**Genetics and Reproduction**

**Review article**

# Some Notes on Genetic Evaluation of Purebred and Crossbred Animals with Emphasis on the Tropics

Alberto Menéndez-Buxadera *, Marco A. Suárez Tronco **

*Independent advisor, Cuba.
**Agrarian University of Havana and Center for Tropical Animal Improvement Research (CIMAGT), Havana, Cuba
Correspondence: marcosuareztronco@gmail.com

Received: July 2024; Accepted: August 2024; Published: September, 2024.

## ABSTRACT

**Aim.** To present some reflections and practical evidence related to the estimation of Genetic Value (GV) in pure and crossbred animals, with special emphasis on its possible use in the genetic evaluation program of Cuba. **Development:** Genetic evaluation of animals is a common practice in any improvement program for which different statistical models are applied. In Cuba, Genetic Value (GV) is generally estimated using an additive effects model, which does not fit the type of crossbred animal between B. taurus (TT) and B. indicus (ZZ), where dominance and epistasis genetic effects also manifest, so the current results may be biased. Reviewing the available references indicates that the GV estimated with this approach has little predictive capacity in the different crosses between TT and ZZ, or in other words, there is genotype-environment interaction. This document shows the basic characteristics of five types of statistical models applied to estimate the GV for dairy cattle, indicating their properties and risks. **Conclusions:** The use of longitudinal models through random regression, although more cumbersome in statistical terms, provides additional information on the general and specific combining ability of the sires, which can have important benefits under Cuban livestock conditions.

**Keywords:** Additive genetic effects, dominance, statistical models, random regression, genetic value (Source: *AIMS)*

## INTRODUCTION

The improvement program carried out in Cuba (Prada 1984) was based on the introduction of specialized genotypes of the B. taurus (TT) type, basically the Holstein breed, which would be used on native B. indicus (ZZ) Zebu females with the purpose of developing an animal with higher dairy potential and adapted to the country's environmental conditions. In this regard, there are databases of the results of these crosses, whose phenotypic expressions reflect additive and non-additive genetic effects; therefore, other statistical models different from the current ones should be used for their genetic evaluation (Hernández, 2019). The aim of this article is to present some reflections

**Citations (APA) Some Notes on Genetic Evaluation of Purebred and Crossbred Animals with Emphasis on the Tropics** (2024). *Journal of Animal Prod.*, *36*(2). https://apm.reduc.edu.cu/index.php/rpa/article/view/e151

and practical evidence related to the estimation of Genetic Value (GV) in pure and crossbred animals, with special emphasis on its possible use in the genetic evaluation program of Cuba.

# DEVELOPMENT

## Genetic Components in Purebred and Crossbreeding

Currently, there is a question about the correlation between the Genetic Value (GV) of purebred and crossbred animals ($\mathbf{r_{pc}}$). According to Vitezica et al. (2016), rpc results are far from unity, meaning they are an unreliable indicator of the behavior of the crossbred population, due, among other factors, to the existence of non-additive genetic effects that can cause differences in GV.

The genotype ($\mathbf{G_i}$) of an individual represents an "aggregate genetic effect," resulting from the action and interaction of countless genes that act individually or in conjunction with other genes or groups of genes. The components of $\mathbf{G_i}$ can be defined as:

**ADDITIVE EFFECT:** It is the effect of a single gene acting independently of the rest of the genotype. Its manifestation should be evaluated as Individual Effect ($\mathbf{g^i}$); Maternal ($\mathbf{g^m}$); maternal grandmother ($\mathbf{g^{mg}}$) and Paternal ($\mathbf{g^p}$).

**DOMINANCE EFFECT:** It is the effect due to the action of a pair of genes within a locus. Similarly, this can manifest at the Individual level ($\mathbf{d^i}$); Maternal ($\mathbf{d^m}$) and Paternal ($\mathbf{d^p}$).

**EPISTATIC EFFECT:** It is the effect due to the joint action of two or more genes at two or more loci. This influence can also manifest at the Individual level ($\mathbf{gg^i}$); Maternal ($\mathbf{gg^m}$) and Paternal ($\mathbf{gg^p}$). Generally, these effects are symbolized as $\mathbf{I_{kk}}$.

Figure 1 exemplifies these effects in practical terms, it shows the milk production results of various genotypes between **ZZ** and **TT** based on an extensive bibliographic compilation presented by Rege (1998).

The extreme values of the figure represent the two pure breeds (**TT** and **ZZ**), whose average genetic components can be represented as:

$$\mathbf{G_{ZZ} = \underbrace{g^i_{ZZ} + gm_{ZZ} + g^{mg}_{ZZ} + g^p_{ZZ}}\ +\ \underbrace{d^i_{ZZ} + d^m_{ZZ} + d^p_{ZZ} + I_{ZZ}}} \qquad \textbf{(1)}$$

**ADDITIVITY.....NO ADDITIVITY**

$$\mathbf{G_{TT} = \overbrace{g^i_{TT} + g^m_{TT} + g^{mg}_{TT} + g^p_{TT}}\ +\ \overbrace{d^i_{TT} + d^m_{TT} + d^p_{TT} + I_{TT}}} \qquad \textbf{(2)}$$

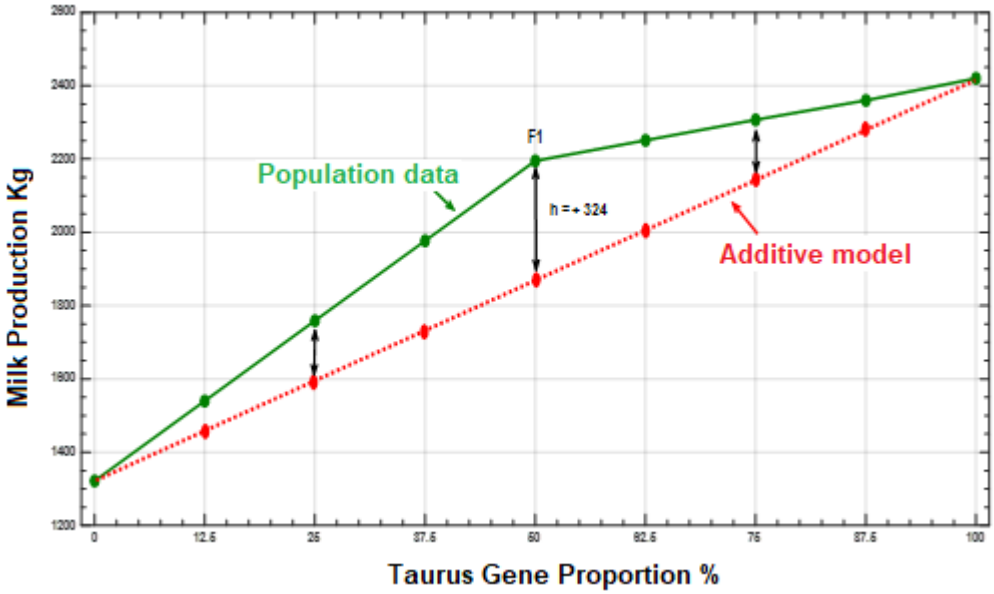


**Figure 1. Milk Production of Crossbred B. taurus and B. indicus Animals**

If we consider that the effect of genes is fundamentally **ADDITIVE**, the behavior of each cross between **TT** and **ZZ** would be situated on the dotted line of Figure 1, where it is evident that a constant amount increases as the proportion of **TT** genes increases. Under such a condition, an **ADDITIVE MODEL** would be the most recommended to describe these different crosses; however, in practical conditions, the hereditary basis is not so simple, and the results of the different genotypes between **TT** and **ZZ** deviate significantly from the purely additive model. For example, the expected value of the first crossbred generation should be $\mathbf{E(G_{TZ}) = 0.5\ (G_{TT} + G_{ZZ})}$, given that this genotype is composed of 50% genes from each parent. However, its performance was much higher than expected (continuous line in Figure 1) due to the existence of other **NON-ADDITIVE** genetic effects.

While the selection process in purebred increases homozygosity (either by functionally identical genes or by a common origin), crossbreeding increases heterozygosity, manifesting in its maximum expression in $\mathbf{F_1}$ since the alleles of each locus come from different breeds. The term heterosis (h) has been introduced to denote the superiority of F1 performance over the average of its ancestors, as can be seen in the figure for the $\mathbf{G_{TZ}}$ genotype. In general, **h** can be estimated as follows:

$$\mathbf{h = G_{TZ} - \frac{(G_{TT} + G_{ZZ})}{2}}$$

From the above, it can be inferred that **h** is the joint manifestation of **DOMINANCE** and **EPISTASIS** genetic effects, so by adequately comparing several types of crossbred animals, it is possible to estimate the importance, magnitude, and genetic origin of the differences in productive behavior of different crossbred genotypes. In this approach, it is considered that $\mathbf{G_{TZ} = G_{ZT}}$, which should receive more attention due to the maternal effects of **ZZ** that can be very different from those of **TT**, such as the duration of lactation.

During the selection process in purebred, not only the frequency of genes with an additive effect ($\mathbf{g^i}$) increases, but also the different epistatic combinations (see models 1 and 2) among non-allelic genes increase. Such favorable gene combinations in F1 are not entirely transmitted from parents to offspring but are reduced during the random segregation process of genes present in the gametes of crossbred progenies. In this way, new gene combinations can be produced that are not present in the parents' generation.

To exemplify these concepts, Figure 2 was created, representing only a pair of alleles at two different loci for the B. taurus ($\mathbf{T_iT_i}$) and *B. indicus* ($\mathbf{Z_iZ_i}$) genotypes.

It is evident that regardless of the random segregation of the gametes of the parental genotypes **TiTi** and **ZiZi**, they will only produce $\mathbf{T_1T_2}$ y $\mathbf{Z_1Z_2}$ alleles, thus the $\mathbf{F_1}$ will have all its genes from two different breeds, i.e., a heterozygosity **H** = 100%, in which case the heterosis is maximum (**h**=100%). At the bottom of the figure, the possible segregation of $\mathbf{F_1}$ individuals' gametes is represented, where new gene combinations not present in the parents' generation are produced, known as gene recombination losses (**r**), with the most significant ones highlighted in the figure. The fact that they are called '*losses*' does not necessarily imply negative economic effects.

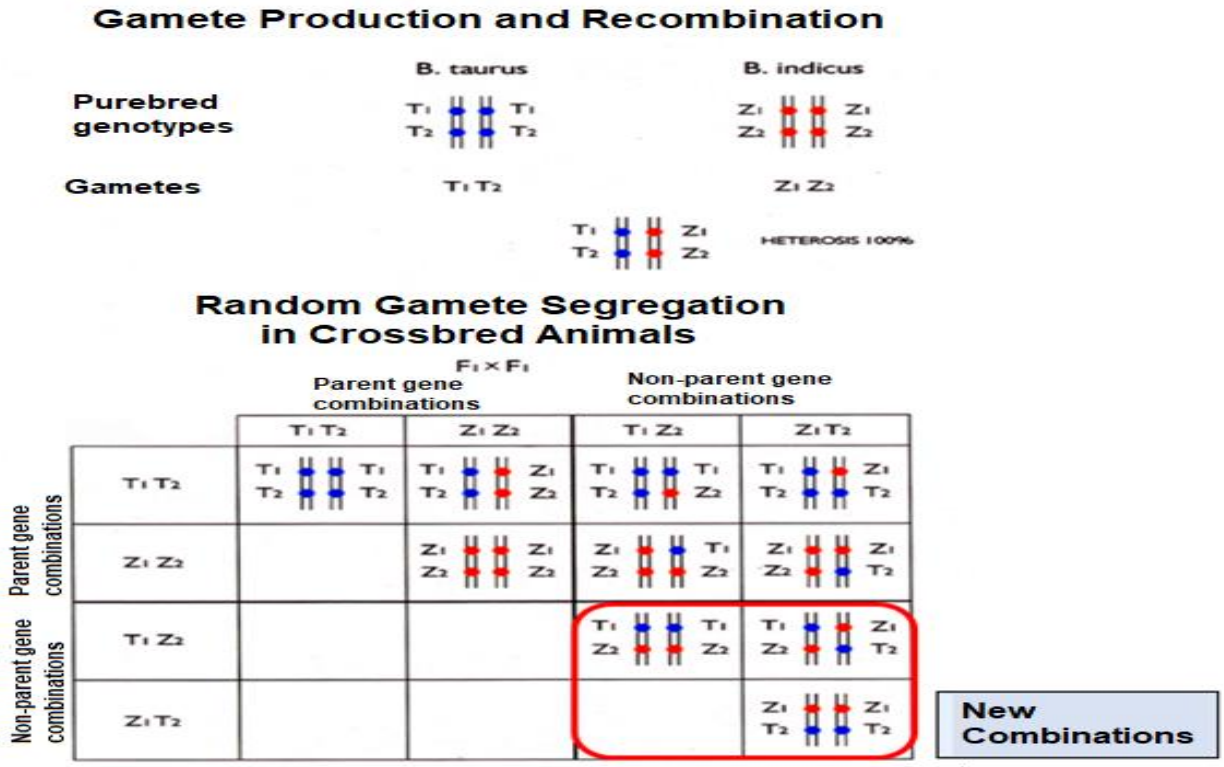


**Figure 2. Production and Recombination of Gametes**

The effects of gene recombination losses represent the limiting factor in the development of new dairy genotypes under tropical conditions (Rutledge, 2001).

**"*The efforts of 100 years of work to produce a dairy cow adapted to the tropics, through the crossing of specialized breeds and Zebu females, have failed, not due to lack of effort, tenacity, or methodological problems, but due to the recombination of incompatible genetic systems, which place a heavy burden on subsequent generations produced by these crossbred animals*."**

In these conclusions, there are several questions that have not been fully examined, particularly in the methods of genetic evaluation in purebred and crossbred animals.

## Estimation of Genetic Value in Purebred TT and Crossbred with ZZ

It is valid to assume that the phenotypic values of each individual are the result (and are controlled) of the action and interaction of a large number of random variables that respond to a polygenic inheritance with an infinite number of genes called minor effect genes that interact with each other and the environment, expressing themselves at different phenotypic levels. In this context, it is almost impossible to evaluate the effects of all loci linked to a particular trait. In such cases, the recommended procedure is to estimate the total cumulative additive effect by estimating the **GV** of the animal producing the record.

It is known that selection in purebred generally does not maximize performance in crossbred animals (Wei and Van der Werf, 1994) and in tropical countries, this can have greater implications as the **GV** results in purebred per se (estimated in the exporting country) have little predictive capacity in **TT x ZZ** crossbreeding in the importing country. In other words, it implies the existence of a double Genotype Environment Interaction, the evidence of which has been reviewed by Menéndez-Buxadera and Mandonnet (2006).

Table 1 shows the results of the effects of **h** and **r** in dairy animals according to several available articles and the first result under Cuban dairy farming conditions.

**Table 1. Estimates of Heterosis and Gene Recombination in Milk Production (kg).**

| Source | Heterosis | Recombination |
|---|---|---|
| Rege (1998); Rutledge (2001) | +258 | -277 |
| Other literature data | +268 | -240 |
| Own estimate of Cuba | +235 | -214 |

In general, trends coincide in indicating that what is gained by **h** is lost by **r**, and this was the argument used by Rutledge (2001) to explain the few successes of dairy cattle crossbreeding programs in the tropics mentioned previously. However, it is necessary to identify new methods to mitigate such contrasting effects. The data of purebred and crossbred animals represented in Figure 1 are not directly comparable in absolute terms as they are manifestations of additive genetic effects (**p**), heterosis (**h**), and gene recombination losses (**r**), whose coefficients can and should be estimated using the classic formula for this type of study:

$$p = 0.5\,(p_s + p_m)\;;\quad h = [\,p_s * (1 - p_m) + p_m * (1 - p_s)\,]\quad \text{y}\quad r = [\,p_s * (1 - p_s) + p_m * (1 - p_m)\,]$$

Where: **p** is the proportion of **TT** breed genes in the father (**s**) and mother (**m**) of each animal. These coefficients are presented in Table 2, for both the crosses in Figure 1 and the systems for creating new breeds that have been applied in Cuba.

**Table 2. Average Genetic Coefficients of Genotypes Represented in the Cuban Crossbreeding Program (expressed as a deviation from TT)**

| GENOTYPE | | | Genetic coefficient | | |
|---|---|---|---|---|---|
| **Father** | **Mother** | **Offspring** | **Activity** | **Dominance** | **Recombination** |
| ZZ | ZZ | ZZ | 0 | 0 | 0 |
| ZZ | TZ | RZ | .25 | .5 | .25 |
| ZZ | TT | ZT | .5 | 1 | 0 |
| TT | TZ | RT1 | .75 | .5 | .25 |
| TT | RT1 | RT2 | .875 | .5 | .1875 |
| TT | RT2 | RT3 | .9375 | .125 | .1875 |
| TT | TT | TT | 1 | 0 | 0 |
| ***Average crosses*** | | | ***0.616*** | ***0.339*** | ***0.125*** |
| TT | RZ | SS1 | .625 | .75 | .1875 |
| SS1 | SS1 | SS2 | .625 | .469 | .469 |
| SS2 | SS2 | SS3 | .625 | .469 | .469 |
| TT | TZ | MM1 | .75 | .5 | .25 |
| MM1 | MM1 | MM2 | .75 | .375 | .375 |
| MM2 | MM2 | MM3 | .75 | .375 | .375 |
| ***Average new breed*** | | | ***0.687*** | ***0.489*** | ***0.354*** |

***The calculation procedure appears in the text. Symbols mean ZZ= B. indicus; TT= B. taurus; RZ= Backcross to ZZ; RT= Backcross to TT, $RT_1$ and $RT_2$ mean absorption crosses towards TT. $SS_1$, $SS_2$, and $SS_3$ refer to the first three generations of the new breed cross called Siboney de Cuba, the same for $MM_1$, $MM_2$, and $MM_3$ named Mambí de Cuba.**

In populations of purebred and crossbred animals, significant variation is evident in their phenotypic values, which depend not only on the environment where they are exploited but also because they maintain a certain degree of relatedness and therefore share genes in common in a proportion depending on their genetic composition (Table 2). This effect can be more evident when using the relationship matrix in estimating the **VG**, as generally more information is available on **TT** ancestors compared to **ZZ**. The causes of variation in that population can be estimated in their genetic components of additivity ($\boldsymbol{\sigma^2_A}$), dominance ($\boldsymbol{\sigma^2_D}$), and epistasis ($\boldsymbol{\sigma^2_I}$, as well as environmental ($\boldsymbol{\sigma^2_E}$), so that the total phenotypic variation is $\boldsymbol{\sigma^2_P = \sigma^2_A + \sigma^2_D + \sigma^2_I + \sigma^2_E}$, from which two basic parameters can be created:

- Narrow-sense heritability as $\boldsymbol{h^2 = \frac{\sigma^2_A}{\sigma^2_A + \sigma^2_D + \sigma^2_I + \sigma^2_E}}$ which explains the proportion of the differences between phenotypic values due to additive genetic effects.
- Broad-sense heritability $H^2 = \frac{\sigma^2_A + \sigma^2_D + \sigma^2_I}{\sigma^2_A + \sigma^2_D + \sigma^2_I + \sigma^2_E}$ which refers to the differences between phenotypic values due to total genetic effects.

These parameters $\mathbf{h^2}$ and $\mathbf{H^2}$ are applicable only to the population and time period in which they were estimated. There is very beneficial evidence in different animal production scenarios,

demonstrating that sustained selection over a period of time according to the **GV** of animals for a single trait (or several based on an index) will increase the frequency of such genes with an additive effect in an amount depending on the selection intensity, existing genetic variance, and the accuracy of **GV** estimation.

**Statistical Models for Data of Purebred and Crossbred Animals**

In general terms, there are two approaches to study the behavior of different types of purebred and crossbred animals:

- Univariate Animal Model (Van der Werf, 1990): In which the **GV** is estimated by incorporating the proportion of genes from different breeds as a fixed genetic group effect or as a covariate.
- **Multiracial Animal Model** (Arnold et al., 1992): Which estimates the **GV** for additive and non-additive effects.

A representation of these models is as follows:

- Additive Multiracial Univariate Model $\mathbf{Y = X\beta + \begin{pmatrix} Q_{fa} \\ b_{fa} \end{pmatrix} + \mu_a + e_{ij}}$
- Non-Additive Multiracial Univariate Model $\mathbf{Y = X\beta + b_{fa} + b_{fh} + b_{fr} + \mu_a + e_{ij}}$
- Multiracial Bivariate Model $\mathbf{Y = X\beta + Z_{u_a} + W_{u_d} + e_{ij}}$

Where:

Y is a vector of observations corresponding to each animal of a given genotype.

**β** is a vector of fixed effects common to all observations (contemporary group, age, etc.).

$\mathbf{b_{fa}}$, $\mathbf{b_{fh}}$, and $\mathbf{b_{fr}}$ are regression coefficients for the proportion of additivity (**fa**); heterosis (**fh**), and gene recombination (**fr**) corresponding to the animal producing the record (Table 2).

$\boldsymbol{\mu_a}$ and $\boldsymbol{\mu_d}$ are vectors of additive ($\boldsymbol{\mu_a}$) and non-additive ($\boldsymbol{\mu_d}$) random genetic effects due to the animal producing the record.

$e_{ij}$ is a random vector of residual effects common to all observations, while in the bivariate multiracial model, it refers to each type of animal (purebred or crossbred).

X, Q, Z, and **W** are incidence matrices to relate the observations to the fixed effects, genetic group, and animal $\mu_a$ and $\mu_d$ with **Y**, respectively.

In the additive and non-additive linear multiracial univariate models (**MMRU**), it is assumed that

$$\mathbf{var(y) \approx N[0, \sigma_y^2 = (G_o \otimes A + \sigma_e^2)]}$$

In this procedure, only the additive genetic effect is exploited, and even when an appropriate statistical procedure has been applied, the **GV** estimates can also be biased. With these results, $\mathbf{h^2}$ can be estimated in both models, where it is assumed that ($\mathbf{h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}}$) is the same for different genotypes. When the genetic group effect $\mathbf{Qf_a}$ is incorporated, it is necessary to have a clear

understanding of its meaning, which reflects the magnitude of the differences between populations subjected to a long selection process. Therefore, these magnitudes should be interpreted as total genetic differences (additive and non-additive genetic effects) between individuals belonging to each genetic group.

In the non-additive animal model, the $\mathbf{Qf_a}$ effect can be replaced by $\mathbf{b_{fa}}$; $\mathbf{b_{fh}}$, and $\mathbf{b_{fr}}$ where $\mathbf{b_i}$ represents the regression of the studied dependent variable on the racial composition of the animal producing the record (see Table 2). Incorporating the random effect of the animal, either with a group effect or as a covariate, improves precision because individual genetic differences are taken into account. Despite this, estimates can be biased as it is assumed *a priori* that genetic (co)variances and the **GV** of animals do not vary along the crossing coefficient trajectory scale shown in Table 2 and included in the model as fixed effects. The magnitude of the bias will depend on the group definition type and will increase in proportion to the genetic distance between both breeds, or in other words, as the level of **h** and **r** increases.

In the multiracial model, the problem of the non-additive genetic component as a random effect has been the critical factor for applying general methods to estimate the genetic merit of purebred ($\boldsymbol{\mu_a}$) and crossbred ($\boldsymbol{\mu_d}$) animals. The works of Rodríguez-Almeida et al. (1997); Newman et *al*. (2002) in beef cattle, and Zumbach et al. (2007) in pigs can be an alternative to the original model by Arnold et al. (1992). These authors consider the results of purebred and crossbred animals as different traits and apply a bivariate multiracial model (**BMRM)**, in which it is assumed:

$$\mathbf{var}(\mathbf{y}) \approx \mathrm{N}[\mathbf{0}, \boldsymbol{\sigma}^2_{\mathbf{y}} = (\mathbf{G_o} = \begin{bmatrix} \boldsymbol{\sigma}^2_{\mathbf{ap}} & \boldsymbol{\sigma}_{\mathbf{apc}} \\ \boldsymbol{\sigma}_{\mathbf{cpa}} & \boldsymbol{\sigma}^2_{\mathbf{ac}} \end{bmatrix} \otimes \mathbf{A}) + \begin{bmatrix} \boldsymbol{\sigma}^2_{\mathbf{ep}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma}^2_{\mathbf{ec}} \end{bmatrix}$$

The genetic variance components $\mathbf{G_o}$ will be calculated in the same way for the classical nonlinear and univariate additive models, **A** is the denominator of the relationship matrix. In the bivariate model, the variance components for purebred ($\boldsymbol{\sigma}^2_{\mathbf{ap}}$) and crossbred ($\boldsymbol{\sigma}^2_{\mathbf{ap}}$) animals, and the covariance between them ($\boldsymbol{\sigma_{apc}} = \boldsymbol{\sigma_{cpa}}$). In this model, $\mathbf{h^2}$ should be considered for each type of animal: $\mathbf{h}^2_{\mathbf{p}} = \frac{\boldsymbol{\sigma}^2_{\mathbf{ap}}}{\boldsymbol{\sigma}^2_{\mathbf{ap}}+\boldsymbol{\sigma}^2_{\mathbf{ep}}}$ for purebred animals and $\mathbf{h}^2_{\mathbf{c}} = \frac{\boldsymbol{\sigma}^2_{\mathbf{ac}}}{\boldsymbol{\sigma}^2_{\mathbf{ac}}+\boldsymbol{\sigma}^2_{\mathbf{ec}}}$ for crossbred animals. Note that the error variances are not the same. On the other hand, the genetic correlation will be $\mathbf{r_{gpc}} = \frac{\boldsymbol{\sigma_{apc}}}{\sqrt{\boldsymbol{\sigma}^2_{\mathbf{ap}} * \boldsymbol{\sigma}^2_{\mathbf{ac}}}}$.

In this **BMRM** model, the **GV** of all animals for additive ($\boldsymbol{\mu_a}$) and non-additive ($\boldsymbol{\mu_d}$) effects from purebred and crossbred animals, respectively, can be obtained, and the total genetic merit is estimated as $\boldsymbol{\mu_T} = \boldsymbol{\mu_a} + \boldsymbol{\mu_d}$.

If the estimated $\mathbf{h^2}$ values are the same for both types of animals and the genetic correlation ($\mathbf{r_{gpc}}$) is equal to 1, it can be inferred that there are no dominance effects. Conversely, if $\mathbf{r_{gpc}}$ is less than one, it implies differences in gene frequency between both parents, thus genetic variances are not the same, and the **GV** of crossbred animals cannot be accurately predicted from purebred results. Some published evidence (Table 3) indicates that this latter case is the rule rather than the exception.

**Table 3. Some Publications on Genetic Parameters in Bivariate Models for Crossing Breeds A (purebred) and B (crossbred)**

| Trait* | Breeds** | $h_A^2$ | $h_B^2$ | $r_{AB}$ | Author |
|---|---|---|---|---|---|
| **Milk RRM model** | **1/2HG-5/8HG** | 0.35 | 0.17 | 0.33 | Pereira Ribeiro *et al*., 2017 (Brasil) |
| | **1/2HG-3/4HG** | 0.35 | 0.15 | 0.41 | |
| | **5/8HG-3/4HG** | 0.17 | 0.15 | 0.88 | |
| **Milk MT model** | **Holst-cross** | 0.12 | 0.29 | 0.63 | Menéndez-Buxadera 2022 (Cuba) |
| **MT model weight 540** | **Cha-5/8Cha** | 0.28 | 0.32 | 0.77 | Menéndez-Buxadera et al., 2022 (Cuba) |

**• RRM: Random Regression Model; MT: Multi-Trait Model. ** Holst = Holstein; Cha = Charolais; HG = Holstein x Gyr.**

In this B**MRM**, the animals in the 'crossbred' group may include different levels of genes from the improved and native breeds, whose genetic coefficients are not the same. Assuming it as a homogeneous fixed effect can be a source of bias. In this regard, the results published by Pereira Ribeiro et al. (2017); Pereira-Ribeiro et al. (2019) and Santos Daltron et al. (2020; 2021) with dairy animals consistently highlight that genetic effects are not constant across different proportions of Holstein and Gyr genes. This evidence indicates a new type of genetic interaction whose importance was previously noted by Martínez et al. (2000) for milk production in this type of animal and by Menéndez-Buxadera and Ayrado (2013) in the fertility of AI sires with Holstein-Zebu crossbred females in Cuba. As noted in Table 3, the best animals for one type of cross may not be the best for another, posing a complex obstacle for the improvement program that requires a meticulous selection process.

The described model can be modified to a multiracial multi-trait model (**MMRM**) if it is considered that the trait measured at each crossbreeding level corresponds to different traits. In this case, solutions for (co)variance components and **GV** for each crossbreeding level included in the genetic group can be obtained. The representation is very similar to the one shown previously:

Multiracial Multi-Trait Model $\mathbf{Y} = \mathbf{X\beta} + \mathbf{ZQ_{ga}} + \mathbf{Z}\boxed{\mathbf{\mu_a} + \mathbf{WS_{gd}}} + \mathbf{W}\boxed{\mathbf{T_{u_d}} + \mathbf{e_{ij:ga}}}$

Additive fixed random component

Non additive fixed and random component

In this new variant, $\boldsymbol{\mu_a}$ and $\boldsymbol{\mu_d}$ depend on a fixed component and a random one, allowing the results to be applied to each combination of genes from the breeds present in the animal producing the record:

$\boldsymbol{\mu_{adi}} = \mathbf{Q_{ga}} + \boldsymbol{\mu_a}$ for additive **GV**.

$\boldsymbol{\mu_{dom}} = \mathbf{S_{gd}} + \mathbf{T_{u_d}}$ for non additive **GV**.

In this case, Q is an incidence matrix that relates the racial group effect to the animal's racial composition ($\mathbf{g_a}$), while $\boldsymbol{\mu_a}$ is the vector of additive genetic effects. $\mathbf{S_{gd}}$ is the matrix that links the vector of fixed dominance effects (**gd**) of the parents with the animal producing the record, and **T**

is the incidence matrix relating the vector of random dominance effects $\boldsymbol{\mu_d}$ with the vector of observations. These **gd** effects are due to the specific cross between the father's and mother's breeds, while the random effects of $\boldsymbol{\mu_d}$ represent the deviation of the animal's record from the average fixed effect of **gd**. This way, the specific contribution of both parents' breeds in the manifestation of **gd** in the progeny is taken into account. In this model, total genetic merit is estimated similarly $\boldsymbol{\mu_t = \mu_{adi} + \mu_{dom}}$.

Estimating these **gd** effects is a very challenging task because it requires a well-connected data structure and a balanced representation of the breeds involved in the program. This is why Pollack and Quaas (2005) have stated their well-known conclusion that under current circumstances, models that assume fixed **gd** effects may be the most recommended. This is the current trend most used in beef cattle, pigs, and poultry.

In the previously indicated **MRMM** model, **W** is a matrix with one on the diagonal corresponding to the crossbred animal, so that $\boldsymbol{WS_{gd} = S_{gd}}$ and $\boldsymbol{WT\mu_d = T\mu_d}$. This strategy has enabled estimating $\boldsymbol{\mu_{adi}}$, which contains the additive genetic effects from purebred effects and the proportion of the same breed of the crossbred ancestor and represents the general combining ability (***gca***) or genetic merit of the animal at each level of the present breed combinations, while $\boldsymbol{\mu_{dom}}$ represents the genetic merit for specific combining ability (***sga***) of one breed with another. This way, results on the effect of different breed proportions on the animal's additive genetic merit are obtained, as presented by Martínez et al. (2000); Pereira-Ribeiro *et al*. (2019) for different crosses between Gir and Holstein.

The characteristics of the $\boldsymbol{Q_{ga}}$ and $\boldsymbol{S_{gd}}$ matrices allow the application of multiracial random regression models (**MRRM**) suggested by Stranden and Mantysaary (2013), which have been applied in various animal genetics studies (Vanderick et al., 2017; Pereira-Ribeiro et al., 2017; Pereira-Ribeiro et al., 2019). In this **MRRM** model, estimates of (co)variance and genetic values are obtained along the proportion of genes from each breed, even in those animals with racial combinations not present in the analyzed data. The general representation of this model is,

Multiracial Random Regression Model

$$\mathbf{Y} = \mathbf{X\beta} + \sum_{r=0}^{1} \boldsymbol{\Phi}_{ra}\,\boldsymbol{\beta}_1\mathbf{Q}_{ga} + \sum_{r=0}^{1} \boldsymbol{\Phi}_{r}\boldsymbol{\lambda}_{ga}\boldsymbol{\mu}_{a} + \sum_{r=0}^{1} \boldsymbol{\Phi}_{rd}\,\boldsymbol{\beta}_2\mathbf{S}_{gd} +$$

Additive fixed random component

Non additive fixed random component

In this case, the $\boldsymbol{Q_{ga}}$ and $\boldsymbol{S_{gd}}$ matrices have been replaced by Legendre polynomial coefficients $\boldsymbol{\Phi}$ of order r = 1. The number of levels for $\boldsymbol{\Phi_{ra} = \Phi_{rd}}$ depends on the data distribution of each cross and their additive and dominance coefficients (Table 2). The terms $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ are the fixed regression coefficients of the additive ($\boldsymbol{Q_{ga}}$) and dominance ($\boldsymbol{S_{gd}}$) racial composition (see Table 2) on the dependent variable. On the other hand, $\boldsymbol{\lambda_{ga}}$ and $\boldsymbol{\lambda_{gd}}$ represent the random regression matrices for additive ($\boldsymbol{\mu_{adi}}$) and dominance ($\boldsymbol{\mu_{dom}}$) genetic effects, respectively, expressed in terms of genetic functions ($\boldsymbol{fg_{adi}}$ and $\boldsymbol{fg_{dom}}$) for these effects in each animal, whose elements (intercept and slope)

depend on the adjustment order **r**. The residual variance $\mathbf{eij_{:ga}}$ is considered heterogeneous for each racial group (**ga**). The expected (co)variance matrix Go of this MRRM model is:

$$\mathbf{var(y) = G_o = [\Phi_{ra}(\lambda_{ga} \otimes A)\Phi'_{ra} + \Phi_{rd}(\lambda_{gd} \otimes A)\Phi'_{rd}] + \sigma^2_{ej:ga}}$$

Assuming that the genetic groups belong to an absorption crossbreeding between **ZZ** and **TT** and data with five cross levels, Table 4 presents the variance estimates for each group, which can be used to estimate heritability for the additive and dominance components.

**Table 4. Procedure for Estimating Additive and Dominance Variances in 5 Crossbreeding Groups, According to a Random Regression Model of Order r = 1**

| Cross* | Polynomial | | Genetic variance | |
|---|---|---|---|---|
| | Intercept | Slope | Additive | dominance |
| $G_1$=0.00TT | $\Phi_{r1}$ = 0.70711 - 1.22474 | | $\sigma^2_{a1} = \Phi_{r1}(\lambda_{ga})\Phi'_{r1}$ | $\sigma^2_{d1} = \Phi_{r1}(\lambda_{gd})\Phi'_{r1}$ |
| $G_2$=0.25TT | $\Phi_{r2}$ = 0.70711 - 0.61237 | | $\sigma^2_{a2} = \Phi_{r2}(\lambda_{ga})\Phi'_{r2}$ | $\sigma^2_{d2} = \Phi_{r2}(\lambda_{gd})\Phi'_{r2}$ |
| $G_3$=0.50TT | $\Phi_{r3}$ = 0.70711 - 0.00000 | | $\sigma^2_{a3} = \Phi_{r3}(\lambda_{ga})\Phi'_{r3}$ | $\sigma^2_{d3} = \Phi_{r3}(\lambda_{gd})\Phi'_{r3}$ |
| $G_4$=0.75TT | $\Phi_{r4}$ = 0.70711 + 0.61237 | | $\sigma^2_{a4} = \Phi_{r4}(\lambda_{ga})\Phi'_{r4}$ | $\sigma^2_{d4} = \Phi_{r4}(\lambda_{gd})\Phi'_{r4}$ |
| $G_5$=1.00TT | $\Phi_{r5}$ = 0.70711 + 1.22474 | | $\sigma^2_{a5} = \Phi_{r5}(\lambda_{ga})\Phi'_{r5}$ | $\sigma^2_{d5} = \Phi_{r5}(\lambda_{gd})\Phi'_{r5}$ |

***The numerical value refers to the percentage of TT genes.**

As an example of parameter estimates for animals for group $\mathbf{G_1}$:

$$\mathbf{h^2_{adi,1}} = \frac{\sigma^2_{a1}}{\sigma^2_{a1}+\sigma^2_{d1}+\sigma^2_{ega1}} \text{ y } \mathbf{h^2_{dom,1}} = \frac{\sigma^2_{d1}}{\sigma^2_{a1}+\sigma^2_{d1}+\sigma^2_{ega1}}$$

For the other groups, it is similar, only the corresponding variances change. This procedure allows estimating the additive genetic ($\mathbf{r_{ga}}$) and dominance ($\mathbf{r_{gd}}$) correlations between any of the groups. For example, for groups $\mathbf{G_1}$ and $\mathbf{G_5}$:

$$\mathbf{r_{ga1,5}} = \frac{\Phi_{r1}(\lambda_{ga})\Phi'_{r5}}{\sqrt{\Phi_{r1}(\lambda_{ga})\Phi'_{r1} * \Phi_{r5}(\lambda_{ga})\Phi'_{r5}}} \text{ y } \mathbf{r_{gd1,5}} = \frac{\Phi_{r1}(\lambda_{gd})\Phi'_{r5}}{\sqrt{\Phi_{r1}(\lambda_{gd})\Phi'_{r1} * \Phi_{r5}(\lambda_{gd})\Phi'_{r5}}}$$

Finally, the genetic values of additive and dominance effects for each $\mathbf{i^{th}}$ animal, for example, for group $\mathbf{G_1}$, will be estimated as:

$$\mathbf{\mu^i_{adi,1} = fg^i_{adi} * \Phi'_{r1}} \text{ y } \mathbf{\mu^i_{dom,1} = fg^i_{dom} * \Phi'_{r1}}$$

The procedure is the same for each group, only the corresponding polynomial coefficients change. Finally, the Total Genetic Value ($\mu_T$) for $\mathbf{G_1}$ will be the next $\mu_{T1} = \mu^i_{adi,1} + \mu^i_{dom,1}$ that corresponds to a **ZZ** animal, whereas **TT** will be $\mu_{T5} = \mu^i_{adi,5} + \mu^i_{dom,5}$. Note that with a single model, **VG** can be estimated for all types of purebred and crossbred animals.

**Generalizing the Results**

The approach presented in this document allows examining the same problem 'the genetic evaluation of purebred and crossbred animals' from various angles that differ in their starting assumptions. As a very general summary, Table 5 shows some characteristics of the presented models, which must be contrasted with the situation of the analyzed data.

If the available data correspond to the genetic groups presented in Table 4, it is preferred to use the Random Regression models, which synthesize all that information into five estimates of additive and non-additive genetic merit so that the relative importance of each component requires additional reflection.

**Table 5. Some Assumptions and Potential Risks of the Presented Models**

| Model | Assumption | Risk |
|---|---|---|
| **Univariate Multiracial** | Additive genetic variances are equal in different crosses. | Biased results, evidence indicates otherwise. |
| **Bivariate Multiracial** | Data in purebred (p) and crossbred (c) as two different traits and correlation equal to one. | Results may be biased as purebred-crossbred correlations are not equal to one. |
| ***Multivariate Multiracial** | Uses results of each breed as different traits. | Less biased than previous, does not consider all cross levels. Highly sensitive to data structure. |
| ***Random Regression Multiracial** | Nothing assumed a priori. | Least risk, but most complex to execute and interpret. |

***Several estimates of the Genetic Value of each animal need to define how they should be weighted.**

There is not much evidence available for this procedure in tropical conditions, but the results of Menéndez-Buxadera and Ayrado (2013) point in a very positive direction.

In particular, it should be noted that the bulls were evaluated with a high level of precision, with an average of 2,340 and 3,752 inseminations for the Robust and Plastic bulls respectively (Figure 3), which allows for the use of the general and specific combining ability of each bull.

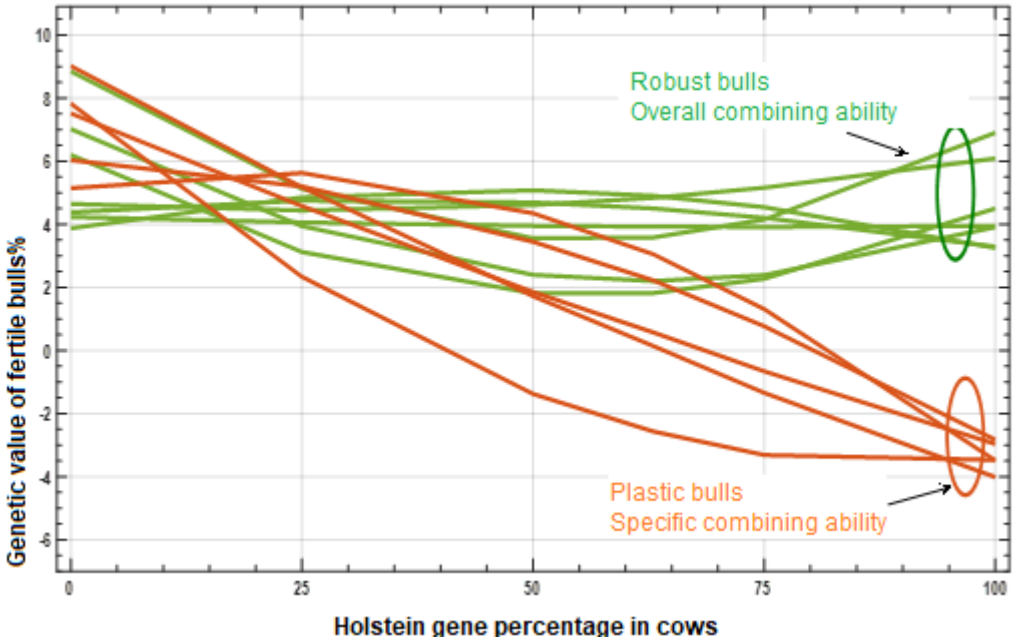


**Figure 3.** Evolution of the Genetic Value of Bulls According to the Racial Proportion of Cows

# CONCLUSIONS

The use of longitudinal models through random regression, although more cumbersome in statistical terms, provides additional information on the general and specific combining ability of the bulls, which can have significant benefits under Cuban livestock conditions.

# REFERENCES

Arnold, J. W., Bertrand, J. K., & Benyshek, L. L. (1992). Animal model for genetic evaluation of multibreed data. *Journal of Animal Science*, *70*(11), 3322-3332. DOI:10.2527/1992.70113322x

Hernández, A. (2019). Selección multicarácter mediante índices de selección en bovinos lecheros. 2019. *Registro CENDA con el*, 1360-4. https://revistaccuba.sld.cu/index.php/revacc/rt/printerFriendly/907/987

Martinez, M. L., Verneque, R. D. S., Teodoro, R. L., Campos, J. D. P., Cruz, M., & Paula, L. R. O. D. (2000). Relações entre as capacidades preditas de transmissão de touros Gir e a produção de leite de suas filhas puras e mestiças. *Revista Brasileira de Zootecnia*, *29*, 692-699. DOI: 10.1590/S1516-35982000000300008

Buxadera, A. M., & Ayrado, Y. (2013). Genetic (co) variance for sire fertility estimated by additive, non-additive and longitudinal models in Holstein–Zebu cross-bred cows. *animal*, *7*(5), 721-728. https://doi.org/10.1017/s1751731112002182

Menéndez-Buxadera A. (2022). Estimación del Valor Genético del vacuno con especial énfasis en condiciones tropicales. VI-Congreso Internacional sobre Mejoramiento Animal, CIMAGT. Palacios de las Convenciones, La Habana, Cuba, 15-18 noviembre 2022, 10 pag. https://www.cubaforevents.com/event/vi-congreso-internacional-sobre-mejoramiento-animal-2022/

Menéndez-Buxadera, A., & Mandonnet, N. (2006). The role and importance of genotype–environment interaction for animal breeding in the tropics. In *Animal Breeding Abstract* (Vol. 74, No. 10, pp. 1-14). https://www.researchgate.net/publication/233865325

Menéndez-Buxadera, A., Rodríguez, M., Mitat, A., Suárez, M. A., & Ramos, F. (2022). Evaluación genética conjunta del peso a 18 meses de edad en las razas Charolais y Chacuba. *Revista investigaciones agropecuarias*, *4*(2), 82-95. https://revistas.up.ac.pa/index.php/investigaciones_agropecuarias/article/view/2930

Newman, S., Reverter, A., & Johnston, D. J. (2002). Purebred-crossbred performance and genetic evaluation of postweaning growth and carcass traits in Bos indicus× Bos taurus crosses in Australia. *Journal of Animal Science*, *80*(7), 1801-1808. DOI:10.2527/2002.8071801x

Ribeiro, V. M. P., Raidan, F. S. S., Barbosa, A. R., Silva, M. V. G. B., Cardoso, F. F., & Toral, F. L. B. (2019). Multiple trait and random regression models using linear splines for genetic evaluation of multiple breed populations. *Journal of dairy science*, *102*(1), 464-475. https://www.journalofdairyscience.org/article/S0022-0302(18)30993-7/pdf

Ribeiro, V. M. P., Merlo, F. A., Gouveia, G. C., Winkelstroter, L. K., Abreu, L. R. A., Silva, M. V. G. B. D., & Toral, F. L. B. (2017). Genetic analysis of productive and reproductive traits in multiple-breed dairy cattle populations. *Pesquisa Agropecuária Brasileira*, *52*(11), 1109-1117. https://doi.org/10.1590/S0100-204X2017001100017

Prada, N. (1984). Programa nacional de mejoramiento genético vacuno. *Rev ACPA*, *3*, 20-26. https://www.eumed.net/rev/caribe/2013/09/ganado-vacuno.pdf

Rege, J. E. O. (1998). Utilization of exotic germplasm for milk production in the tropics. https://hdl.handle.net/10568/51304

Rodríguez-Almeida, F. A., Van Vleck, L. D., & Gregory, K. E. (1997). Estimation of direct and maternal breed effects for prediction of expected progeny differences for birth and weaning weights in three multibreed populations. *Journal of Animal Science*, *75*(5), 1203-1212. DOI:10.2527/1997.7551203x

Rutledge, J. J. (2001). Greek temples, tropical kine and recombination load. *Livestock Production Science*, *68*(2-3), 171-179. http://tiesmexico.cals.cornell.edu/courses/shortcourse4/minisite/pdf/Rutledge%20tropical%20kine.pdf

Santos Daltro, D., Silva, M. V. G. B. D., Telo da Gama, L., Machado, J. D., Kern, E. L., Campos, G. S., & Cobuci, J. A. (2020). Estimates of genetic and crossbreeding parameters for 305-

day milk yield of Girolando cows. *Italian Journal of Animal Science*, *19*(1), 86-94. http://www.alice.cnptia.embrapa.br/alice/handle/doc/1133496

Santos Daltro, D., Negri, R., & Cobuci, J. A. (2021). Heterosis effects on 305-day milk yield in a Girolando dairy cattle population in different lactation orders. *Livestock Science*, *245*, 104428. DOI:10.1016/j.livsci.2021.104428

Strandén, I., & Mäntysaari, E. A. (2013). Use of random regression model as an alternative for multibreed relationship matrix. *Journal of Animal Breeding and Genetics*, *130*(1), 4-9. DOI:10.1111/jbg.12014

Van der Werf, J. H. (1990). *Models to estimate genetic parameters in crossbred dairy cattle populations under selection*. Wageningen University and Research. https://edepot.wur.nl/202222

Vanderick, S., Gillon, A., Glorieux, G., Mayeres, P., Mota, R. R., & Gengler, N. (2017). Usefulness of multi-breed models in genetic evaluation of direct and maternal calving ease in Holstein and Belgian Blue Walloon purebreds and crossbreds. *Livestock Science*, *198*, 129-137. DOI:10.1016/j.livsci.2017.02.019

Vitezica, Z. G., Varona, L., Elsen, J. M., Misztal, I., Herring, W., & Legarra, A. (2016). Genomic BLUP including additive and dominant variation in purebreds and F1 crossbreds, with an application in pigs. *Genetics Selection Evolution*, *48*, 1-8. https://doi.org/10.1186/s12711-016-0185-1

Wei, M., & van der Werf, J. J. (1995). Genetic correlation and heritabilities for purebred and crossbred performance in poultry egg production traits. *Journal of animal science*, *73*(8), 2220-2226. DOI:10.2527/1995.7382220x

Zumbach, B., Misztal, I., Tsuruta, S., Holl, J., Herring, W., & Long, T. (2007). Genetic correlations between two strains of Durocs and crossbreds from differing production environments for slaughter traits. *Journal of animal science*, *85*(4), 901-908. DOI:10.2527/jas.2006-499

## AUTHOR CONTRIBUTION STATEMENT

Research conception and design: AMB; data analysis and data interpretation: AMB; redaction of the manuscript: AMB, MST.

## CONFLICT OF INTEREST STATEMENT

The authors state there are no conflicts of interest whatsoever.