

Article

Not peer-reviewed version

---

# Structural Equation Modelling of Additive Genetic and Residual Covariance Matrices in Beef Cattle

---

[Marcos Jun-Iti Yokoo](#)\*, Gustavo de los Campos, [Vinicius Silva Junqueira](#), Fernando Flores Cardoso, [Guilherme Jordão Magalhães Rosa](#), Lucia Galvão Albuquerque

Posted Date: 22 December 2025

doi: 10.20944/preprints202512.1910.v1

Keywords: Bayesian method; factor analysis; genetic evaluation; mixed model; recursive relationships



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Structural Equation Modelling of Additive Genetic and Residual Covariance Matrices in Beef Cattle

Marcos Jun-Iti Yokoo<sup>1,\*</sup>, Gustavo de los Campos<sup>2</sup>, Vinícius Silva Junqueira<sup>3</sup>, Fernando Flores Cardoso<sup>4</sup>, Guilherme Jordão Magalhães Rosa<sup>5</sup> and Lucia Galvão Albuquerque<sup>6</sup>

<sup>1</sup> Embrapa Southeastern Livestock (CPPSE), Brazilian Agricultural Research Corporation (Embrapa), Rod. Washington Luis, km 234, São Carlos, SP, 13560-970, Brazil

<sup>2</sup> Department of Epidemiology & Biostatistics, Statistics & Probability, Institute for Quantitative Health Sciences and Engineering, Michigan State University

<sup>3</sup> Bayer Crop Science; Bayer, Uberlândia, MG, Brazil

<sup>4</sup> Embrapa South Livestock (CPPSUL), Brazilian Agricultural Research Corporation (Embrapa)

<sup>5</sup> Department of Animal and Dairy Sciences, University of Wisconsin

<sup>6</sup> Department of Animal Science, São Paulo State University (UNESP)

\* Correspondence: marcos.yokoo@embrapa.br

## Abstract

The continuous increase in the number of records collected and the amount of traits available for beef cattle genetic evaluations poses statistical and computational challenges when estimating the genetic and environmental covariance matrices needed to predict breeding values. Structural equation models (SEM) using either factor analysis (FA) or recursive models (REC) can be used to structure genetic and environmental covariance matrices and to obtain more parsimonious and efficient parameterizations. In this article, we use SEM to estimate parameters for growth and ultrasound carcass traits in beef cattle. Data consisted of 2,942 animals, and six traits were analyzed using standard multiple-trait mixed models with either unstructured covariance matrices (SMTM) or structured covariance matrices (SEM). For the latter, we considered FA and REC models implemented with six alternative parameterizations, in which random effects were represented as linear combinations of fewer unobservable random variables. Comparing with SMTM, all heritability estimates from 2-factor SEM for the additive genetic matrix (FA2G) and the model with six recursive effects zeroed out at the residual covariance matrix (REC1) were within one standard error of those obtained by SMTM. The correlations between estimated breeding values (EBV) for all traits and models ranged from 0.94 to 1.00. The most parsimonious model in terms of number of estimated parameters ( $p_D$ ) was FA2G, with 431.2  $p_D$ , and 25.3  $p_D$  fewer than the traditional model SMTM. The REC1 model showed as a good alternative for this kind of dataset, as it had a smaller  $p_D$  (443.6) than the SMTM model (456.5) and a better deviance information criterion than all other models tested (e.g., 37,868.6 for REC1 and 37,874.7 for SMTM). Results from this study indicate that mixed-effects multi-trait models in beef cattle can be successfully analyzed with FA or REC models. These models offer a parsimonious representation of the underlying covariance patterns and offer an interesting option for breeding value prediction.

**Keywords:** Bayesian method; factor analysis; genetic evaluation; mixed model; recursive relationships

## 1. Introduction

In animal breeding, standard multiple-trait mixed models (SMTM) are often applied to analyze multiple traits [1], estimating unstructured covariance matrices. The number of (co)variance parameters involved grows quadratically with the number of traits, and when phenotypes are highly correlated, the residual and genetic covariance matrices can be singular, making the convergence of estimation algorithms complex [1,2] and inferences unreliable.

Structural equation models (SEM) are often used to obtain more parsimonious (and sometimes more interpretable) parametrization of multivariate models. This approach has been used in quanti-

tative genetics to structure the genetic and environmental covariance matrices of multivariate linear mixed models [2–6], commonly used for genetic evaluations. Primarily due to its ability to consider causal relationships between latent variables, allowing for the modeling of complex phenomena while simultaneously reducing the dimensionality of the data [7–9]. SEM methodology comprises many different techniques and procedures used together to model covariance structures. These models can be viewed as an extension of the multivariate linear mixed model [10] that can express functional networks among traits [2]. Gianola and Sorensen [3] discussed the use of recursive (REC) and simultaneous equation models (both special cases of SEM) acting on phenotypes. Valente et al. [11] proposed a methodology to search for recursive causal structures among phenotypes within a multivariate linear mixed model, which has been applied to various species and traits, including carcass and meat quality traits in beef cattle [12]. Some authors discussed the use of REC models operating on genetic covariance matrices [6,8,13]. In certain cases, modeling the covariance matrices helps infer or analyze causality between the breeding values of different traits or examine the ratio of two correlated traits, thereby reducing the model's dimension with more parsimonious and efficient parameterizations [8,14].

Alternatively, factor analysis (FA), another specific case of SEM, can be used to reduce the dimension of the estimated genetic covariance matrix [15,16] and achieve a more parsimonious model of genetic effects within a multivariate linear mixed model, without decreasing the original number of traits and records. In the FA framework, the observed traits are modeled as linear combinations of fewer unobservable latent factors and trait-specific residual components, effectively capturing the common sources of variation underlying correlated traits while preserving individual trait identities. This approach has been successfully applied in quantitative genetics to structure genetic and environmental covariance matrices [16,17], proving particularly useful when analyzing highly correlated traits where traditional unstructured covariance matrices may suffer from convergence issues or poor estimability. The reduction in the number of parameters not only improves computational efficiency but can also lead to more stable parameter estimates and better model convergence, especially in datasets with limited sample sizes relative to the number of traits being analyzed [7]. Moreover, FA models can provide biological insights by identifying latent factors that may represent underlying biological processes or developmental pathways common to multiple phenotypes.

While previous studies have demonstrated the utility of SEM in various contexts, their application to structuring both genetic and residual covariance matrices independently in beef cattle breeding remains underexplored. In the current omics era, breeding programs are increasingly collecting larger numbers of correlated phenotypes, making computationally efficient, statistically robust methods for multi-trait genetic evaluation even more critical. These more parsimonious models offer practical advantages for routine genetic evaluations in breeding programs, enabling the simultaneous analysis of multiple traits across large populations with improved computational efficiency and numerical stability compared to fully parameterized models. Therefore, the objective of this study was to evaluate the performance of factor analysis and recursive models applied independently to additive genetic and residual covariance matrices for growth and ultrasound carcass traits in Nellore cattle, comparing them with standard multiple-trait mixed models in terms of parameter estimation, breeding value prediction, and model fit criteria.

## 2. Materials and Methods

### 2.1. Dataset Description

The present study used data from 2,942 Nellore animals collected over 3 years from an existing database of 10 herds participating in the Brazilian Nellore Breeding Program across 6 States. On-farm data collection was conducted according to routine management procedures and in accordance with the [18] guidelines. To collect carcass data without slaughtering the animals, the ultrasound methodology was applied [19]. Real-time ultrasound measurements of carcass traits included the longissimus muscle area (LMA, cm<sup>2</sup>) and backfat thickness (BF, mm), both obtained from a cross-

sectional image of the longissimus dorsi muscle between the 12th and 13th ribs. Rump fat thickness (RF, mm) was measured at the intersection of the gluteus medius and biceps femoris muscles, located between the hooks and pin bones. Backfat thickness was measured at the 3/4 position from the chine bone end of the longissimus muscle, using the cross-sectional ribeye image. Other traits recorded included body weight (BW, kg) and hip height (HH, cm), both measured at the date of ultrasound scanning, and 450-day age-standardized scrotal circumference (SC, mm). Age-standardized SC was obtained as described in the Guidelines for Uniform Beef Improvement Programs [20]. The traits were assessed in animals aged between 480 and 629 days.

Contemporary groups (CG) were defined as animals of the same sex, born in the same herd, year, and season (dry or rainy), and reared within the same management group. A total of 302 contemporary groups were formed, with an average of 9.7 animals per group (range: 3 to 45 animals). Additional effects considered in the analyses included age of animal at scanning (linear covariate for all traits except SC) and age of dam at calving (linear and quadratic covariates for BF, RF, HH, and BW; classes for LMA and SC). A description of the traits and data structure for Brazilian Nelore cattle is presented in Table 1.

**Table 1.** Descriptive statistics of ultrasound carcass and growth traits in Nelore cattle

Traits	No. of records	Mean $\pm$ SD	No. of sires	No. of dams	No. CG
LMA	2,770	48.05 $\pm$ 8.36	231	2,552	243
BF	2,577	1.87 $\pm$ 1.07	226	2,397	253
RF	2,566	2.95 $\pm$ 1.94	226	2,384	252
SC	1,340	245.87 $\pm$ 30.22	106	1,009	88
HH	2,349	136.06 $\pm$ 5.04	226	2,308	250
BW	2,942	339.69 $\pm$ 65.98	236	2,683	302

LMA = longissimus muscle area (cm<sup>2</sup>); BF = backfat thickness (mm); RF = rump fat thickness (mm); SC = standardized scrotal circumference (mm) at 450 days of age; BW and HH = body weight (kg) and hip height (cm) obtained at the time of scanning, respectively. SD = standard deviation; No. CG = number of contemporary groups.

## 2.2. Standard Multi-Trait Mixed Model

Genetic parameters and model solutions were estimated using Bayesian inference implemented via Gibbs sampling. In the following sections, we first describe the standard multi-trait mixed model commonly used in genetic evaluation, then present extensions based on factor analytic (FA) and recursive (REC) model structures.

In the standard multi-trait mixed model (SMTM), a sire model was fitted for  $p$  traits measured on  $n$  individuals. The mixed-model equation can be described as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon}, \quad (1)$$

where  $\mathbf{y}$  is the vector of phenotypic observations,  $\boldsymbol{\beta}$  is the vector of fixed effects,  $\mathbf{u}$  is the vector of random additive sire effects,  $\boldsymbol{\epsilon}$  is the vector of residual effects, and  $\mathbf{X}$  and  $\mathbf{Z}$  are the corresponding incidence matrices. Under the Bayesian framework, the likelihood function was specified as:

$$\mathbf{y} | \boldsymbol{\beta}, \mathbf{u}, \mathbf{G}_0, \mathbf{R}_0 \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}, \mathbf{I}_n \otimes \mathbf{R}_0),$$

with prior distributions  $\mathbf{u} | \mathbf{A}, \mathbf{G}_0 \sim N(\mathbf{0}, \mathbf{A} \otimes \mathbf{G}_0)$  and  $\boldsymbol{\epsilon} | \mathbf{R}_0 \sim N(\mathbf{0}, \mathbf{I}_n \otimes \mathbf{R}_0)$ , where  $\mathbf{G}_0$  is the  $p \times p$  additive genetic (co)variance matrix among sire effects,  $\mathbf{R}_0$  is the  $p \times p$  residual (co)variance matrix,  $\mathbf{A}$  is Wright's numerator relationship matrix among sires,  $\mathbf{I}_n$  is an identity matrix of order  $n$ , and  $\otimes$  denotes the Kronecker product operator.

Prior distributions for model parameters were specified as follows. Systematic effects included contemporary groups, animal age at scanning (linear effect, except for SC), and dam age (linear and quadratic effects for BF, RF, HH, and BW). For the vector of systematic effects  $\boldsymbol{\beta}$  with vague information [21], prior distributions were assigned as:

$$\boldsymbol{\beta} \sim N(\mathbf{0}, \mathbf{I} \otimes 100^2),$$

For the (co)variance matrices, scaled inverse Wishart distributions were assigned as prior distributions:

$$\mathbf{G}_0 \sim IW(\mathbf{S}_g, \nu_g) \quad \text{and} \quad \mathbf{R}_0 \sim IW(\mathbf{S}_e, \nu_e),$$

where  $\mathbf{S}_g$  and  $\mathbf{S}_e$  are  $p \times p$  scale matrices and  $\nu_g$  and  $\nu_e$  are degrees of freedom parameters. The joint posterior distribution of all model parameters is proportional to the product of the likelihood and prior distributions:

$$p(\boldsymbol{\beta}, \mathbf{u}, \mathbf{G}_0, \mathbf{R}_0 | \mathbf{y}) \propto p(\mathbf{y} | \boldsymbol{\beta}, \mathbf{u}, \mathbf{G}_0, \mathbf{R}_0) \times p(\mathbf{u} | \mathbf{A}, \mathbf{G}_0) \times p(\boldsymbol{\beta}) \times p(\mathbf{G}_0) \times p(\mathbf{R}_0).$$

### 2.3. Factor Analytic Model

Factor analysis is a multivariate statistical technique that simplifies complex and correlated data structures [22,23]. It can be viewed as an extension of principal component analysis, where the factor-analytic variance-covariance structure serves as a parsimonious approximation to a fully unstructured variance-covariance matrix, enabling more efficient parameterization of complex correlation patterns. In a standard FA, a vector of random variables (say  $\mathbf{u}$ ) can be decomposed as a linear combination of fewer unobservable random variables called common factors ( $\mathbf{f}$ ), with an unobservable incidence matrix ( $\boldsymbol{\Lambda}$ ) of factor loadings, plus a vector of trait-specific errors ( $\boldsymbol{\delta}$ ) to each individual  $i$  that represent the lack of fit of the model. In compact notation:

$$\mathbf{u}_i = \boldsymbol{\Lambda}_u \mathbf{f}_i + \delta_{u,i}, \quad (2)$$

In vector notation, the factor analytic model for these effects can be written as:

$$\mathbf{u} = (\mathbf{I}_n \otimes \boldsymbol{\Lambda}) \mathbf{f} + \boldsymbol{\delta}, \quad (3)$$

where  $\mathbf{u} = (\mathbf{u}'_1, \dots, \mathbf{u}'_n)'$ ,  $\mathbf{f} = (\mathbf{f}'_1, \dots, \mathbf{f}'_n)'$ , and  $\boldsymbol{\delta} = (\boldsymbol{\delta}'_1, \dots, \boldsymbol{\delta}'_n)'$ . This results in the following covariance matrix of  $\mathbf{u}$ :

$$\text{Cov}(\mathbf{u}_i) = \mathbf{G}_0 = \boldsymbol{\Lambda}_u \boldsymbol{\Lambda}'_u + \boldsymbol{\Psi}_u,$$

where  $\boldsymbol{\Psi}_u = \text{Diag}\{\psi_{u,1}, \dots, \psi_{u,p}\}$  is a diagonal matrix of specific variances. Likewise, the residual vector ( $\boldsymbol{\epsilon}$ ) can be decomposed into common and trait-specific factors. Consequently, the marginal distributions of  $\mathbf{u}_i$  and  $\epsilon_i$  are:

$$\mathbf{u}_i \stackrel{\text{iid}}{\sim} N(\mathbf{0}, \boldsymbol{\Lambda}_u \boldsymbol{\Lambda}'_u + \boldsymbol{\Psi}_u) \quad \text{and} \quad \epsilon_i \stackrel{\text{iid}}{\sim} N(\mathbf{0}, \boldsymbol{\Lambda}_\epsilon \boldsymbol{\Lambda}'_\epsilon + \boldsymbol{\Psi}_\epsilon),$$

where “iid” stands for “independent and identically distributed”.

To model the additive genetic and residual (co)variance matrices using FA, it was assumed that Equation 2 holds for the vector of random additive genetic effects ( $\mathbf{u}$ ) in Equation 1 and, likewise, for the vector of random residual effects ( $\boldsymbol{\epsilon}$ ) such that

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}[(\mathbf{I}_u \otimes \boldsymbol{\Lambda}_u)\mathbf{f}_u] + \mathbf{Z}\boldsymbol{\delta}_u + [(\mathbf{I}_\epsilon \otimes \boldsymbol{\Lambda}_\epsilon)\mathbf{f}_\epsilon] + \boldsymbol{\delta}_\epsilon,$$

where  $\boldsymbol{\Lambda}$ ,  $\mathbf{X}$ ,  $\mathbf{Z}$  and  $\boldsymbol{\beta}$  are as before and  $\mathbf{f}$  and  $\boldsymbol{\delta}$  are interpreted as vectors of common and specific additive genetic effects, respectively, similarly for the random residual effects ( $\boldsymbol{\epsilon}$ ). Combining the assumptions of the FA model described above with those of SMTM leads to the following random effects joint distribution,

$$p(\boldsymbol{\epsilon}, \mathbf{u}) = N[\boldsymbol{\epsilon} | \mathbf{0}, \mathbf{I}_\epsilon \otimes (\boldsymbol{\Lambda}_\epsilon \boldsymbol{\Lambda}'_\epsilon + \boldsymbol{\Psi}_\epsilon)] N[\mathbf{u} | \mathbf{0}, \mathbf{A} \otimes (\boldsymbol{\Lambda}_u \boldsymbol{\Lambda}'_u + \boldsymbol{\Psi}_u)],$$

where  $\Lambda_u$  and  $\Lambda_\epsilon$  are the matrices of additive genetic and residual factor loads, respectively,  $\Psi_u$  and  $\Psi_\epsilon$  are the diagonal matrices of specific additive genetic and residual variances, respectively. More details on the implementation of this model can be found in [17].

#### 2.4. Recursive Models

Recursive models are a category of SEMs that postulate causal and unidirectional relationships between latent variables, unlike simultaneous models, which admit mutual feedback effects between traits. For more information on the procedure for REC effects decomposition, see [17,21]. Using REC, the vector  $\mathbf{u}$  can be decomposed as

$$\mathbf{u}_i = \Pi_u \mathbf{u}_i + \boldsymbol{\epsilon}_{u,i},$$

where  $\mathbf{u}_i$  is the random effect of the  $i$ th sire,  $\Pi_u = \{\pi_{u,ij}\}$  is a strictly lower-triangular matrix; i.e.,  $\pi_{u,ij} = 0$  for all  $i \leq j$ , whose non-zero entries define recursive effects, and  $\boldsymbol{\epsilon}_{u,i}$  is a vector of random effects whose (co)variance matrix is:  $\text{Cov}(\boldsymbol{\epsilon}_u) = \mathbf{A} \otimes \Gamma_u$  where  $\Gamma_u = \text{Cov}(\boldsymbol{\epsilon}_{u,i})$  is a diagonal matrix. The reduced form model is  $\mathbf{u}_i = (\mathbf{I} - \Pi_u)^{-1} \boldsymbol{\epsilon}_{u,i}$  (Equation 4).

Similarly to the FA model, to implement the REC to model the additive genetic and the residual (co)variance matrices, it was assumed that Equation 4 holds the vector of random additive genetic effects ( $\mathbf{u}$ ) in Equation 1, likewise for the vector of random residual effects ( $\boldsymbol{\epsilon}$ ). Therefore, the additive genetic covariance matrix is equal to  $\mathbf{G}_0$ , as  $\text{Cov}(\mathbf{u}_i) = \mathbf{G}_0 = (\mathbf{I} - \Pi_u)^{-1} \Gamma_u (\mathbf{I} - \Pi_u)^{-1'}$ . Similarly, using REC for residuals,  $\mathbf{R}_0$  can be represented as  $\text{Cov}(\boldsymbol{\epsilon}_i) = \mathbf{R}_0 = (\mathbf{I} - \Pi_\epsilon)^{-1} \Gamma_\epsilon (\mathbf{I} - \Pi_\epsilon)^{-1'}$ , where  $\Pi_\epsilon$  is a strictly lower-triangular matrix defining recursive effects between model residuals, and  $\Gamma_\epsilon$  is diagonal. A fully recursive model (FRM) occurs when all lower-triangular entries of  $\Pi_u$  and of  $\Pi_\epsilon$  are non-zero, i.e., all  $\pi_{u,ij}$  and  $\pi_{\epsilon,ij}$  with  $i > j$  are parameters to be estimated. This model has as many (co)dispersion parameters as the SMTM, and there is a one to one map from the unknowns in  $\{\Pi_u, \Gamma_u, \Pi_\epsilon, \Gamma_\epsilon\}$  to those in  $\{\mathbf{G}_0, \mathbf{R}_0\}$ . Therefore, the FRM is just a re-parameterization of SMTM; however, various degrees of parsimony can be obtained by setting some of the lower-triangular entries of  $\Pi_u$  and  $\Pi_\epsilon$  to zero. Models in this study differed in how  $\Pi_u$  and  $\Pi_\epsilon$  were structured.

#### 2.5. Statistical Analyses

**Bayesian inference and MCMC implementation:** All models were implemented in a Bayesian framework using Gibbs sampling via the MTM package [24] in R software [25]. An initial chain of 1,500,000 iterations was run to assess convergence. The final chain length, burn-in period, and thinning interval were determined using the criteria of [26], implemented in the Bayesian Output Analysis (BOA) package [27]. Convergence was further evaluated through visual inspection of trace plots for variance components. Based on these diagnostics, posterior inference was based on 146,000 samples obtained after discarding the first 40,000 iterations as burn-in and retaining every 10th iteration thereafter.

**Model comparison:** Multiple alternative models were fitted to assess the effectiveness of parsimonious covariance structures. The standard multi-trait mixed model (SMTM) with unstructured (co)variance matrices served as the baseline, requiring estimation of 42 (co)variance parameters (21 for  $\mathbf{G}_0$  and 21 for  $\mathbf{R}_0$ ).

**Factor analytic models:** Four FA models were evaluated. The FA2F model specified two common factors for both additive genetic and residual (co)variance matrices, reducing the total number of parameters to 24 (12 genetic + 12 residual). The FA3F model extended this to three factors, increasing the parameter count to 36 (18 genetic + 18 residual). To evaluate the benefit of parsimony in only one (co)variance component, two additional models were fitted: FA2G applied two factors to  $\mathbf{G}_0$  while maintaining the unstructured  $\mathbf{R}_0$  (33 parameters total), whereas FA2R applied two factors to  $\mathbf{R}_0$  while maintaining the unstructured  $\mathbf{G}_0$  (33 parameters total).

**Recursive models:** Three recursive models were evaluated. First, a fully recursive model (FRM) was fitted, which is equivalent to SMTM in terms of number of parameters but reparameterizes

the covariance structures using recursive paths. Based on FRM results, two reduced models were constructed. The REC1 model set the  $ij$ th entry of  $\mathbf{\Pi}_e$  to zero if  $|\hat{\pi}_{ij}|/sd_{ij} < 1.96$  or  $i \leq j$ , where  $\hat{\pi}_{ij}$  and  $sd_{ij}$  denote the posterior mean and standard deviation of the  $ij$ th entry of  $\mathbf{\Pi}_e$  from FRM, respectively. This criterion resulted in six recursive effects being removed from  $\mathbf{\Pi}_e$ . The REC2 model removed any effect in  $\mathbf{\Pi}_u$  or  $\mathbf{\Pi}_e$  whose posterior mean from FRM had absolute value less than 0.15 (an arbitrary threshold chosen to evaluate a more parsimonious structure), resulting in removal of 15 recursive effects (10 from  $\mathbf{\Pi}_e$  and 5 from  $\mathbf{\Pi}_u$ ).

**Model selection criteria:** Models were compared using the deviance information criterion [DIC; 28,29], the posterior mean of the log-likelihood [Mean( $\mathcal{L}$ )], and the effective number of parameters [ $p_D$ ; 28,29]. Lower DIC values indicate better model fit penalized for complexity, while  $p_D$  quantifies model complexity. Additionally, Spearman and Pearson correlations between predicted sire breeding values across models were calculated to assess ranking stability.

### 3. Results

#### 3.1. Genetic Parameters

Tables 2, 3, and 4 present estimates of heritabilities (diagonal), genetic (above diagonal), and residual (below diagonal) correlations, obtained from the models SMTM, FA2G, and REC1, respectively. Standard errors obtained by these three models ranged from 0.05 to 0.12 for estimates of heritabilities. The Spearman and Pearson correlations of the estimated breeding values of all traits, comparing the same trait across models, ranged from 0.94 to 1.00.

**Table 2.** Estimates of heritability (diagonal), genetic (above diagonal) and residual (below diagonal) correlations, with standard errors, from the standard multi-trait mixed model with sire model (SMTM)

Traits	LMA	BF	RF	BW	HH	SC
LMA	0.29 ± 0.06	0.13 ± 0.15	0.05 ± 0.16	0.35 ± 0.13	-0.14 ± 0.15	0.19 ± 0.17
BF	0.14 ± 0.02	0.47 ± 0.10	0.62 ± 0.10	0.03 ± 0.17	-0.52 ± 0.12	0.09 ± 0.22
RF	0.11 ± 0.02	0.57 ± 0.01	0.35 ± 0.08	0.07 ± 0.17	-0.42 ± 0.14	0.05 ± 0.19
BW	0.49 ± 0.02	0.22 ± 0.02	0.15 ± 0.02	0.28 ± 0.06	0.24 ± 0.15	0.09 ± 0.17
HH	0.11 ± 0.02	-0.01 ± 0.02	-0.03 ± 0.02	0.42 ± 0.02	0.38 ± 0.08	-0.17 ± 0.21
SC	0.21 ± 0.04	-0.02 ± 0.07	0.02 ± 0.07	0.39 ± 0.04	0.20 ± 0.05	0.40 ± 0.10

LMA = longissimus muscle area (cm<sup>2</sup>); BF = backfat thickness (mm); RF = rump fat thickness (mm); SC = standardized scrotal circumference (mm) at 450 days; BW = body weight (kg); HH = hip height (cm).

For the genetic correlations, the standard errors obtained by these three models presented in Tables 2, 3, and 4 ranged from 0.07 to 0.23. Considering the overlap within  $\pm 1.0$  standard errors, differences between genetic correlation estimates obtained by SMTM (Table 2) and the REC1 (Table 4) model were not important. Differences between the genetic correlations estimates obtained by SMTM (Table 2) and FA2G (Table 3) models were observed for LMA and BW, BF and RF, BF and HH, RF and HH, and BW and HH.

**Table 3.** Estimates of heritability (diagonal), genetic (above diagonal) and residual (below diagonal) correlations, with standard errors, from model FA2G (model with two factors only for the matrix of additive genetic loading factors and in the residual matrix it was considered as in the standard multi-trait mixed models)

Traits	LMA	BF	RF	BW	HH	SC
LMA	0.25 ± 0.05	0.10 ± 0.12	0.07 ± 0.10	0.03 ± 0.08	-0.08 ± 0.11	0.03 ± 0.08
BF	0.14 ± 0.02	0.31 ± 0.08	0.31 ± 0.13	-0.01 ± 0.14	-0.33 ± 0.12	0.08 ± 0.15
RF	0.11 ± 0.02	0.58 ± 0.01	0.23 ± 0.06	0.00 ± 0.11	-0.25 ± 0.12	0.05 ± 0.12
BW	0.49 ± 0.02	0.22 ± 0.02	0.15 ± 0.02	0.22 ± 0.05	0.03 ± 0.12	0.01 ± 0.07
HH	0.11 ± 0.02	-0.01 ± 0.02	-0.03 ± 0.02	0.42 ± 0.02	0.30 ± 0.08	-0.07 ± 0.13
SC	0.21 ± 0.04	-0.02 ± 0.07	0.01 ± 0.07	0.39 ± 0.04	0.20 ± 0.05	0.40 ± 0.11

LMA = longissimus muscle area (cm<sup>2</sup>); BF = backfat thickness (mm); RF = rump fat thickness (mm); SC = standardized scrotal circumference (mm) at 450 days; BW = body weight (kg); HH = hip height (cm).

For the residual correlations, the standard errors obtained by these three models presented in Tables 2, 3, and 4 ranged from 0.01 to 0.07. Considering these standard error ranges, differences in residual correlation estimates obtained with the SMTM (Table 2) and REC1 (Table 4) models were observed only for SC and BF. The differences between the residual correlation estimates obtained by the SMTM (Table 2) and FA2G (Table 3) models were not significant.

**Table 4.** Estimates of heritability (diagonal), genetic (above diagonal) and residual (below diagonal) correlations, with standard errors, from model REC1 (six recursive effects zeroed-out at residual covariance matrix)

Traits	LMA	BF	RF	BW	HH	SC
LMA	0.25 ± 0.05	0.20 ± 0.19	0.09 ± 0.19	0.43 ± 0.14	-0.16 ± 0.17	0.25 ± 0.20
BF	0.14 ± 0.02	0.45 ± 0.09	0.65 ± 0.10	0.05 ± 0.18	-0.49 ± 0.13	0.05 ± 0.23
RF	0.08 ± 0.01	0.57 ± 0.01	0.35 ± 0.08	0.12 ± 0.17	-0.40 ± 0.15	0.04 ± 0.19
BW	0.49 ± 0.02	0.22 ± 0.02	0.13 ± 0.01	0.31 ± 0.07	0.38 ± 0.15	0.12 ± 0.19
HH	0.11 ± 0.02	0.00 ± 0.02	-0.03 ± 0.02	0.42 ± 0.02	0.44 ± 0.09	-0.19 ± 0.23
SC	0.20 ± 0.02	0.09 ± 0.01	0.05 ± 0.01	0.40 ± 0.04	0.17 ± 0.02	0.49 ± 0.12

LMA = longissimus muscle area (cm<sup>2</sup>); BF = backfat thickness (mm); RF = rump fat thickness (mm); SC = standardized scrotal circumference (mm) at 450 days; BW = body weight (kg); HH = hip height (cm).

### 3.2. Recursive Effects

Estimates of genetic (above diagonal) and residual (below diagonal) recursive effects from the FRM in Nellore cattle are shown in Table 5, assuming that the causation flows from the left (trait of the first column) to the right. In other words, a consequent variable never exerts causal influence (either directly or indirectly) on an antecedent variable that first exerts causal influence on it.

**Table 5.** Estimates of genetic (above diagonal) and residual (below diagonal) recursive effects, with standard errors, from the fully-recursive model (FRM) in Nellore cattle

Traits	LMA	BF	RF	BW	HH	SC
LMA	—	0.40 ± 0.40	-0.13 ± 0.27	0.48 ± 0.18	-0.54 ± 0.27	0.23 ± 0.48
BF	0.21 ± 0.03	—	0.56 ± 0.12	-0.08 ± 0.12	-0.23 ± 0.16	-0.09 ± 0.32
RF	0.04 ± 0.03	0.55 ± 0.02	—	0.10 ± 0.14	-0.24 ± 0.19	-0.06 ± 0.30
BW	0.45 ± 0.02	0.09 ± 0.02	0.01 ± 0.02	—	0.83 ± 0.25	0.34 ± 0.49
HH	-0.14 ± 0.03	-0.05 ± 0.02	-0.04 ± 0.02	0.60 ± 0.03	—	-0.41 ± 0.40
SC	0.03 ± 0.06	-0.10 ± 0.06	0.02 ± 0.05	0.48 ± 0.06	0.03 ± 0.06	—

LMA = longissimus muscle area (cm<sup>2</sup>); BF = backfat thickness (mm); RF = rump fat thickness (mm); SC = standardized scrotal circumference (mm) at 450 days; BW = body weight (kg); HH = hip height (cm).

### 3.3. Model comparison

The deviance information criterion (DIC), the posterior mean of the log-likelihood (Mean(L)), and the estimated numbers of effective parameters ( $p_D$ ) used to compare the proposed models are shown

in Table 6. Although the SMTM model is considered fully parameterized for the additive genetic and residual matrices, the  $p_D$  were largest in models that have many (co)dispersion parameters, as in the FA3F and FA2R, compared to the SMTM model; thus, the model with the smallest  $p_D$  was FA2G.

**Table 6.** Model comparison criteria to assess performance of different Bayesian models

Models	DIC	$p_D$	Mean(L)
SMTM	37,874.7	456.5	−18,709.2
FA2F	38,036.8	455.2	−18,790.8
FA3F	38,915.2	564.1	−19,175.6
FA2G	37,898.3	<b>431.2</b>	−18,733.5
FA2R	38,043.5	499.7	−18,771.9
FRM	37,871.8	447.5	−18,712.1
REC1	<b>37,868.6</b>	443.6	−18,712.5
REC2	37,937.4	440.2	−18,748.6

SMTM = standard multi-trait mixed model; FA2F = model with two factors for additive genetic and residual loading matrices; FA3F = model with three factors for additive genetic and residual loading matrices; FA2G = model with two factors for additive genetic loading matrix only (residual as in SMTM); FA2R = model with two factors for residual loading matrix only (additive genetic as in SMTM); FRM = fully-recursive model; REC1 = model with six recursive effects zeroed-out at residual covariance matrix; REC2 = model with ten recursive effects zeroed-out at residual covariance matrix and five at additive genetic covariance matrix. DIC = Deviance Information Criterion;  $p_D$  = estimated number of effective parameters; Mean(L) = posterior mean of log-likelihood.

The models with the best (smallest) DIC and Mean(L) were REC1 and SMTM, respectively. Conversely, FA3F was the worst model for these three comparison criteria (DIC, Mean(L), and  $p_D$ ).

#### 4. Discussion

This study aimed to evaluate the potential advantages of using structured equation models, such as recursive and factor-analytic models, for sire modeling in beef cattle genetic evaluations of ultrasound carcass and growth traits.

Given the magnitude of standard errors, differences between heritability estimates obtained by SMTM and the two best models (REC1 and FA2G) were only observed for BF and RF traits in the FA2G model, which yielded slightly lower estimates, except for SC in the REC1 model, but with the same pattern and within one standard deviation. The correlations (Spearman and Pearson) between estimated breeding values for each trait, across models were high and close to the unit, showing that selection decision would be basically the same using any of the methodologies considered here, despite the lower genetic variability (heritability estimates) of some traits in some models, because when estimating genetic breeding values the variances are generally fixed.

The FA2G model estimated genetic correlations with the same pattern as the SMTM model, but with a lower magnitude. It is suggested that only when genetic correlations are of lower magnitude can FA models properly estimate the true covariance structure; therefore, FA models are not a good alternative for estimating high correlations, as discussed by [30]. On the other hand, most of the genetic correlations estimated with the REC1 model were of slightly larger magnitude than those estimated by the SMTM, but yet very similar, indicating a good alternative for modeling additive genetic covariance matrices in beef cattle.

As expected, REC1 and FA2G modeled residual covariance matrices of beef cattle data very similarly to the SMTM model, except for the association between SC and BF in the REC1 model. Nevertheless, the residual correlation in the REC1 model was basically null ( $0.09 \pm 0.01$ ). Considering the reasonable agreement between covariance parameters and breeding values estimated by the REC1 and FA2G models with the traditional model (SMTM), it could be suggested that these structural equations can be successfully used to reduce dimensionality of genetic evaluations, optimizing processing by a single analysis including all traits of interest and, consequently, providing faster and reliable estimation of breeding values compared to the traditional univariate analyses for each selection criteria. Thus,

beef cattle breeding programs could run multiple-trait mixed-model analyses of large datasets nearly every day, such that as soon as the farmers send in new field data, they can get prompt results for selection decisions on their herds.

Recursive models act separately on  $\mathbf{u}$  and  $\epsilon$ , allowing specific patterns to be captured in each component (Table 3). The different patterns of genetic and residual correlations can be interpreted using estimates of recursive effects, and based on this information, irrelevant effects can be suppressed from the covariance matrices. The six recursive effects zeroed out in  $\Pi_\epsilon$  for the REC1 model were between RF and LMA, SC and LMA, BW and RF, HH and RF, SC and RF, and SC and HH. Therefore, the REC1 model saves these six recursive effects, which were virtually null, out of the total 21 parameters to be estimated, suggesting that this type of structural equation model can facilitate the estimation of breeding values for multiple traits in a single analysis.

In terms of the  $p_D$  criterion, all recursive models (FRM, REC1, and REC2) had smaller  $p_D$  values than the SMTM model; even for FRM, which has the same dimension as the SMTM, some benefits were observed when using this kind of structural equation model. For the FA models,  $p_D$  favored the FA2G (two factors) model over all other FA and REC models. However, when we tested a model with three factors (FA3F), the  $p_D$  was the largest. This was expected due to the number of parameters ( $p$ ) of  $\text{Cov}(\mathbf{u}) = \mathbf{G}_0$  in the FA, which can be calculated as  $p = q + mq - m(m - 1)/2$  parameters, where  $q \times m$  are the dimensions of the  $\Lambda$  matrix of factor loads (Equation 2). FA3F has 21  $p$  in each matrix (additive genetic and residual), while FA2F has 17  $p$ . In the SMTM, the  $p$  of  $\mathbf{G}_0$  can be calculated as  $p = q(q + 1)/2$  parameters, where  $q \times q$  are the dimensions of the  $\mathbf{G}_0$  matrix, thus yielding 21  $p$  in each matrix (additive genetic and residual) in the SMTM model. Depending on the data set, i.e., the number of records, a greater number of traits, and the most correlated traits, the FA models may be used as a special case of SEM to reduce the covariance matrices' rank in the model and provide a more parsimonious estimation of genetic parameters compared to SMTM. In this paper, the FA3F model had a poorer fit to the data than the other models because this dataset has only six traits. Nevertheless, if the dataset had more traits, a model with more factors would probably work better at reducing the  $p_D$ . Generally, when more factors are included in the model, it better explains the relationships between traits; however, with fewer traits, fewer factors can be precisely estimated. Because we have only six traits and the number of records of SC used is limited, just because it is collected only in males ( $n = 1,340$ ), this could at least partially explain the worst performance in terms of DIC,  $p_D$ , and Mean(L) of the FA3F model. As discussed by [31] and [30], the FA model usually reduces the rank of covariance matrices. Still, the  $\Psi$  matrix has to be close to zero  $\text{Cov}(\mathbf{u}) = \mathbf{G}_0 = \Lambda\Lambda' + \Psi$ , when specific effects are assumed absent. This results in a mixed-model formulation with covariance matrices that are not of full rank.

In the same sense, the Mean(L) and DIC demonstrated considerably better fits to the data for the models SMTM and REC1 than for all the other FA and REC models. Since DIC also accounts for dimensionality through a penalty ( $p_D$ ), the REC1 model might be an alternative to the SMTM in beef cattle genetic evaluation schemes that involve multiple-trait estimates.

In general, even though FA2G was effective in terms of  $p_D$ , comparing all the criteria and parameters, the REC1 model showed a good alternative for this kind of dataset (e.g., six traits) because it had a smaller  $p_D$  compared to the SMTM model, better DIC, and a reasonable Mean(L) compared with all other models tested. This model, when applied to the residual effect, can be used to obtain more parsimonious representations of the residual covariance matrix. This reduces the number of parameters to be estimated, and may yield useful interpretations of the underlying patterns of (co)variability. Exploration of genetic and residual relationships between traits can be detected using the REC1 model, helping to understand how selecting certain criteria can affect another trait. Moreover, unlike recursive models acting on phenotypes, these models do not impose cross-restrictions between the genetic and residual covariance matrices. Results from this study indicate that beef cattle growth and ultrasound carcass traits can be successfully analyzed using recursive models to estimate breeding

values and identify a possible causal effect between latent variables, allowing for the modeling of complex phenomena while simultaneously reducing the dimensionality of the data.

## 5. Conclusions

Results from this study indicate that beef cattle growth and ultrasound carcass traits can be successfully analyzed using factor analysis or recursive models to estimate breeding values for multiple traits in a single analysis. Depending on the size of the data set, namely the number of records and traits, the factor analyses and recursive models applied independently to genetic and residual effects can reduce covariance matrices' ranks and provide a more parsimonious estimation of genetic dispersion parameters compared to standard multiple-trait mixed models, especially if the covariances between the traits are of low magnitude. Exploration of genetic and residual relationships between traits via recursive models provides information to better understand how selection based on specific criteria can affect another trait through a causal effect.

**Author Contributions:** MJY, GC, LGA, and GJMR designed the project. MJY and GC carried out the computational tasks and prepared the first draft of the manuscript. GJMR, FFC, VSJ, and LGA provided insights and reviewed the manuscript. All authors read and approved the final manuscript.

**Funding:** This work was supported by the National Council for Scientific and Technological Development of Brazil (CNPq-Brasil) [grant number 482277/2012-2] and São Paulo Research Foundation (FAPESP) [grant number 2009/11977-0].

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data used in this study can be made available upon request to the authors.

**Acknowledgments:** The authors would like to thank The Nelore Breeding Program - Nelore Brazil from The National Association of Breeders and Researchers of Ribeirão Preto, Brazil for providing the phenotypes and the pedigree used in this study. The authors acknowledge the Bacuri, Guaporé Pecuária, Grupo Hofig Ramos, Rancho da Matinha, Colonial, São Dimas, Remanso, Passa Quatro, and Santa Marta farms for providing the animals and data to conduct this research. "During the preparation of this manuscript/study, the authors used phenotypes for the purposes of genetic evaluations. The authors have reviewed and edited the output and take full responsibility for the content of this publication."

**Conflicts of Interest:** The authors declare that they have no competing interests.

## Abbreviations

The following abbreviations are used in this manuscript:

BF	Backfat thickness
BOA	Bayesian Output Analysis
BW	Body weight
CG	Contemporary group
DIC	Deviance information criterion
EBV	Estimated breeding values
FA	Factor analytic model
FA3F	Model with three factors for genetic and residual matrices
FA2G	Model with two factors for the genetic matrix
FA2R	Model with two factors for residual matrix
FRM	Fully recursive model
HH	Hip height

LMA	Longissimus muscle area
MCMC	Markov Chain Monte Carlo
Mean(L)	Posterior mean of log-likelihood
MTM	Multivariate Gaussian Mixed Effects Model
$p_D$	Number of estimated parameters
REC	Recursive model
REC1	Model with six recursive effects zeroed out at the residual matrix
REC2	Model with ten recursive effects zeroed-out at residual and five at genetic matrices
RF	Rump fat thickness
SC	Scrotal circumference
SEM	Structural equation models
SMTM	Standard multi-trait mixed model

## References

1. Meyer, K. Genetic principal components for live ultrasound scan traits of Angus cattle. *Animal Science* **2005**, *81*, 337–345. <https://doi.org/10.1079/ASC50850337>.
2. Rosa, G.J.; Valente, B.D.; de los Campos, G.; Wu, X.L.; Gianola, D.; Silva, M.A. Inferring causal phenotype networks using structural equation models. *Genetics Selection Evolution* **2011**, *43*, 6. <https://doi.org/10.1186/1297-9686-43-6>.
3. Gianola, D.; Sorensen, D. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics* **2004**, *167*, 1407–1424. <https://doi.org/10.1534/genetics.103.025734>.
4. De los Campos, G.; Gianola, D.; Heringstad, B. A structural equation model for describing relationships between somatic cell score and milk yield in first-lactation dairy cows. *Journal of Dairy Science* **2006**, *89*, 4445–4455. [https://doi.org/10.3168/jds.S0022-0302\(06\)72493-6](https://doi.org/10.3168/jds.S0022-0302(06)72493-6).
5. Rosa, G.J.; Felipe, V.P.; Peñagaricano, F. Applications of graphical models in quantitative genetics and genomics. In *Systems Biology in Animal Production and Health, Vol. 1*; Springer, 2016; pp. 95–116.
6. Inoue, K. Application of Bayesian causal inference and structural equation model to animal breeding. *Animal Science Journal* **2020**, *91*, e13359.
7. Peñagaricano, F.; Valente, B.; Steibel, J.; Bates, R.; Ernst, C.; Khatib, H.; Rosa, G. Searching for causal networks involving latent variables in complex traits: Application to growth, carcass, and meat quality traits in pigs. *Journal of Animal Science* **2015**, *93*, 4617–4623. <https://doi.org/10.2527/jas.2015-9213>.
8. Varona, L.; González-Recio, O. Invited review: recursive models in animal breeding: interpretation, limitations, and extensions. *Journal of Dairy Science* **2023**, *106*, 2198–2212.
9. Nuñez, P.; Martínez-Boggio, G.; Casellas, J.; Varona, L.; Peñagaricano, F.; Ibáñez-Escriche, N. Applying recursive modelling to assess the role of the host genome and the gut microbiome on feed efficiency in pigs. *Animal* **2025**, *19*, 101453. <https://doi.org/10.1016/j.animal.2025.101453>.
10. Henderson, C.; Quaas, R. Multiple trait evaluation using relatives' records. *Journal of Animal Science* **1976**, *43*, 1188–1197.
11. Valente, B.D.; Rosa, G.J.; de Los Campos, G.; Gianola, D.; Silva, M.A. Searching for recursive causal structures in multivariate quantitative genetics mixed models. *Genetics* **2010**, *185*, 633–644.
12. Inoue, K.; Valente, B.; Shoji, N.; Honda, T.; Oyama, K.; Rosa, G. Inferring phenotypic causal structures among meat quality traits and the application of a structural equation model in Japanese Black cattle. *Journal of Animal Science* **2016**, *94*, 4133–4142.
13. Bello, N.M.; Ferreira, V.C.; Gianola, D.; Rosa, G.J. Conceptual framework for investigating causal effects from observational data in livestock. *Journal of Animal Science* **2018**, *96*, 4045–4062.
14. Jamrozik, J.; Johnston, J.; Sullivan, P.; Miglior, F. Recursive model approach to traits defined as ratios: Genetic parameters and breeding values. *Journal of Dairy Science* **2017**, *100*, 3767–3772.
15. Jöreskog, K.G. A general method for analysis of covariance structures. *Biometrika* **1970**, *57*, 239–251.
16. Meyer, K. Factor-analytic models for genotype × environment type problems and structured covariance matrices. *Genetics Selection Evolution* **2009**, *41*, 21. <https://doi.org/10.1186/1297-9686-41-21>.
17. de Los Campos, G.; Gianola, D. Factor analysis models for structuring covariance matrices of additive genetic effects: a Bayesian implementation. *Genetics Selection Evolution* **2007**, *39*, 481.
18. National Research Council (US) Committee for the Update of the Guide for the Care and Use of Laboratory Animals. *Guide for the Care and Use of Laboratory Animals*, eighth edition ed.;

- <https://www.nationalacademies.org/publications/5140>: Washington, DC, 2011. <https://doi.org/10.17226/12910>.
19. Meškinytė, E.; Jukna, V.; Zigmantaitė, V.; Ilina, O.; Kučinskas, A. The Effectiveness of the Use of Ultrasound Methodology (Applied to Live Animals) to Assess the Quality of Meat. *Animals* **2025**, *15*. <https://doi.org/10.3390/ani15060872>.
  20. (BIF), B. *Guidelines for Uniform Beef Improvement Programs*. Athens, GA, USA: Univ. of Georgia, 8th ed. ed., 2002. Available at: <https://www.mertolenga.com/BIF%20Guidelines%20Eighth%20Edition.PDF>.
  21. Wu, X.L.; Heringstad, B.; Chang, Y.M.; de los Campos, G.; Gianola, D. Inferring Relationships Between Somatic Cell Score and Milk Yield Using Simultaneous and Recursive Models. *Journal of Dairy Science* **2007**, *90*, 3508–3521. <https://doi.org/https://doi.org/10.3168/jds.2006-762>.
  22. Lawley, D.N.; Maxwell, A.E. *Factor Analysis as a Statistical Tool*, 2nd ed.; Butterworths: London, 1971; p. 448.
  23. Mardia, K.V.; Kent, J.T.; Bibby, J.M. *Multivariate Analysis*; Academic Press: London, 1988; p. 521.
  24. de los Campos, G. *MTM: Multivariate Gaussian Mixed Effects Model*, 2014. R package version 1.0.0.
  25. Team, R.C. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014, 2019.
  26. Gelman, A.; Rubin, D.B. Inference from iterative simulation using multiple sequences. *Statistical Science* **1992**, *7*, 457–472.
  27. Smith, B.J. boa: an R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software* **2007**, *21*, 1–37.
  28. Spiegelhalter, D.J.; Best, N.G.; Carlin, B.P.; Van Der Linde, A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **2002**, *64*, 583–639.
  29. Spiegelhalter, D.J.; Best, N.G.; Carlin, B.P.; Linde, A. The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **2014**, *76*, 485–493.
  30. Kirkpatrick, M.; Meyer, K. Direct estimation of genetic principal components: simplified analysis of complex phenotypes. *Genetics* **2004**, *168*, 2295–2306. <https://doi.org/10.1534/genetics.104.029181>.
  31. Smith, A.; Cullis, B.; Thompson, R. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* **2001**, *57*, 1138–1147. <https://doi.org/10.1111/j.0006-341X.2001.01138.x>.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.