438
	M	158.3	8.064	142.500	174.100	11.4	.98	.84	.47	598
	AM	-169.0	7.608	-183.900	-154.000	16.7	.97	.80	.34	636
	MPe	82.98	3.194	76.720	89.240	23.0	.97	.79	.36	360
	R	540.7	4.074	532.700	548.700	63.5	.81	.57	.25	158
683	A	373.2	9.142	355.200	391.100	24.6	.97	.82	.48	266
	M	182.2	5.642	171.200	193.300	22.4	.94	.74	.48	318
	AM	-236.4	9.704	-255.400	-217.300	15.2	.97	.86	.66	360
	MPe	109.1	4.498	100.300	117.900	13.4	.97	.85	.58	550
	R	687.8	4.433	679.100	696.500	42.6	.87	.67	.33	160

[1] Variance components: additive direct (D), additive maternal (M), direct and maternal additive (AM), maternal permanent environment (MPe), and residual (R).

TABLE 12 – Mean, standard deviation (SD), high posterior density interval, effective number size, auto correlations and independent chain size with multiple trait model with no missing traits

Mean age (days)	VC[1]	Mean	SD	High posterior density interval (95%)		Effective number size	Auto correlations lag			Independent Chain size
							1	10	50	
1	A	1.320	0.023	1.275	1.364	77.8	.96	.67	.23	260
	M	0.9458	0.044	0.860	1.032	21.5	.98	.82	.47	536
	AM	-0.729	0.021	-0.771	-0.689	62.2	.97	.75	.32	172
	MPe	0.2065	0.013	0.181	0.232	27.1	.99	.89	.60	342
	R	3.381	0.040	3.303	3.460	21.5	.82	.65	.40	608
60	A	10.49	0.505	9.497	11.480	20.0	.98	.83	.53	746
	M	17.63	0.747	16.170	19.100	24.7	.98	.84	.50	528
	AM	-2.366	0.254	-2.366	-2.864	54.7	.97	.80	.38	162
	MPe	5.905	0.798	4.341	7.469	9.2	.99	.93	.75	1480
	R	56.54	0.392	55.770	57.310	53.8	.77	.60	.29	266
152	A	28.50	0.684	27.160	29.840	64.3	.97	.79	.37	192
	M	57.52	1.272	55.030	60.010	50.3	.96	.77	.44	182
	AM	-0.289	2.199	-4.601	4.021	9.7	.98	.87	.61	1670
	MPe	20.41	2.345	15.820	25.010	10.6	.99	.93	.73	1422
	R	121.2	0.839	119.600	122.800	76.0	.78	.49	.23	230
243	A	56.60	1.491	53.680	59.520	49.0	.97	.79	.42	198
	M	86.40	2.613	81.280	91.520	25.7	.96	.74	.45	712
	AM	-0.976	3.406	-6.948	4.995	11.4	.98	.84	.55	1270

	MPe	22.01	2.388	17.330	26.690	12.8	.99	.93	.75	904
	R	215.6	1.195	213.300	218.000	92.3	.74	.46	.22	174
333	A	71.70	2.088	67.610	75.790	40.8	.97	.78	.46	256
	M	82.51	4.134	74.410	90.620	15.6	.97	.82	.56	716
	AM	-9.872	4.106	-17.920	-1.824	9.8	.98	.86	.61	1300
	MPe	28.93	1.356	26.270	31.590	37.8	.99	.88	.58	204
	R	281.9	2.750	276.500	287.300	26.6	.75	.53	.27	900
426	A	83.35	3.307	76.870	89.830	20.8	.97	.79	.43	828
	M	65.70	1.913	61.950	69.450	30.2	.96	.75	.40	526
	AM	-17.36	2.163	-21.600	-13.120	18.5	.97	.80	.45	962
	MPe	27.40	1.771	23.930	30.870	25.2	.99	.89	.59	410
	R	321.4	3.469	314.600	328.200	22.1	.76	.52	.22	1160
517	A	136.9	2.992	131.000	142.700	53.7	.96	.73	.29	246
	M	81.25	2.779	75.810	86.700	27.2	.97	.77	.44	468
	AM	-7.975	1.949	-11.790	-4.155	47.4	.97	.78	.42	206
	MPe	38.76	2.848	33.170	44.340	23.0	.99	.91	.68	382
	R	350.0	2.078	345.900	354.000	88.3	.78	.53	.16	212
601	A	188.1	3.344	181.600	194.700	71.7	.96	.71	.22	246
	M	107.8	4.560	98.870	116.700	19.4	.97	.79	.40	644
	AM	-30.31	4.336	-38.810	-21.810	20.2	.97	.79	.43	748
	MPe	37.08	2.887	31.420	42.740	21.3	.99	.92	.68	618
	R	394.3	2.270	389.800	398.700	109.5	.80	.55	.14	208
683	A	198.2	2.292	193.700	202.600	146.2	.95	.65	.08	62
	M	83.65	6.003	71.890	95.420	10.4	.98	.87	.63	918
	AM	-3.059	5.320	-13.490	7.368	14.4	.98	.85	.57	750
	MPe	49.24	3.776	41.840	56.640	23.2	.99	.91	.67	442
	R	462.3	2.228	458.000	466.700	138.1	.78	.50	.07	202

[1] Variance components: additive direct (D), additive maternal (M), direct and maternal additive (AM), maternal permanent environment (MPe), and residual (R).

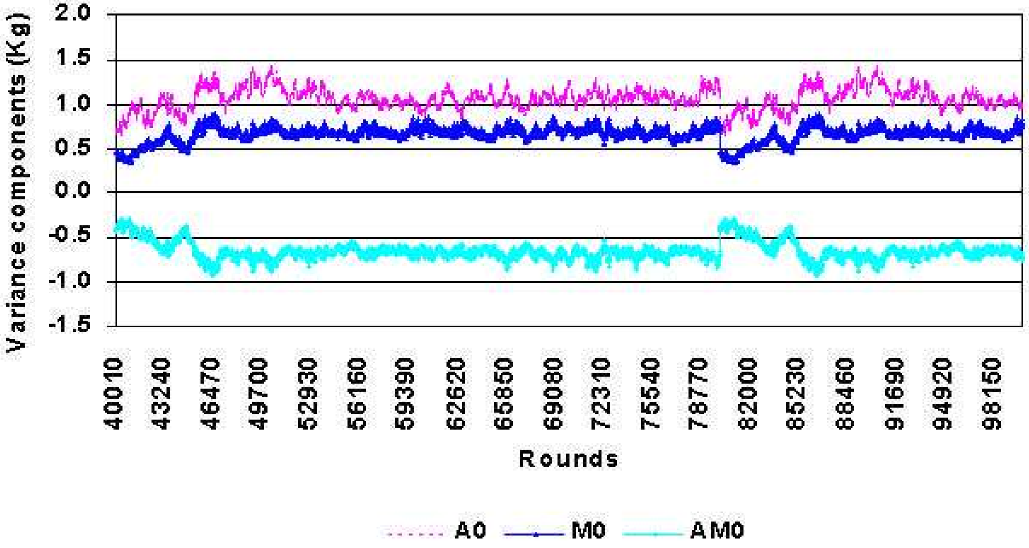

FIGURE 21 – Additive direct (A0), additive maternal (M0) and direct and maternal additive (AM0) components at different ages for birth weight with multiple trait model for herds with missing traits.

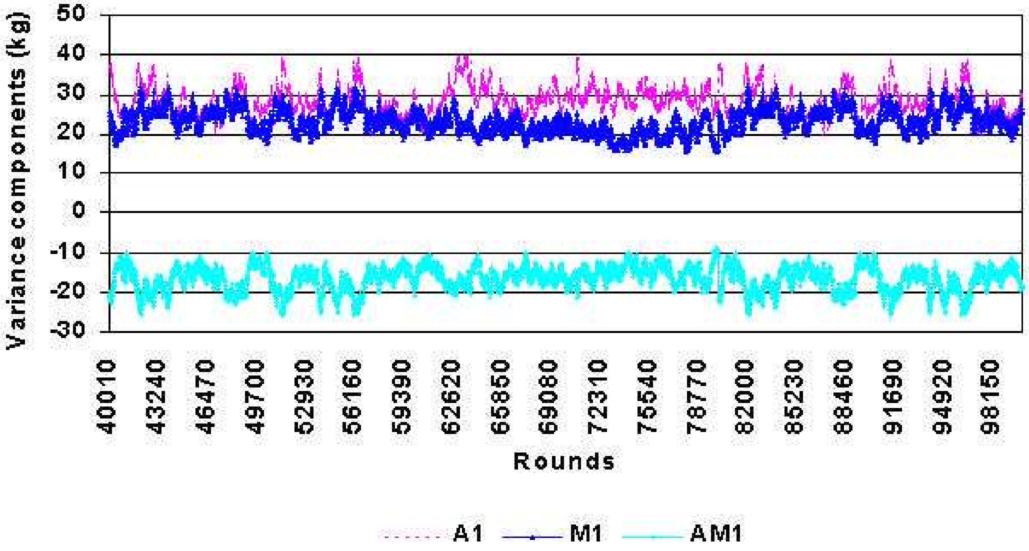

FIGURE 22 – Additive direct (A1), additive maternal (M1) and direct and maternal additive (AM1) components at different ages for W1 with multiple trait model for herds with missing traits.

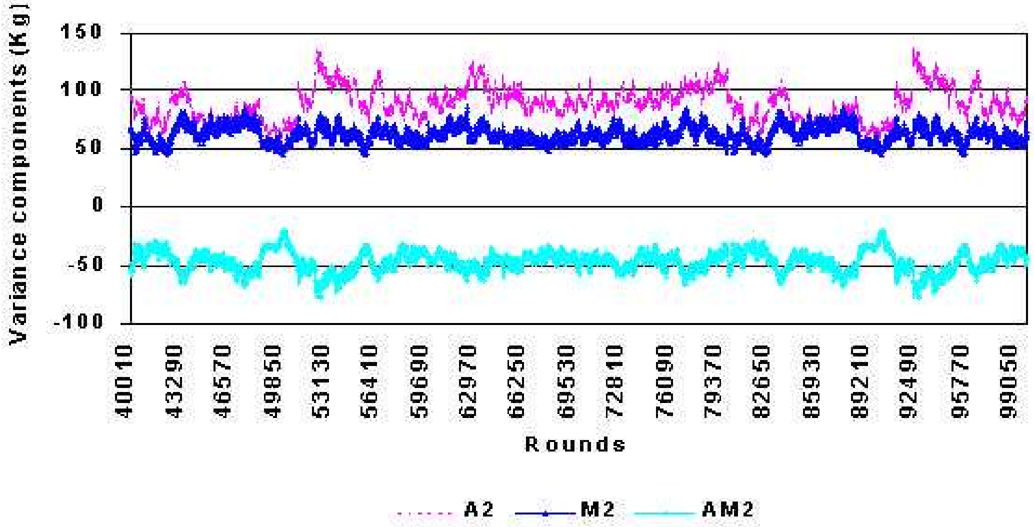

FIGURE 23 – Additive direct (A2), additive maternal (M2) and direct and maternal additive (AM2) components at different ages for W2 with multiple trait model for herds with missing traits.

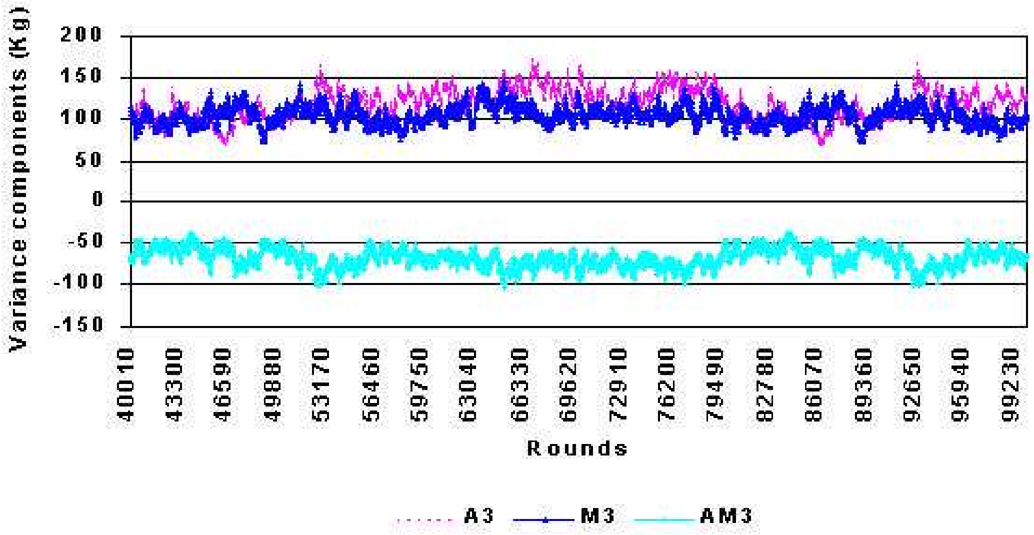

FIGURE 24 – Additive direct (A3), additive maternal (M3) and direct and maternal additive (AM3) components at different ages for W3 with multiple trait model for herds with missing traits.

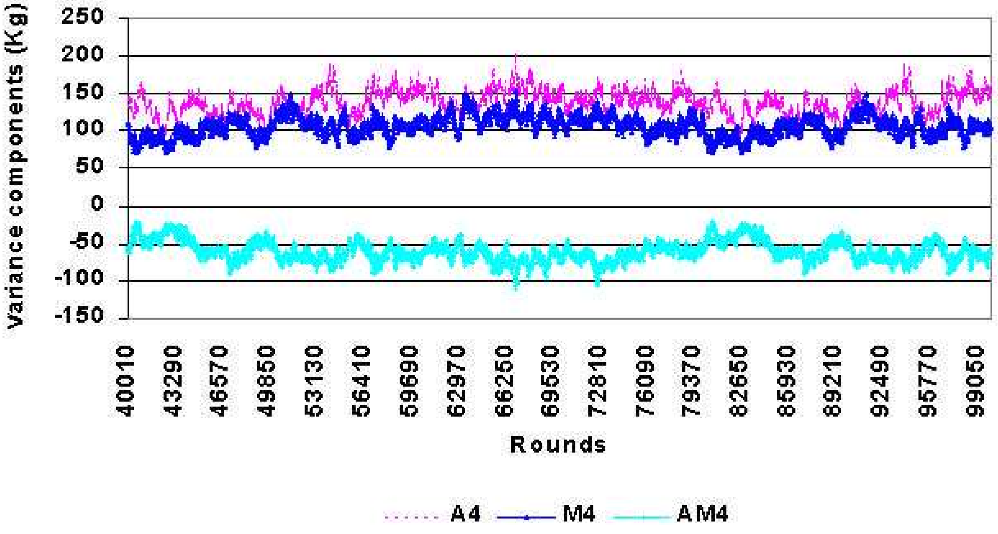

FIGURE 25 – Additive direct (A4), additive maternal (M4) and direct and maternal additive (AM4) components at different ages for W4 with multiple trait model for herds with missing traits.

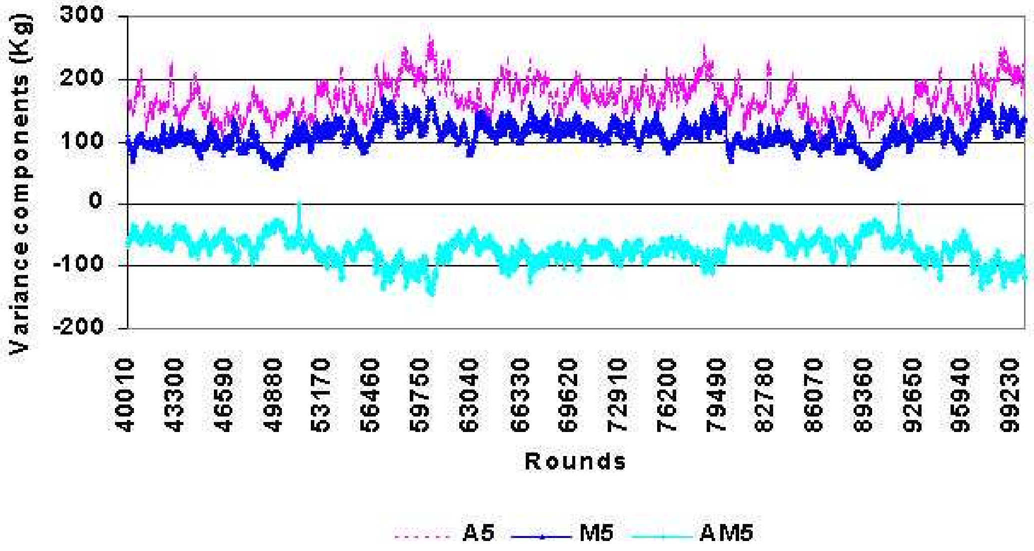

FIGURE 26 – Additive direct (A5), additive maternal (M5) and direct and maternal additive (AM5) components at different ages for W5 with multiple trait model for herds with missing traits.

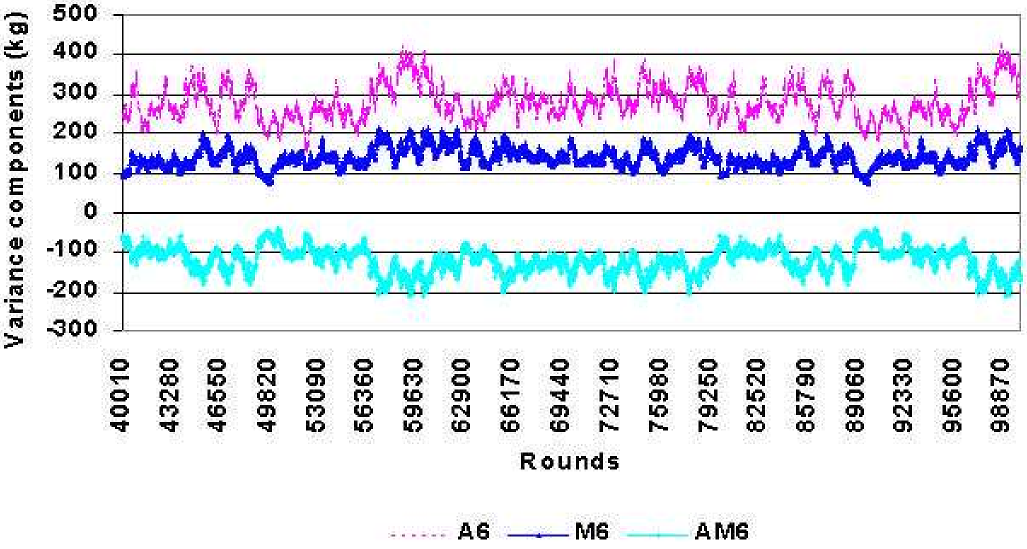

FIGURE 27 – Additive direct (A6), additive maternal (M6) and direct and maternal additive (AM6) components at different ages for W6 with multiple trait model for herds with missing traits.

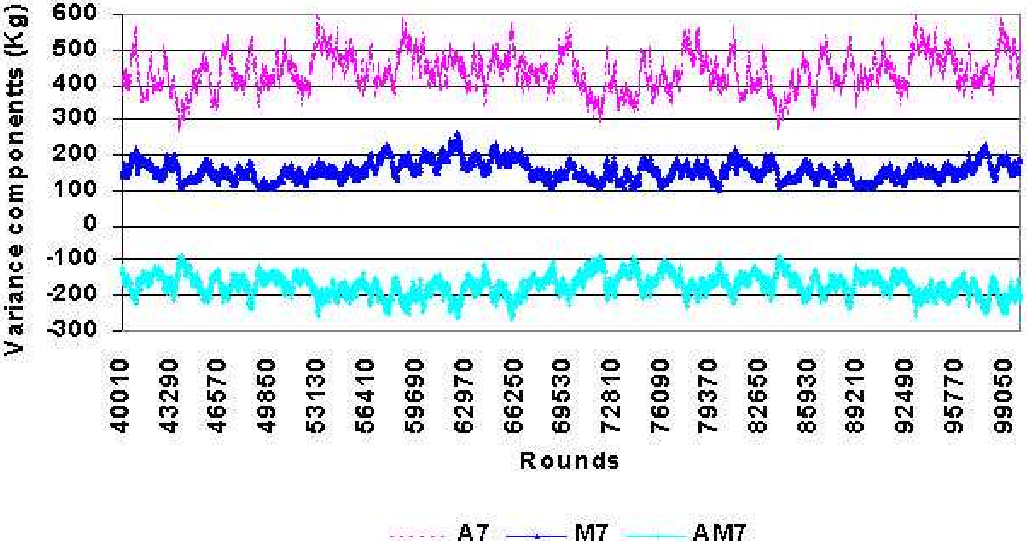

FIGURE 28 – Additive direct (A7), additive maternal (M7) and direct and maternal additive (AM7) components at different ages for W7 with multiple trait model for herds with missing traits.

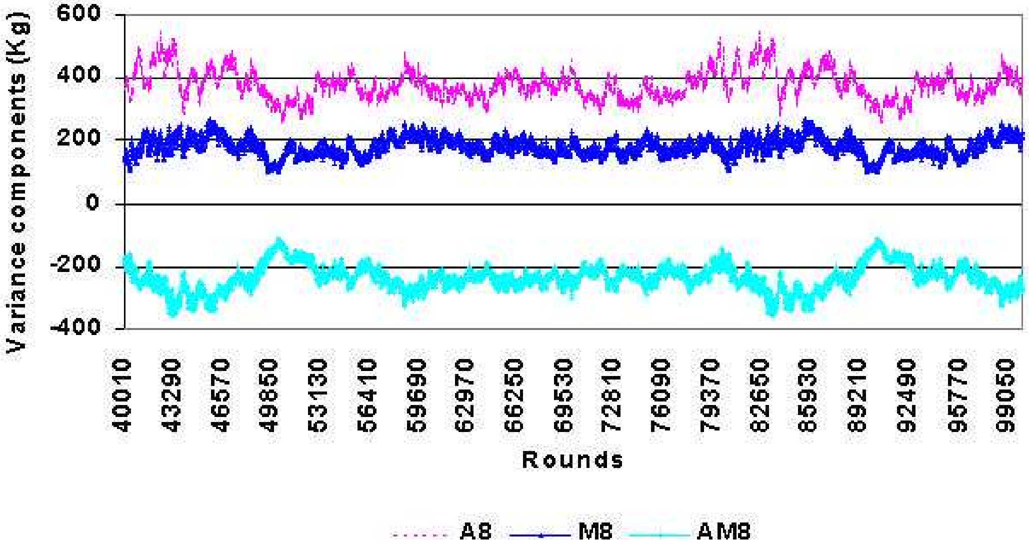

FIGURE 29 – Additive direct (A8), additive maternal (M8) and direct and maternal additive (AM8) components at different ages for W8 with multiple trait model for herds with missing traits.

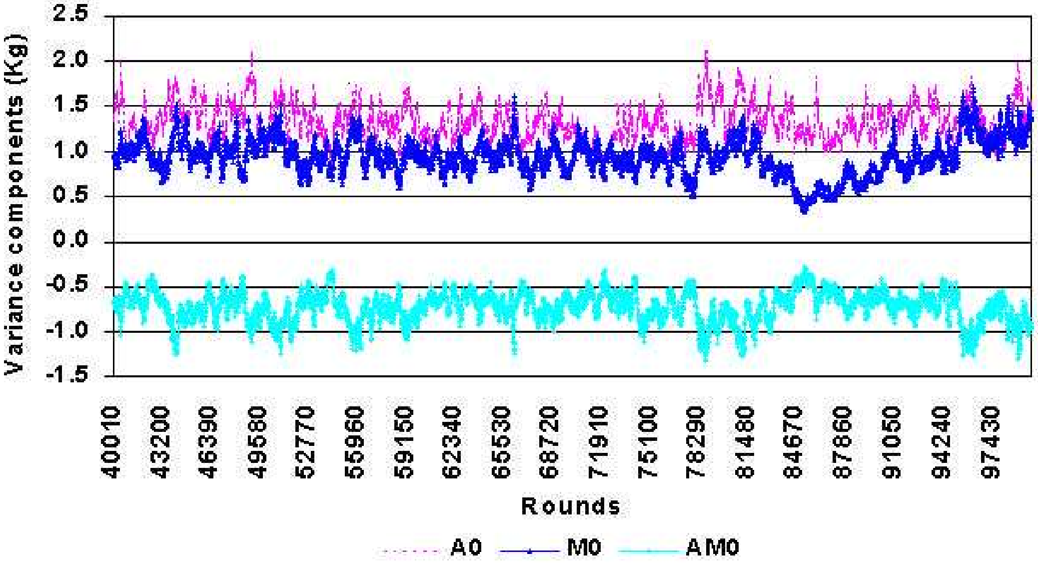

FIGURE 30 – Additive direct (A0), additive maternal (M0) and direct and maternal additive (AM0) components at different ages for birth weight with multiple trait model for herds with no missing traits.

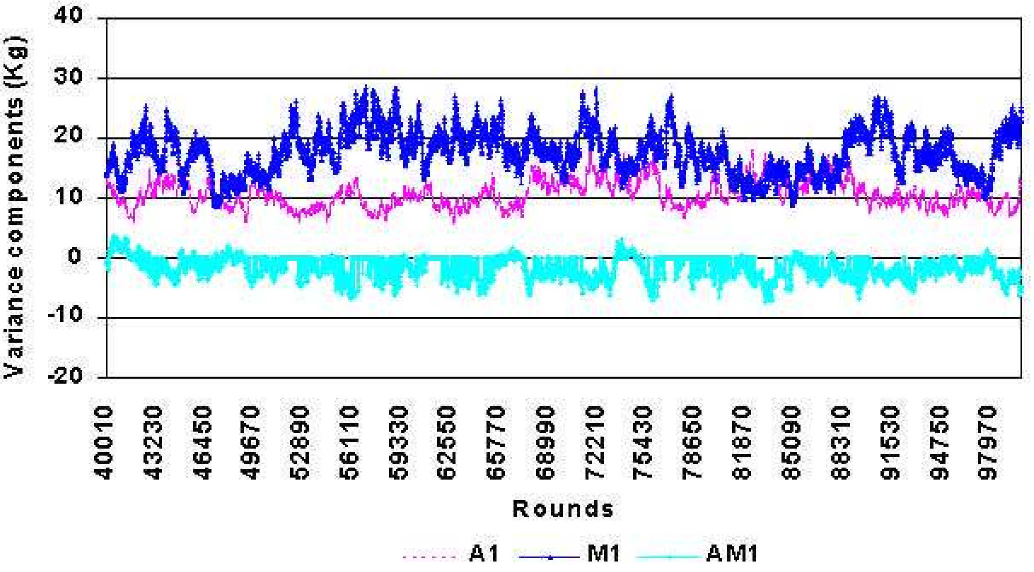

FIGURE 31 – Additive direct (A1), additive maternal (M1) and direct and maternal additive (AM1) components at different ages for W1 with multiple trait model for herds with no missing traits.

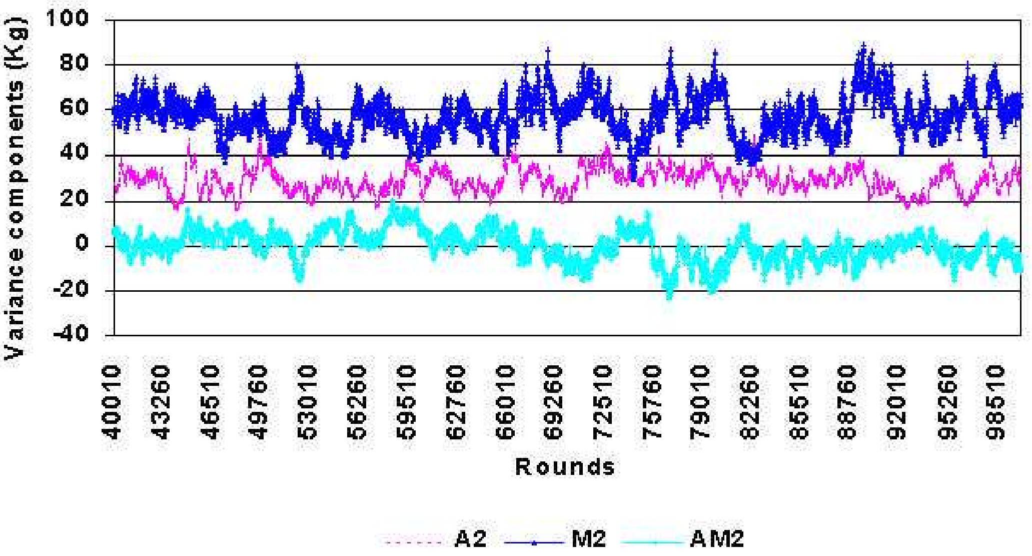

FIGURE 32 – Additive direct (A2), additive maternal (M2) and direct and maternal additive (AM2) components at different ages for W2 with multiple trait model for herdswith no missing traits.

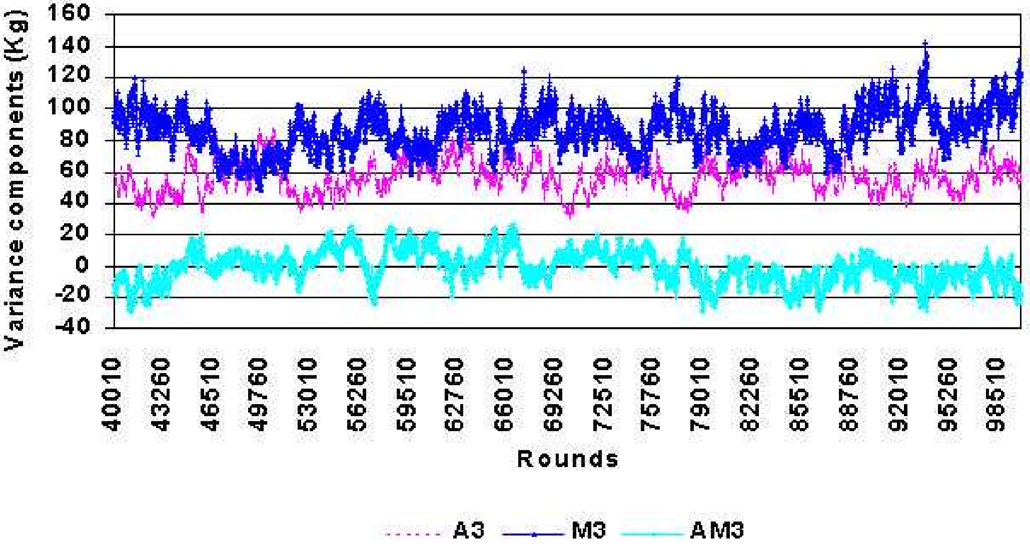

FIGURE 33 – Additive direct (A3), additive maternal (M3) and direct and maternal additive (AM3) components at different ages for W3 with multiple trait model for herds with no missing traits.

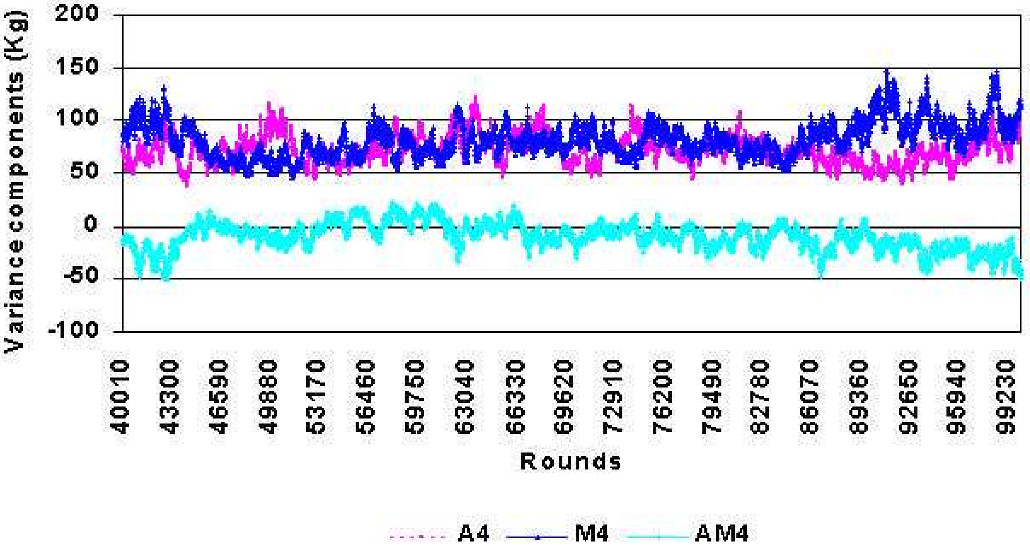

FIGURE 34 – Additive direct (A4), additive maternal (M4) and direct and maternal additive (AM4) components at different ages for W4 with multiple trait model for herds with no missing traits.

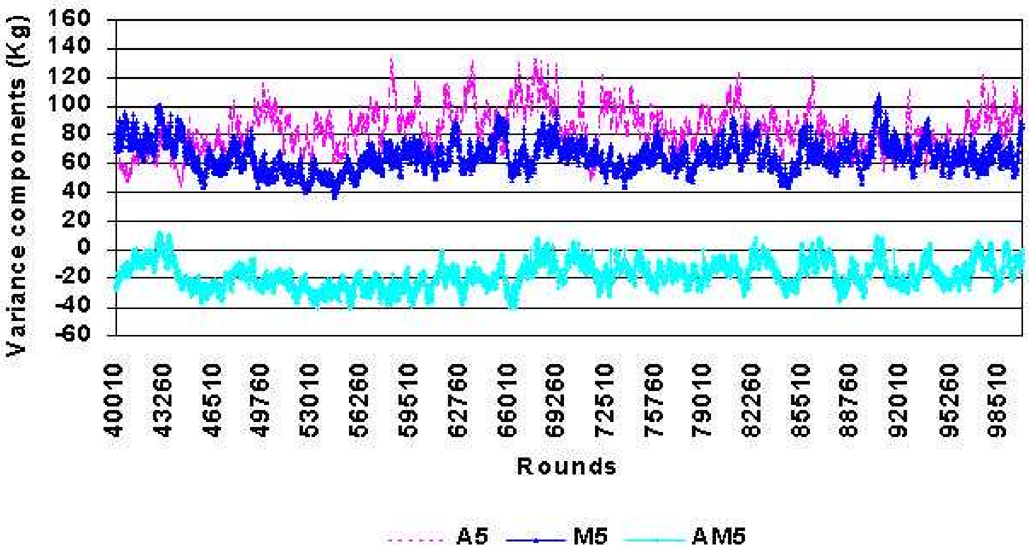

FIGURE 35 – Additive direct (A5), additive maternal (M5) and direct and maternal additive (AM5) components at different ages for W5 with multiple trait for herds with no with no missing traits.

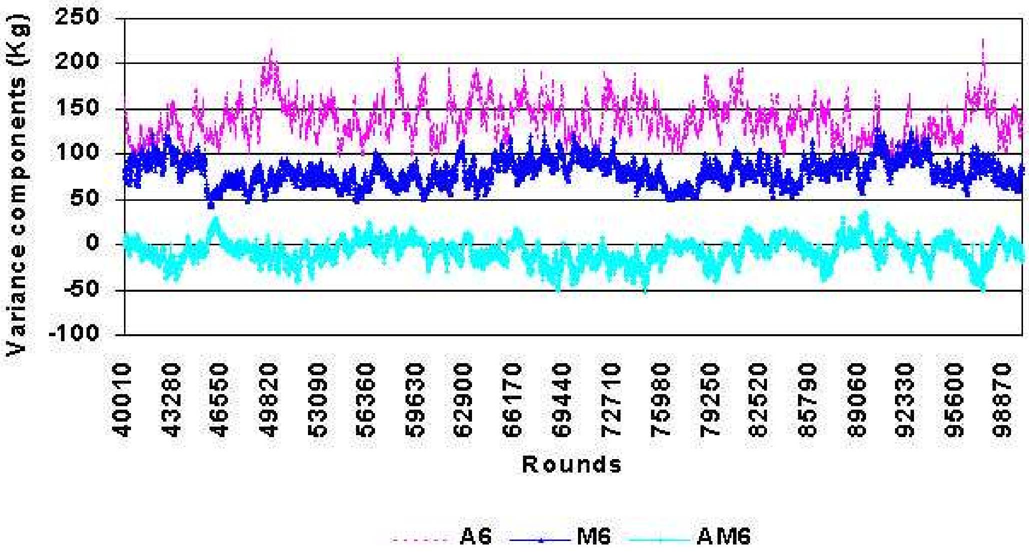

FIGURE 36 – Additive direct (A6), additive maternal (M6) and direct and maternal additive (AM6) components at different ages for W6 with multiple trait model for herds with no missing traits.

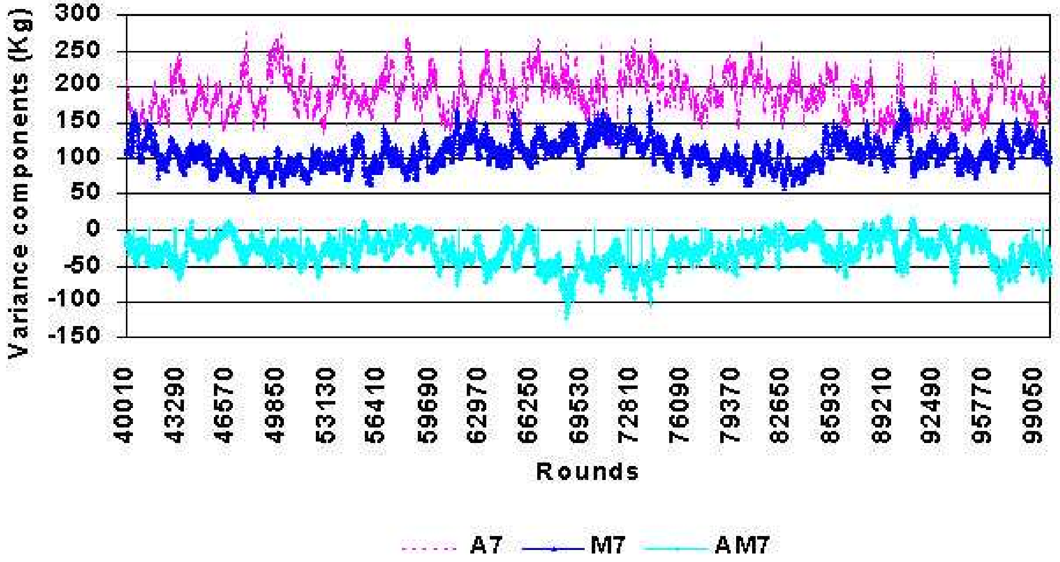

FIGURE 37 – Additive direct (A7), additive maternal (M7) and direct and maternal additive (AM7) components at different ages for W7 with multiple trait model for herds with no missing traits.

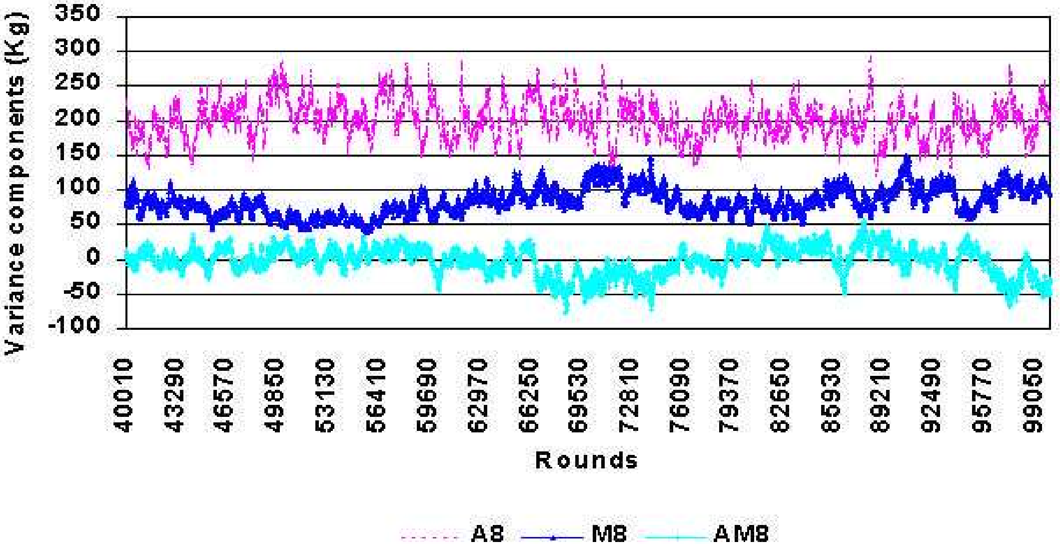

FIGURE 38 – Additive direct (A8), additive maternal (M8) and direct and maternal additive (AM8) components at different ages for W8 with multiple trait model for herds with no missing traits.

IMPLICATIONS

A complete multiple trait model with the number of traits equal to the number of ages would result in a highly overparameterised analysis. As a consequence, this would be likely to impose unnecessary computational demands. Also, more memory requirement in RRM is demanded as the number of covariates in the model is increased.

There are several advantages of Gibbs sampling to estimate variance components. First, the Gibbs sampling algorithm is easy for programming. Second, Bayesian inference using Gibbs sampling has potentially a low cost. Third, the Gibbs sampling allows estimation of densities directly as well as any function of interest (e.g., mean, variance, standard deviation, high posterior density interval). However, Gibbs sampling may be more computationally intensive.

In this study the memory requirements of an optimized Bayesian analysis via Gibbs sampling was low enough for a nine or more traits multiple trait analysis. However, running time was very long. A chain size of 100,000 seemed sufficient for a multiple trait analysis of a sample without missing traits but insufficient for a sample with missing traits. A chain size of 50,000 seemed insufficient for a cubic random regression model. The analyses of convergence as done visually for the burn in and formulas for the number of effective samples may be misleading with the result that the end of burn-in may be pronounced too early and the number of effective samples may be larger for analysis of more complete than incomplete data. Analyses using Bayesian methodology via Gibbs sampling need to be designed and scrutinized carefully. For analyzes with standard models where inferences on shapes of parameters are not important, analyses by REML may be more robust.

CONCLUSIONS

The growth curves by MTM for samples with missing and without missing data were similar, which indicates that the MT methodology via Gibbs sampling is resistant to missing trait if the missing pattern is sequential, as did via REML method.

Growth curves for the sample for herds with no missing traits were similar to those by REML.

Genetic parameters by GS and REML were different. This most likely was caused by insufficient chain length.

LITERATURE CITED

Alenda, R. and T.G. Martin. 1987. Genetic parameters and consequences of selection for growth traits in a beef herd selected for yearling weight. J. Anim. Sci. 70:734-741.

Bertrand, J.K. and L.L. Benyshek. 1987. Variance and covariance estimates for maternally influenced beef growth traits. J. Anim. Sci. 64:728-734.

Eler, J.P., L.D. Van Vleck, J.B.S. Ferraz and R.B. Lôbo. 1995. Estimation of variances due to direct and maternal effects for growth traits of Nelore cattle. J. Anim. Sci. 73:3253-3258.

Garrick, D.J., E.J. Pollack, R.L. Quaas and L.D. Van Vleck. 1989. Variance heterogeneity in direct and maternal weight traits by sex and percent purebred for Simmental-sired calves. J. Anim. Sci. 67:2515-2518.

Grignola, F.E., J. Jamrozik, L. Varona, I. Misztal and K. Bertrand. 1998. A random regression approach to estimate variance components for longitudinal records in beef cattle via Gibbs sampling. J. Anim. Sci. 76 (Suppl. 1): 57 (Abstr.).

Mattos, D., J.K. Bertrand and I. Misztal. 2000. Investigation of genotype by environment interactions for weaning weight for Herefords in three countries. J. Anim. Sci. 78:2121-2126.

Meyer, K. 1992. Variance components due to direct and maternal effects for growth traits of Australian beef cattle. Livest. Prod. Sci. 31:179-204.

Meyer, K. 1995. Estimates of genetic parameters and breeding values for New Zealand and Australian Angus cattle. Aust. J. Agric. Res. 46:1219-1229.

Meyer, K. 1999. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. J. Anim. Breed. Genet 116:181-205.

Meyer, K. 2000. Random regression to model phenotypic variation in monthly weights of Australian beef cows. Livest. Prod. Sci. 65:19-38.

Misztal, I. 2001. GIBBS2F90 manual. Available at: ftp://nce.ads.uga.edu/pub/ignacy/blupf90/. Accessed Mar 22, 01.

Robinson, D.L. 1996a. Estimation and interpretation of direct and maternal genetic parameters for weights of Australian Angus cattle. Livest. Prod. Sci. 45:1-11.

Robinson, D.L. 1996b. Models which might explain negative correlations between direct and maternal genetic effects. Livest Prod. Sci. 45:111-122.

Schaeffer, L.R. 1998. Variance Component Estimation Course Notes. University of New England, Armidale.

Sorensen, D.A. and B.W. Kennedy. 1984. Estimation of response to selection using least-squares and mixed model methodology. J. Anim. Sci. 58:1097-1106.

Thompson, R. 1973. The estimation of variance and covariance components with an application when records are subject to culling. Biometrics. 29:527-550.

Trus, D. and J.W. Wilton. 1988. Genetic parameters for maternal traits in beef cattle. Canad. J. Anim. Sci. 68:119-128.

Tsuruta, S. 2001. POSTGIBBS2F90. In: Misztal, I. 2001. GIBBS2F90 manual: Available at: ftp://nce.ads.uga.edu/pub/ignacy/blupf90/. Accessed Mar 22, 01.

Vesely, J.A. and O.W. Robinson. 1971. Genetic and maternal effect on preweaning growth and type score in beef calves. J. Anim.Sci. 32:825-831.

Waldron, D.F., C.A. Morris, R.L. Baker and D.L. Robinson. 1993. Maternal effects for growth traits in beef cattle. Livest. Prod. Sci. 34:57-70.

Willham , R.L. 1980. Problems in estimating maternal effects. Livest. Prod. Sci. 7:405-418.

CHAPTER V – EXPECTED PROGENY DIFFERENCES (EPD) OF GROWTH TRAITS OF NELLORE CATTLE USING MULTIPLE TRAIT AND RANDOM REGRESSION MODELS

ABSTRACT

Data of 812,393 Nellore cattle contained nine sequential weights measured at ages ranging from birth weight to 733 days for a total of 2,946,847 records were used to estimate expected progeny differences (EPD). EPDs were obtained by two models: a five-trait multiple trait model (MTM) and a random regression model (RRM). Effects in the MTM included contemporary group, age of dam class, additive direct, additive maternal, and permanent environment. The RRM included cubic regressions on age of animal, contemporary group, age of dam class, additive direct, permanent environment, additive maternal, maternal permanent environment, and error effects. EPDs were computed by two methods: a finite method using sparse factorization (SF) and interating (IT) by preconditioned conjugate gradient (PCG). The correlations between EPDs from MTM and RRM by IT were ≤ .43 until the random regressions were orthogonalized. After orthogonalization high computing requirements of RRM were reduced by removing regressions corresponding to very low eigenvalues and by replacing the random error effects with weights. Correlations between EPDs from MTM and RRM for the additive direct effect were .87, .89, .89, .87, and .86 for W1(weight at 60 days), W2 (weight at 252 days), W3 (weight at 243 days), W5 (weight at 426 days), and W7 (weight at 601 days), respectively. The corresponding correlations for the additive maternal effect were .85, .86, .88, .85 and .84, respectively. These low correlations were mostly due to differences in variances between the models and, to a lesser degree, due to better accounting for environmental effects and more data by RRM. The RRM applied to beef weights may be poorly conditioned numerically.

Key Words: beef cattle, expected progeny differences.

RESUMO

Informações de 812.393 animais Nelore contendo nove pesos seqüenciais coletados, do nascer aos 733 dias de idade, totalizando 2.946.847 pesos, foram usadas para estimar a diferença esperada na progênie (DEP). As DEPs foram estimadas por meio de dois modelos; o de características múltiplas (MTM), com cinco características, e o de regressões aleatórias (RRM). Os efeitos incluídos no MTM foram grupo contemporâneo, classe de idade da vaca, aditivo, materno e ambiente materno permanente. No RRM incluíram-se regressões cúbicas nos efeitos de idade do animal, grupo contemporâneo, classe de idade da vaca, aditivo direto, ambiente permanente, aditivo materno, ambiente materno permanente e resíduo. Dois métodos foram considerados nas análises: um método finito, FSPAKF90 (*Factorization sparse matrix package*), e o de iteração nos dados, PCG (*Preconditioned conjugate gradient*). As correlações entre as diferenças esperadas nas progênies (DEP), estimadas pelo MTM e pelo RRM, pelo método de iteração nos dados, foram muito baixas antes de se terem as regressões aleatórias ortogonais. Grande demanda computacional dos RRM foi reduzida pela remoção das regressões correspondentes a pequenas variâncias e também pela substituição dos efeitos aleatórios do erro por específica ponderação. Correlações entre DEPs, estimadas pelo MTM e pelo RRM, para o efeito aditivo direto foram .87, .89, .89, .87 e .86 para W1 (peso aos 60 dias), W2 (peso aos 152 dias), W3 (peso aos 243 dias), W5 (peso aos 426 dias) e W7 (peso aos 601 dias), respectivamente. As correlações correspondentes para o efeito aditivo materno foram .85, .86, .88, .85 e .84, respectivamente. Estimativas obtidas pelos RRM, em informações ponderais de gado de corte, podem não ser adequadas em virtude das propriedades numéricas desses modelos. Em geral, baixas correlações são devidas a diferenças de variâncias entre modelos, número insuficiente de graus de liberdade para estimar os efeitos de ambiente e informações perdidas nos RRM.

Palavras-chave: bovinos de corte, diferença esperada na progênie.

INTRODUCTION

In the past, genetic evaluation of beef cattle was done using traits defined at certain ages of animals (ABCZ, 2001; BIF, 1996 and CNPGC, 2001), i.e., at birth, 205, 365, and 550 days. Because ages at actual weights rarely fit the defined traits, weights obtained at ages close to those defined were preadjusted, and the other weights were discarded. The preadjustment was only for the mean but not for variances.

Currently, there is an interest in longitudinal models for beef, where weights at all ages can be accomodated, and expected progeny differences (EPD) can be obtained at any age of life. Varona et al. (1997) fit a growth curve to a Bayesian model. Their procedure was very time consuming and was suitable for evaluation only for small populations. Meyer (1999) analyzed continuous growth using random regression models. The author looked at the degree of fitting using various degrees of polynomials. Covariances between the direct and maternal effect were set to 0. Also, the author found that polynomials greater than the 4th order introduce too many artifacts into covariance functions while the polynomials below the 4th order may not model the covariances adequately. In Chapter III parameters of sequential weights of Nellore cattle were analyzed. Covariances between the direct and maternal effect were estimated. The authors have found that cubic polynomials provided satisfactory fit for covariances, and that estimated parameters were sensitive to the number of missing traits.

The analyzes of continuous traits may result in increased accuracy of evaluation by eliminating the need for preadjustment, by its ability to incorporate all weights with appropriate covariances. However the actual gains are unknown as none of the above studies attempted the evaluation with large data sets. Random regression models result in intensive computations (Misztal, 1999). Computer programs used for random regression models in dairy are complicated and extensively optimized (Jamrozik and Schaeffer, 2000). In beef cattle, the models are more complicated than in dairy because of correlated direct and maternal additive effects. Recently, Misztal (1999) developed a BLUP program that is simple but supports a large number of models including general random regression models. Tsuruta et al. (2001) has extended that program to support large data sets by iteration on data technique using the preconditioned conjugate gradient (PCG) iteration. The last program called BLUP90iod has sufficiently low memory requirements to support national genetic evaluations.

The first goal of this study was to implement the genetic evaluation of weights for a large population of beef cattle using the random regression model. The second goal was to compare these evaluations with those obtained from a multitrait evaluation.

MATERIALS AND METHODS

Data

Data were collected by the Brazilian Zebu Breeders Association (ABCZ) and provided by the Brazilian Agricultural Research Corporation (EMBRAPA). The data consisted of records on 619,989 Nellore animals; progeny of 11,847 sires, and 273,263 dams Nellore raised under Brazilian pasture conditions. The records were collected from 1975 to 1999.

Traits considered were birth weight (BW), 10 to 110 days weight (W1 or weight at 60 days), 102 to 202 days weight (W2 or weight at 152 days), 193 to 293 days weight (W3 or weight at 243 days), 283 to 383 days weight (W4 or weight at 333 days), 376 to 476 days weight (W5 or weight at 426 days), 467 to 567 days weight (W6 or weight at 517 days), 551 to 651 days weight (W7 or weight at 601 days), and 633 to 733 days weight (W8 or weight at 683 days).

Edits included eliminating records of animals outside the range of three standard deviations from the overall mean for each weight, and eliminating records outside of the range in age classes provided above. Table 1 presents characteristics of the data.

TABLE 1 – Characteristics of the data set by traits

Traits [1]	N [2]	Sires [3]	Dams [3]	Herds [3]	Mean (kg)	SD (kg)
BW (1)	608,870	11,847	273,263	4,747	29.81	2.65
W1 (60)	519,664	10,862	245,078	4,547	74.09	22.21
W2 (152)	451,273	10,191	224,496	4,329	133.22	28.64
W3 (243)	421,347	10,033	215,785	4,153	180.18	33.43
W4 (333)	312,869	8,916	173,148	3,702	206.11	37.96
W5 (426)	258,460	8,014	148,972	3,424	234.64	44.57
W6 (517)	216,486	7,280	129,260	3,166	270.30	51.88
W7 (601)	159,853	6,250	102,544	2,829	298.58	56.50
W8 (683)	37,565	3,505	32,813	1,642	328.73	61.37
Records	2,986,387					

[1] Numbers within parenthesis refer to the mean age (days); [2] Number of records; [3] With progeny in the data set.

Some analyzes used a sample of the complete data set consisting of 21,055 animals with 71,867 records. This sample was formed by randomly sampling from herds with missing traits with more than 500 birth weight records, and an average group size greater than 5 within each herd. Single record contemporary groups were eliminated from the sample, and then 5% of the herds that remained were sampled.

Models

General procedures

Two models were used for analyses: a multiple trait model (MTM) and a random regression model (RRM). The MTM was:

$$y = X\beta + Z_1 d + Z_2 m + Z_3 mp + e$$

where:

y was a vector of records preadjusted to fixed age; β was a vector of fixed effects (contemporary group and age of dam class); d was a vector of additive direct genetic random effects of the animal; m was a vector of additive maternal genetic random effects; mp was a vector of random effects of maternal permanent environment; X was the incidence matrix for fixed effects; Z_1 , Z_2 , and Z_3 were incidence matrices for animal, maternal, and maternal permanent environmental effects, respectively; and e was the vector of residual random effects.

The variances and covariances were defined as follows:

$$V = Var = \begin{bmatrix} d \\ m \\ mp \\ e \end{bmatrix} = \begin{bmatrix} G\otimes A & & 0 & 0 \\ & & 0 & 0 \\ 0 & 0 & MP\otimes I_c & 0 \\ 0 & 0 & 0 & R\otimes I_n \end{bmatrix}$$

where:

G was a 10 x 10 covariance matrix of random genetic effects; *MP* was a 5 x 5 covariance matrix of random maternal permanent environmental effects; *R* was a 5 x 5 covariance matrix of random residual effects; *A* was the additive genetic relationship matrix; I_c was an identity matrix whose order was the number of dams; I_n was an identity matrix whose order was the number of animals; and $\otimes$ was the direct product operator.

The random regression model (RRM) was defined as follows:

$$y_{ijklm} = \sum_{d=1}^{3} \hat{a}_d z_d + \sum_{d=0}^{3} cg_{di} z_{di} + \sum_{d=0}^{3} cad_{dj} z_{dj} + \sum_{d=0}^{3} d_{dk} z_{dk} + \sum_{d=0}^{3} p_{dk} z_{dk}$$
$$+ \sum_{d=0}^{3} m_{dl} z_{dl} + \sum_{d=0}^{3} mp_{dl} z_{dl} + \sum_{d=0}^{3} r_{dm} z_{dm} + \varepsilon_{ijklm}$$

where:

y_{ijklm} was the observation on contemporary group *i, age of dam class j,* animal *k,* dam *l,* and record *m;* β_d was the fixed regression coefficient *d* for age of animal*;* cg_{di} was the fixed regression coefficient *d* for contemporary group *i*; cad_{dj} was the fixed regression coefficient *d* for age of dam class *j;* d_{dk} and p_{dk} were random regression coefficients *d* for additive direct and permanent environmental effects of animal *k;* m_{dl} and mp_{dl} were random regression coefficients *d* for additive maternal and maternal permanent environmental effects of dam *l;* r_{dm} was the random regression coefficient *d* for residual effect of record *m;* z_d , z_{di} , z_{dj} , z_{dk} , z_{dl} , and z_{dm} were Legendre polynomials; and ε_{ijklm} were residual effects.

The mixed model equation with random regressions in matrix notation was:

$$y = X\beta + Z_1 d + Z_2 p + Z_3 m + Z_4 mp + Z_5 r + e$$

where:

y was the vector of records; β was the vector of fixed regressions; *d, p, m, mp,* and *r* were vectors for additive direct genetic, permanent environment, additive maternal genetic, maternal permanent environmental and residual effects, respectively; *X* was the incidence matrix for fixed effects; and Z_1, Z_2, Z_3,

Z_4, and Z_5 were incidence covariate matrices for additive direct genetic, permanent environment, additive maternal genetic, maternal permanent environment and residual random effects, respectively; and e was residual random effects.

The variances and covariances were defined as follows:

$$V = Var \begin{bmatrix} d \\ m \\ p \\ mp \\ r \\ e \end{bmatrix} = \begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{G \otimes A}} & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ 0 & 0 & P \otimes I_k & 0 & 0 & 0 \\ 0 & 0 & 0 & MP \otimes I_l & 0 & 0 \\ 0 & 0 & 0 & 0 & R \otimes I_m & 0 \\ 0 & 0 & 0 & 0 & 0 & I_n \sigma_e^2 \end{bmatrix}$$

where:

G was an 8 x 8 covariance matrix of random regression for genetic effects; P, MP and R were 4 x 4 covariance matrices of random regression for permanent environment, maternal permanent environment and residual effects, respectively; σ_e^2 was assumed constant residual variance; A was additive genetic relationship matrix; I_k was an identity matrix whose order was the number of animals; I_l was an identity matrix whose order was the number of dams; I_m was an identity matrix whose order was the number of records; I_n was the number of records; and $\otimes$ was the direct product operator.

Covariance components for both models were as estimated in Chapter III from a sample with missing traits.

Diagonalization

Legendre polynomials generally result in less correlated effects than regular polynomials but the degree of correlations is dependent on actual variances and distribution of ages in records. Any correlated regressions can be reperameterized to uncorrelated using the following transformations (Misztal et al., 2000). Considering the random regressions part of a model for one effect as:

$$\sum_{i=1}^{n} z_i a_{ij} = z' a_j \ , \ Var(a_j) = G_o$$

where:

z_i is the the i-th covariable,

a_{ij} is i^{th} regression for effect j,

G_o is the variance of the animal effect for the effect *j.*

Find eigenvectors *V* and eigenvalues *D* of G_o:

$G_o = V' D V$

Then, random regressions can be transformed as:

$z' a_j = z' V V' a_j = z^{*\prime} a^*_j$

where:

$z^* = V' z$ are new covariables, $a_j^* = V' a_j$ is the new regression vector for effect *j*, and the covariance structure for the effect *j* is diagonal:

$Var\ (a_j^*) = D$

Such a transformation can be done separately for each effect. In this case, covariables for each effect are no longer identical because they depend on G_o for a particular effect. After the diagonalization solutions are computed on a new scale. The solutions on the original scale can be obtained as $a_j = V a^*_j$.

In the reparameterization above, some eigenvalues may be very close to zero. Regressions corresponding to those covariables have values close to 0. Consequently, these regressions can be dropped from the model with negligible decrease in accuracy but at noticeable savings in computations.

In the RRM above, there was an extra residual effect *r*. The combined value of the residual + error variance was:

$$\hat{\sigma}_e^{2*} = z' R z + \hat{\sigma}_e^2$$

where:

R was covariance matrix of residual effects,

$\hat{\sigma}_e^2$ was residual variance,

z was the vector of covariables.

The effect r can be eliminated if the residual variance is computed for each observation as above and implemented in the model as a weight.

Computing Procedures

EPDs were obtained by program BLUPF90 (Misztal, 2001) with solutions obtained by the sparse-matrix factorization (Misztal and Perez-Enciso, 1998) and by BLUP90iod with the preconditioned conjugate gradient solver (Tsuruta et al., 2001). The first program computes exact solutions in the absence of numerical errors. The second program is iterative and computes increasingly more accurate solutions as the iteration progresses. The convergence criterion for that program was defined as the relative average squared differences between consecutive solutions; two criteria were used: 10^{-10} and 10^{-12}.

Computations

Initially, EPDs were obtained by programs BLUPF90 and BLUP90iod for the sample data set with the MT and RR models. Solutions by RRM were calculated before and after diagonalization and with lower and higher accuracy for BLUP90iod. Subsequently, the computations were repeated for the complete data set but only with program BLUP90iod because the computing requirements for BLUPF90 were excessive.

RESULTS AND DISCUSSION

Table 4 presents correlations between expected progeny differences obtained with multiple trait and random regression models for five traits based on the complete data set before diagonalization with BLUP90iod, which solutions were computed at the lower accuracy.

TABLE 4 – Estimated correlations between expected progeny differences with multiple trait and random regression models

	Traits [1]				
Effect	W1	W2	W3	W5	W7
Additive direct	.43	.40	.33	.24	.15
Additive maternal	.27	.26	.23	.20	.14

[1] W1 = weight at 60 days, W2 = weight at 152 days, W3 = weight at 243 days, W5 = weight at 426 days and W7 = weight at 601days.

The estimated correlations were very low, indicating numerical problems in calculation of solutions; i. e., bad numerical properties of Blup90iod, matrices poorly conditioned and accuracy level.

Table 5 shows correlations between estimates obtained by BLUPF90 and BLUP90iod at both accuracy levels for the sample data sets and the MT model.

TABLE 5 – Estimated correlations between genetic effects solution with multiple trait model for the sample data set

	Traits				
Programs	W1	W2	W3	W5	W7
BLUPF90 x BLUP90iod[1]	.99870	.99903	.99934	.99934	.99938
BLUPF90 x BLUP90iod[2]	.99970	.99903	.99994	.99994	.99998

[1] Convergence criterion = 10^{-10}.

[2] Convergence criterion = 10^{-12}.

While the correlations for the MT model were close to 1.0, the correlations between expected progeny differences from the two models were low, again indicating the convergence problem. Computing problems often occur when covariance matrices are poorly conditioned. Table 6 shows the values of eigenvalues corresponding to covariances in each effect of the RRM.

TABLE 6 – Eigenvalues for genetic (G), permanent environment (Pe), maternal permanent environment (MPe), and residual (R) effects with random regression model

Effects	Eigenvalues							
G	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	.99787	15.23651	22.55033	36.18708	160.652681
Pe	$< 10^{-4}$	6.92246	39.47664	240.07977	---	---	---	---
MPe	$< 10^{-4}$	1.79353	3.65663	63.686211	---	---	---	---
R	$< 10^{-2}$	13.17468	40.65412	72.11852	---	---	---	---

For the additive effects, three eigenvalues were close to 0, and one eigenvalue was close to 0 in each of the remaining effects. Small eigenvalues indicated that parameters of RRM were poorly conditioned and also indicated potential of reducing the number of effects in the model.

The diagonalization was done for three different data sets. The first one was for the birth weight sample. The second was for the complete sample, which contained all weights, and the last was for the complete data. Also, the reparameterization was done in three steps. In the first step, all covariance matrices were diagonalized (complete model with 31 effects). In the second step, the error effect was reduced to weights (weighed model with 27 effects). In the third step, effects with eigenvalues close to 0 were dropped from the model (reduced model with 22 effects). Low correlations remained until the covariance matrices were recreated as VD^*V', where D^* was like D but with small eigenvalues set to 0. Table 7 shows correlations between estimates obtained by BLUPF90 and BLUP90iod at both accuracy levels for the sample data sets for the RRM after the reparameterization with “recreated” covariances.

TABLE 7 – Estimated correlations between genetic effects solution with random regression model for the sample data set

Number of effects in RRM[1]	Program	Correlation
31	BLUPF90 x BLUP90iod[2]	.979
	BLUPF90 x BLUP90iod [3]	.998
27	BLUPF90 x BLUP90iod [2]	.979
	BLUPF90 x BLUP90iod[3]	.998
22	BLUPF90 x BLUP90iod[2]	.979
	BLUPF90 x BLUP90iod[3]	.998

[1] 31 effects = complete model, 27 effects = weighed model and 22 effects = reduced model.

[2] Convergence criterion = 10^{-10}; [3] Convergence criterion = 10^{-12}.

In all cases, the correlations between the estimates from BLUPF90 and BLUP90iod were .979 for the lower accuracy and .998 for the higher accuracy. This indicated that the reparameterization was successful and that higher accuracy of BLUP90iod was essential in obtaining EPDs from RRM.

Misztal et al. (2000) reported that the goal of parameter estimation in test-day models is not to obtain “true” estimates because, as such, they do not exist, but to achieve an acceptable compromise between model complexity and estimation accuracy. The estimated correlations in Table 7 were not equal in magnitude for different convergence criterion with BLUP90iod program; however, they were close to 1.0 when the convergence criterion was 10^{-12}. Lidauer et al. (1999) reported that estimation of breeding values with RRM required greater accuracy in the solutions of mixed model equations than with the single trait animal model. According to the authors, a breeding value of an animal in RRM is a function of breeding values coefficients rather than a single solution from mixed model equations.

Table 8 presents estimated correlations between expected progeny differences with multiple trait and random regression models for sample data set with birth weight as the only trait, memory requirement and running time with BLUPF90 and BLUP90iod programs.

TABLE 8 – Estimated correlations between expected progeny differences with multiple trait and random regression models based on the birth weight sample

Number of effects in RRM[1]	Program	Correlation Additive direct	 Additive maternal	Memory requirements (M)	Running time (minutes)
31	BLUPF90	.86	.70	482-618	16
	BLUP90iod	.86	.70	16	10
27	BLUPF90	.86	.70	349-465	13
	BLUP90iod	.86	.70	13	8
22	BLUPF90	.86	.70	213-273	9
	BLUP90iod	.86	.70	9.6	2

[1] 31 effects = complete model, 27 effects = weighed model and 22 effects = reduced model.

The estimated correlations in Table 8 were equal in magnitude despite the fact that the model had a different number of effects. Thus simplifications in models decreased the computing cost drastically without decreasing the accuracies of solutions. However, the correlations were below .99 indicating differences in the models.

Estimated correlations between expected progeny differences with multiple trait and random regression models by BLUP90iod program with high convergence criterion (10^{-12}) for the complete sample, which contained all weights, with different numbers of effects in RRM are reported in Table 9.

TABLE 9 – Estimated correlations between expected progeny differences with multiple trait and random regression models by BLUP90iod program based on the complete sample

Number of effects in RRM[1]	Traits						Memory requirement (M)	Running time (minutes)
	Effects	W1	W2	W3	W5	W7		
31	Additive direct	.86	.86	.86	.87	.85	24	75
	Additive maternal	.80	.81	.80	.85	.85		
27	Additive direct	.86	.86	.86	.87	.85	13	60
	Additive maternal	.80	.81	.80	.85	.85		
22	Additive direct	.86	.86	.86	.87	.85	9.6	40
	Additive maternal	.80	.81	.80	.85	.85		

[1] 31 effects = complete model, 27 effects = weighed model and 22 effects = reduced model.

The results in this table (Table 9) confirm the results for the previous table (Table 8) indicating that numerical problems have been solved and that there were other problems that caused the correlations to be < .90.

For the complete data set, 812,393 animals and 2,946,847 records, the convergence criterion was 10^{-12}. Estimates of correlations for additive direct and additive maternal effects based on the complete data set are presented in Table 10. Also, memory requirement and running time are reported for these analyses.

TABLE 10 – Estimated correlations between expected progeny differences with multiple trait and random regression models by BLUP90iod program based on the complete data set

Number of effects in RRM[1]	Traits						Memory requirement (M)	Running time (hours)
	Effects	W1	W2	W3	W5	W7		
31	Additive direct	.87	.89	.89	.87	.86	935 [2]	96
	Additive maternal	.85	.86	.88	.85	.84		
27	Additive direct	.87	.89	.89	.87	.86	486 [3]	89
	Additive maternal	.85	.86	.88	.85	.84		
22	Additive direct	.87	.89	.89	.87	.86	339 [4]	36
	Additive maternal	.85	.86	.88	.85	.84		

[1] 31 effects = complete model, 27 effects = weighed model and 22 effects = reduced model.

[2] Convergence was with 1891 rounds, 10^{-12}.

[3] Convergence was with 1297 rounds, 10^{-12}.

[4] Convergence was with 887 rounds, 10^{-12}.

The results presented in Table 10 show that estimated correlations based on expected progeny differences with MTM and RRM were of the same magnitude for both effects, additive direct and additive maternal, with different number of effects in RRM. Memory requirement and time running were reduced as well to 63.7% and 62.5%, respectively, when residual effects were replaced by a weight and covariates corresponding to very small eigenvalues were eliminated.

The definition of birth weight does not involve any preadjustment. Therefore, the correlations for BW between RR and MT should be very high.

However, these correlations were only .86 and .70 for additive and maternal effects, respectively (Table 8). Because no numerical problems were involved, these differences must be due to different parameters (covariances) used in both models. Analyzing parameters estimated for these models, presented in Chapter III, they show differences in the shapes of variances and correlations. Some of these differences are due to sampling errors and some due to artifacts of the random regressions models.

The importance of genetic parameters decreases when the amount of data per animal increases. Table 11 shows additive direct correlations for sires with at least 20 progenies.

TABLE 11 – Estimated correlations between expected progeny differences with multiple trait and random regression models for sires with as least 20 progenies

	Traits				
Effect	W1	W2	W3	W5	W7
Additive direct	.94	.94	.93	.92	.92

These correlations were all >.90. While some of these differences are due to not very accurate parameters, some of these may indicate the superiority of the RR because of its ability to handle more records, no need for preadjustment, and better accounting for changed variances.

A more complete comparison would involve parameters estimated initially, presented in Chapter III, but later refined for smoothness and reality of correlations; for instance the estimated additive maternal correlation below .70 seems excessively low. The new parameters will be made equivalent for both MT and RR models. Subsequent analysis will no longer test the numerical accuracies, or differences in parameters, but the real superiority of one model over the other.

Memory requirement and running time for the three different data sets of the analyses with MTM are presented in Table 12.

TABLE 12 – Memory requirement and running time for different data sets by BLUP90iod program

Data	Memory requirement (M)	Running time
BW sample [1]	3.3	one minute
BW, … , W8 sample[2]	13	two hours
BW, … , W8 complete data set[3]	478	five days

[1] BW sample = 21,055 animals, 14,572 records.

[2] BW, … ,WB sample = 21,055 animals, 71,867 records.

[3] complete data set = 812,393 animals, 2,946,847 records.

Analysis with a five multiple trait model took more memory and time than random regression model with the complete data set. In addition, analyses with repeated measurements by RRM offer the possibility of analyzing continuous changes.

Estimated genetic parameters by RRM can be inaccurate for various reasons; data set size, data selection, model and methodology applied. However, the parameters can be estimated more accurately after improvements in methodologies, making computations more reliable and less expensive (Misztal et al., 2000). In this study missing traits and estimates of direct maternal correlations for some ages equal to -1.0 were important factors in memory requirements and running time.

CONCLUSIONS

RRM could be a useful model in beef cattle because all records can be used, and EPDs can be estimated for any age. On the other hand, MTM is unable to use all records and requires preadjustments to constant ages, resulting in loss of accuracy relative to RRM.

EPDs for growth in beef cattle by RRM can be poor because of poor numerical properties of these models. These properties can be improved by orthogonalizing covariables for each random effect. Thereafter, a computing package using the preconditioned conjugate gradient (PCG) iteration can compute EPDs for national data sets in reasonable time. Even if the numerical properties of the RRM are adequate, EPDs may not be better than MTM if parameter estimates used in the RRM are poor. It seems that best parameters for the RRM may be obtained by conversion of parameters from MTM with additional ”smoothing”.

Due to the fact that weights at all ages can be accommodated by RRM, and EPDs can be obtained at any age of the animals, weights can be measured at larger intervals, resulting in a lower cost of recording.

LITERATURE CITED

ABCZ. 2001. Associação Brasileira dos Criadores de Zebu. CPD.– Controle de Desenvolvimento Ponderal. Available at: http://www.abcz.org.br. Acessed Jan 10, 01.

BIF. 1996. Guidelines for uniform beef improvement programs. Beef Improvement Federation.

CNPGC. 2001. MA/ABCZ/EMBRAPA. Sumário das raças zebuínas de corte – 2000. Available at: http://www.cnpgc.embrapa.br. Accessed Jan. 10, 01.

Jamrozik, J. and L.R. SCHAEFFER. 2000. Comparison of two computing algorithms for solving mixed model equations for multiple trait random regression test day models. Livest. Prod. Sci. 67:143-153.

Jamrozik, J., L.R. Schaeffer and J.C.M. Dekkers, 1997. Genetic evaluation of dairy cattle using test day yields and a random regression model. J. Dairy Sci. 80:1217-1226.

Kaps, M., W.O. Herring and W.R. Lamberson. 2000. Genetic and environmental parameters for traits derived from the Brody growth curve and their relationship with weaning weight in Angus cattle. J. Anim. Sci. 78:1436-1442.

Kirkpatrick, M., D. Lofsvold and M. Bulmer, 1990. Analysis of the inheritance, selection and evolution of growth trajectories. Genetics. 124:979-993.

Lidauer, M., I. Stradén, E.A. Mäntysaari, J. Pösö and A. Kettunen. 1999. Solving large test-day models by iteration on data and preconditioned conjugate gradient. J. Dairy Sci. 82:2788-2796.

Meyer, K. 1998. Estimating covariance functions for data using a random regression model. Genet. Sel. Evol. 30:221-240.

Meyer, K. 2000. Random regression to model phenotypic variation in monthly weights of Australian beef cows. Livest. Prod. Sci 65:19-38.

Misztal, I. 1999. Complex models, more data: simpler programming? In: Computational Cattle Breeding'99 Abstracts. Finland. Workshop: 18-20/03/1999.

Misztal, I. 2001. REML90 manual: Available at: ftp://nce.ads.uga.edu/pub/ignacy/blupf90/. Accessed Mar 22, 01.

Misztal, I. and Perez-Enciso. 1998. A Fortran 90 interface to sparse matrix package FSPAK with dynamic memory allocation and sparse matrix structure. Proc. 6th World Congress Genet. Appl. Livest. Prod. 22:77-78.

Misztal, I., T. Strabel, J. Jamrozik, E.A. Mäntysaari and T.H.E. Menwissen. 2000. Strategies for estimating the parameters needed for different test-day models. J. Dairy Sci. 83:1125-1134.

Pool, M.H. and T.H.E. MEUWISSEN. 2000. Reduction of the number of parameters needed for a polynomial random regression test day model. Livest. Prod. Sci. 64:133-145.

Schaeffer, L.R. and J.C.M. Dekkers 1994. Random regressions in animal models for test-day production in dairy cattle. In: Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph. 18:443-446.

Swalve, H.H. 1998. Use of test day records for genetic evaluation. In: Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale. 23:295-301.

Tsuruta, S. 2001. BLUP90IOD. In: MISZTAL, I. 2001. BLUP90IOD manual: Available at: ftp://nce.ads.uga.edu/pub/ignacy/blupf90/. Accessed Jun 22, 01.

Tsuruta, S., I. Misztal and I. Strandén. 2001. Use of preconditioned conjugate gradient algorithm as a generic solver for mixed-model equations in animal breeding applications. J. Anim. Sci. 79:116-1172.

VanRaden, P. M., G. R. Wiggans, and C.A. ERNST. 1991. Expansion of projected lactation yields to stabilise genetic variance. J. Dairy Sci. 74:4344-4349.

Varona, L., C. Moreno, L.A. Garcia Cortes and J. Altarriba. 1997. Multiple trait genetic analysis of underlying biological variables of production functions. Livest. Prod. Sci. 47:201-209.

CHAPTER VI – ABSTRACT AND CONCLUSIONS

Data collected by the Brazilian Zebu Breeders Association (ABCZ) and provided by the Brazilian Agricultural Research Corporation (EMBRAPA) were analyzed. The data consisted of records on 619,989 Nellore animals; progeny of 11,847 sires, and 273, 263 dams raised under Brazilian pasture conditions. The records were collected from 1975 to 1999. Traits considered were birth weight (BW), 10 to 110 days weight (W1 or weight at 60 days), 102 to 202 days weight (W2 or weight at 152 days), 193 to 293 days weight (W3 or weight at 243 days), 283 to 383 days weight (W4 or weight at 333 days), 376 to 476 days weight (W5 or weight at 426 days), 467 to 567 days weight (W6 or weight at 517 days), 551 to 651 days weight (W7 or weight at 601 days), and 633 to 733 days weight (W8 or weight at 683 days). Two methodologies, restricted maximum likelihood (REML) and a Bayesian method by Gibbs sampling (GS) were used. Also, two models, multiple trait model (MTM) and random regression model (RRM), were studied. Data sets with missing traits and with no missing traits were sampled. With REML, estimates of covariances by MTM were similar for both data sets, although those from the missing data set showed more variability from age to age. The estimates from RRM were similar to those from MTM only for the complete-trait case and showed large artifacts for the case of missing traits. Estimates of direct maternal correlations under RRM for some ages approached -1.0, and most likely contained artifacts. If many traits are missing, the best approach to obtaining parameters for RRM would be conversion from smoothed MTM estimates. The memory requirements of an optimized Bayesian analysis via Gibbs sampling was low enough for a nine or more traits multiple trait analysis. However, running time was very long. A chain size of 100,000 seemed sufficient for a multiple trait analysis of a sample without missing traits but insufficient for a sample with missing traits. A chain size of 50,000 seemed insufficient for a cubic random regression model. Due to the fact that covariance components based on RRM were inflated for herds with missing traits, estimates obtained with MTM should be used and converted to covariance functions. Expected progeny differences (EPD) for growth in beef cattle by random regression models can be poor because of poor numerical properties of these models. These properties can be improved by orthogonalizing covariables for each random effect. Thereafter, a generic procedure using the preconditioned conjugate gradient (PCG) iteration can compute EPDs for national data sets in reasonable time. Even if the numerical properties of the RRM are adequate, EPDs may not be better than in the MT model if parameters used in the RRM are poor. It seems that best parameters for the RR model may be obtained by conversion of parameters from MT model with additional ”smoothing”.